



Σχολή Εφαρμοσμένων Τεχνών και Βιώσιμου Σχεδιασμού

Περιβαλλοντικός Σχεδιασμός για την Αντιμετώπιση της
Κλιματικής Αλλαγής

Διπλωματική Εργασία

Βέλτιστη κατανομή ατμοσφαιρικών ρύπων στις πόλεις της Στερεάς
Ελλάδας

Ιωάννα Ντζάθα

Επιβλέπων καθηγητής: Νικόλαος Ραγομανίκης

Πάτρα, Ιούλιος 2026

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του/της φοιτητή/φοιτήτριας («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοικτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού. Ο συγγραφέας/δημιουργός διατηρεί το σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.



Βέλτιστη κατανομή ατμοσφαιρικών ρύπων στις πόλεις της Στερεάς Ελλάδας

Ιωάννα Ντζάθα

Επιτροπή Διπλωματικής Εργασίας

Επιβλέπων Καθηγητής:
Νικόλαος Ραψομανίκης
ΣΕΠ, Ε.Α.Π

Συν-Επιβλέπων Καθηγητής:
Δημήτριος Ναλμπάντης
ΣΕΠ, Ε.Α.Π

Πάτρα, Ιούλιος 2026

Ευχαριστίες - Αφιέρωση

Η παρούσα διπλωματική εργασία δεν αποτελεί μόνο το αποτέλεσμα μιας ακαδημαϊκής διαδικασίας. Είναι το αποτύπωμα μιας διαδρομής που είχε κόπο, επιμονή, σιωπηλές μάχες, στιγμές αμφιβολίας και στιγμές βαθιάς προσωπικής υπέρβασης. Κάθε σελίδα της κουβαλά όχι μόνο γνώση και έρευνα, αλλά και χρόνο, προσπάθεια και πίστη σε έναν στόχο που κάποτε έμοιαζε μακρινός.

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κ. Νικόλαο Ραψομανίκη, για την καθοδήγηση, την εμπιστοσύνη και την υποστήριξή του καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Πάνω απ' όλα, ευχαριστώ τους ανθρώπους που στάθηκαν δίπλα μου όχι μόνο στις εύκολες στιγμές, αλλά κυρίως στις δύσκολες. Εκείνους που με στήριζαν όταν κουραζόμουν, που με άκουσαν όταν αμφέβαλλα και που μου θύμισαν την αξία μου όταν εγώ την ξεχνούσα.

Και υπάρχει ένας άνθρωπος που δεν χρειάζεται να κατονομαστεί, γιατί ξέρει. Ξέρει τις σιωπές, τις αγωνίες, τα βλέμματα, τις μικρές νίκες και τις στιγμές που η παρουσία του ήταν αρκετή για να γίνει το βάρος λίγο πιο ελαφρύ. Σε εκείνη τη σταθερή, αθόρυβη δύναμη κρατώ μέσα σε αυτές τις γραμμές ένα ξεχωριστό ευχαριστώ.

Αφιερώνω την εργασία αυτή σε όσους πίστεψαν σε εμένα αθόρυβα, σταθερά και αληθινά. Και, τελικά, την αφιερώνω και στον ίδιο μου τον εαυτό για όλες τις φορές που δεν τα παράτησα.

Περίληψη

Η παρούσα διπλωματική εργασία διερευνά τη στατιστική συμπεριφορά των συγκεντρώσεων αιωρούμενων σωματιδίων PM_{10} σε επιλεγμένες πόλεις της Στερεάς Ελλάδας, με στόχο τον προσδιορισμό της θεωρητικής κατανομής πιθανότητας που αποδίδει καταλληλότερα την εμπειρική μορφή των δεδομένων. Η ανάλυση βασίστηκε σε διαθέσιμες χρονοσειρές PM_{10} από σταθμούς παρακολούθησης της ποιότητας του αέρα και περιέλαβε στάδια ελέγχου πληρότητας, προεπεξεργασίας, περιγραφικής στατιστικής και εφαρμογής θεωρητικών κατανομών.

Στο πλαίσιο της εργασίας εξετάστηκαν η λογαριθμοκανονική κατανομή, η κατανομή Gamma, η κατανομή Weibull και η κανονική κατανομή ως συγκριτική βάση. Η αξιολόγηση της προσαρμογής πραγματοποιήθηκε μέσω της σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων, διαγνωστικών διαγραμμάτων, διαγραμμάτων υπολοίπων, διαγραμμάτων Q-Q και αριθμητικής αποτίμησης με βάση το συγκριτικό στατιστικό χ^2 . Η επιλογή της καταλληλότερης κατανομής δεν βασίστηκε σε ένα μόνο κριτήριο, αλλά στη συνδυαστική αξιολόγηση γραφικών και αριθμητικών ενδείξεων.

Τα αποτελέσματα έδειξαν ότι οι συγκεντρώσεις PM_{10} παρουσιάζουν μη αρνητική και θετικά ασύμμετρη συμπεριφορά, στοιχείο που υποστηρίζει την εξέταση κατανομών κατάλληλων για περιβαλλοντικά δεδομένα με δεξιά ουρά. Η λογαριθμοκανονική κατανομή παρουσίασε την καλύτερη συνολική προσαρμογή στις τρεις από τις τέσσερις αναλύσιμες χρονοσειρές, ενώ στην περίπτωση της Χαλκίδας επιλέχθηκε η κατανομή Weibull. Το εύρημα αυτό δείχνει ότι δεν υπάρχει μία καθολικά βέλτιστη κατανομή για όλες τις περιπτώσεις, αλλά ότι η τελική επιλογή εξαρτάται από τα ιδιαίτερα χαρακτηριστικά κάθε χρονοσειράς.

Συνολικά, η εργασία αναδεικνύει τη σημασία της στατιστικής μοντελοποίησης για την κατανόηση της κατανομικής συμπεριφοράς των συγκεντρώσεων PM_{10} σε περιφερειακές πόλεις και συμβάλλει στη συστηματική αξιοποίηση περιβαλλοντικών δεδομένων ποιότητας του αέρα. Παράλληλα, επισημαίνει την ανάγκη επαρκούς πληρότητας και αξιοπιστίας των χρονοσειρών, καθώς και τη χρησιμότητα μελλοντικών αναλύσεων που θα περιλαμβάνουν περισσότερους ρύπους, μετεωρολογικές παραμέτρους και μεγαλύτερες χρονικές περιόδους.

Λέξεις-Κλειδιά: PM_{10} , ατμοσφαιρική ρύπανση, Στερεά Ελλάδα, θεωρητικές κατανομές, λογαριθμοκανονική κατανομή, στατιστική μοντελοποίηση

Best-Fit Distribution of Air Pollutants in the Cities of Central Greece

Ioanna Ntzatha

Abstract

This diploma thesis investigates the statistical behavior of PM_{10} concentrations in selected cities of Central Greece, with the aim of identifying the theoretical probability distribution that best represents the empirical structure of the data. The analysis was based on available PM_{10} time series from air quality monitoring stations and included data completeness assessment, preprocessing, descriptive statistical analysis, and the application of theoretical probability distributions.

The distributions examined were the Lognormal, Gamma, Weibull, and Normal distributions, with the Normal distribution used as a comparative reference. Goodness of fit was assessed through comparisons of observed and expected frequencies, diagnostic plots, residual plots, Q-Q plots, and numerical evaluation based on the comparative chi-square statistic. The selection of the most appropriate distribution was not based on a single criterion, but on a combined assessment of graphical and numerical evidence.

The results showed that PM_{10} concentrations exhibit non-negative and positively skewed behavior, supporting the examination of distributions suitable for right-tailed environmental data. The Lognormal distribution provided the best overall fit for three of the four analyzable time series, while the Weibull distribution was selected for Chalkida. This finding indicates that no single distribution is universally optimal, as the final selection depends on the specific characteristics of each time series.

Overall, this thesis highlights the importance of statistical modeling for understanding the distributional behavior of PM_{10} concentrations in regional cities and contributes to the systematic use of air quality data. It also emphasizes the need for adequate time-series completeness and reliability, as well as the value of future analyses incorporating additional pollutants, meteorological parameters, and longer observation periods.

Keywords: PM_{10} , air pollution, Central Greece, theoretical distributions, Lognormal distribution, statistical modeling

Περιεχόμενα

Περίληψη.....	v
Abstract.....	vi
Περιεχόμενα.....	vii
Κατάλογος Σχημάτων.....	x
Κατάλογος Πινάκων.....	xv
Συντομογραφίες & Ακρωνύμια.....	xvi
Τα 5 σημαντικότερα σημεία της εργασίας.....	xvii
Συνεισφορά Εργασίας.....	xviii
1. Εισαγωγή.....	1
1.1 Αντικείμενο της εργασίας.....	1
1.2 Σκοπός της εργασίας.....	2
1.3 Συνεισφορά της εργασίας.....	4
1.4 Ερευνητικά Ερωτήματα.....	5
1.5 Δομή της εργασίας.....	6
2. Θεωρητικό υπόβαθρο: ατμοσφαιρική ρύπανση, PM_{10} και στατιστική μοντελοποίηση.....	9
2.1 Ατμοσφαιρική ρύπανση και αιωρούμενα σωματίδια PM_{10}	9
2.1.1 Ορισμός ατμοσφαιρικής ρύπανσης.....	10
2.1.2 Κατηγορίες ατμοσφαιρικών ρύπων.....	12
2.1.3 Αιωρούμενα σωματίδια: PM_{10} και $PM_{2.5}$	14
2.1.4 Φυσικοχημικά χαρακτηριστικά των PM_{10}	16
2.1.5 Επιπτώσεις των PM_{10} στην υγεία και το περιβάλλον.....	19
2.2 Πηγές και παράγοντες μεταβλητότητας των PM_{10}	21
2.2.1 Φυσικές πηγές PM_{10}	22
2.2.2 Ανθρωπογενείς πηγές PM_{10}	23
2.2.3 Πηγές PM_{10} στην Ελλάδα και στη Στερεά Ελλάδα.....	26
2.2.4 Μετεωρολογικοί και τοπογραφικοί παράγοντες.....	27
2.3 Θεσμικό πλαίσιο και παρακολούθηση PM_{10}	29
2.3.1 Ευρωπαϊκά όρια και κατευθυντήριες τιμές.....	29
2.3.2 Μέθοδοι μέτρησης PM_{10}	31
2.3.3 Αβεβαιότητα, πληρότητα και αξιοπιστία δεδομένων.....	33
2.4 Στατιστική μοντελοποίηση περιβαλλοντικών δεδομένων.....	35
2.4.1 Ρόλος των κατανομών πιθανότητας.....	35
2.4.2 Συνάρτηση πυκνότητας πιθανότητας και αθροιστική συνάρτηση κατανομής.....	36
2.4.3 Κανονική κατανομή.....	39
2.4.4 Λογαριθμοκανονική κατανομή.....	40
2.4.5 Κατανομή Gamma.....	42
2.4.6 Κατανομή Weibull.....	43
2.4.7 Αξιολόγηση προσαρμογής και επιλογή βέλτιστης κατανομής.....	45
2.5 Προηγούμενες μελέτες για κατανομές.....	48
3. Μεθοδολογία.....	53
3.1 Πηγή και χαρακτηριστικά των δεδομένων.....	54

3.2 Προεπεξεργασία και καθαρισμός δεδομένων PM_{10}	57
3.3 Κριτήρια επιλογής αναλύσιμων χρονοσειρών.....	58
3.4 Δημιουργία εμπειρικών συχνοτήτων.....	59
3.5 Εφαρμογή θεωρητικών κατανομών	61
3.5.1 Γενική λογική εφαρμογής των θεωρητικών κατανομών.....	62
3.5.2 Λογαριθμοκανονική κατανομή.....	62
3.5.3 Κατανομή Gamma.....	64
3.5.4 Κατανομή Weibull.....	65
3.5.5 Κανονική κατανομή ως συγκριτική βάση.....	66
3.6 Κριτήριο και αξιολόγηση προσαρμογής.....	66
3.7 Γραφική αξιολόγηση της προσαρμογής των θεωρητικών κατανομών.....	68
3.8 Επιλογή καταλληλότερης κατανομής ανά πόλη και έτος.....	69
4. Αποτελέσματα.....	71
4.1 Τελικό αναλυτικό σύνολο δεδομένων.....	71
4.2 Περιγραφική στατιστική των αναλύσιμων χρονοσειρών PM_{10}	74
4.3 Εμπειρικές κατανομές συχνοτήτων των PM_{10}	76
4.4 Προσαρμογή θεωρητικών κατανομών στην Άμφισσα για το έτος 2018.....	81
4.4.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων.....	82
4.4.2 Διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων.....	88
4.4.3 Διαγράμματα υπολοίπων.....	94
4.4.4 Διαγράμματα Q-Q.....	100
4.4.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2	105
4.4.6 Επιλογή καταλληλότερης κατανομής για την Άμφισσα 2018.....	109
4.5 Προσαρμογή θεωρητικών κατανομών στη Λαμία το έτος 2017.....	111
4.5.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων.....	112
4.5.2 Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων για την πόλη της Λαμίας.....	118
4.5.3 Διαγράμματα υπολοίπων.....	125
4.5.4 Διαγράμματα Q-Q.....	133
4.5.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2	140
4.5.6 Επιλογή καταλληλότερης κατανομής για τη Λαμία 2017.....	143
4.6 Προσαρμογή θεωρητικών κατανομών στη Χαλκίδα το 2017.....	146
4.6.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων.....	147
4.6.2 Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων.....	154
4.6.3 Διαγράμματα υπολοίπων.....	162
4.6.4 Διαγράμματα Q-Q.....	169
4.6.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2	177
4.6.6 Επιλογή καταλληλότερης κατανομής για τη Χαλκίδα 2017.....	180
4.7 Προσαρμογή θεωρητικών κατανομών στη Λειβαδιά για το έτος 2019.....	183
4.7.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων.....	184
4.7.2 Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων.....	190
4.7.3 Διαγράμματα υπολοίπων.....	195
4.7.4 Διαγράμματα Q-Q.....	202

4.7.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2	208
4.7.6 Επιλογή καταλληλότερης κατανομής για τη Λειβαδιά το 2019.....	211
4.8 Περίπτωση Καρπενησίου: γραφική απεικόνιση της πληρότητας των δεδομένων.....	213
4.8.1 Έλεγχος πληρότητας των δεδομένων Καρπενησίου.....	215
4.8.2 Γραφική απεικόνιση της πληρότητας ανά έτος.....	217
4.8.3 Μεθοδολογική αιτιολόγηση αποκλεισμού.....	219
4.9 Συγκριτική αξιολόγηση των θεωρητικών κατανομών.....	220
4.9.1 Συγκριτική παρουσίαση των τιμών χ^2 και των τελικών επιλογών.....	222
4.9.2 Συνολική ερμηνεία της απόδοσης των κατανομών.....	224
4.9.3 Διαφοροποίηση της Χαλκίδας και σημασία της πόλης.....	227
4.9.4 Συμπέρασμα συγκριτικής αξιολόγησης.....	228
5. Συμπεράσματα.....	232
5.1 Σύνοψη της εργασίας.....	232
5.2 Βασικά ευρήματα της ανάλυσης.....	234
5.3 Απάντηση στα ερευνητικά ερωτήματα.....	236
5.4 Μεθοδολογική αποτίμηση της προσέγγισης.....	239
5.5 Περιορισμοί της εργασίας.....	240
5.6 Προτάσεις για μελλοντική έρευνα.....	242
5.7 Τελικό συμπέρασμα.....	243
Βιβλιογραφία.....	245

Κατάλογος Σχημάτων

Σχήμα 2.1 Εννοιολογική απεικόνιση της σχέσης μεταξύ πηγών εκπομπών, ατμοσφαιρικών ρύπων, διεργασιών μεταφοράς/μετασχηματισμού, έκθεσης και επιπτώσεων.....	10
Σχήμα 2.2 Ποσοστιαία κατανομή των εκπομπών αέριων ρύπων και PM_{10} κατά τομέα οικονομικής δραστηριότητας και νοικοκυριά στην Ελλάδα, 2022-2023.	12
Σχήμα 2.3. Συγκριτική απεικόνιση του μεγέθους των αιωρούμενων σωματιδίων PM_{10} και $PM_{2.5}$ σε σχέση με την ανθρώπινη τρίχα και κόκκους λεπτής άμμου.....	14
Σχήμα 2.4. Ενδεικτικές επιπτώσεις της ατμοσφαιρικής ρύπανσης στην ανθρώπινη υγεία.....	19
Σχήμα 2.5 Ποσοστό σταθμών παρακολούθησης στην Ευρώπη που κατέγραψαν υπερβάσεις των οριακών τιμών της ΕΕ και των κατευθυντήριων τιμών του WHO για τα PM_{10} , με βάση επικυρωμένα δεδομένα για το 2023 και προκαταρκτικά δεδομένα για το 2024.....	30
Σχήμα 3.1 Χωρική θέση των σταθμών παρακολούθησης PM_{10} στην Περιφέρεια Στερεάς Ελλάδας.....	55
Σχήμα 4.3.α Ιστόγραμμα συχνοτήτων των ωριαίων συγκεντρώσεων PM_{10} στην Άμφισσα για το έτος 2018.....	77
Σχήμα 4.3.β Ιστόγραμμα συχνοτήτων των ωριαίων συγκεντρώσεων PM_{10} στη Λαμία για το έτος 2017.....	78
Σχήμα 4.3.γ Ιστόγραμμα συχνοτήτων των ωριαίων συγκεντρώσεων PM_{10} στη Λειβαδιά για το έτος 2019.....	79
Σχήμα 4.3.δ Ιστόγραμμα συχνοτήτων των ωριαίων συγκεντρώσεων PM_{10} στη Χαλκίδα για το έτος 2017.....	80
Σχήμα 4.4.1α. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κανονικής κατανομής στην Άμφισσα για το έτος 2018.	83
Σχήμα 4.4.1β. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κατανομής Gamma στην Άμφισσα για το έτος 2018.	84
Σχήμα 4.4.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κατανομής Weibull στην Άμφισσα για το έτος 2018.....	85
Σχήμα 4.4.1δ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κατανομής Lognormal στην Άμφισσα για το έτος 2018.....	86
Σχήμα 4.4.2α .Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.	89
Σχήμα 4.4.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.	90

Σχήμα 4.4.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.....	91
Σχήμα 4.4.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.....	92
Σχήμα 4.4.3α. Διάγραμμα υπολοίπων (Residual plot) της κανονικής κατανομής (Normal) για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018.....	95
Σχήμα 4.4.3β. Διάγραμμα υπολοίπων (Residual plot) της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018.....	96
Σχήμα 4.4.3γ. Διάγραμμα υπολοίπων (Residual plot) της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018.....	97
Σχήμα 4.4.3δ. Διάγραμμα υπολοίπων (Residual plot) της λογαριθμοκανονικής κατανομής (Lognormal) για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018.....	98
Σχήμα 4.4.4α. Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.	101
Σχήμα 4.4.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.	102
Σχήμα 4.4.4γ. Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.	103
Σχήμα 4.4.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.	104
Σχήμα 4.5.1α. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	113
Σχήμα 4.5.1β. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	114
Σχήμα 4.5.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	115
Σχήμα 4.5.1δ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.	116
Σχήμα 4.5.2α. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	119
Σχήμα 4.5.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	120
Σχήμα 4.5.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	121

Σχήμα 4.5.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	122
Σχήμα 4.5.3α. Διάγραμμα υπολοίπων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.	126
Σχήμα 4.5.3β. Διάγραμμα υπολοίπων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.	127
Σχήμα 4.5.3γ. Διάγραμμα υπολοίπων (Residual plot) της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.....	128
Σχήμα 4.5.3δ. Διάγραμμα υπολοίπων (Residual plot) της λογαριθμοκανονικής κατανομής (Lognormal) για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.....	129
Σχήμα 4.5.4α. Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	134
Σχήμα 4.5.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	135
Σχήμα 4.5.4γ. Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	136
Σχήμα 4.5.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.....	137
Σχήμα 4.6.1α. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.....	148
Σχήμα 4.6.1β. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.....	149
Σχήμα 4.6.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.....	150
Σχήμα 4.6.1δ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.....	152
Σχήμα 4.6.2α. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.....	156
Σχήμα 4.6.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.....	157
Σχήμα 4.6.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.	158
Σχήμα 4.6.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.	160

Σχήμα 4.6.3α. Διάγραμμα υπολοίπων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας το 2017.	163
Σχήμα 4.6.3β. Διάγραμμα υπολοίπων (Residual plot) της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας κατά το έτος 2017.....	164
Σχήμα 4.6.3γ. Διάγραμμα υπολοίπων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας το 2017.	165
Σχήμα 4.6.3δ. Διάγραμμα υπολοίπων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας το 2017.....	166
Σχήμα 4.6.4α. Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.	170
Σχήμα 4.6.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.	171
Σχήμα 4.6.4γ. Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.	172
Σχήμα 4.6.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.	174
Σχήμα 4.7.1α. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	184
Σχήμα 4.7.1β. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	185
Σχήμα 4.7.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	186
Σχήμα 4.7.1δ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.....	187
Σχήμα 4.7.2α. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	190
Σχήμα 4.7.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.....	191
Σχήμα 4.7.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.....	192
Σχήμα 4.7.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019. ...	193
Σχήμα 4.7.3α. Διάγραμμα υπολοίπων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	197
Σχήμα 4.7.3β. Διάγραμμα υπολοίπων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	198

Σχήμα 4.7.3γ. Διάγραμμα υπολοίπων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.....	199
Σχήμα 4.7.3δ. Διάγραμμα υπολοίπων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	200
Σχήμα 4.7.4α. Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	203
Σχήμα 4.7.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	204
Σχήμα 4.7.4γ. Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	205
Σχήμα 4.7.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	206
Σχήμα 4.8.2. Ποσοστό διαθέσιμων ωριαίων μετρήσεων PM_{10} ανά έτος στον σταθμό Καρπενησίου και σύγκριση με το εσωτερικό όριο πληρότητας 90%.....	218

Κατάλογος Πινάκων

Πίνακας 2.1 Ενδεικτική φυσικοχημική σύσταση των PM_{10} , πιθανές πηγές προέλευσης και σχέση με τη χωροχρονική μεταβλητότητα των συγκεντρώσεων.....	17
Πίνακας 2.2. Κύριες ανθρωπογενείς πηγές PM_{10} και μηχανισμοί συμβολής τους στις συγκεντρώσεις.....	23
Πίνακας 2.3 Σύγκριση ευρωπαϊκών οριακών τιμών και κατευθυντήριων τιμών του WHO για τα PM_{10}	29
Πίνακας 2.4 Ενδεικτικές προηγούμενες μελέτες προσαρμογής κατανομών σε δεδομένα PM_{10}	51
Πίνακας 3.1. Μεταδεδομένα σταθμών μέτρησης ατμοσφαιρικής ρύπανσης του ΕΔΠΑΡ που την ευθύνη λειτουργία τους έχει η Περιφέρεια Στερεάς Ελλάδας. Τα στοιχεία προέρχονται από τον αντίστοιχο πίνακα της ιστοσελίδας του ΥΠΕΝ.....	54
Πίνακας 3.2. Τελικό σώμα δεδομένων που χρησιμοποιήθηκε στην ανάλυση.....	56
Πίνακας 4.1 Τελικό αναλυτικό σύνολο δεδομένων PM_{10} που χρησιμοποιήθηκε στις στατιστικές αναλύσεις.....	72
Πίνακας 4.2 Περιγραφικά στατιστικά χαρακτηριστικά των αναλύσιμων χρονοσειρών PM_{10}	74
Πίνακας 4.4.5.1. Τιμές του στατιστικού χ^2 και κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα κατά το έτος 2018.	106
Πίνακας 4.5.5. Τιμές του στατιστικού χ^2 και σχετική κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.	141
Πίνακας 4.6.5. Τιμές του στατιστικού χ^2 και σχετική κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.	177
Πίνακας 4.7.5.1. Τιμές του στατιστικού χ^2 και σχετική κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.	209
Πίνακας 4.8.1. Έλεγχος πληρότητας των ωριαίων συγκεντρώσεων PM_{10} στον σταθμό του Καρπενησίου για την περίοδο 2017-2024.....	215
Πίνακας 4.9 Συγκριτική αξιολόγηση των θεωρητικών κατανομών στις αναλύσιμες χρονοσειρές PM_{10}	222

Συντομογραφίες & Ακρωνύμια

CDF	Αθροιστική συνάρτηση κατανομής
EEA	Ευρωπαϊκός Οργανισμός Περιβάλλοντος - European Environment Agency
PDF	Συνάρτηση πυκνότητας πιθανότητας
PM	Αιωρούμενα σωματίδια - Particulate matter
PM ₁₀	Αιωρούμενα σωματίδια με αεροδυναμική διάμετρο μικρότερη ή ίση με 10 μm
PM _{2.5}	Αιωρούμενα σωματίδια με αεροδυναμική διάμετρο μικρότερη ή ίση με 2,5 μm
Q-Q plot	Διάγραμμα ποσοστημορίων-ποσοστημορίων
U.S. EPA	Υπηρεσία Προστασίας Περιβάλλοντος των Ηνωμένων Πολιτειών - U.S. Environmental Protection Agency
WHO	Παγκόσμιος Οργανισμός Υγείας - World Health Organization
ΕΛΣΤΑΤ	Ελληνική Στατιστική Αρχή
ΥΠΕΚΑ	Υπουργείο Περιβάλλοντος, Ενέργειας και Κλιματικής Αλλαγής
ΥΠΕΝ	Υπουργείο Περιβάλλοντος και Ενέργειας

Τα 5 σημαντικότερα σημεία της εργασίας

- Αναλύθηκαν ωριαίες συγκεντρώσεις PM_{10} σε πέντε πόλεις της Στερεάς Ελλάδας.
- Η επιλογή των χρονοσειρών βασίστηκε σε πληρότητα και ποιότητα δεδομένων $\geq 90\%$.
- Η Lognormal έδωσε την καλύτερη προσαρμογή σε Άμφισσα, Λαμία και Λειβαδιά.
- Στη Χαλκίδα, η Weibull υπερίσχυσε με βάση αριθμητικά και γραφικά κριτήρια.
- Η καταλληλότερη κατανομή διαφοροποιήθηκε ανά πόλη και εξεταζόμενη χρονοσειρά.

Συνεισφορά Εργασίας

Η παρούσα διπλωματική εργασία συμβάλλει στη στατιστική διερεύνηση της ατμοσφαιρικής ρύπανσης σε περιφερειακή κλίμακα, μέσω της ανάλυσης ωριαίων συγκεντρώσεων PM_{10} σε επιλεγμένες πόλεις της Στερεάς Ελλάδας. Η συνεισφορά της έγκειται στη συστηματική αξιοποίηση των διαθέσιμων δεδομένων ποιότητας του αέρα και στον έλεγχο της πληρότητας και της αξιοπιστίας των χρονοσειρών. Παράλληλα, συγκρίνονται η κανονική, η λογαριθμοκανονική, η Gamma και η Weibull κατανομή με αριθμητικά και γραφικά κριτήρια. Η προσέγγιση αυτή επιτρέπει την τεκμηριωμένη επιλογή της καταλληλότερης κατανομής για κάθε χρονοσειρά και προσφέρει χρήσιμο μεθοδολογικό πλαίσιο για μελλοντικές αναλύσεις ποιότητας του αέρα.

ΒΕΛΤΙΣΤΗ ΚΑΤΑΝΟΜΗ ΑΤΜΟΣΦΑΙΡΙΚΩΝ ΡΥΠΩΝ ΣΤΙΣ ΠΟΛΕΙΣ ΤΗΣ ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ

Ιωάννα Ντζάθα¹, Νικόλαος Ραψομανίκης²

^{1,2} Σχολή Εφαρμοσμένων Τεχνών και Βιώσιμου Σχεδιασμού,
Ελληνικό Ανοικτό Πανεπιστήμιο

email: std153496@ac.eap.gr , nikolaos.rapsomanikis@ac.eap.gr

ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία διερευνά τη στατιστική συμπεριφορά των ωριαίων συγκεντρώσεων αιωρούμενων σωματιδίων PM_{10} σε πέντε πόλεις της Στερεάς Ελλάδας, με στόχο τον προσδιορισμό της θεωρητικής κατανομής πιθανότητας που περιγράφει αποτελεσματικότερα τις διαθέσιμες χρονοσειρές. Χρησιμοποιήθηκαν δεδομένα από σταθμούς παρακολούθησης της ποιότητας του αέρα στην Άμφισσα, τη Λαμία, τη Λειβαδιά, τη Χαλκίδα και το Καρπενήσι. Μετά τον έλεγχο πληρότητας και χρονικής συνέχειας επιλέχθηκαν τέσσερις ετήσιες χρονοσειρές: Άμφισσα 2018, Λαμία 2017, Λειβαδιά 2019 και Χαλκίδα 2017, ενώ το Καρπενήσι αποκλείστηκε λόγω εκτεταμένων ελλείψεων.

Στις επιλεγμένες χρονοσειρές προσαρμόστηκαν οι κατανομές Normal, Gamma, Weibull και Lognormal. Η αξιολόγηση βασίστηκε σε ιστογράμματα, συγκρίσεις παρατηρούμενων και αναμενόμενων συχνοτήτων, διαγράμματα υπολοίπων, διαγράμματα Q-Q και στο στατιστικό χ^2 . Οι συγκεντρώσεις παρουσίασαν έντονη θετική ασυμμετρία και εκτεταμένες δεξιές ουρές. Η Lognormal έδωσε την καλύτερη συνολική προσαρμογή στην Άμφισσα, τη Λαμία και τη Λειβαδιά, με τιμές χ^2 96,26, 227,85 και 192,77, αντίστοιχα. Στη Χαλκίδα υπερίσχυσε η Weibull, με τιμή χ^2 46,09, ενώ η Normal παρουσίασε τη χαμηλότερη επίδοση σε όλες τις περιπτώσεις. Τα αποτελέσματα δείχνουν ότι η επιλογή της κατανομής πρέπει να πραγματοποιείται ανά σταθμό και χρονική περίοδο, με συνδυασμό αριθμητικών και γραφικών κριτηρίων.

Λέξεις κλειδιά: PM_{10} , ατμοσφαιρική ρύπανση, Στερεά Ελλάδα, κατανομές πιθανότητας, στατιστική μοντελοποίηση

ABSTRACT

This study investigates the statistical behaviour of hourly PM_{10} concentrations in five cities of Central Greece, aiming to identify the probability distribution that most effectively describes the available time series. Air-quality monitoring data from Amfissa, Lamia, Livadia, Chalkida and Karpenisi were examined. Following completeness and temporal-continuity screening, four annual time series were retained: Amfissa 2018, Lamia 2017, Livadia 2019 and Chalkida 2017, whereas Karpenisi was excluded because of extensive missing data. The Normal, Gamma, Weibull and Lognormal distributions were fitted to the selected time series. Goodness of fit was evaluated using histograms, observed-versus-expected frequencies, residual plots, Q-Q plots and the chi-square statistic. The concentrations exhibited pronounced positive skewness and extended right tails. The Lognormal distribution provided the best overall fit for Amfissa, Lamia and Livadia, with chi-square values of 96.26, 227.85 and 192.77, respectively. In Chalkida, the Weibull distribution provided the best fit, with a chi-square value of 46.09, whereas the Normal distribution showed the poorest performance in all cases. The findings demonstrate that distribution selection should be performed separately for each monitoring station and period, using a combination of numerical and graphical goodness-of-fit criteria.

Keywords: PM_{10} , air pollution, Central Greece, probability distributions, statistical modelling

1. ΕΙΣΑΓΩΓΗ

Τα αιωρούμενα σωματίδια PM_{10} αποτελούν βασικό δείκτη ποιότητας του αέρα, καθώς συνδέονται με επιπτώσεις στην ανθρώπινη υγεία, την ορατότητα και τα οικοσυστήματα. Οι συγκεντρώσεις τους επηρεάζονται από την κυκλοφορία, την οικιακή θέρμανση, τις βιομηχανικές δραστηριότητες, την επαναιώρηση σκόνης, τη μεταφορά αερίων μαζών και τις μετεωρολογικές συνθήκες. Η συνύπαρξη πολλών πηγών και μηχανισμών μεταφοράς δημιουργεί συχνά θετικά ασύμμετρες χρονοσειρές, με μεγάλο αριθμό χαμηλών ή μεσαίων τιμών και περιορισμένα επεισόδια υψηλής συγκέντρωσης. Για τον λόγο αυτό, η επιλογή κατάλληλης θεωρητικής κατανομής πιθανότητας είναι σημαντική για την περιγραφή της μεταβλητότητας, της δεξιάς ουράς και των υψηλών ποσοστημορίων των δεδομένων.

Η ορθή κατανομική περιγραφή έχει ιδιαίτερη σημασία, επειδή διαφορετικά θεωρητικά μοντέλα μπορεί να αποδίδουν με διαφορετική ακρίβεια το κύριο σώμα και τις ακραίες τιμές μιας χρονοσειράς. Η εφαρμογή μίας προεπιλεγμένης κατανομής χωρίς προηγούμενο έλεγχο μπορεί επομένως να οδηγήσει σε εσφαλμένη εκτίμηση της πιθανότητας εμφάνισης αυξημένων συγκεντρώσεων.

Στην παρούσα εργασία συγκρίνονται οι κατανομές Normal, Gamma, Weibull και Lognormal σε ετήσιες χρονοσειρές PM_{10} από πόλεις της Στερεάς Ελλάδας. Σκοπός είναι να προσδιοριστεί η κατανομή που αποδίδει καλύτερα κάθε χρονοσειρά και να διερευνηθεί αν υπάρχει κοινό κατανομικό πρότυπο ή αν απαιτείται διαφορετική επιλογή ανά σταθμό και χρονική περίοδο. Η προσέγγιση αυτή επιτρέπει την αξιολόγηση της κατανομικής ομοιογένειας μεταξύ των πόλεων και την ανάδειξη πιθανών τοπικών διαφοροποιήσεων.

2. ΜΕΘΟΔΟΛΟΓΙΑ - ΠΗΓΕΣ ΔΕΔΟΜΕΝΩΝ

Χρησιμοποιήθηκαν ωριαίες συγκεντρώσεις PM_{10} από σταθμούς παρακολούθησης στην Άμφισσα, τη Λαμία, τη Λειβαδιά, τη Χαλκίδα και το Καρπενήσι. Οι χρονοσειρές ελέγχθηκαν ως προς την εγκυρότητα, τη χρονική συνέχεια και την πληρότητά τους. Ως κριτήριο ένταξης στην ανάλυση ορίστηκε ετήσια πληρότητα τουλάχιστον 90%. Με βάση το κριτήριο αυτό επιλέχθηκαν οι χρονοσειρές Άμφισσα 2018, Λαμία 2017, Λειβαδιά 2019 και Χαλκίδα 2017, ενώ το Καρπενήσι αποκλείστηκε λόγω ανεπαρκών δεδομένων.

Για κάθε χρονοσειρά υπολογίστηκαν βασικά περιγραφικά στατιστικά και προσαρμόστηκαν οι κατανομές Normal, Gamma, Weibull και Lognormal. Οι παράμετροι των θεωρητικών κατανομών εκτιμήθηκαν από τις διαθέσιμες παρατηρήσεις και οι αναμενόμενες συχνότητες υπολογίστηκαν για τις ίδιες

κλάσεις που χρησιμοποιήθηκαν στην εμπειρική κατανομή. Με τον τρόπο αυτό κατέστη δυνατή η άμεση σύγκριση της παρατηρούμενης με τη θεωρητική μορφή κάθε χρονοσειράς. Η αξιολόγηση πραγματοποιήθηκε με ιστογράμματα, σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων, διαγράμματα υπολοίπων, Q-Q plots και το στατιστικό χ^2 . Η τελική επιλογή βασίστηκε στη συνδυαστική αξιολόγηση αριθμητικών και γραφικών κριτηρίων, ώστε να περιοριστεί η εξάρτηση από ένα μόνο μέτρο καλής προσαρμογής.

Πίνακας 1. Τελικό σύνολο αναλύσιμων χρονοσειρών PM_{10} .

Πόλη	Έτος	Ώρες	Πληρότητα
Άμφισσα	2018	8750	99,89%
Λαμία	2017	8059	92,00%
Λειβαδιά	2019	8745	99,83%
Χαλκίδα	2017	8702	99,34%
Καρπενήσι	-	-	Ανεπαρκής πληρότητα

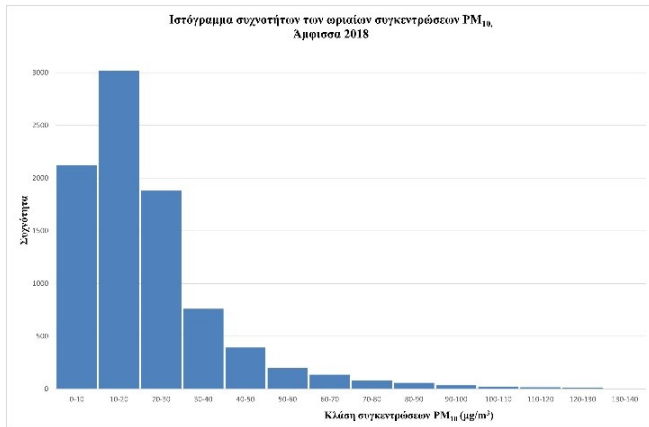
3. ΑΠΟΤΕΛΕΣΜΑΤΑ

3.1 Περιγραφικά χαρακτηριστικά

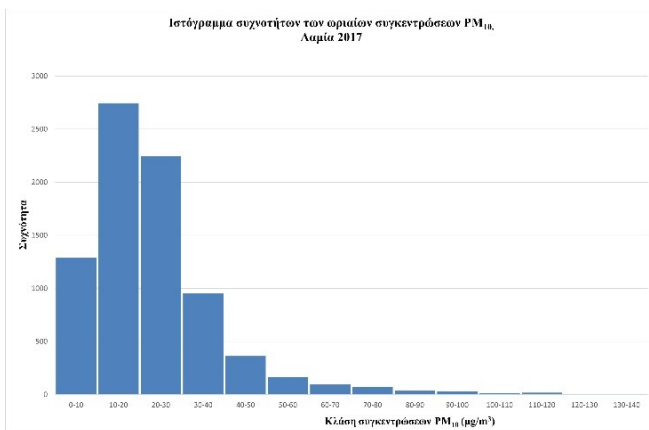
Οι τέσσερις αναλύσιμες χρονοσειρές παρουσίασαν μέση τιμή μεγαλύτερη από τη διάμεσο και θετική ασυμμετρία, γεγονός που δείχνει συγκέντρωση των περισσότερων παρατηρήσεων στις χαμηλές και μεσαίες τιμές και παρουσία εκτεταμένης δεξιάς ουράς. Η Χαλκίδα εμφάνισε τη μεγαλύτερη μέση συγκέντρωση, τυπική απόκλιση και μέγιστη τιμή, ενώ η Λειβαδιά και η Χαλκίδα παρουσίασαν την εντονότερη ασυμμετρία. Τα χαρακτηριστικά αυτά ενισχύουν την καταλληλότητα ασύμμετρων και θετικά ορισμένων κατανομών για τη μοντελοποίηση των συγκεκριμένων χρονοσειρών.

Πίνακας 2. Βασικά περιγραφικά στατιστικά των ωριαίων συγκεντρώσεων PM_{10} ($\mu\text{g}/\text{m}^3$).

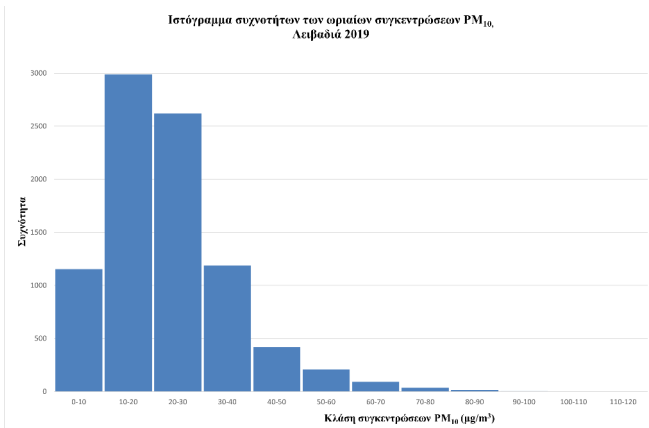
Πόλη	Μέση Τιμή	Διάμεσος	Τυπική Απόκλιση	Ασυμμετρία	Μέγιστη τιμή
Άμφισσα	22,21	18,00	18,78	3,051	260
Λαμία	23,86	20,00	17,57	3,348	250
Λειβαδιά	23,57	21,00	15,62	6,528	438
Χαλκίδα	36,48	31,00	30,01	7,946	891



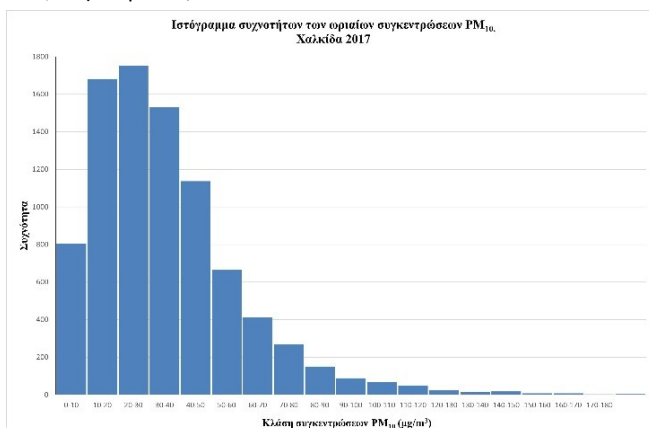
Σχήμα 1α. Εμπειρική κατανομή των ωριαίων συγκεντρώσεων PM_{10} στην Άμφισσα, 2018.



Σχήμα 1β. Εμπειρική κατανομή των ωριαίων συγκεντρώσεων PM_{10} στη Λαμία, 2017.



Σχήμα 1γ. Εμπειρική κατανομή των ωριαίων συγκεντρώσεων PM_{10} στη Λειβαδιά, 2019.

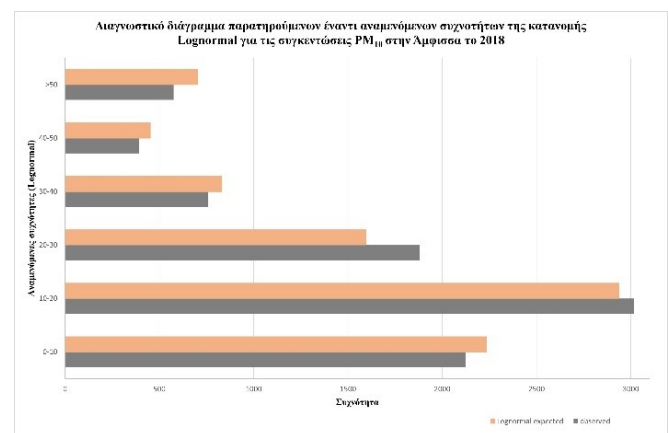


Σχήμα 1δ. Εμπειρική κατανομή των ωριαίων συγκεντρώσεων PM_{10} στη Χαλκίδα, 2017.

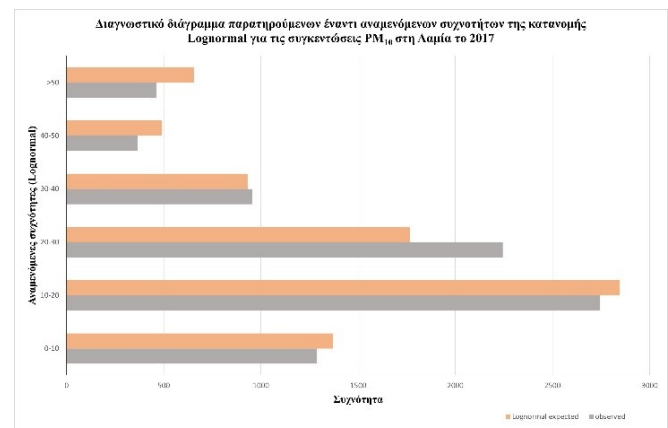
Τα ιστογράμματα επιβεβαιώνουν τη θετική ασυμμετρία των χρονοσειρών και την περιορισμένη καταλληλότητα της Normal. Η Άμφισσα, η Λαμία και η Λειβαδιά εμφανίζουν έντονη συγκέντρωση στις χαμηλές και μεσαίες κλάσεις, ενώ η Χαλκίδα παρουσιάζει ευρύτερη κατανομή και σταδιακότερη μείωση των συχνοτήτων. Η εικόνα αυτή προϋποθέτει για καλύτερη προσαρμογή θετικά ορισμένων και ασύμμετρων κατανομών σε σχέση με τη Normal.

3.2 Σύγκριση θεωρητικών κατανομών

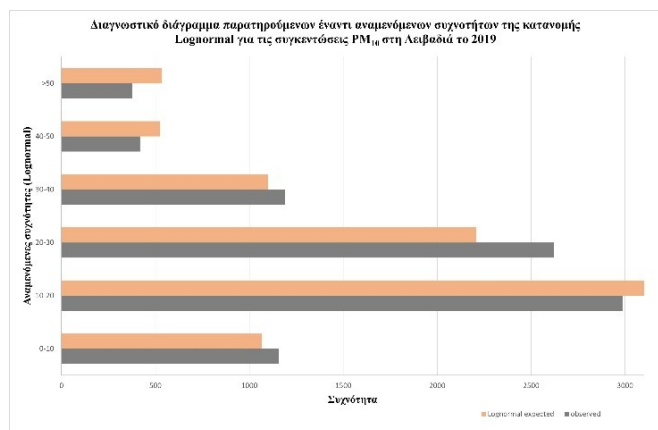
Η σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων ανέδειξε σαφή υπεροχή των ασύμμετρων κατανομών έναντι της Normal. Στην Άμφισσα, τη Λαμία και τη Λειβαδιά, η Lognormal απέδωσε καλύτερα την κύρια κορυφή και τη σταδιακή εξασθένηση της δεξιάς ουράς. Στη Χαλκίδα, αντίθετα, η Weibull παρουσίασε καλύτερη συμφωνία κυρίως στις χαμηλές και μεσαίες κλάσεις, όπου συγκεντρώνεται το μεγαλύτερο μέρος των παρατηρήσεων. Οι αποκλίσεις που παρατηρήθηκαν στις υψηλότερες κλάσεις δείχνουν, ωστόσο, ότι καμία κατανομή δεν αναπαριστά πλήρως όλα τα ακραία επεισόδια. Η συνολική εικόνα ενισχύει την ανάγκη επιλογής της κατανομής ξεχωριστά για κάθε χρονοσειρά.



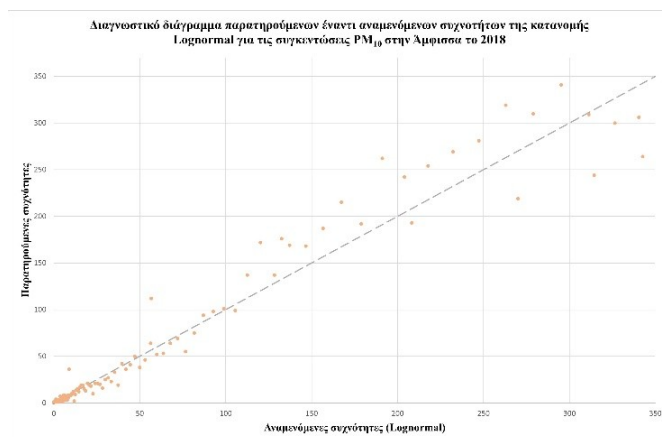
Σχήμα 2α. Παρατηρούμενες και αναμενόμενες συχνότητες για την προσαρμογή της Lognormal στην Άμφισσα, 2018.



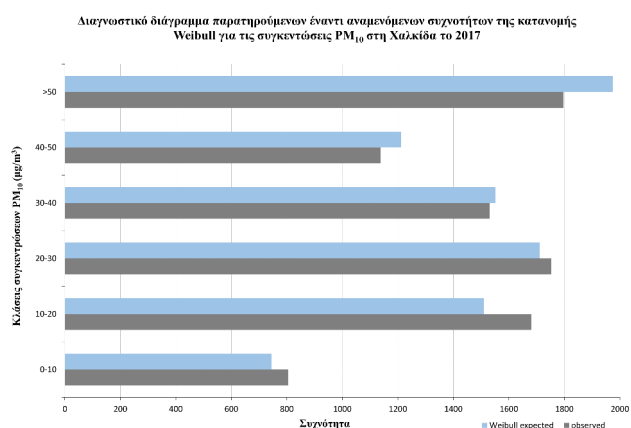
Σχήμα 2β. Παρατηρούμενες και αναμενόμενες συχνότητες για την προσαρμογή της Lognormal στη Λαμία, 2017.



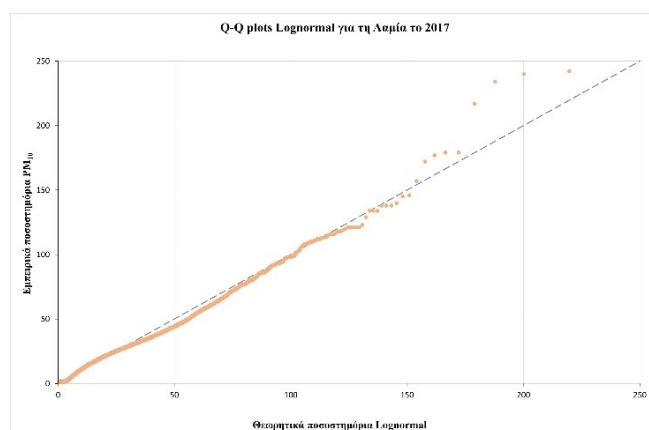
Σχήμα 2γ. Παρατηρούμενες και αναμενόμενες συχνότητες για την προσαρμογή της Lognormal στη Λειβαδιά, 2019.



Σχήμα 3α. Q-Q plot για την προσαρμογή της Lognormal στην Αμφίσσα, 2018.



Σχήμα 2δ. Παρατηρούμενες και αναμενόμενες συχνότητες για την προσαρμογή της Weibull στη Χαλκίδα, 2017.

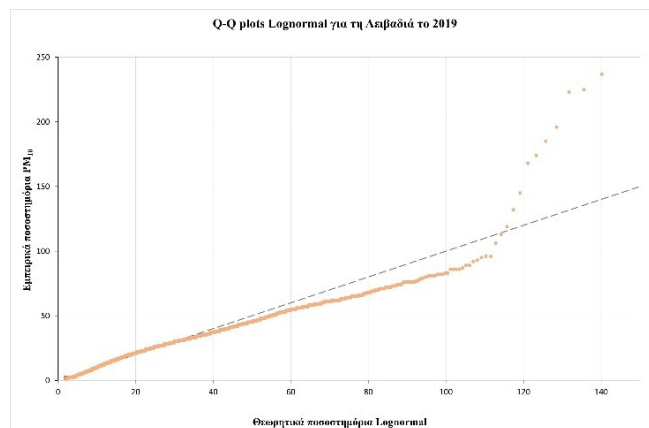


Σχήμα 3β. Q-Q plot της Lognormal στη Λαμία, 2017.

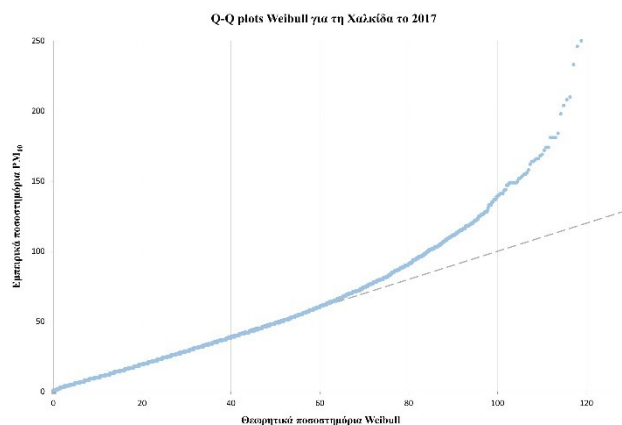
Η συνολική εικόνα δείχνει ότι η Lognormal περιγράφει αποτελεσματικότερα τις τρεις πρώτες χρονοσειρές, ενώ η Χαλκίδα εμφανίζει διαφορετική κατανομική μορφή. Η διαφοροποίηση αυτή αποτελεί βασική ένδειξη ότι η επιλογή μοντέλου πρέπει να γίνεται ξεχωριστά για κάθε σταθμό και χρονική περίοδο.

3.3 Διαγράμματα Q-Q των βέλτιστων κατανομών

Τα διαγράμματα Q-Q χρησιμοποιήθηκαν για την αξιολόγηση της συμφωνίας μεταξύ εμπειρικών και θεωρητικών ποσοστημορίων. Όσο περισσότερο τα σημεία ακολουθούν τη διαγώνιο αναφοράς, τόσο καλύτερα η κατανομή αποδίδει το κύριο σώμα και τις ουρές των δεδομένων. Οι συστηματικές αποκλίσεις από τη διαγώνιο, ιδιαίτερα στα ανώτερα ποσοστημόρια, υποδηλώνουν αδυναμία του μοντέλου να περιγράψει πλήρως τις ακραίες συγκεντρώσεις. Η συνολική μορφή των σημείων αξιολογήθηκε με έμφαση τόσο στη γραμμικότητα του κεντρικού τμήματος όσο και στη συμπεριφορά των υψηλότερων ποσοστημορίων. Για τον λόγο αυτό, η ερμηνεία τους πραγματοποιήθηκε συμπληρωματικά προς τα υπόλοιπα γραφικά και αριθμητικά κριτήρια προσαρμογής.



Σχήμα 3γ. Q-Q plot της Lognormal στη Λειβαδιά, 2019.



Σχήμα 3δ. Q-Q plot της Weibull στη Χαλκίδα, 2017.

Στην Άμφισσα, τη Λαμία και τη Λειβαδιά, τα σημεία της Lognormal παραμένουν σχετικά κοντά στη γραμμή αναφοράς στο μεγαλύτερο μέρος των ποσοστημορίων, με εντονότερες αποκλίσεις στο ανώτερο άκρο. Οι αποκλίσεις αυτές δείχνουν ότι η κατανομή αποδίδει ικανοποιητικά το κύριο σώμα των δεδομένων, αλλά δεν αναπαριστά πλήρως τις υψηλότερες συγκεντρώσεις.

Στη Χαλκίδα, η Weibull παρουσιάζει επίσης αποκλίσεις στα υψηλά ποσοστημόρια και δεν αποδίδει πλήρως την άνω ουρά της χρονοσειράς. Η επιλογή της ως καταλληλότερης κατανομής δεν βασίστηκε αποκλειστικά στη μορφή του Q-Q plot, αλλά κυρίως στην καλύτερη συμφωνία στις ομαδοποιημένες συχνότητες και στη σημαντικά χαμηλότερη τιμή χ^2 . Η συνδυαστική αυτή αξιολόγηση περιορίζει τον κίνδυνο επιλογής μοντέλου αποκλειστικά βάσει ενός μόνο διαγνωστικού κριτηρίου.

3.4 Αριθμητική αξιολόγηση με το χ^2

Η αριθμητική κατάταξη επιβεβαίωσε τα γραφικά ευρήματα. Η Lognormal παρουσίασε τη μικρότερη τιμή χ^2 στην Άμφισσα, τη Λαμία και τη Λειβαδιά, ενώ στη Χαλκίδα η Weibull υπερίσχυσε με σαφώς χαμηλότερη τιμή. Η Normal εμφάνισε τη χειρότερη προσαρμογή σε όλες τις περιπτώσεις. Η συμφωνία αριθμητικών και γραφικών κριτηρίων ενισχύει την αξιοπιστία της τελικής επιλογής για κάθε χρονοσειρά.

Πίνακας 3. Τιμές του στατιστικού χ^2 και επιλεγμένη θεωρητική κατανομή ανά χρονοσειρά.

Χρονοσειρά	Normal	Gamma	Weibull	Lognormal	Βέλτιστη κατανομή
Άμφισσα 2018	1634,93	466,61	476,42	96,26	Lognormal
Λαμία 2017	1568,23	574,16	346,07	227,85	Lognormal
Λειβαδιά 2019	1273,76	480,65	445,54	192,77	Lognormal
Χαλκίδα 2017	2062,81	591,19	46,09	234,53	Weibull

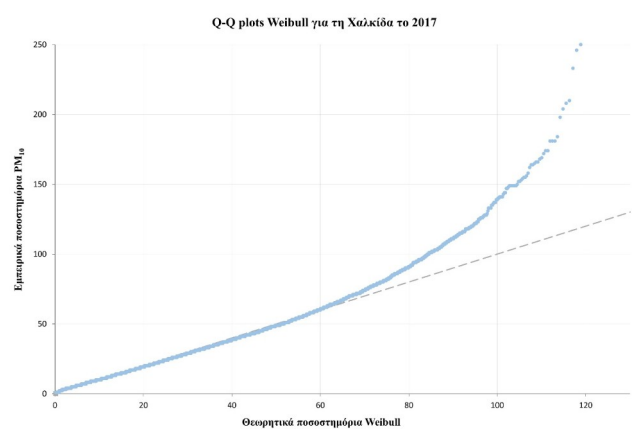
Η αριθμητική υπεροχή της Weibull στη Χαλκίδα είναι ιδιαίτερα έντονη, καθώς η τιμή χ^2 είναι πολύ χαμηλότερη από εκείνη της Lognormal. Αντίστοιχα, στις άλλες τρεις πόλεις η Lognormal κατατάσσεται σταθερά πρώτη, επιβεβαιώνοντας ότι αποδίδει αποτελεσματικότερα τη μορφή των αντίστοιχων χρονοσειρών.

3.5 Η διαφοροποίηση της Χαλκίδας

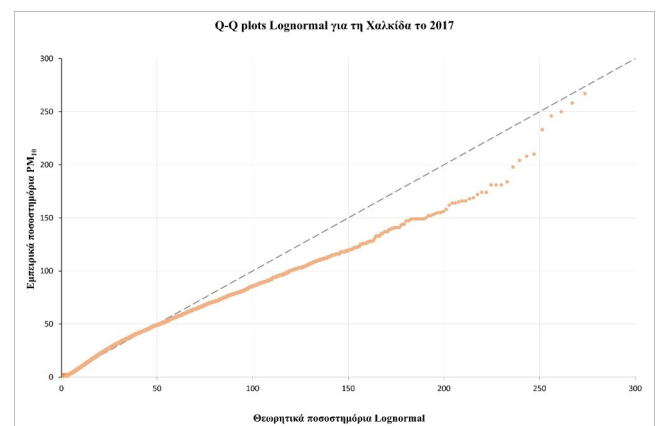
Η Χαλκίδα αποτελεί το βασικό διαφοροποιητικό εύρημα της μελέτης. Σε αντίθεση με την Άμφισσα, τη Λαμία και τη Λειβαδιά, όπου υπερίσχυσε η Lognormal, στη Χαλκίδα η Weibull παρουσίασε

σαφώς καλύτερη συνολική προσαρμογή. Η υπεροχή της τεκμηριώνεται κυρίως από τη σημαντικά χαμηλότερη τιμή χ^2 , την καλύτερη συμφωνία μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων και τη μικρότερη συστηματική δομή των υπολοίπων.

Η εμπειρική κατανομή της Χαλκίδας είναι ευρύτερη, με μεγαλύτερη διασπορά και σταδιακότερη μείωση των συχνοτήτων. Το αποτέλεσμα αυτό δείχνει ότι η καταλληλότερη θεωρητική κατανομή δεν είναι ενιαία για όλες τις πόλεις, αλλά εξαρτάται από τα ιδιαίτερα χαρακτηριστικά κάθε χρονοσειράς. Επειδή η ανάλυση αφορά ένα μόνο έτος, η διαφοροποίηση πρέπει να θεωρηθεί ένδειξη τοπικής και χρονικά συγκεκριμένης συμπεριφοράς και όχι μόνιμο χαρακτηριστικό της πόλης.



Σχήμα 4α. Q-Q plot της Weibull για τη Χαλκίδα, 2017.



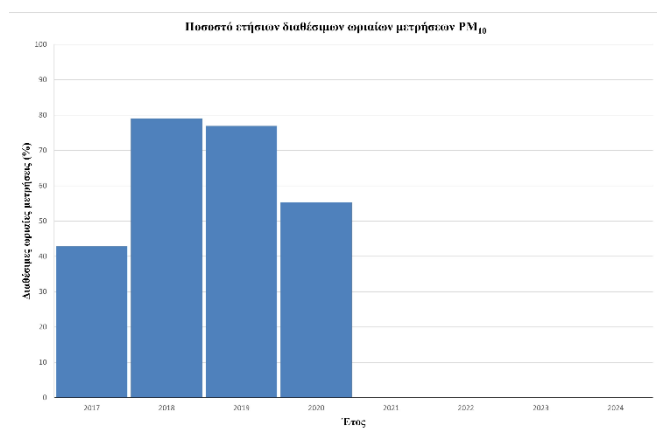
Σχήμα 4β. Q-Q plot της Lognormal για τη Χαλκίδα, 2017.

Η σύγκριση των δύο Q-Q plots δείχνει ότι και οι δύο κατανομές παρουσιάζουν αποκλίσεις στα υψηλά ποσοστημόρια και δεν αναπαριστούν πλήρως την άνω ουρά. Η Weibull, ωστόσο, επιλέχθηκε ως καταλληλότερη για τη Χαλκίδα λόγω της καλύτερης συμφωνίας στις ομαδοποιημένες συχνότητες και της σαφώς χαμηλότερης τιμής χ^2 . Η περίπτωση αυτή αναδεικνύει τη σημασία της ξεχωριστής αξιολόγησης κάθε χρονοσειράς πριν από την επιλογή του τελικού θεωρητικού μοντέλου. Παράλληλα, επιβεβαιώνει ότι η μικρότερη τιμή χ^2 πρέπει να ερμηνεύεται σε

συνδυασμό με τη γραφική συμπεριφορά του μοντέλου.

3.6 Η περίπτωση του Καρπενησίου

Το Καρπενήσι δεν εντάχθηκε στη στατιστική προσαρμογή, επειδή κανένα διαθέσιμο έτος δεν ικανοποιούσε το κριτήριο ετήσιας πληρότητας 90%. Η καλύτερη διαθεσιμότητα καταγράφηκε το 2018, αλλά παρέμεινε ανεπαρκής για αξιόπιστη μοντελοποίηση. Ο αποκλεισμός του σταθμού περιορίζει τον κίνδυνο εξαγωγής μη αντιπροσωπευτικών συμπερασμάτων και υπογραμμίζει ότι η ποιότητα των δεδομένων προηγείται της εφαρμογής οποιασδήποτε θεωρητικής κατανομής. Η ανεπαρκής πληρότητα υποδηλώνει σημαντικά κενά στη χρονική κάλυψη, τα οποία θα μπορούσαν να επηρεάσουν την εκτίμηση των παραμέτρων και τη μορφή της προσαρμοσμένης κατανομής.



Σχήμα 5. Ποσοστό διαθέσιμων ωριαίων μετρήσεων PM₁₀ ανά έτος στον σταθμό Καρπενησίου και σύγκριση με το κριτήριο πληρότητας 90%.

Κανένα διαθέσιμο έτος δεν έφθασε το απαιτούμενο όριο πληρότητας 90%. Η υψηλότερη διαθεσιμότητα καταγράφηκε το 2018, με 79,03%, ενώ για τα επόμενα έτη η πληρότητα ήταν χαμηλότερη ή δεν υπήρχαν διαθέσιμα δεδομένα. Για τον λόγο αυτό, το Καρπενήσι αποκλείστηκε από τη στατιστική προσαρμογή, ώστε να αποφευχθούν μη αντιπροσωπευτικά αποτελέσματα. Η επιλογή αυτή ενισχύει την αξιοπιστία της ανάλυσης, καθώς βασίζεται μόνο σε χρονοσειρές με επαρκή χρονική κάλυψη.

4. ΣΥΖΗΤΗΣΗ

Τα αποτελέσματα ανέδειξαν σαφείς διαφορές στην κατανομική συμπεριφορά των χρονοσειρών PM₁₀ μεταξύ των εξεταζόμενων πόλεων. Η επικράτηση της Lognormal στην Άμφισσα, τη Λαμία και τη Λειβαδιά, σε συνδυασμό με τη διαφοροποίηση της Χαλκίδας, όπου υπερίσχυσε η Weibull, επιβεβαιώνει ότι δεν

υπάρχει μία ενιαία θεωρητική κατανομή κατάλληλη για όλες τις περιπτώσεις.

Η διαφοροποίηση αυτή πιθανότατα συνδέεται με τις ιδιαίτερες συνθήκες κάθε περιοχής, όπως η ένταση και το είδος των τοπικών εκπομπών, η μορφολογία, η κυκλοφορία των αερίων μαζών και η συχνότητα επεισοδίων μεταφοράς σκόνης. Επομένως, η κατανομική μορφή των συγκεντρώσεων δεν πρέπει να ερμηνεύεται ως σταθερό χαρακτηριστικό μιας πόλης, αλλά ως αποτέλεσμα της αλληλεπίδρασης τοπικών και χρονικά μεταβαλλόμενων παραγόντων.

Παράλληλα, η περίπτωση του Καρπενησίου υπογραμμίζει ότι η πληρότητα και η ποιότητα των δεδομένων αποτελούν απαραίτητες προϋποθέσεις πριν από οποιαδήποτε στατιστική μοντελοποίηση. Η αξιοπιστία των συμπερασμάτων εξαρτάται όχι μόνο από την επιλογή της κατάλληλης κατανομής, αλλά και από την αντιπροσωπευτικότητα της ίδιας της χρονοσειράς.

Συνεπώς, η ερμηνεία των αποτελεσμάτων απαιτεί ταυτόχρονη συνεκτίμηση της τοπικής περιβαλλοντικής μεταβλητότητας και της ποιότητας των διαθέσιμων δεδομένων.

4.1 Ερμηνεία και συνολική αξιολόγηση

Η επιλογή της βέλτιστης κατανομής δεν βασίστηκε αποκλειστικά στη μικρότερη τιμή χ^2 , αλλά στη συνδυαστική αξιολόγηση όλων των διαθέσιμων ενδείξεων. Τα ιστογράμματα ανέδειξαν τη θετική ασυμμετρία των χρονοσειρών, τα διαγράμματα παρατηρούμενων και αναμενόμενων συχνοτήτων εντόπισαν τις αποκλίσεις ανά κλάση, τα διαγράμματα υπολοίπων αποκάλυψαν πιθανή συστηματική δομή και τα Q-Q plots έδειξαν τη συμφωνία των εμπειρικών με τα θεωρητικά ποσοστημόρια.

Στην Άμφισσα, τη Λαμία και τη Λειβαδιά, η Lognormal παρουσίασε ικανοποιητική προσαρμογή στο μεγαλύτερο μέρος της κατανομής, με εντονότερες αποκλίσεις κυρίως στην άνω ουρά. Στη Χαλκίδα, καμία κατανομή δεν απέδωσε πλήρως τις ακραίες τιμές, ωστόσο η Weibull παρουσίασε σαφώς καλύτερη συμφωνία στις ομαδοποιημένες συχνότητες και σημαντικά χαμηλότερη τιμή χ^2 . Η σύγκλιση αριθμητικών και γραφικών κριτηρίων θεωρήθηκε ισχυρότερη ένδειξη από οποιοδήποτε μεμονωμένο μέτρο.

Η διαφοροποίηση αυτή δείχνει ότι η καταλληλότητα ενός θεωρητικού μοντέλου εξαρτάται τόσο από το κύριο σώμα της κατανομής όσο και από τη συμπεριφορά της δεξιάς ουράς. Επομένως, μία κατανομή μπορεί να εμφανίζει ικανοποιητική συνολική προσαρμογή, αλλά να παρουσιάζει σημαντικές αποκλίσεις στις υψηλές συγκεντρώσεις, οι οποίες έχουν ιδιαίτερο ενδιαφέρον στην περιβαλλοντική αξιολόγηση.

Πίνακας 4. Συνδυαστική αξιολόγηση των βέλτιστων κατανομών.

Χρονοσειρά	Βέλτιστη	Κύρια γραφική ένδειξη	χ^2
Άμφισσα 2018	Lognormal	Καλή ευθυγράμμιση, απόκλιση μόνο στην άνω ουρά	96,26
Λαμία 2017	Lognormal	Ικανοποιητικό κέντρο, μέτριες αποκλίσεις στα υψηλά ποσοστημόρια	227,85
Λειβαδιά 2019	Lognormal	Καλή συνολική γραμμικότητα, λίγες ακραίες αποκλίσεις	192,77
Χαλκίδα 2017	Weibull	Καλύτερη συμφωνία στις ομαδοποιημένες συχνότητες και μικρότερη συστηματική δομή υπολοίπων	46,09

Η συνδυαστική αξιολόγηση επιβεβαιώνει την επικράτηση της Lognormal στις τρεις χρονοσειρές και τη σαφή διαφοροποίηση της Χαλκίδας, όπου η Weibull απέδωσε καλύτερα τη συνολική μορφή των δεδομένων. Παράλληλα, καταδεικνύει ότι η επιλογή του καταλληλότερου μοντέλου πρέπει να βασίζεται στη σύγκλιση αριθμητικών και γραφικών ενδείξεων.

4.2 Περιβαλλοντική σημασία

Η επιλογή κατάλληλης θεωρητικής κατανομής έχει πρακτική σημασία, καθώς μπορεί να επηρεάσει την εκτίμηση υψηλών ποσοστημορίων και της πιθανότητας εμφάνισης αυξημένων συγκεντρώσεων PM_{10} . Ένα μοντέλο που δεν αποδίδει σωστά τη δεξιά ουρά ενδέχεται να υποεκτιμήσει ή να υπερεκτιμήσει τη συχνότητα ακραίων επεισοδίων, τα οποία παρουσιάζουν ιδιαίτερο ενδιαφέρον για την αξιολόγηση της ποιότητας του αέρα και της ανθρώπινης έκθεσης.

Η αξιόπιστη κατανομική περιγραφή μπορεί επομένως να υποστηρίξει τη σύγκριση διαφορετικών σταθμών και χρονικών περιόδων, καθώς και την εκτίμηση της πιθανότητας ασυνήθιστα υψηλών συγκεντρώσεων. Παράλληλα, μπορεί να λειτουργήσει συμπληρωματικά προς τις πραγματικές μετρήσεις, χωρίς να υποκαθιστά την άμεση παρακολούθηση και την κανονιστική αξιολόγηση της ποιότητας του αέρα. Η εφαρμογή της κατάλληλης κατανομής μπορεί επίσης να συμβάλει στην πιο αξιόπιστη εκτίμηση της συχνότητας επεισοδίων υψηλής ρύπανσης και στη συγκριτική αξιολόγηση της μεταβλητότητας μεταξύ διαφορετικών περιοχών.

Με τον τρόπο αυτό, η στατιστική μοντελοποίηση μπορεί να ενισχύσει την ερμηνεία των επεισοδίων ρύπανσης και να υποστηρίξει πιο τεκμηριωμένες συγκρίσεις μεταξύ περιοχών.

4.3 Περιορισμοί και κατευθύνσεις περαιτέρω έρευνας

Η αξιοπιστία των αποτελεσμάτων ενισχύεται από το μεγάλο πλήθος ωριαίων παρατηρήσεων και τη χρήση πολλαπλών διαγνωστικών εργαλείων. Ωστόσο, κάθε χρονοσειρά αφορά ένα συγκεκριμένο έτος, επομένως η βέλτιστη κατανομή δεν πρέπει να θεωρείται αμετάβλητο χαρακτηριστικό της πόλης. Διαφορετικές μετεωρολογικές συνθήκες, επεισόδια μεταφοράς σκόνης ή μεταβολές στις τοπικές εκπομπές μπορούν να οδηγήσουν σε διαφορετική κατανομική μορφή. Επιπλέον, το χ^2 επηρεάζεται από το πλήθος και τα όρια των κλάσεων, ενώ τα Q–Q plots είναι ευαίσθητα στις ακραίες παρατηρήσεις. Μελλοντικές αναλύσεις θα μπορούσαν να περιλαμβάνουν πολυτεείς και εποχικές χρονοσειρές, ημερήσιες μέσες τιμές, πρόσθετους ρύπους, μετεωρολογικές μεταβλητές και εναλλακτικά κριτήρια, όπως AIC, BIC, Kolmogorov–Smirnov και Anderson–Darling. Ιδιαίτερα για τη Χαλκίδα, η συνεκτίμηση θαλάσσιων, βιομηχανικών και μετεωρολογικών παραγόντων θα μπορούσε να ερμηνεύσει πληρέστερα την υπεροχή της Weibull.

4.4 Σύγκριση με τη διεθνή βιβλιογραφία

Η υπεροχή της Lognormal σε τρεις από τις τέσσερις αναλύσιμες χρονοσειρές συμφωνεί με προηγούμενες μελέτες για τις συγκεντρώσεις PM_{10} στην Ελλάδα και διεθνώς (Giavis et al., 2009· Lu & Fang, 2003· Wang et al., 2013). Η λογαριθμοκανονική μορφή θεωρείται συμβατή με τη μη αρνητικότητα, τη θετική ασυμμετρία και τη συνδυασμένη επίδραση μηχανισμών εκπομπής, μεταφοράς και ατμοσφαιρικής ανάμιξης.

Ωστόσο, η βιβλιογραφία δεν υποστηρίζει μία καθολικά βέλτιστη κατανομή. Σε διαφορετικές περιοχές και χρονικές κλίμακες έχουν αναδειχθεί επίσης η Gamma, η Weibull και μοντέλα μιγμάτων (Plocoste et al., 2020· Bağcı, 2023). Οι διαφοροποιήσεις αυτές μπορούν να συνδέονται με τη χρονική ανάλυση των δεδομένων, τη διάρκεια της εξεταζόμενης περιόδου, τη θέση του σταθμού και τη συνύπαρξη διαφορετικών πηγών ρύπανσης.

Η περίπτωση της Χαλκίδας επιβεβαιώνει αυτή τη μη καθολικότητα, καθώς η Weibull παρουσίασε πολύ μικρότερη τιμή χ^2 και καλύτερη συμφωνία στις ομαδοποιημένες συχνότητες. Το εύρημα δεν αποδεικνύει ότι η Weibull είναι γενικά καταλληλότερη για παράκτιες πόλεις, αλλά ότι η συγκεκριμένη χρονοσειρά του 2017 παρουσιάζει μορφή που περιγράφεται αποτελεσματικότερα από τη Weibull.

Η χρήση πολλαπλών κριτηρίων αποτελεί βασικό πλεονέκτημα της μελέτης. Η ταυτόχρονη εξέταση συχνοτήτων, υπολοίπων, Q–Q plots και χ^2 επιτρέπει να εντοπίζονται περιπτώσεις στις οποίες ένα μοντέλο αποδίδει ικανοποιητικά το κύριο σώμα των

δεδομένων, αλλά παρουσιάζει αποκλίσεις στις ουρές. Παράλληλα, ο αποκλεισμός του Καρπενησίου λόγω ανεπαρκούς πληρότητας είναι συμβατός με τη διεθνή πρακτική ποιοτικού ελέγχου πριν από τη στατιστική μοντελοποίηση.

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Η ανάλυση έδειξε ότι οι ωριαίες συγκεντρώσεις PM_{10} στις εξεταζόμενες πόλεις είναι μη αρνητικές, έντονα ασύμμετρες και χαρακτηρίζονται από εκτεταμένες δεξιές ουρές. Η μορφή αυτή επιβεβαιώνει ότι η Normal κατανομή δεν περιγράφει ικανοποιητικά τα δεδομένα, ιδιαίτερα στις υψηλές συγκεντρώσεις. Αντίθετα, οι θετικά ορισμένες και ασύμμετρες κατανομές παρουσίασαν σαφώς καλύτερη προσαρμογή. Η Lognormal αποτέλεσε την καταλληλότερη επιλογή για την Άμφισσα 2018, τη Λαμία 2017 και τη Λειβαδιά 2019, ενώ στη Χαλκίδα 2017 η Weibull παρουσίασε την καλύτερη συνολική επίδοση. Η Gamma έδωσε ενδιάμεσα αποτελέσματα, χωρίς να υπερισχύσει σε κάποια από τις εξεταζόμενες χρονοσειρές.

Τα αποτελέσματα επιβεβαιώνουν ότι δεν υπάρχει μία καθολικά βέλτιστη θεωρητική κατανομή για όλες τις χρονοσειρές PM_{10} της Στερεάς Ελλάδας. Η επιλογή πρέπει να πραγματοποιείται χωριστά για κάθε σταθμό και χρονική περίοδο, λαμβάνοντας υπόψη τόσο τα αριθμητικά όσο και τα γραφικά κριτήρια προσαρμογής. Η περίπτωση της Χαλκίδας αποτελεί το σημαντικότερο διαφοροποιητικό εύρημα της μελέτης, καθώς δείχνει ότι μία κατανομή που αποδίδει ικανοποιητικά σε αρκετές πόλεις δεν πρέπει να εφαρμόζεται μηχανιστικά σε όλες τις περιοχές. Παράλληλα, ο αποκλεισμός του Καρπενησίου υπογραμμίζει ότι η επάρκεια, η χρονική συνέχεια και η ποιότητα των πρωτογενών δεδομένων αποτελούν αναγκαίες προϋποθέσεις πριν από κάθε στατιστική μοντελοποίηση.

Η συνδυαστική χρήση ιστογραμμάτων, παρατηρούμενων και αναμενόμενων συχνοτήτων, διαγραμμάτων υπολοίπων, Q-Q plots και του στατιστικού χ^2 ενίσχυσε την αξιοπιστία της τελικής επιλογής. Καμία μεμονωμένη μέθοδος δεν αποδείχθηκε επαρκής για την πλήρη αξιολόγηση της προσαρμογής, ιδιαίτερα στις ουρές των κατανομών, όπου εμφανίζονται τα επεισόδια υψηλών συγκεντρώσεων. Για τον λόγο αυτό, η τελική απόφαση πρέπει να βασίζεται στη σύγκλιση περισσότερων ενδείξεων και όχι αποκλειστικά στη μικρότερη αριθμητική τιμή ενός κριτηρίου.

Σε επίπεδο εφαρμογής, η κατάλληλη θεωρητική κατανομή μπορεί να υποστηρίξει την εκτίμηση υψηλών ποσοστημορίων, την αξιολόγηση της πιθανότητας εμφάνισης ασυνήθιστα αυξημένων συγκεντρώσεων και τη σύγκριση διαφορετικών

σταθμών ή ετών. Οι εκτιμήσεις αυτές δεν υποκαθιστούν τις πραγματικές μετρήσεις ούτε την κανονιστική αξιολόγηση της ποιότητας του αέρα, μπορούν όμως να λειτουργήσουν συμπληρωματικά στην περιγραφή της μεταβλητότητας και των ακραίων επεισοδίων.

Η μελέτη μπορεί να αποτελέσει βάση για ένα επαναλήψιμο πλαίσιο ανάλυσης δεδομένων ποιότητας του αέρα, το οποίο περιλαμβάνει αρχικό έλεγχο πληρότητας και συνέχειας, περιγραφική διερεύνηση, εφαρμογή εναλλακτικών κατανομών και τεκμηριωμένη επιλογή του τελικού μοντέλου. Μελλοντικές έρευνες μπορούν να αξιοποιήσουν πολυετείς και εποχικές χρονοσειρές, ημερήσιες μέσες τιμές, πρόσθετους ρύπους, μετεωρολογικές παραμέτρους και εναλλακτικά κριτήρια προσαρμογής. Με τον τρόπο αυτό θα είναι δυνατή η αξιολόγηση της χρονικής σταθερότητας των ευρημάτων και η πληρέστερη ερμηνεία των διαφορών μεταξύ των πόλεων, ιδιαίτερα της διαφορετικής κατανομικής συμπεριφοράς που παρατηρήθηκε στη Χαλκίδα.

BIBΛΙΟΓΡΑΦΙΑ

- Bağcı, K. (2023). Application of statistical distributions to PM_{10} concentrations: Van, Türkiye. *Yüzüncü Yıl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 60, 87–95. <https://doi.org/10.53568/yyusbed.1214637>
- Giavis, G. M., Kambezidis, H. D., & Lykoudis, S. P. (2009). Frequency distribution of particulate matter (PM_{10}) in urban environments. *International Journal of Environment and Pollution*, 36(1–3), 99–109. <https://doi.org/10.1504/ijep.2009.021819>
- Lu, H.-C., & Fang, G.-C. (2003). Predicting the exceedances of a critical PM_{10} concentration—A case study in Taiwan. *Atmospheric Environment*, 37(25), 3491–3499. [https://doi.org/10.1016/s1352-2310\(03\)00453-9](https://doi.org/10.1016/s1352-2310(03)00453-9)
- Plocoste, T., Calif, R., Euphrasie-Clotilde, L., & Brute, F.-N. (2020). The statistical behavior of PM_{10} events over Guadeloupean archipelago: Stationarity, modelling and extreme events. *Atmospheric Research*, 241, 104956. <https://doi.org/10.1016/j.atmosres.2020.104956>
- Wang, X., Chen, R.-J., Chen, B.-H., & Kan, H.-D. (2013). Application of statistical distribution of PM_{10} concentration in air quality management in 5 representative cities of China. *Biomedical and Environmental Sciences*, 26(8), 638–646. <https://doi.org/10.3967/0895-3988.2013.08.002>

Βέλτιστη κατανομή ατμοσφαιρικών ρύπων στις πόλεις της Στερεάς Ελλάδας

Συγγραφέας: Ιωάννα Ντζάθα

Επιβλέπων καθηγητής: Νικόλαος Ραφομανίκης

Συνεπιβλέπων καθηγητής: Δημήτριος Ναλμπάντης

Ερευνητικό πλαίσιο και σκοπός

Τα αιωρούμενα σωματίδια PM_{10} αποτελούν έναν από τους σημαντικότερους ατμοσφαιρικούς ρύπους, καθώς επηρεάζουν την ανθρώπινη υγεία, το περιβάλλον και την ποιότητα του αέρα.

Παρότι οι συγκεντρώσεις PM_{10} παρακολουθούνται συστηματικά, οι πόλεις της ελληνικής περιφέρειας δεν έχουν αποτελέσει εξίσου συχνό αντικείμενο στατιστικής διερεύνησης σε σύγκριση με τα μεγάλα αστικά κέντρα.

Στόχος: Προσδιορισμός της θεωρητικής κατανομής που αποδίδει καταλληλότερα την εμπειρική συμπεριφορά των ωριαίων συγκεντρώσεων PM_{10} στις πόλεις της Στερεάς Ελλάδας.

Περιοχή μελέτης και δεδομένα

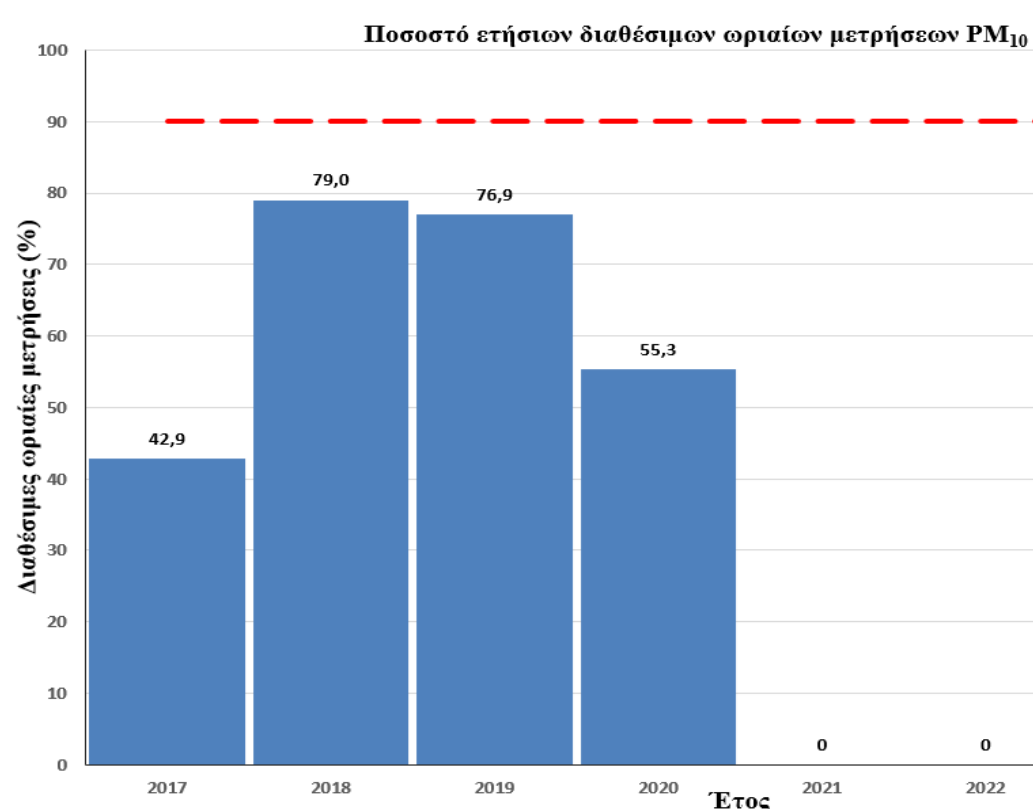


- Χρονική περίοδος: 2017-2024
- 5 σταθμοί παρακολούθησης
- Δεδομένα: ωριαίες συγκεντρώσεις αιωρούμενων σωματιδίων PM_{10}
- Πηγή: Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης (ΕΔΠΑΡ)

Μεθοδολογία

1. **Συλλογή** ωριαίων συγκεντρώσεων PM_{10} από το ΕΔΠΑΡ.
2. **Καθαρισμός και έλεγχος** πληρότητας.
3. **Επιλογή** ετήσιων χρονοσειρών προς ανάλυση.
4. **Προσαρμογή** των κατανομών Normal, Gamma, Weibull και Lognormal.
5. **Αξιολόγηση** με βάση το κριτήριο χ^2 και τα διαγνωστικά γραφήματα.

Έλεγχος πληρότητας δεδομένων: Καρπενήσι



01 Κριτήριο ένταξης:

Ετήσια πληρότητα $\geq 90\%$.

02 Καρπενήσι: Καμία ετήσια χρονοσειρά δεν πληροί το κριτήριο.

03 Αποτέλεσμα: Δεν εντάχθηκε στη διαδικασία προσαρμογής των κατανομών Normal, Gamma, Weibull και Lognormal

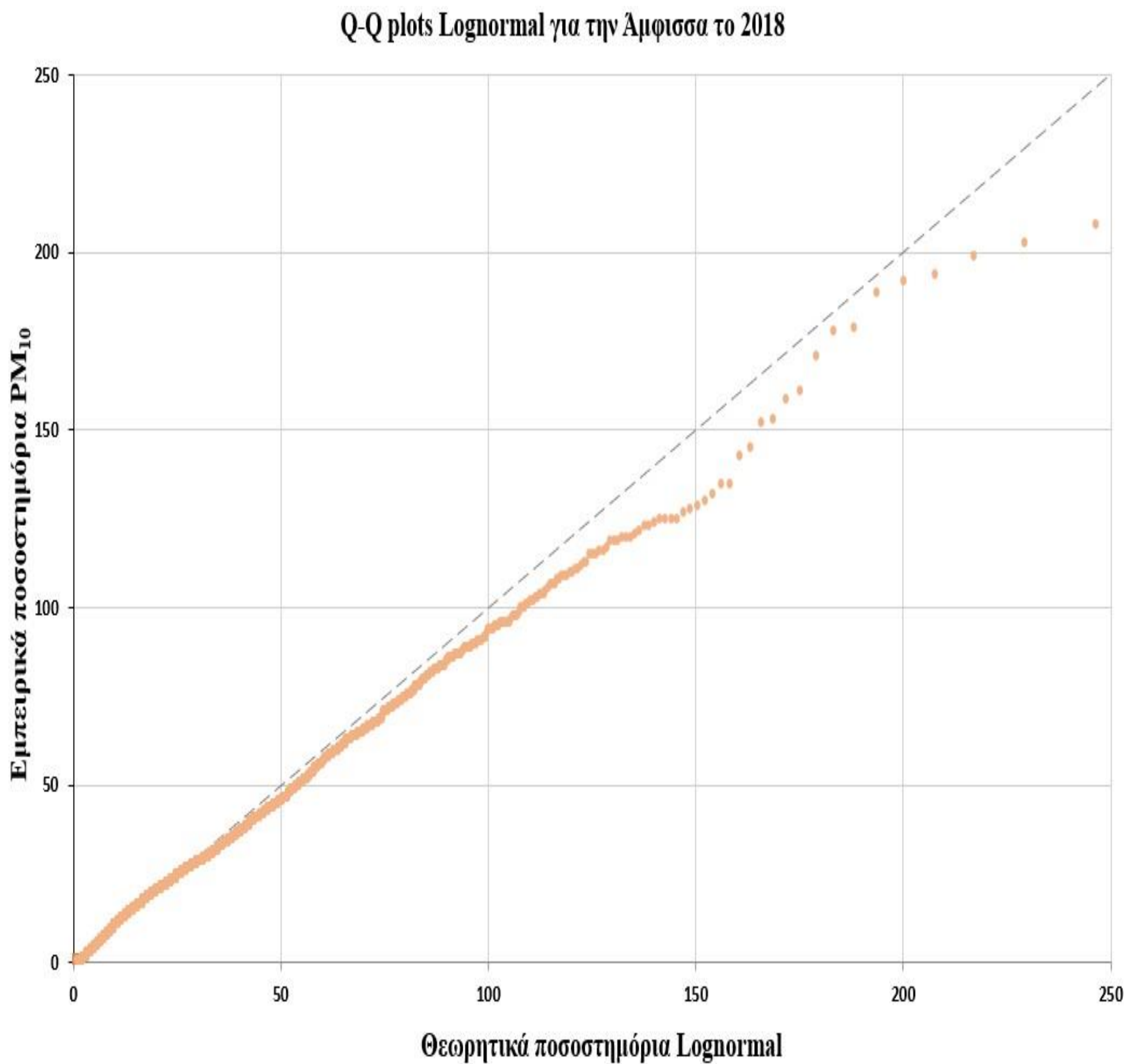
Τελική επιλογή χρονοσειρών PM_{10}

Μετά τον έλεγχο πληρότητας και χρονικής συνέχειας επιλέχθηκε προς ανάλυση μία ετήσια χρονοσειρά ανά πόλη.

Πόλη	Έτος	Έγκυρες ώρες	Πληρότητα (%)
Αμφισσα	2018	8750	99,89
Λαμία	2017	8059	92,00
Λειβαδιά	2019	8745	99,83
Χαλκίδα	2017	8702	99,34
Καρπενήσι	-	-	Μη αναλύσιμο

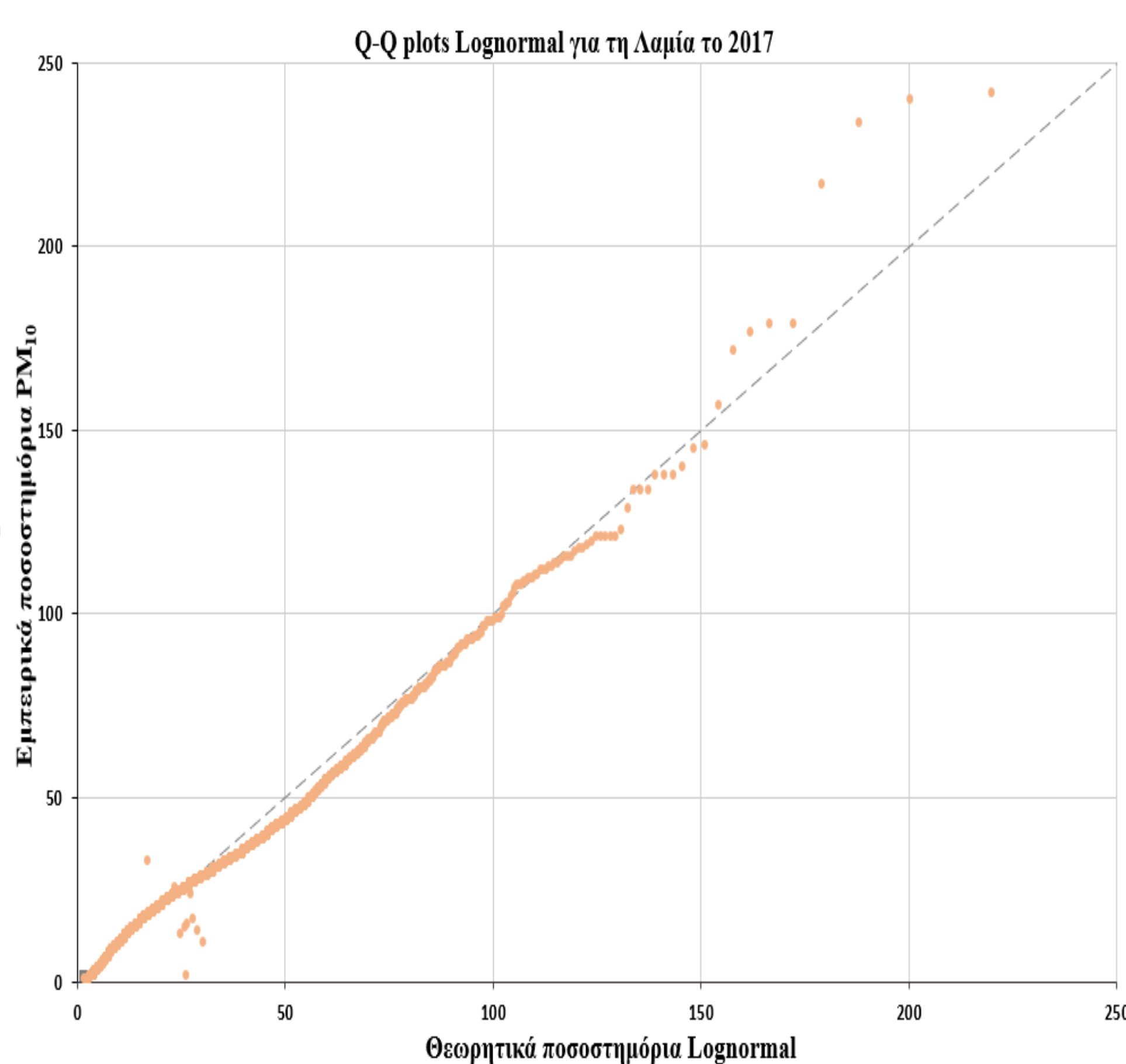
Κύρια αποτελέσματα προσαρμογής των θεωρητικών κατανομών

Α. Άμφισσα, 2018



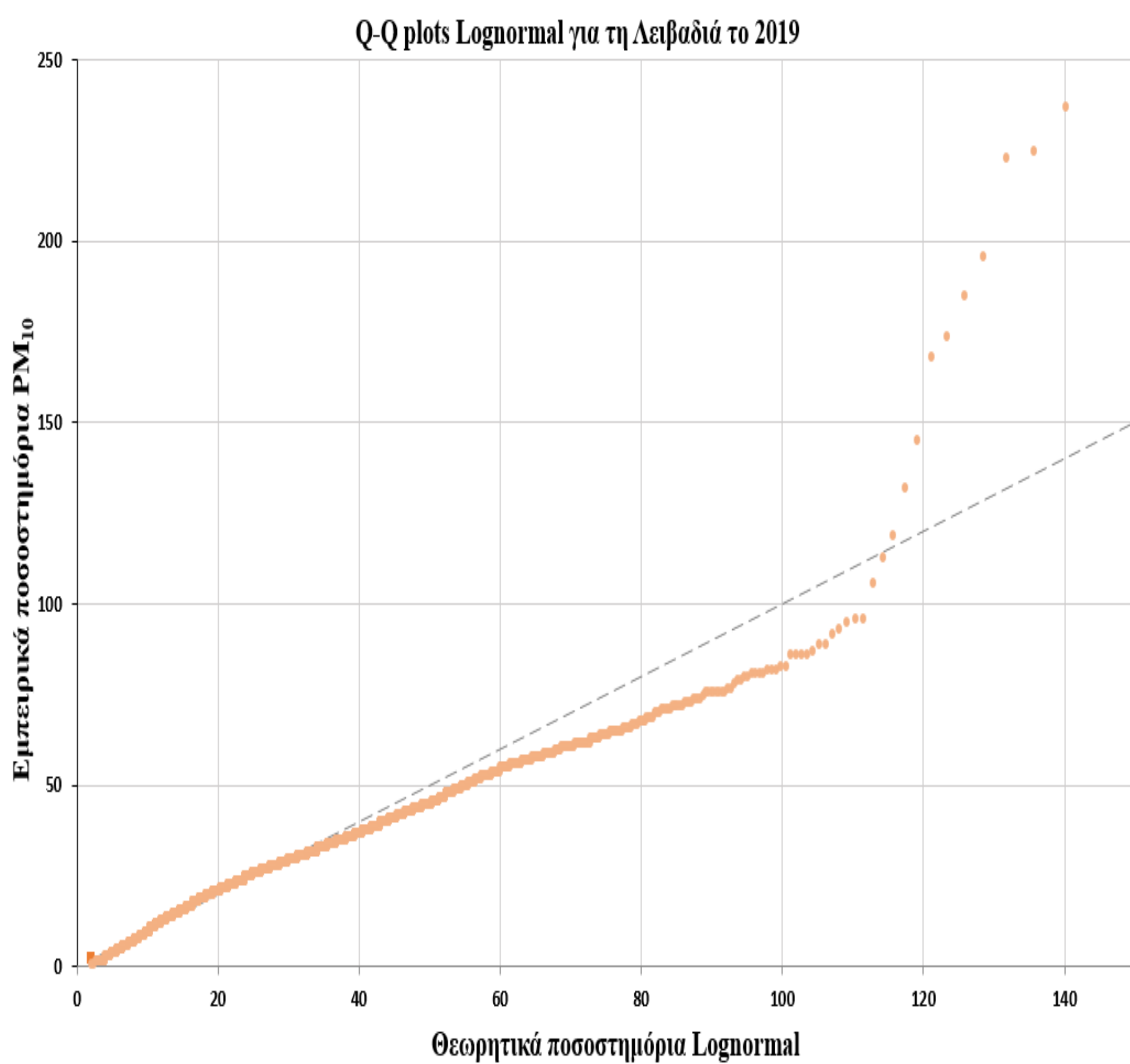
Η Lognormal εμφανίζει την καλύτερη συνολικά συμφωνία με τις παρατηρούμενες συχνότητες.

Β. Λαμία, 2017



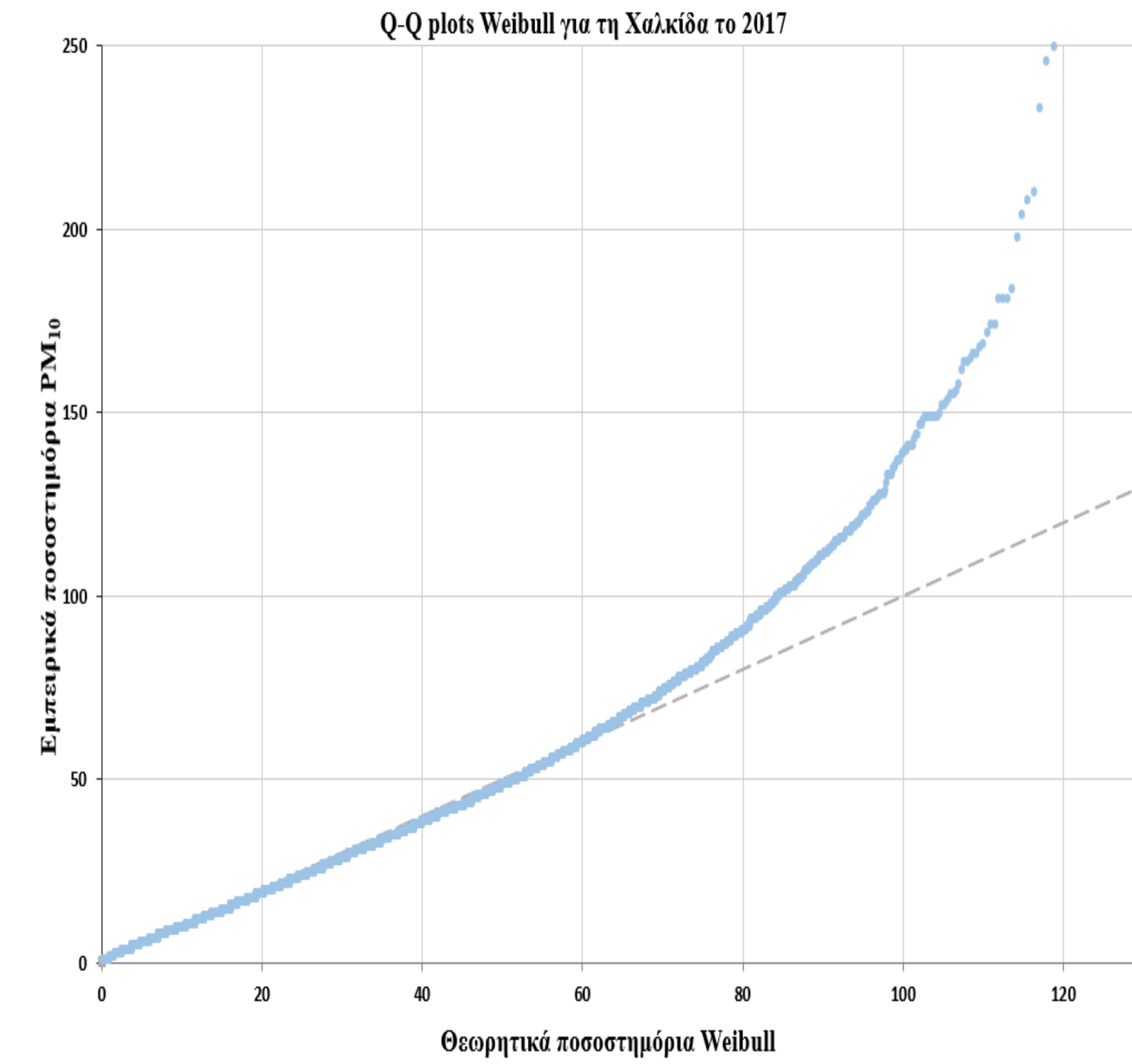
Η Lognormal εμφανίζει την πλησιέστερη ευθυγράμμιση των ποσοστημορίων στο μεγαλύτερο μέρος της κατανομής.

Γ. Λειβαδιά, 2019



Η Lognormal παρουσιάζει καλή συνολική συμφωνία, με αποκλίσεις κυρίως στις υψηλότερες συγκεντρώσεις.

Δ. Χαλκίδα, 2017



Η Weibull παρέχει την καλύτερη συνολικά προσαρμογή στα εμπειρικά ποσοστημόρια.

Κύριο εύρημα: Οι ασύμμετρες θεωρητικές κατανομές περιέγραψαν αποτελεσματικότερα τις ωριαίες συγκεντρώσεις PM_{10} από την κανονική (Normal) κατανομή.

Η Lognormal αναδείχθηκε ως η καταλληλότερη κατανομή σε 3 από τις 4 αναλύσιμες χρονοσειρές, ενώ στη Χαλκίδα υπερίσχυσε η Weibull.

Καταλληλότερη θεωρητική κατανομή

Πόλη	Έτος	χ^2	Καταλληλότερη κατανομή
Αμφισσα	2018	96,26	Lognormal
Λαμία	2017	227,85	Lognormal
Λειβαδιά	2019	192,77	Lognormal
Χαλκίδα	2017	46,09	Weibull

Αριθμητική αξιολόγηση προσαρμογής

01 - Κριτήριο χ^2 . Βασικό αριθμητικό κριτήριο κατάταξης.

02 - Διαγνωστικά γραφήματα. Σύγκριση:

- ✓ Observed-Expected
- ✓ Residual plots
- ✓ Q-Q plots

Αξιολόγηση τοπικών και συστηματικών αποκλίσεων της προσαρμογής

03 - Τελική επιλογή κατανομής. Συνεκτίμηση αριθμητικών και γραφικών ενδείξεων.

Κύρια ευρήματα

01 - Οι κατανομές Gamma, Weibull και Lognormal παρουσίασαν σαφώς καλύτερη προσαρμογή στις ωριαίες συγκεντρώσεις PM_{10} σε σχέση με την κανονική κατανομή (Normal).

02 - Η κατανομή Lognormal αναδείχθηκε ως η καταλληλότερη θεωρητική κατανομή στις 3 από τις 4 αναλύσιμες χρονοσειρές: Άμφισσα, Λαμία και Λειβαδιά, παρουσιάζοντας τη μεγαλύτερη συνολική συνέπεια.

03 - Στη Χαλκίδα, η κατανομή Weibull παρουσίασε την καλύτερη συνολική προσαρμογή.

Συμπεράσματα και προοπτικές

01 - Ερμηνεία

Η διαφοροποίηση της καταλληλότερης κατανομής μεταξύ των χρονοσειρών δείχνει ότι η στατιστική συμπεριφορά των συγκεντρώσεων PM_{10} δεν είναι ενιαία σε όλες τις πόλεις. Η επιλογή του θεωρητικού μοντέλου πρέπει, επομένως, να πραγματοποιείται ξεχωριστά για κάθε χρονοσειρά και όχι να προϋποτίθεται μία κοινή κατανομή.

02 - Μεθοδολογική συμβολή

Η συνεκτίμηση του κριτηρίου χ^2 και των διαγνωστικών γραφημάτων Observed-Expected, Residuals και Q-Q plots παρείχε τεκμηριωμένη και αξιόπιστη αξιολόγηση των υποψήφιων κατανομών. Επίσης περιόρισε την εξάρτηση της τελικής επιλογής από ένα μόνο αριθμητικό κριτήριο.

03 - Προοπτικές

Η μεθοδολογία μπορεί να επεκταθεί σε περισσότερα έτη, επιπλέον ρύπους, μετεωρολογικές μεταβλητές και ευρύτερες γεωγραφικές περιοχές, ώστε να διερευνηθεί συστηματικότερα η χρονική και χωρική διαφοροποίηση των κατανομών.

Βιβλιογραφία

- Loy, A., Follett, L., & Hofmann, H. (2016). Variations of Q-Q Plots: The Power of Our Eyes! The American Statistician, 70(2), 202-214. <https://doi.org/10.1080/00031305.2015.1077728>.
- Lu, H.-C. (2002). The statistical characters of PM_{10} concentration in Taiwan area. Atmospheric Environment, 36(3), 491-502. [https://doi.org/10.1016/S1352-2310\(01\)00245-X](https://doi.org/10.1016/S1352-2310(01)00245-X).
- NIST/SEMATECH. (2012). e-Handbook of Statistical Methods: Chi-Square Goodness-of-Fit Test. National Institute of Standards and Technology. <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35f.htm>.
- Papanastasiou, D. K. & Melas, D. (2010). Application of PM_{10} 's Statistical Distribution to Air Quality Management—A Case Study in Central Greece. , 207(1-4), 115-122. doi:10.1007/s11270-009-0123-8.
- World Health Organization. (2021). WHO global air quality guidelines: Particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. <https://www.who.int/publications/item/9789240034228>.
- Υπουργείο Περιβάλλοντος και Ενέργειας. (χ.χ.). Δεδομένα μετρήσεων ατμοσφαιρικής ρύπανσης. Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης <https://ypen.gov.gr/perivallon/poiotita-tis-atmosfairas/dedomena-metriseon-atmosfairikis-rypansis/>.

1. Εισαγωγή

1.1 Αντικείμενο της εργασίας

Η ατμοσφαιρική ρύπανση αποτελεί ένα από τα σημαντικότερα περιβαλλοντικά προβλήματα της σύγχρονης εποχής, καθώς συνδέεται με σοβαρές επιπτώσεις στην ανθρώπινη υγεία, στο φυσικό περιβάλλον και στην ποιότητα ζωής. Προκύπτει από την παρουσία αέριων και σωματιδιακών ρύπων στην ατμόσφαιρα, οι οποίοι προέρχονται τόσο από ανθρωπογενείς δραστηριότητες, όπως η βιομηχανική παραγωγή, η κυκλοφορία, η θέρμανση και η καύση καυσίμων, όσο και από φυσικές διεργασίες (Ζέρβας & Σέμπος, 2024). Η συστηματική παρακολούθηση και στατιστική ανάλυση των συγκεντρώσεων των ρύπων αποτελεί βασική προϋπόθεση για την κατανόηση της ποιότητας του αέρα και την τεκμηριωμένη αξιολόγηση της περιβαλλοντικής επιβάρυνσης (Κουϊμτζή, Φυτιάνου & Σαμαρά-Κωνσταντίνου, 1998· WHO, 2021).

Στην παρούσα διπλωματική εργασία το ενδιαφέρον επικεντρώνεται στα αιωρούμενα σωματίδια με αεροδυναμική διάμετρο μικρότερη ή ίση με 10 μm , γνωστά ως PM_{10} . Τα PM_{10} αποτελούν εισπνεύσιμο κλάσμα της σωματιδιακής ρύπανσης και παρουσιάζουν ιδιαίτερη σημασία για τη δημόσια υγεία, καθώς μπορούν να εισέλθουν στο αναπνευστικό σύστημα και έχουν συσχετιστεί με αναπνευστικές και καρδιαγγειακές επιβαρύνσεις, ιδιαίτερα σε ευάλωτες ομάδες του πληθυσμού (Κουϊμτζή, Φυτιάνου & Σαμαρά-Κωνσταντίνου, 1998· WHO, 2021· Ζέρβας & Σέμπος, 2024). Παράλληλα, επηρεάζουν την ορατότητα, συμμετέχουν σε διεργασίες εναπόθεσης και συμβάλλουν στην υποβάθμιση της ποιότητας του ατμοσφαιρικού περιβάλλοντος (Seinfeld & Pandis, 2016).

Αντικείμενο της εργασίας είναι η στατιστική διερεύνηση των ωριαίων συγκεντρώσεων PM_{10} σε επιλεγμένες πόλεις της Στερεάς Ελλάδας, με σκοπό την περιγραφή των βασικών χαρακτηριστικών τους και την αξιολόγηση θεωρητικών κατανομών πιθανότητας ως προς την ικανότητά τους να αναπαράγουν την εμπειρική μορφή των δεδομένων. Η ανάλυση δεν περιορίζεται στην αποτύπωση μέτρων κεντρικής τάσης και διασποράς, αλλά εξετάζει τη συνολική κατανομική συμπεριφορά των παρατηρήσεων, συμπεριλαμβανομένης της ασυμμετρίας και της παρουσίας υψηλών τιμών (Wilks, 2019· Lu & Fang, 2003· Sansuddin et al., 2010).

Η επιλογή της Στερεάς Ελλάδας ως πεδίου μελέτης συνδέεται με το γεγονός ότι η περιοχή περιλαμβάνει πόλεις διαφορετικού μεγέθους, γεωγραφικής θέσης και λειτουργικού χαρακτήρα. Σε αντίθεση με μεγάλα μητροπολιτικά κέντρα, όπως η Αθήνα και η

Θεσσαλονίκη, οι περιφερειακές πόλεις δεν έχουν διερευνηθεί στον ίδιο βαθμό ως προς τη στατιστική συμπεριφορά των συγκεντρώσεων PM_{10} . Παράλληλα, μπορεί να επηρεάζονται από διαφορετικούς συνδυασμούς τοπικών εκπομπών, κυκλοφοριακής δραστηριότητας, οικιακής θέρμανσης, παραγωγικών ή αγροτικών δραστηριοτήτων και φυσικών διεργασιών, όπως η μεταφορά σκόνης (Diapoulis et al., 2017· ΥΠΕΝ, χ.χ.). Η εξέταση των διαθέσιμων μετρήσεων συμβάλλει, επομένως, στην ανάδειξη της περιφερειακής διάστασης της ατμοσφαιρικής ρύπανσης.

Η εστίαση στα PM_{10} είναι επιστημονικά και μεθοδολογικά τεκμηριωμένη. Αφενός, αποτελούν βασικό δείκτη ποιότητας του αέρα λόγω των επιπτώσεών τους στην υγεία και στο περιβάλλον. Αφετέρου, κατά την προκαταρκτική διερεύνηση των διαθέσιμων δεδομένων διαπιστώθηκε ότι οι χρονοσειρές PM_{10} παρείχαν μεγαλύτερη δυνατότητα συστηματικής στατιστικής αξιοποίησης σε σχέση με άλλους ρύπους. Η επιλογή ενός ενιαίου ρύπου ενισχύει τη συνοχή της ανάλυσης και επιτρέπει την εφαρμογή κοινής μεθοδολογίας στις επιλεγμένες χρονοσειρές.

Δεδομένου ότι οι συγκεντρώσεις PM_{10} είναι μη αρνητικές και συχνά παρουσιάζουν θετική ασυμμετρία, εξετάζονται η λογαριθμοκανονική, η Gamma και η Weibull, καθώς και η κανονική κατανομή ως συγκριτική βάση. Οι κατανομές αυτές αξιολογούνται ως προς τη συμφωνία τους με τα εμπειρικά δεδομένα, χωρίς καμία από αυτές να θεωρείται εκ των προτέρων καταλληλότερη (Wilks, 2019· Lu & Fang, 2003· Sansuddin et al., 2010· Plocoste et al., 2020). Με τον τρόπο αυτό, η εργασία συνδέει την περιβαλλοντική παρακολούθηση με τη στατιστική μοντελοποίηση και επιδιώκει να συμβάλει στην καλύτερη κατανόηση της κατανομικής συμπεριφοράς των PM_{10} σε περιφερειακές πόλεις.

1.2 Σκοπός της εργασίας

Σκοπός της παρούσας διπλωματικής εργασίας είναι η στατιστική διερεύνηση των ωριαίων συγκεντρώσεων PM_{10} σε επιλεγμένες πόλεις της Στερεάς Ελλάδας και ο προσδιορισμός της θεωρητικής κατανομής πιθανότητας που παρουσιάζει την καλύτερη προσαρμογή στα εμπειρικά δεδομένα, μεταξύ των μοντέλων που εξετάζονται. Η εργασία επικεντρώνεται στην κατανομική μορφή των συγκεντρώσεων και στη διερεύνηση χαρακτηριστικών όπως

η κεντρική τάση, η διασπορά, η ασυμμετρία και η παρουσία υψηλών παρατηρήσεων (Wilks, 2019· Lu & Fang, 2003· Sansuddin et al., 2010).

Η ανάλυση βασίζεται σε διαθέσιμες χρονοσειρές PM_{10} από σταθμούς παρακολούθησης της ποιότητας του αέρα στις εξεταζόμενες πόλεις, οι οποίες αντλήθηκαν από το Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης (ΥΠΕΝ, χ.χ.). Ιδιαίτερη έμφαση δίνεται στον έλεγχο της πληρότητας, της συνέχειας και της εγκυρότητας των δεδομένων, καθώς η επάρκεια των χρονοσειρών αποτελεί βασική προϋπόθεση για την εφαρμογή της στατιστικής ανάλυσης και την εξαγωγή αξιόπιστων συμπερασμάτων. Για κάθε πόλη επιλέγεται το πληρέστερο διαθέσιμο έτος που ικανοποιεί το καθορισμένο κριτήριο πληρότητας, χωρίς να θεωρείται ότι το συγκεκριμένο έτος αντιπροσωπεύει κατ' ανάγκη τη μακροχρόνια συμπεριφορά της περιοχής.

Οι επιμέρους στόχοι της εργασίας είναι:

- α) η συλλογή, οργάνωση, προετοιμασία και ομογενοποίηση των διαθέσιμων δεδομένων PM_{10} για τις εξεταζόμενες πόλεις και χρονικές περιόδους,
- β) ο έλεγχος της πληρότητας και της ποιότητας των χρονοσειρών και η επιλογή εκείνων που μπορούν να ενταχθούν αξιόπιστα στην ανάλυση,
- γ) η περιγραφική στατιστική διερεύνηση των συγκεντρώσεων PM_{10} , με σκοπό την αποτύπωση της κεντρικής τάσης, της διασποράς, της ασυμμετρίας και των υψηλών παρατηρήσεων ανά πόλη και έτος,
- δ) η εφαρμογή της κανονικής, της λογαριθμοκανονικής, της Gamma και της Weibull κατανομής στις επιλεγμένες χρονοσειρές,
- ε) ο υπολογισμός των θεωρητικά αναμενόμενων συχνοτήτων και η σύγκρισή τους με τις αντίστοιχες εμπειρικές συχνότητες,
- στ) η αξιολόγηση της προσαρμογής μέσω του συγκριτικού στατιστικού χ^2 , των Q-Q διαγραμμάτων, των διαγραμμάτων παρατηρούμενων-αναμενόμενων συχνοτήτων και των διαγραμμάτων υπολοίπων,
- ζ) η επιλογή της καταλληλότερης κατανομής για κάθε αναλύσιμη χρονοσειρά με βάση τη συνδυαστική αξιολόγηση των αριθμητικών και γραφικών ενδείξεων,
- η) η συγκριτική αποτίμηση των αποτελεσμάτων μεταξύ των εξεταζόμενων πόλεων, στο μέτρο που το επιτρέπουν η πληρότητα, η ποιότητα και η συγκρισιμότητα των διαθέσιμων δεδομένων.

Μέσω των παραπάνω στόχων, η εργασία επιδιώκει να διαμορφώσει ένα συνεκτικό και διαφανές πλαίσιο στατιστικής ανάλυσης των PM_{10} σε περιφερειακή κλίμακα. Η επιλογή της καταλληλότερης κατανομής δεν αντιμετωπίζεται ως απόδειξη μιας μοναδικής ή «αληθινής» κατανομής των συγκεντρώσεων, αλλά ως συγκριτική αξιολόγηση των συγκεκριμένων μοντέλων στα διαθέσιμα δεδομένα. Παράλληλα, η σύνδεση της ποιότητας των χρονοσειρών με την ερμηνεία των αποτελεσμάτων περιορίζει την εξαγωγή γενικεύσεων που δεν υποστηρίζονται από το εύρος και την πληρότητα των παρατηρήσεων.

1.3 Συνεισφορά της εργασίας

Η παρούσα διπλωματική εργασία συμβάλλει στη μελέτη της ατμοσφαιρικής ρύπανσης σε περιφερειακή κλίμακα, εστιάζοντας στις ωριαίες συγκεντρώσεις αιωρούμενων σωματιδίων PM_{10} σε επιλεγμένες πόλεις της Στερεάς Ελλάδας. Η συνεισφορά της εντοπίζεται αρχικά στη μετατόπιση του ερευνητικού ενδιαφέροντος από τα μεγάλα μητροπολιτικά και έντονα αστικοποιημένα κέντρα της χώρας, τα οποία έχουν αποτελέσει συχνότερο πεδίο μελέτης της σωματιδιακής ρύπανσης, προς μικρομεσαίες πόλεις της περιφέρειας, οι οποίες δεν έχουν διερευνηθεί στον ίδιο βαθμό ως προς τη στατιστική συμπεριφορά των συγκεντρώσεων PM_{10} (Diarouli et al., 2017· Argyropoulos et al., 2017). Με τον τρόπο αυτό, η εργασία αναδεικνύει την ανάγκη συστηματικότερης αξιοποίησης των δεδομένων ποιότητας του αέρα σε περιοχές που, παρότι δεν αποτελούν μεγάλους αστικούς πυρήνες, μπορεί να επηρεάζονται από διαφορετικούς συνδυασμούς τοπικών εκπομπών, κυκλοφοριακής δραστηριότητας, θέρμανσης, παραγωγικών δραστηριοτήτων, φυσικών πηγών και ατμοσφαιρικών συνθηκών.

Πέρα από τη γεωγραφική εστίαση, η εργασία παρουσιάζει και μεθοδολογική συνεισφορά, καθώς εξετάζει την κατανομική μορφή των ωριαίων συγκεντρώσεων PM_{10} . Η προσέγγιση αυτή υπερβαίνει την αποκλειστική χρήση βασικών περιγραφικών δεικτών και επιτρέπει τη διερεύνηση χαρακτηριστικών όπως η ασυμμετρία, η διασπορά και η παρουσία υψηλών τιμών. Για τον σκοπό αυτό εφαρμόζονται η κανονική, η λογαριθμοκανονική, η Gamma και η Weibull κατανομή, χωρίς καμία να θεωρείται εκ των προτέρων καταλληλότερη. Η συγκριτική αξιολόγησή τους προσφέρει ένα πιο ολοκληρωμένο πλαίσιο περιγραφής της

εμπειρικής συμπεριφοράς των συγκεντρώσεων PM_{10} (Lu, 2002· Lu & Fang, 2003· Sansuddin et al., 2010· Plocoste et al., 2020).

Σημαντική συμβολή αποτελεί επίσης η συστηματική αξιοποίηση των διαθέσιμων χρονοσειρών PM_{10} , με έμφαση στον έλεγχο της πληρότητας, της συνέχειας και της ποιότητας των δεδομένων (ΥΠΕΝ, χ.χ· Wilks, 2019). Η διαδικασία αυτή είναι κρίσιμη, καθώς η αξιοπιστία της στατιστικής ανάλυσης εξαρτάται άμεσα από την επάρκεια και τη συγκρισιμότητα των πρωτογενών παρατηρήσεων. Η εργασία δεν εντάσσει αυτομάτως όλες τις διαθέσιμες χρονοσειρές στην ανάλυση, αλλά εφαρμόζει σαφές κριτήριο πληρότητας και επιλέγει μόνο εκείνες που παρέχουν επαρκή βάση για συγκριτική αξιολόγηση. Η περίπτωση του Καρπενησίου αναδεικνύει ιδιαίτερα τη σημασία αυτής της μεθοδολογικής επιλογής.

Τέλος, η εργασία διαμορφώνει ένα συνεκτικό πλαίσιο συνδυαστικής αξιολόγησης θεωρητικών κατανομών σε περιβαλλοντικά δεδομένα. Η επιλογή της καταλληλότερης κατανομής πραγματοποιείται με βάση το συγκριτικό στατιστικό χ^2 και τη γραφική αξιολόγηση μέσω διαγραμμάτων παρατηρούμενων-αναμενόμενων συχνοτήτων, υπολοίπων και Q-Q. Με τον τρόπο αυτό, η τελική κρίση δεν βασίζεται σε ένα μόνο αριθμητικό μέτρο, αλλά στη συνολική εικόνα της προσαρμογής. Η προσέγγιση αυτή προσφέρει χρήσιμο μεθοδολογικό υπόβαθρο για μελλοντικές αναλύσεις ποιότητας του αέρα σε περιφερειακές περιοχές, χωρίς να υποστηρίζει ότι μία κατανομή αποτελεί καθολικά ορθή περιγραφή των συγκεντρώσεων PM_{10} .

1.4 Ερευνητικά Ερωτήματα

Με βάση το αντικείμενο, τον σκοπό και τη συνεισφορά της εργασίας, η έρευνα οργανώνεται γύρω από τέσσερα βασικά ερευνητικά ερωτήματα. Τα ερωτήματα καλύπτουν διαδοχικά τον έλεγχο της καταλληλότητας των διαθέσιμων δεδομένων, την περιγραφική διερεύνηση των συγκεντρώσεων, την εφαρμογή των θεωρητικών κατανομών και τη συγκριτική αξιολόγηση της προσαρμογής τους. Παράλληλα, οριοθετούν σαφώς το πεδίο της ανάλυσης, το οποίο αφορά την κατανομική συμπεριφορά των ωριαίων

συγκεντρώσεων PM_{10} και όχι την αιτιώδη διερεύνηση των πηγών ή των μετεωρολογικών μηχανισμών που τις διαμορφώνουν.

Ειδικότερα, η εργασία επιχειρεί να απαντήσει στα ακόλουθα ερευνητικά ερωτήματα:

1. Ποιες από τις διαθέσιμες χρονοσειρές ωριαίων συγκεντρώσεων PM_{10} στις εξεταζόμενες πόλεις της Στερεάς Ελλάδας διαθέτουν επαρκή πληρότητα και ποιότητα ώστε να ενταχθούν αξιόπιστα στη στατιστική ανάλυση;
2. Ποια είναι τα βασικά περιγραφικά χαρακτηριστικά των επιλεγμένων χρονοσειρών PM_{10} ως προς την κεντρική τάση, τη διασπορά, την ασυμμετρία και την παρουσία υψηλών τιμών;
3. Ποια από τις εξεταζόμενες κατανομές, κανονική, λογαριθμοκανονική, Gamma και Weibull, παρουσιάζει την καλύτερη προσαρμογή σε καθεμία από τις επιλεγμένες χρονοσειρές, με βάση τη συνδυαστική αξιολόγηση των γραφικών και αριθμητικών ενδείξεων;
4. Διαφοροποιείται η καταλληλότερη κατανομή μεταξύ των επιλεγμένων χρονοσειρών πόλης-έτους και τι υποδηλώνει η διαφοροποίηση αυτή για την κατανομική συμπεριφορά των συγκεντρώσεων PM_{10} ;

Η διατύπωση των ερωτημάτων αυτών επιτρέπει τη συνεπή σύνδεση της εισαγωγής με τη μεθοδολογία, την παρουσίαση των αποτελεσμάτων και τα τελικά συμπεράσματα. Το πρώτο ερώτημα αφορά την επιλογή των δεδομένων, το δεύτερο την περιγραφική τους αποτύπωση, το τρίτο τη συγκριτική εφαρμογή των θεωρητικών μοντέλων και το τέταρτο τη σύνθεση των διαφορών μεταξύ των εξεταζόμενων χρονοσειρών. Έτσι, κάθε ερευνητικό ερώτημα αντιστοιχεί άμεσα σε συγκεκριμένο στάδιο της ανάλυσης και μπορεί να απαντηθεί με σαφήνεια στα επόμενα κεφάλαια.

1.5 Δομή της εργασίας

Η παρούσα διπλωματική εργασία οργανώνεται σε πέντε κεφάλαια, τα οποία ακολουθούν τη λογική πορεία της ερευνητικής διαδικασίας, από τη θεωρητική τεκμηρίωση του αντικειμένου και την παρουσίαση της μεθοδολογίας έως τη στατιστική ανάλυση, την ερμηνεία των αποτελεσμάτων και τη διατύπωση των τελικών συμπερασμάτων. Η

διάρθρωση αυτή επιτρέπει τη σταδιακή μετάβαση από το γενικότερο πλαίσιο της ατμοσφαιρικής ρύπανσης στη συγκεκριμένη διερεύνηση της κατανομικής συμπεριφοράς των ωριαίων συγκεντρώσεων PM_{10} στις επιλεγμένες πόλεις της Στερεάς Ελλάδας.

Το πρώτο κεφάλαιο αποτελεί την εισαγωγή της εργασίας. Σε αυτό παρουσιάζονται το αντικείμενο της μελέτης, ο σκοπός, η συνεισφορά, τα ερευνητικά ερωτήματα και η συνολική διάρθρωση της διπλωματικής εργασίας. Με τον τρόπο αυτό οριοθετείται το ερευνητικό πεδίο και τεκμηριώνεται η ανάγκη στατιστικής διερεύνησης των συγκεντρώσεων PM_{10} σε περιφερειακή κλίμακα.

Το δεύτερο κεφάλαιο περιλαμβάνει το θεωρητικό και βιβλιογραφικό υπόβαθρο της έρευνας. Παρουσιάζονται βασικές έννοιες που αφορούν την ατμοσφαιρική ρύπανση, την ποιότητα του αέρα και τα αιωρούμενα σωματίδια PM_{10} , καθώς και ζητήματα σχετικά με τις πηγές, τις επιπτώσεις και τους παράγοντες μεταβλητότητας των συγκεντρώσεών τους. Παράλληλα, αναπτύσσεται το στατιστικό υπόβαθρο της εργασίας, παρουσιάζονται η κανονική, η λογαριθμοκανονική, η Gamma και η Weibull κατανομή και εξετάζονται οι μέθοδοι αξιολόγησης της προσαρμογής τους στα εμπειρικά δεδομένα. Το κεφάλαιο ολοκληρώνεται με ανασκόπηση προηγούμενων μελετών που έχουν εφαρμόσει θεωρητικές κατανομές σε δεδομένα PM_{10} .

Το τρίτο κεφάλαιο παρουσιάζει τα δεδομένα και τη μεθοδολογία της έρευνας. Περιγράφονται η προέλευση των μετρήσεων από το Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης, η διαδικασία καθαρισμού και οργάνωσης των δεδομένων, ο έλεγχος πληρότητας και τα κριτήρια επιλογής των αναλύσιμων χρονοσειρών. Στη συνέχεια παρουσιάζονται η δημιουργία των εμπειρικών συχνοτήτων, η εφαρμογή των τεσσάρων θεωρητικών κατανομών και η διαδικασία αξιολόγησης και επιλογής της καταλληλότερης κατανομής για κάθε εξεταζόμενη χρονοσειρά.

Το τέταρτο κεφάλαιο περιλαμβάνει την παρουσίαση και ανάλυση των αποτελεσμάτων. Αρχικά παρουσιάζονται η πληρότητα των διαθέσιμων δεδομένων, τα βασικά περιγραφικά χαρακτηριστικά των επιλεγμένων χρονοσειρών και οι εμπειρικές κατανομές συχνοτήτων. Στη συνέχεια αξιολογείται η προσαρμογή της κανονικής, της λογαριθμοκανονικής, της Gamma και της Weibull κατανομής μέσω της σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων, των διαγραμμάτων παρατηρούμενων-αναμενόμενων τιμών, των διαγραμμάτων υπολοίπων, των Q-Q διαγραμμάτων και του συγκριτικού στατιστικού χ^2 . Τα αποτελέσματα παρουσιάζονται χωριστά για κάθε επιλεγμένη χρονοσειρά πόλης-

έτους και ακολουθεί συνολική συγκριτική αποτίμηση. Ιδιαίτερη αναφορά γίνεται στην περίπτωση του Καρπενησίου, το οποίο αποκλείστηκε από την κατανομική ανάλυση λόγω ανεπαρκούς πληρότητας των διαθέσιμων δεδομένων.

Το πέμπτο κεφάλαιο περιλαμβάνει τα συμπεράσματα της εργασίας. Συνοψίζονται τα βασικά ευρήματα, δίνονται απαντήσεις στα ερευνητικά ερωτήματα και αποτιμάται η καταλληλότητα των εξεταζόμενων θεωρητικών κατανομών για καθεμία από τις επιλεγμένες χρονοσειρές. Παράλληλα, παρουσιάζονται η μεθοδολογική αποτίμηση της προσέγγισης, οι βασικοί περιορισμοί της έρευνας και προτάσεις για μελλοντική διερεύνηση με περισσότερες χρονικές περιόδους, πρόσθετους ρύπους και συμπληρωματικές περιβαλλοντικές μεταβλητές.

2. Θεωρητικό υπόβαθρο: ατμοσφαιρική ρύπανση, PM_{10} και στατιστική μοντελοποίηση

Η διερεύνηση των συγκεντρώσεων PM_{10} στις επιλεγμένες πόλεις της Στερεάς Ελλάδας προϋποθέτει τη διαμόρφωση ενός σαφούς θεωρητικού και βιβλιογραφικού πλαισίου, το οποίο συνδέει την περιβαλλοντική διάσταση της ατμοσφαιρικής ρύπανσης με τη στατιστική ανάλυση των δεδομένων. Αρχικά παρουσιάζονται οι βασικές έννοιες που αφορούν την ατμοσφαιρική ρύπανση, την ποιότητα του αέρα και τα αιωρούμενα σωματίδια, με ιδιαίτερη έμφαση στα PM_{10} . Στη συνέχεια εξετάζονται οι κύριες πηγές προέλευσής τους, οι επιπτώσεις τους στην ανθρώπινη υγεία και στο περιβάλλον, καθώς και οι μετεωρολογικοί, τοπογραφικοί και ανθρωπογενείς παράγοντες που επηρεάζουν τη χωρική και χρονική μεταβλητότητα των συγκεντρώσεων.

Παράλληλα, αναπτύσσεται το στατιστικό υπόβαθρο που απαιτείται για την ανάλυση των ωριαίων δεδομένων PM_{10} . Ιδιαίτερη έμφαση δίνεται στη χρήση θεωρητικών κατανομών πιθανότητας για την περιγραφή της εμπειρικής μορφής των συγκεντρώσεων, καθώς οι τιμές είναι μη αρνητικές και συχνά παρουσιάζουν θετική ασυμμετρία, αυξημένη μεταβλητότητα και περιορισμένο αριθμό υψηλών παρατηρήσεων. Με τον τρόπο αυτό, το κεφάλαιο λειτουργεί ως σύνδεση μεταξύ του περιβαλλοντικού αντικειμένου της εργασίας και της στατιστικής μεθοδολογίας που εφαρμόζεται στα επόμενα κεφάλαια.

2.1 Ατμοσφαιρική ρύπανση και αιωρούμενα σωματίδια PM_{10}

Η ατμοσφαιρική ρύπανση αποτελεί ένα από τα σημαντικότερα και πιο σύνθετα περιβαλλοντικά προβλήματα της σύγχρονης εποχής, καθώς συνδέεται με επιπτώσεις στην ανθρώπινη υγεία, στην ποιότητα ζωής και στο φυσικό περιβάλλον. Η σημασία της εξαρτάται όχι μόνο από την παρουσία ρύπων στην ατμόσφαιρα, αλλά και από τη συγκέντρωσή τους, τη διάρκεια της έκθεσης, τη χημική τους σύσταση και τη δυνατότητα μεταφοράς ή μετασχηματισμού τους. Οι επιπτώσεις της περιλαμβάνουν αναπνευστικές και καρδιαγγειακές επιβαρύνσεις, υποβάθμιση της περιβαλλοντικής ποιότητας και συμμετοχή σε ατμοσφαιρικές και κλιματικές διεργασίες (Ζέρβας & Σέμπος, 2024· WHO, 2021· EEA, 2026). Μεταξύ των βασικών ατμοσφαιρικών ρύπων, τα αιωρούμενα σωματίδια (Particulate Matter - PM) παρουσιάζουν ιδιαίτερη υγειονομική και

περιβαλλοντική σημασία. Πρόκειται για στερεά σωματίδια και υγρά σταγονίδια που αιωρούνται στην ατμόσφαιρα και προέρχονται τόσο από φυσικές πηγές, όπως η εδαφική σκόνη και το θαλάσσιο αλάτι, όσο και από ανθρωπογενείς δραστηριότητες, όπως η καύση καυσίμων, η κυκλοφορία και η βιομηχανική παραγωγή (Κατσαφάδος & Μαυροματίδης, 2024· U.S. EPA, 2026b).

Η παρούσα εργασία επικεντρώνεται στα PM₁₀, δηλαδή στα σωματίδια με αεροδυναμική διάμετρο έως 10 μm. Τα PM₁₀ αποτελούν βασικό δείκτη σωματιδιακής ρύπανσης και παρουσιάζουν έντονη χωρική και χρονική μεταβλητότητα, η οποία συνδέεται με διαφορετικές πηγές και ατμοσφαιρικές συνθήκες (WHO, 2021· EEA, 2026). Η στατιστική διερεύνηση της κατανομής των συγκεντρώσεών τους συμβάλλει στην κατανόηση της συμπεριφοράς των δεδομένων και στη διαμόρφωση κατάλληλου πλαισίου για την αξιολόγηση της ποιότητας του αέρα.

2.1.1 Ορισμός ατμοσφαιρικής ρύπανσης

Η ατμοσφαιρική ρύπανση συνιστά υποβάθμιση της ποιότητας του αέρα και σχετίζεται με την παρουσία χημικών, φυσικών ή βιολογικών παραγόντων σε συγκεντρώσεις ή για χρονικές διάρκειες ικανές να προκαλέσουν δυσμενείς επιπτώσεις στην ανθρώπινη υγεία, στους ζωντανούς οργανισμούς, στα οικοσυστήματα και στο περιβάλλον (Tiwary, Williams & Colls, 2019).

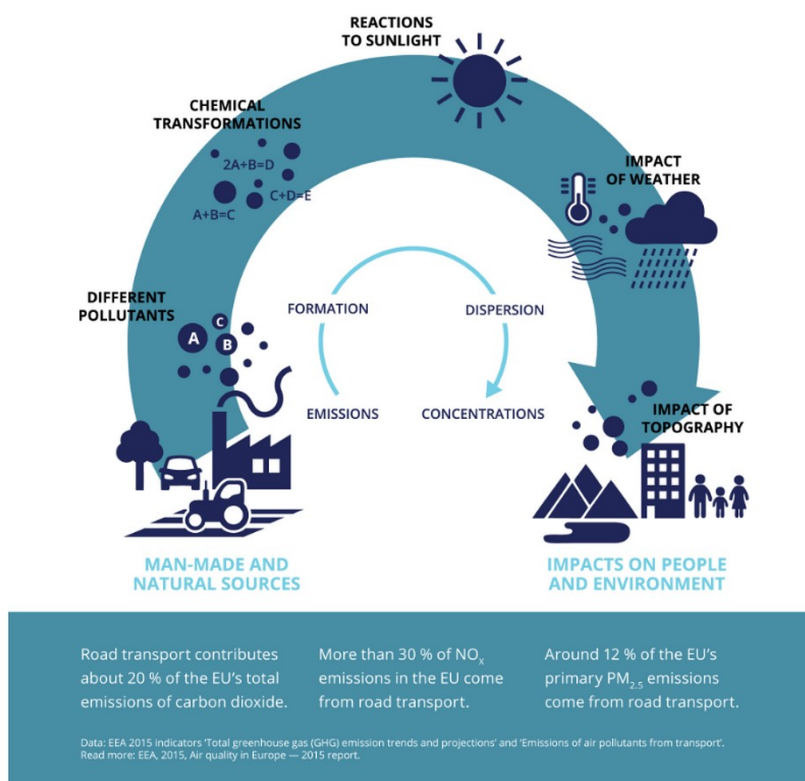
Αντίστοιχα, το Υπουργείο Περιβάλλοντος και Ενέργειας συνδέει την ατμοσφαιρική ρύπανση με την παρουσία ουσιών στην ατμόσφαιρα σε επίπεδα ή για χρονική διάρκεια που μπορούν να προκαλέσουν αρνητικές επιδράσεις στην υγεία και στο περιβάλλον. Ο Παγκόσμιος Οργανισμός Υγείας την περιγράφει ως επιβάρυνση του εσωτερικού ή εξωτερικού αέρα από χημικούς, φυσικούς ή βιολογικούς παράγοντες που μεταβάλλουν τα φυσικά χαρακτηριστικά της ατμόσφαιρας (WHO, 2021).

Στο ευρωπαϊκό θεσμικό πλαίσιο, η έννοια εξειδικεύεται μέσω των ορισμών του «ρύπου» και του «επιπέδου». Η Οδηγία 2008/50/EK ορίζει ως ρύπο κάθε ουσία που εμφανίζεται στον ατμοσφαιρικό αέρα και μπορεί να έχει αρνητικές επιπτώσεις στην ανθρώπινη υγεία ή στο περιβάλλον, ενώ ως επίπεδο νοείται η συγκέντρωση ενός ρύπου στον αέρα ή η εναπόθεσή του σε επιφάνεια σε δεδομένη χρονική στιγμή (ΥΠΕΚΑ, 2011). Η Οδηγία (ΕΕ) 2024/2881 διατηρεί τη βασική αυτή προσέγγιση και την εντάσσει σε αυστηρότερο

πλαίσιο προστασίας της ποιότητας του αέρα και επίτευξης του στόχου μηδενικής ρύπανσης έως το 2050 (European Parliament and Council of the European Union, 2024).

Οι πηγές της ατμοσφαιρικής ρύπανσης διακρίνονται γενικά σε φυσικές και ανθρωπογενείς. Στις ανθρωπογενείς περιλαμβάνονται η καύση καυσίμων, οι μεταφορές, η βιομηχανία, η παραγωγή ενέργειας και η γεωργία. Στις φυσικές εντάσσονται η εδαφική σκόνη, το θαλάσσιο αλάτι, οι πυρκαγιές, οι ηφαιστειακές εκπομπές και η μεταφορά σκόνης. Για τα αιωρούμενα σωματίδια ιδιαίτερη σημασία έχουν τόσο οι άμεσες εκπομπές όσο και ο δευτερογενής σχηματισμός τους στην ατμόσφαιρα (EEA, 2019).

Όπως παρουσιάζεται στο Σχήμα 2.1, οι εκπομπές από φυσικές και ανθρωπογενείς πηγές υπόκεινται σε διεργασίες σχηματισμού, μεταφοράς, χημικού μετασχηματισμού και διασποράς. Οι διεργασίες αυτές επηρεάζονται από την ηλιακή ακτινοβολία, τις μετεωρολογικές συνθήκες και την τοπογραφία, διαμορφώνοντας τελικά τα επίπεδα συγκέντρωσης και τις επιπτώσεις στον άνθρωπο και στο περιβάλλον. Επομένως, η ατμοσφαιρική ρύπανση αποτελεί δυναμική διαδικασία που προκύπτει από την αλληλεπίδραση πολλών παραγόντων.



Σχήμα 2.1 Εννοιολογική απεικόνιση της σχέσης μεταξύ πηγών εκπομπών, ατμοσφαιρικών ρύπων, διεργασιών μεταφοράς/μετασχηματισμού, έκθεσης και επιπτώσεων.

Πηγή: [Προσαρμογή από European Environment Agency, 2016.](#)

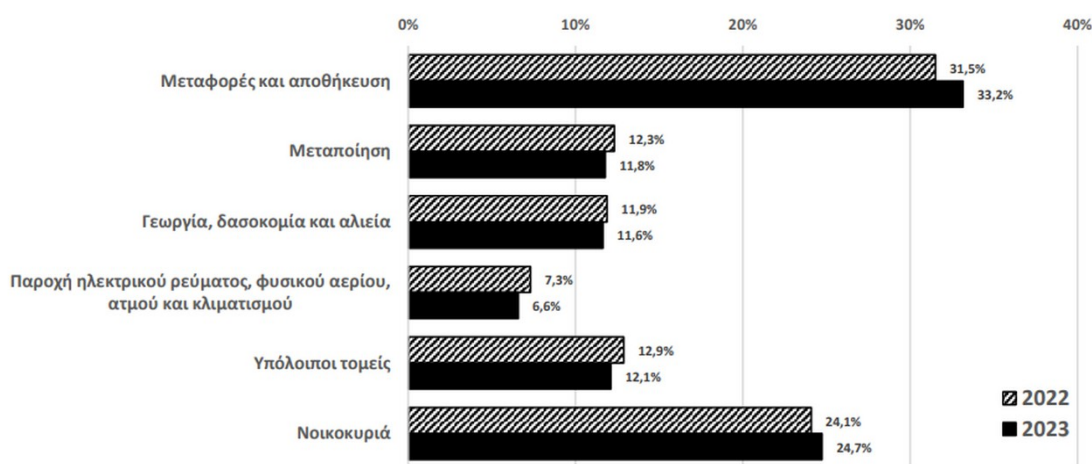
Στο πλαίσιο της παρούσας εργασίας, η ατμοσφαιρική ρύπανση προσεγγίζεται κυρίως μέσα από τη μελέτη των αιωρούμενων σωματιδίων PM_{10} . Τα PM_{10} αποτελούν βασικό δείκτη επιβάρυνσης της ποιότητας του αέρα, παρουσιάζουν έντονη χωρική και χρονική μεταβλητότητα και παρακολουθούνται συστηματικά στο ευρωπαϊκό κανονιστικό πλαίσιο (European Parliament and Council of the European Union, 2008· WHO, 2021· EEA, 2026).

Η αξιολόγησή τους αφορά όχι μόνο το ύψος των συγκεντρώσεων, αλλά και τον τρόπο με τον οποίο οι τιμές κατανέμονται στον χρόνο και διαφοροποιούνται μεταξύ περιοχών. Για τον λόγο αυτό, η παρούσα μελέτη συνδέει τον έλεγχο ποιότητας των δεδομένων με τη στατιστική διερεύνηση της κατανομικής συμπεριφοράς των συγκεντρώσεων PM_{10} .

2.1.2 Κατηγορίες ατμοσφαιρικών ρύπων

Οι ατμοσφαιρικοί ρύποι ταξινομούνται με βάση τη φυσική τους μορφή, την προέλευση και τον τρόπο σχηματισμού τους στην ατμόσφαιρα. Ως προς τη φυσική μορφή διακρίνονται σε αέριους ρύπους, υγρά σταγονίδια, στερεά σωματίδια και αερολύματα, δηλαδή μίγματα σωματιδίων και σταγονιδίων (Tiwary, Williams & Colls, 2019). Στους κύριους ρύπους που παρακολουθούνται συστηματικά περιλαμβάνονται τα αιωρούμενα σωματίδια (PM), το διοξείδιο του αζώτου (NO_2), το διοξείδιο του θείου (SO_2), το μονοξείδιο του άνθρακα (CO) και το όζον (O_3) (WHO, 2021). Σημαντικό ρόλο έχουν επίσης πρόδρομες ενώσεις, όπως τα οξείδια του αζώτου (NO_x), οι μη μεθανιούχες πτητικές οργανικές ενώσεις (NMVOCs) και η αμμωνία (NH_3), καθώς συμμετέχουν στον σχηματισμό δευτερογενών σωματιδίων και τροποσφαιρικού όζοντος (EEA, 2021).

Η συμβολή διαφορετικών δραστηριοτήτων στις συνολικές εκπομπές αποτυπώνεται ενδεικτικά στα διαθέσιμα στοιχεία για την Ελλάδα. Όπως παρουσιάζεται στο Σχήμα 2.2, οι μεταφορές και η αποθήκευση και τα νοικοκυριά εμφανίζουν τα μεγαλύτερα ποσοστά, ενώ ακολουθούν η μεταποίηση, η γεωργία, η παραγωγή ενέργειας και οι υπόλοιποι τομείς. Η κατανομή αυτή δείχνει ότι οι εκπομπές δεν προέρχονται από μία ενιαία πηγή, αλλά διαμορφώνονται από τη συνδυασμένη επίδραση διαφορετικών οικονομικών δραστηριοτήτων.



*Προσωρινά στοιχεία.

Σχήμα 2.2 Ποσοστιαία κατανομή των εκπομπών αέριων ρύπων και PM_{10} κατά τομέα οικονομικής δραστηριότητας και νοικοκυριά στην Ελλάδα, 2022-2023. Πηγή: [ΕΛΣΤΑΤ, 2026](#).

Μια βασική διάκριση στην ατμοσφαιρική επιστήμη είναι εκείνη μεταξύ πρωτογενών και δευτερογενών ρύπων. Πρωτογενείς είναι οι ρύποι που εκπέμπονται απευθείας στην ατμόσφαιρα από φυσικές ή ανθρωπογενείς πηγές, όπως οι καύσεις, οι βιομηχανικές διεργασίες, οι πυρκαγιές και η επαναιώρηση σκόνης (Tiwary, Williams & Colls, 2019). Αντίθετα, οι δευτερογενείς ρύποι σχηματίζονται στην ατμόσφαιρα μέσω χημικών ή φωτοχημικών αντιδράσεων μεταξύ πρωτογενών ρύπων και άλλων ατμοσφαιρικών συστατικών.

Το τροποσφαιρικό όζον αποτελεί χαρακτηριστικό δευτερογενή ρύπο, καθώς σχηματίζεται μέσω αντιδράσεων μεταξύ NO_x και NMVOCs υπό την επίδραση της ηλιακής ακτινοβολίας. Αντίστοιχα, μέρος των αιωρούμενων σωματιδίων αποτελείται από δευτερογενή συστατικά, όπως θειικά, νιτρικά και αμμωνιακά άλατα, τα οποία προκύπτουν από την ατμοσφαιρική μετατροπή πρόδρομων ενώσεων, όπως το SO_2 , τα NO_x και η NH_3 (Tiwary, Williams & Colls, 2019· EEA, 2021).

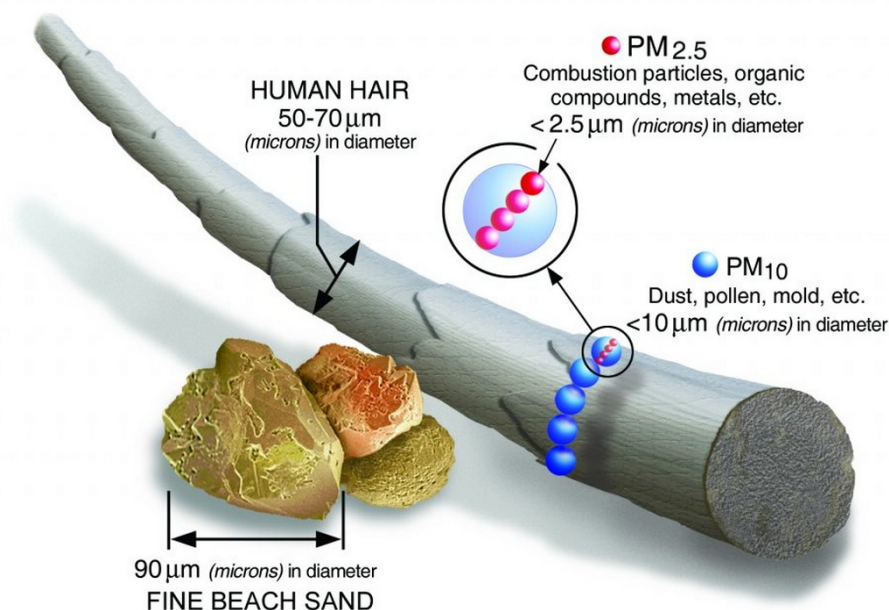
Τα αιωρούμενα σωματίδια παρουσιάζουν ιδιαίτερο ενδιαφέρον, επειδή δεν αποτελούν ενιαίο χημικό είδος, αλλά ετερογενές μίγμα στερεών σωματιδίων και υγρών σταγονιδίων με διαφορετικό μέγεθος, σχήμα, σύσταση και προέλευση. Ένα μέρος τους εκπέμπεται άμεσα από πηγές, όπως η εδαφική και οδική σκόνη, η καύση βιομάζας, η βιομηχανία και οι μεταφορές, ενώ άλλο μέρος σχηματίζεται δευτερογενώς στην ατμόσφαιρα (Tiwary, Williams & Colls, 2019· WHO, 2021· EEA, 2026).

Για τον λόγο αυτό, η σημασία των PM δεν εξαρτάται μόνο από τη συνολική συγκέντρωση, αλλά και από το μέγεθος, τη χημική σύσταση, την προέλευση και την ατμοσφαιρική τους συμπεριφορά. Τα συχνότερα χρησιμοποιούμενα κλάσματα είναι τα PM₁₀, με αεροδυναμική διάμετρο έως 10 μm, και τα PM_{2.5}, με διάμετρο έως 2,5 μm. Η λεπτομερέστερη παρουσίαση των δύο κλασμάτων ακολουθεί στην επόμενη υποενότητα.

2.1.3 Αιωρούμενα σωματίδια: PM₁₀ και PM_{2.5}

Τα αιωρούμενα σωματίδια, διεθνώς γνωστά ως Particulate Matter (PM), αποτελούν μίγμα στερεών σωματιδίων και υγρών σταγονιδίων που βρίσκονται σε αιώρηση στον ατμοσφαιρικό αέρα. Δεν συνιστούν ενιαίο χημικό είδος, αλλά ετερογενή κατηγορία σωματιδιακής ύλης, της οποίας τα χαρακτηριστικά εξαρτώνται από το μέγεθος, τη χημική σύσταση, την προέλευση και τις ατμοσφαιρικές διεργασίες σχηματισμού και μετασχηματισμού (Seinfeld & Pandis, 2016· Saxena & Sonwani, 2019· U.S. EPA, 2026b). Η βασική ταξινόμησή τους πραγματοποιείται με κριτήριο την αεροδυναμική διάμετρο, η οποία εκφράζει τη συμπεριφορά των σωματιδίων κατά την κίνηση, τη μεταφορά και την εναπόθεσή τους στον αέρα. Ως PM₁₀ ορίζονται τα εισπνεύσιμα σωματίδια με αεροδυναμική διάμετρο έως 10 μm. Λόγω του μεγέθους τους, μπορούν να εισέλθουν στο αναπνευστικό σύστημα και να εναποτεθούν σε τμήματα της ανώτερης και κατώτερης αναπνευστικής οδού, γεγονός που τα καθιστά σημαντικό δείκτη σωματιδιακής ρύπανσης και ποιότητας του αέρα (Saxena & Sonwani, 2019· U.S. EPA, 2026b· WHO, 2021).

Η διαφορά μεγέθους μεταξύ των PM₁₀ και PM_{2.5} γίνεται σαφέστερη μέσω της σύγκρισής τους με οικεία μεγέθη. Όπως παρουσιάζεται στο Σχήμα 2.3, η ανθρώπινη τρίχα έχει διάμετρο περίπου 50-70 μm και ένας κόκκος λεπτής άμμου περίπου 90 μm, ενώ τα PM₁₀ και PM_{2.5} είναι σημαντικά μικρότερα. Η σύγκριση αυτή αναδεικνύει τη δυνατότητα των αιωρούμενων σωματιδίων να διεισδύουν στο αναπνευστικό σύστημα και να παραμένουν αιωρούμενα για διαφορετικά χρονικά διαστήματα.



Σχήμα 2.3. Συγκριτική απεικόνιση του μεγέθους των αιωρούμενων σωματιδίων PM₁₀ και PM_{2.5} σε σχέση με την ανθρώπινη τρίχα και κόκκους λεπτής άμμου.

Πηγή: [U.S. Environmental Protection Agency, Particulate Matter \(PM\) Basics](#). (U.S. EPA, 2026b)

Ως PM_{2.5} ορίζονται τα λεπτά εισπνεύσιμα σωματίδια με αεροδυναμική διάμετρο έως 2,5 μm. Εξαιτίας του μικρότερου μεγέθους τους, μπορούν να διεισδύσουν βαθύτερα στο αναπνευστικό σύστημα, έως και τις κυψελίδες, και συνδέονται με αυξημένη επιβάρυνση του αναπνευστικού και καρδιαγγειακού συστήματος (Saxena & Sonwani, 2019· WHO, 2021· U.S. EPA, 2026a· EEA, 2026). Επιπλέον, τα λεπτότερα σωματίδια έχουν συχνά μεγαλύτερο χρόνο παραμονής στην ατμόσφαιρα και αυξημένη δυνατότητα μεταφοράς σε μεγαλύτερες αποστάσεις.

Η σχέση μεταξύ των δύο κλασμάτων απαιτεί σαφή αποσαφήνιση. Τα PM₁₀ δεν αποτελούν απλώς «μεγάλα σωματίδια», αλλά περιλαμβάνουν όλα τα σωματίδια με αεροδυναμική διάμετρο έως 10 μm και, επομένως, περιέχουν και το λεπτότερο κλάσμα PM_{2.5}. Το κλάσμα PM₁₀ περιλαμβάνει τόσο τα PM_{2.5} όσο και το αδρότερο κλάσμα PM₁₀-PM_{2.5}. Τα δύο κλάσματα διαφοροποιούνται ως προς τις κυριότερες πηγές προέλευσης, τη χημική

σύσταση, τον χρόνο παραμονής στην ατμόσφαιρα, τη δυνατότητα μεταφοράς και τη θέση εναπόθεσης στο αναπνευστικό σύστημα (Saxena & Sonwani, 2019· U.S. EPA, 2026a).

Το αδρότερο κλάσμα $PM_{10-PM_{2.5}}$ συνδέεται συχνότερα με μηχανικές διεργασίες, όπως η επαναιώρηση εδαφικής και οδικής σκόνης, η φθορά υλικών, οι κατασκευαστικές δραστηριότητες και το θαλάσσιο αερόλυμα. Αντίθετα, τα $PM_{2.5}$ συνδέονται σε μεγαλύτερο βαθμό με καύσεις, κυκλοφοριακές και βιομηχανικές εκπομπές, καθώς και με τον δευτερογενή σχηματισμό σωματιδίων από αέριες πρόδρομες ενώσεις. Η διάκριση αυτή δεν είναι απόλυτη, αλλά βοηθά στην κατανόηση της διαφορετικής περιβαλλοντικής και υγειονομικής σημασίας των δύο κλασμάτων.

Στην παρούσα εργασία η έμφαση δίνεται στα PM_{10} , επειδή αποτελούν βασικό δείκτη σωματιδιακής ρύπανσης και παρακολουθούνται συστηματικά από τα δίκτυα μέτρησης ποιότητας του αέρα. Παρουσιάζουν έντονη χωρική και χρονική μεταβλητότητα και μπορούν να προέρχονται από φυσικές και ανθρωπογενείς πηγές, όπως η εδαφική και οδική σκόνη, το θαλάσσιο αλάτι, η μεταφορά σκόνης, η οικιακή θέρμανση, η βιομηχανία και οι μεταφορές, καθώς και από δευτερογενή σχηματισμό στην ατμόσφαιρα (Saxena & Sonwani, 2019· WHO, 2021· EEA, 2026).

Η στατιστική διερεύνηση των συγκεντρώσεων PM_{10} επιτρέπει την αποτύπωση της κατανομής, της μεταβλητότητας και της συχνότητας εμφάνισης υψηλών τιμών, παρέχοντας τη βάση για τη σύγκριση των επιλεγμένων χρονοσειρών.

2.1.4 Φυσικοχημικά χαρακτηριστικά των PM_{10}

Τα PM_{10} αποτελούν κλάσμα της αιωρούμενης σωματιδιακής ύλης που ορίζεται με βάση την αεροδυναμική διάμετρο και όχι αποκλειστικά το γεωμετρικό μέγεθος. Ως PM_{10} χαρακτηρίζονται τα σωματίδια με αεροδυναμική διάμετρο έως 10 μm , δηλαδή σωματίδια που, ως προς την κίνησή τους στον αέρα, συμπεριφέρονται όπως σφαίρες μοναδιαίας πυκνότητας αντίστοιχου μεγέθους (Ρεμουντάκη, 2003· Seinfeld & Pandis, 2016).

Η αεροδυναμική διάμετρος επηρεάζει τον χρόνο παραμονής των σωματιδίων στην ατμόσφαιρα, τη δυνατότητα μεταφοράς τους, τους μηχανισμούς απομάκρυνσής τους μέσω ξηρής ή υγρής απόθεσης και τη θέση εναπόθεσής τους στο αναπνευστικό σύστημα. Παράλληλα, καθορίζει σε σημαντικό βαθμό την αποτελεσματικότητα των

δειγματοληπτικών συστημάτων και, επομένως, τον τρόπο μέτρησης των συγκεντρώσεων (Ρεμουντάκη, 2003).

Η χημική σύσταση των PM_{10} είναι ιδιαίτερα ετερογενής. Μπορεί να περιλαμβάνει ορυκτογενή στοιχεία, όπως Si, Al, Ca, Mg και Fe, θαλάσσια άλατα, όπως NaCl, $MgCl_2$ και KCl, καθώς και θειικά, νιτρικά και αμμωνιακά άλατα. Τα τελευταία σχηματίζονται συχνά δευτερογενώς στην ατμόσφαιρα από πρόδρομες αέριες ενώσεις, όπως το SO_2 , τα NO_x και η NH_3 . Επιπλέον, τα PM_{10} μπορεί να περιέχουν στοιχειακό και οργανικό άνθρακα, black carbon, μέταλλα, ιχνοστοιχεία και βιολογικά σωματίδια, όπως γύρη, σπόρια και μικροβιακό υλικό (Ρεμουντάκη, 2003· Κούγκολος, 2022· Seinfeld & Pandis, 2016· Mukherjee & Agrawal, 2017).

Η σύστασή τους συνδέεται άμεσα με τις πηγές και τους μηχανισμούς σχηματισμού τους. Η εδαφική και οδική σκόνη χαρακτηρίζεται κυρίως από ορυκτογενή στοιχεία και προέρχεται από αιολική διάβρωση, μηχανική θραύση και επαναιώρηση υλικού. Τα θαλάσσια αερολύματα είναι πλούσια σε Na, Cl και Mg και αποκτούν ιδιαίτερη σημασία σε παράκτιες περιοχές. Αντίθετα, ο στοιχειακός και οργανικός άνθρακας συνδέονται κυρίως με διεργασίες καύσης, όπως η καύση ορυκτών καυσίμων και βιομάζας και οι εκπομπές από την κυκλοφορία.

Τα θειικά, νιτρικά και αμμωνιακά άλατα αποτελούν κυρίως προϊόντα δευτερογενών ατμοσφαιρικών αντιδράσεων. Παρότι συγκεντρώνονται συχνότερα στο λεπτότερο κλάσμα, συμβάλλουν και στη συνολική μάζα των PM_{10} , καθώς το κλάσμα PM_{10} περιλαμβάνει και τα μικρότερα σωματίδια (Ρεμουντάκη, 2003· Κούγκολος, 2022· Remoundaki et al., 2011· EEA, 2026).

Κύρια συστατικά PM ₁₀	Ενδεικτική προέλευση	Σχέση με τη μεταβλητότητα
Ορυκτογενή στοιχεία: Si, Al, Ca, Mg, Fe	Εδαφική και οδική σκόνη, επαναιώρηση, διάβρωση	Επηρεάζονται από τον άνεμο, την ξηρότητα, την κυκλοφορία και τη μεταφορά σκόνης
Θαλάσσια άλατα: Na, Cl, Mg	Θαλάσσιο αερόλυμα	Σημαντικότερα σε παράκτιες περιοχές και υπό συγκεκριμένες διευθύνσεις ανέμου
Θειικά, νιτρικά, αμμωνιακά άλατα	Δευτερογενής σχηματισμός από SO ₂ , NO _x , NH ₃	Επηρεάζονται από την ατμοσφαιρική χημεία, την υγρασία και τη μεταφορά αερίων μαζών
Ανθρακούχο υλικό: elemental carbon, organic carbon, black carbon	Καύσεις, κυκλοφορία, βιομάζα, οικιακή θέρμανση	Παρουσιάζει εποχική διαφοροποίηση και εξάρτηση από δραστηριότητες καύσης
Βιολογικά σωματίδια: γύρη, σπόρια, μικροβιακό υλικό	Φυσικές και βιογενείς πηγές	Εξαρτώνται από την εποχή, τη βλάστηση και τις μετεωρολογικές συνθήκες

Πίνακας 2.1 Ενδεικτική φυσικοχημική σύσταση των PM₁₀, πιθανές πηγές προέλευσης και σχέση με τη χωροχρονική μεταβλητότητα των συγκεντρώσεων.

Πηγές: EEA, 2026· U.S. EPA, 2026b· Ρεμουντάκη, 2003· Κούγκολος, 2022.

Η φυσικοχημική σύσταση των PM₁₀ μεταβάλλεται ανάλογα με τις τοπικές πηγές, τις φυσικές διεργασίες και τις ατμοσφαιρικές συνθήκες. Επομένως, οι συγκεντρώσεις τους δεν παρουσιάζουν ενιαία ή σταθερή συμπεριφορά, αλλά διαφοροποιούνται χωρικά και χρονικά. Τα PM₁₀ δεν πρέπει, συνεπώς, να αντιμετωπίζονται ως χημικά ομοιογενής ρύπος, αλλά ως δυναμικό μίγμα σωματιδίων διαφορετικής προέλευσης, σύστασης, μεγέθους και ατμοσφαιρικής συμπεριφοράς. Η ετερογένεια αυτή συμβάλλει στη μεταβλητότητα των συγκεντρώσεών τους και υποστηρίζει την ανάγκη χρήσης κατάλληλων στατιστικών

μεθόδων για την περιγραφή και τη μοντελοποίηση της κατανομής τους (Grivas et al., 2018· Mukherjee & Agrawal, 2017).

2.1.5 Επιπτώσεις των PM_{10} στην υγεία και το περιβάλλον

Τα PM_{10} αποτελούν σημαντικό δείκτη ποιότητας του ατμοσφαιρικού αέρα, καθώς η έκθεση σε αυξημένες συγκεντρώσεις τους συνδέεται με επιπτώσεις στην ανθρώπινη υγεία και στο περιβάλλον. Λόγω της αεροδυναμικής τους διαμέτρου, μπορούν να εισέλθουν στο αναπνευστικό σύστημα και να εναποτεθούν σε τμήματα της ανώτερης και κατώτερης αναπνευστικής οδού. Δεδομένου ότι το κλάσμα PM_{10} περιλαμβάνει και τα λεπτότερα $PM_{2.5}$, μέρος των σωματιδίων μπορεί να διεισδύσει βαθύτερα στους πνεύμονες (Pope & Dockery, 2006· U.S. EPA, 2026a).

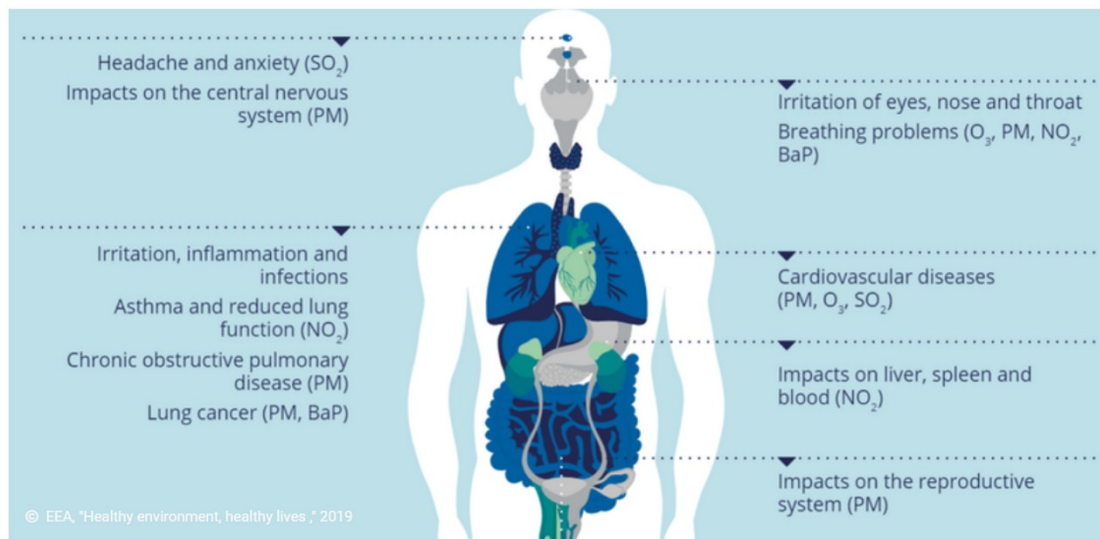
Η έκθεση σε σωματιδιακή ρύπανση έχει συσχετιστεί με ερεθισμό των αεραγωγών, επιδείνωση του άσθματος, μείωση της πνευμονικής λειτουργίας και παροξύνσεις χρόνιων αναπνευστικών παθήσεων. Παράλληλα, η βιβλιογραφία αναφέρει συσχετίσεις με καρδιαγγειακές επιπτώσεις, όπως αρρυθμίες, επιδείνωση καρδιακής ανεπάρκειας και αυξημένο κίνδυνο καρδιαγγειακής και συνολικής θνησιμότητας, ιδίως σε άτομα με προϋπάρχοντα νοσήματα (Pope & Dockery, 2006· Brook et al., 2010· U.S. EPA, 2026a).

Οι επιπτώσεις διαφοροποιούνται ανάλογα με τη διάρκεια και την ένταση της έκθεσης. Η βραχυχρόνια έκθεση σε αυξημένες συγκεντρώσεις μπορεί να συνδέεται με οξείες αναπνευστικές επιβαρύνσεις, όπως βήχα, δύσπνοια, επιδείνωση άσθματος και αύξηση των επισκέψεων σε υπηρεσίες υγείας. Αντίστοιχα, η μακροχρόνια έκθεση έχει συσχετιστεί με αυξημένο κίνδυνο εμφάνισης ή επιδείνωσης χρόνιων αναπνευστικών και καρδιαγγειακών νοσημάτων. Η μετα-ανάλυση των Orellano, Reynoso, Quaranta, Bardach, & Ciapponi ανέδειξε συσχέτιση της βραχυχρόνιας έκθεσης σε PM_{10} και $PM_{2.5}$ με αυξημένο κίνδυνο ολικής, καρδιαγγειακής και αναπνευστικής θνησιμότητας (Orellano et al., 2020). Οι κατευθυντήριες οδηγίες του Παγκόσμιου Οργανισμού Υγείας επισημαίνουν επίσης ότι δυσμενείς επιδράσεις μπορούν να εμφανιστούν και σε σχετικά χαμηλά επίπεδα συγκεντρώσεων (WHO, 2021).

Οι επιπτώσεις της ατμοσφαιρικής ρύπανσης δεν περιορίζονται στο αναπνευστικό σύστημα, αλλά μπορεί να αφορούν και άλλα συστήματα του ανθρώπινου οργανισμού, όπως συνοψίζεται στο Σχήμα 2.4. Το σχήμα αναφέρεται στην ατμοσφαιρική ρύπανση

γενικότερα και όχι αποκλειστικά στα PM_{10} , επομένως χρησιμοποιείται ως συνοπτική απεικόνιση των πιθανών επιπτώσεων διαφορετικών ρύπων.

Health impacts of air pollution



Note: Particulate matter with a diameter of $2.5\ \mu m$ or less ($PM_{2.5}$), particulate matter with a diameter of $10\ \mu m$ or less (PM_{10}), ozone (O_3), nitrogen dioxide (NO_2), benzo[a]pyrene (BaP) and sulphur dioxide (SO_2).

Σχήμα 2.4. Ενδεικτικές επιπτώσεις της ατμοσφαιρικής ρύπανσης στην ανθρώπινη υγεία.

Πηγή: [European Environment Agency, 2019](https://www.eea.europa.eu/en/press/2019/04/04).

Ιδιαίτερη σημασία έχει η διαφοροποίηση της ευαλωτότητας μεταξύ πληθυσμιακών ομάδων. Παιδιά, ηλικιωμένοι και άτομα με άσθμα, χρόνια αποφρακτική πνευμονοπάθεια, καρδιαγγειακά νοσήματα ή άλλα χρόνια προβλήματα υγείας εμφανίζουν μεγαλύτερο κίνδυνο δυσμενών επιπτώσεων. Τα παιδιά είναι περισσότερο ευάλωτα επειδή το αναπνευστικό και το ανοσοποιητικό τους σύστημα βρίσκονται ακόμη σε ανάπτυξη, ενώ οι ηλικιωμένοι και τα άτομα με προϋπάρχοντα νοσήματα διαθέτουν μικρότερη φυσιολογική εφεδρεία (EEA, 2023· U.S. EPA, 2026a).

Πέρα από την ανθρώπινη υγεία, τα PM_{10} έχουν σημαντικές περιβαλλοντικές συνέπειες. Μπορούν να μεταφερθούν και να εναποτεθούν σε εδάφη, φυτά, υδάτινα σώματα και δομημένες επιφάνειες, μεταφέροντας ανόργανες και οργανικές ενώσεις. Ανάλογα με τη σύστασή τους, ενδέχεται να συμβάλλουν στην οξίνιση εδαφών και υδάτων, στη μεταβολή της θρεπτικής ισορροπίας των οικοσυστημάτων, στη φθορά υλικών και κτιρίων και στην υποβάθμιση του αστικού περιβάλλοντος. Επιπλέον, η σωματιδιακή ρύπανση μπορεί να

μειώσει την ορατότητα και να επηρεάσει την ακτινοβολιακή ισορροπία και τις μικροφυσικές ιδιότητες των νεφών (U.S. EPA, 2026b· Seinfeld & Pandis, 2016).

Τα PM_{10} αποτελούν, επομένως, σημαντικό περιβαλλοντικό και υγειονομικό δείκτη. Η αξιόπιστη αποτύπωση της μεταβλητότητας και της κατανομής των συγκεντρώσεών τους είναι απαραίτητη για την αξιολόγηση της ποιότητας του αέρα. Στην παρούσα εργασία, η στατιστική τους διερεύνηση συμβάλλει στην περιγραφή της συχνότητας εμφάνισης χαμηλών, μέσων και υψηλών τιμών και στη σύγκριση της κατανομικής συμπεριφοράς μεταξύ των εξεταζόμενων χρονοσειρών.

2.2 Πηγές και παράγοντες μεταβλητότητας των PM_{10}

Τα PM_{10} αποτελούν ετερογενές μίγμα στερεών σωματιδίων και υγρών σταγονιδίων, των οποίων η παρουσία και η παραμονή στην ατμόσφαιρα επηρεάζονται από το μέγεθος, τη σύσταση και τις επικρατούσες ατμοσφαιρικές συνθήκες (U.S. EPA, 2026b· EEA, 2026). Μπορεί να είναι πρωτογενή, όταν εκπέμπονται απευθείας από φυσικές ή ανθρωπογενείς πηγές, ή δευτερογενή, όταν σχηματίζονται στην ατμόσφαιρα από χημικές αντιδράσεις αέριων πρόδρομων ενώσεων, όπως το SO_2 , τα NO_x και η NH_3 . Η διάκριση αυτή εξηγεί γιατί οι συγκεντρώσεις PM_{10} δεν αντανakλούν πάντοτε μία μόνο εμφανή πηγή, αλλά προκύπτουν από τον συνδυασμό εκπομπών, μεταφοράς, μετασχηματισμού και απομάκρυνσης των σωματιδίων (U.S. EPA, 2026b· European Parliament and Council of the European Union, 2024· Κούγκολος, 2022).

Οι βασικές πηγές και οι παράγοντες μεταβλητότητας των PM_{10} μπορούν να διακριθούν σε τέσσερις κατηγορίες. Οι φυσικές πηγές περιλαμβάνουν την εδαφική και ορυκτή σκόνη, το θαλάσσιο αερόλυμα, τις πυρκαγιές και τη μεταφορά σωματιδίων μέσω αερίων μαζών, ιδίως σε περιοχές της Μεσογείου (Remoundaki et al., 2011· Copernicus, 2024). Οι ανθρωπογενείς πηγές περιλαμβάνουν τη βιομηχανία, τις μεταφορές, την οικιακή θέρμανση, τις κατασκευαστικές δραστηριότητες και τη γεωργία, οι οποίες μπορούν να επηρεάζουν σημαντικά τις αστικές και περιαστικές συγκεντρώσεις (Grivas et al., 2018· Tsiaousidis et al., 2023).

Παράλληλα, οι τοπικές και περιφερειακές ιδιαιτερότητες κάθε περιοχής μπορούν να διαφοροποιούν το επίπεδο και τη χρονική κατανομή των συγκεντρώσεων. Παράγοντες

όπως η ένταση της κυκλοφορίας, η χρήση οικιακής θέρμανσης, οι τοπικές παραγωγικές δραστηριότητες, η γειτνίαση με αγροτικές ή ημιορεινές περιοχές και η μεταφορά ρύπων από γειτονικές ζώνες ενδέχεται να επηρεάζουν διαφορετικά κάθε πόλη. Ωστόσο, η πραγματική συμβολή των παραγόντων αυτών δεν μπορεί να προσδιοριστεί χωρίς ειδική ανάλυση πηγών και πρόσθετα περιβαλλοντικά δεδομένα.

Ιδιαίτερη σημασία έχουν επίσης οι μετεωρολογικοί και τοπογραφικοί παράγοντες. Η ταχύτητα και η διεύθυνση του ανέμου, η θερμοκρασία, οι θερμοκρασιακές αναστροφές, η υγρασία, η βροχόπτωση και το ανάγλυφο επηρεάζουν τη διασπορά, τη συσσώρευση, τη μεταφορά και την απομάκρυνση των σωματιδίων από την ατμόσφαιρα (Seinfeld & Pandis, 2016· Vardoulakis & Kassomenos, 2008· Ρεμουντάκη, 2003).

Η μεταβλητότητα των συγκεντρώσεων PM_{10} στις πόλεις της Στερεάς Ελλάδας πρέπει, επομένως, να αντιμετωπίζεται ως αποτέλεσμα σύνθετης αλληλεπίδρασης μεταξύ φυσικών και ανθρωπογενών πηγών, τοπικής γεωγραφίας, εποχικότητας, μετεωρολογικών συνθηκών και μεταφοράς σωματιδίων. Η παρούσα εργασία δεν επιχειρεί να αποδώσει αιτιολογικά τις παρατηρούμενες συγκεντρώσεις σε συγκεκριμένους μηχανισμούς, αλλά χρησιμοποιεί το παραπάνω πλαίσιο για την ερμηνεία της χωρικής και χρονικής διαφοροποίησης των δεδομένων. Η πολυπαραγοντική αυτή φύση ενισχύει την ανάγκη στατιστικής διερεύνησης της κατανομικής συμπεριφοράς των συγκεντρώσεων PM_{10} .

2.2.1 Φυσικές πηγές PM_{10}

Οι φυσικές πηγές των PM_{10} είναι εκείνες που δεν συνδέονται άμεσα με ανθρώπινες δραστηριότητες. Σε αυτές περιλαμβάνονται η εδαφική και ορυκτή σκόνη, το θαλάσσιο άλας, η γύρη και άλλα βιογενή σωματίδια, η ηφαιστειακή τέφρα, οι δασικές πυρκαγιές όταν αντιμετωπίζονται ως φυσικά ή μεικτά φαινόμενα, καθώς και η μεταφορά ερημικής σκόνης (Viana et al., 2011· Viana et al., 2014). Ο Ευρωπαϊκός Οργανισμός Περιβάλλοντος αναγνωρίζει επίσης τη σκόνη της Σαχάρας, το θαλάσσιο άλας και τις ηφαιστειακές εκπομπές ως σημαντικές φυσικές πηγές αιωρούμενων σωματιδίων (EEA, 2026).

Το αδρότερο κλάσμα των PM_{10} συνδέεται συχνά με ορυκτά και γεωλογικά σωματίδια που προέρχονται από ξηρά εδάφη, χωμάτινες επιφάνειες και αγροτικές εκτάσεις (Viana et al., 2014· Κούγκολος, 2022). Η δράση του ανέμου μπορεί να προκαλέσει επαναιώρηση αυτών

των σωματιδίων και να αυξήσει προσωρινά τις συγκεντρώσεις τους. Η διεργασία αυτή μπορεί να είναι φυσική, αλλά συχνά ενισχύεται από την κυκλοφορία, τα εργοτάξια και την κατάσταση των οδοστρωμάτων.

Στις παράκτιες περιοχές, σημαντική φυσική πηγή αποτελεί το θαλάσσιο αερόλυμα, το οποίο παράγεται από σταγονίδια θαλασσινού νερού μέσω του κυματισμού και της δράσης του ανέμου (Viana et al., 2014). Η συμβολή του μπορεί να είναι σημαντικότερη σε πόλεις όπως η Χαλκίδα, χωρίς όμως να μπορεί να ποσοτικοποιηθεί στην παρούσα εργασία χωρίς ειδική ανάλυση πηγών.

Ιδιαίτερη σημασία για την Ελλάδα έχει και η μεταφορά ερημικής σκόνης από τη Βόρεια Αφρική. Η γεωγραφική θέση της χώρας στην Ανατολική Μεσόγειο την καθιστά συχνά εκτεθειμένη σε αέριες μάζες που μεταφέρουν σωματίδια από τη Σαχάρα (Aleksandropoulou & Lazaridis, 2013· Viana et al., 2014). Η υπηρεσία CAMS/Copernicus έχει καταγράψει τέτοια φαινόμενα στην Ελλάδα και στην ευρύτερη περιοχή, τα οποία μπορούν να αυξήσουν σημαντικά τις επιφανειακές συγκεντρώσεις PM_{10} (Copernicus, 2024).

Μελέτες στην Ελλάδα έχουν δείξει ότι οι υψηλές συγκεντρώσεις PM_{10} μπορεί να συνδέονται είτε με τοπικές πηγές είτε με μεταφορά αφρικανικής σκόνης, ανάλογα με τις επικρατούσες μετεωρολογικές συνθήκες (Matthaios et al., 2016). Επομένως, οι φυσικές πηγές αποτελούν δυναμικό παράγοντα που μπορεί να συμβάλλει σε σημαντικές ημερήσιες και εποχικές μεταβολές των συγκεντρώσεων.

2.2.2 Ανθρωπογενείς πηγές PM_{10}

Οι ανθρωπογενείς πηγές των PM_{10} περιλαμβάνουν δραστηριότητες όπως οι καύσεις, η μηχανική φθορά, η επαναιώρηση σκόνης, η βιομηχανία, οι μεταφορές, η γεωργία και ο δευτερογενής σχηματισμός σωματιδίων στην ατμόσφαιρα. Η συμβολή τους διαφοροποιείται ανάλογα με το είδος και την ένταση της δραστηριότητας, τη θέση των πηγών και τις επικρατούσες μετεωρολογικές συνθήκες.

Σημαντική πηγή αποτελεί η οικιακή θέρμανση, ιδιαίτερα όταν χρησιμοποιούνται στερεά καύσιμα και βιομάζα. Οι εκπομπές από καύσεις περιλαμβάνουν τόσο πρωτογενή σωματίδια όσο και αέριες ενώσεις που μπορούν να συμβάλουν στον σχηματισμό

δευτερογενών σωματιδίων. Η επίδρασή τους είναι συνήθως εντονότερη κατά τους ψυχρούς μήνες. Οι οδικές μεταφορές συμβάλλουν μέσω των καυσαερίων, αλλά και μέσω μη εξαντλητικών εκπομπών, όπως η φθορά ελαστικών και φρένων, η αποτριβή του οδοστρώματος και η επαναιώρηση οδικής σκόνης. Αντίστοιχα, οι βιομηχανικές, παραγωγικές και κατασκευαστικές δραστηριότητες μπορούν να αυξήσουν τις συγκεντρώσεις PM_{10} μέσω καύσεων, επεξεργασίας υλικών και παραγωγής σκόνης.

Η γεωργία επηρεάζει τις συγκεντρώσεις τόσο άμεσα, μέσω της κατεργασίας του εδάφους και της κίνησης μηχανημάτων, όσο και έμμεσα μέσω εκπομπών αμμωνίας, η οποία συμμετέχει στον σχηματισμό δευτερογενών ανόργανων σωματιδίων.

Κατηγορία πηγής	Παραδείγματα	Μηχανισμός συμβολής στα PM_{10}	Χρονική/χωρική ιδιαιτερότητα
Οδική κυκλοφορία	Καυσαέρια, φθορά φρένων, φθορά ελαστικών, αποτριβή οδοστρώματος	Πρωτογενείς εκπομπές, μηχανική φθορά, επαναιώρηση οδικής σκόνης	Εντονότερη σε αστικές περιοχές και κοντά σε οδικούς άξονες
Οικιακή θέρμανση	Καύση ξύλου, πέλετ, στερεών καυσίμων	Εκπομπή ανθρακούχων και οργανικών σωματιδίων	Κυρίως χειμερινή επιβάρυνση
Βιομηχανία και παραγωγή ενέργειας	Καπνοδόχοι, καύσεις, παραγωγικές διεργασίες	Πρωτογενή PM_{10} και πρόδρομοι αέριοι ρύποι	Εξαρτάται από τη χωροθέτηση και το είδος δραστηριότητας
Κατασκευές και λατομεία	Εργοτάξια, εκσκαφές και χύδην υλικά	Χονδρόκοκκη σκόνη, μηχανική τριβή, επαναιώρηση	Κυρίως τοπική και παροδική επιβάρυνση
Γεωργία	Αγροτικές εργασίες, εδαφική σκόνη, αμμωνία	Επαναιώρηση εδαφικής σκόνης και δευτερογενής σχηματισμός σωματιδίων	Εντονότερη σε αγροτικές και περιαστικές περιοχές

Πίνακας 2.2. Κύριες ανθρωπογενείς πηγές PM_{10} και μηχανισμοί συμβολής τους στις συγκεντρώσεις.

Ο Πίνακας 2.2 συνοψίζει τις βασικές ανθρωπογενείς πηγές PM_{10} και διακρίνει τους κύριους μηχανισμούς μέσω των οποίων συμβάλλουν στις συγκεντρώσεις. Η οδική κυκλοφορία αποτελεί ιδιαίτερα σύνθετη πηγή, καθώς συνδυάζει εκπομπές καύσης, μηχανική φθορά ελαστικών και φρένων, αποτριβή του οδοστρώματος και επαναιώρηση σκόνης. Η οικιακή θέρμανση εμφανίζει συνήθως έντονη εποχικότητα, ενώ η βιομηχανία, οι κατασκευές, τα λατομεία και οι αγροτικές δραστηριότητες μπορούν να επηρεάζουν περισσότερο συγκεκριμένες περιοχές ή χρονικές περιόδους, ανάλογα με τη θέση και την ένταση της δραστηριότητας (Kassomenos et al., 2014· Grivas et al., 2018· Tsiaousidis et al., 2023).

Οι διαφορετικοί μηχανισμοί εκπομπής είναι δυνατόν να επηρεάζουν και τη στατιστική μορφή των συγκεντρώσεων. Πηγές με σχετικά σταθερή παρουσία, όπως η κυκλοφορία ή ορισμένες παραγωγικές δραστηριότητες, μπορεί να συμβάλλουν στη διαμόρφωση του βασικού επιπέδου των τιμών. Αντίθετα, δραστηριότητες με έντονη εποχικότητα ή περιορισμένη χρονική διάρκεια, όπως η οικιακή θέρμανση, οι κατασκευές και η αυξημένη επαναιώρηση σκόνης, ενδέχεται να συνδέονται με την εμφάνιση υψηλότερων συγκεντρώσεων και εντονότερης θετικής ασυμμετρίας. Η σχέση αυτή, ωστόσο, δεν είναι άμεση, καθώς η τελική συγκέντρωση επηρεάζεται παράλληλα από τις συνθήκες διασποράς, μεταφοράς και απομάκρυνσης των σωματιδίων.

Η πραγματική συμβολή κάθε ανθρωπογενούς πηγής στις πόλεις της Στερεάς Ελλάδας δεν μπορεί να προσδιοριστεί μόνο από τις μετρήσεις συγκέντρωσης PM_{10} . Για μια τέτοια εκτίμηση θα απαιτούνταν πρόσθετα δεδομένα εκπομπών, χημικής σύστασης, κυκλοφοριακού φόρτου, χρήσεων γης και μετεωρολογικών συνθηκών, καθώς και εφαρμογή ειδικών μεθόδων κατανομής πηγών.

Στην παρούσα εργασία, επομένως, οι ανθρωπογενείς πηγές χρησιμοποιούνται ως θεωρητικό πλαίσιο για την ερμηνεία της μεταβλητότητας και της κατανομικής συμπεριφοράς των συγκεντρώσεων, χωρίς να επιχειρείται αιτιώδης ή ποσοτική απόδοση των παρατηρούμενων τιμών σε συγκεκριμένες δραστηριότητες.

2.2.3 Πηγές PM₁₀ στην Ελλάδα και στη Στερεά Ελλάδα

Η Ελλάδα χαρακτηρίζεται από ένα σύνθετο σωματιδιακό περιβάλλον, στο οποίο συνυπάρχουν αστικές, βιομηχανικές, αγροτικές και παράκτιες δραστηριότητες, ενώ παράλληλα ασκούνται επιδράσεις από φυσικές και υπερτοπικές πηγές. Στις ελληνικές και γενικότερα στις μεσογειακές πόλεις, οι συγκεντρώσεις PM₁₀ έχουν συνδεθεί με την οδική κυκλοφορία, την οικιακή θέρμανση, τη βιομηχανική δραστηριότητα, την επαναιώρηση εδαφικής και οδικής σκόνης και τη μεταφορά σωματιδίων σε μεγάλες αποστάσεις (Κοζακ, Mihalopoulos & Kubilay, 2009· Diapouli et al., 2017· Grivas et al., 2018). Επομένως, η σωματιδιακή ρύπανση δεν αποτελεί αποτέλεσμα μιας μόνο κατηγορίας πηγών, αλλά της αλληλεπίδρασης τοπικών, περιφερειακών και υπερτοπικών διεργασιών.

Ιδιαίτερη σημασία για την Ελλάδα έχει η μεταφορά ερημικής σκόνης από τη Βόρεια Αφρική. Τα φαινόμενα σαχαριανής σκόνης μπορούν να αυξήσουν σημαντικά τις συγκεντρώσεις PM₁₀ και να μεταβάλουν τη φυσικοχημική τους σύσταση, κυρίως στη Νότια και Ανατολική Μεσόγειο (Aleksandropoulou & Lazaridis, 2013· Dimitriou & Kassomenos, 2019· Copernicus, 2024). Μελέτες στην Ελλάδα έχουν δείξει ότι υψηλές τιμές PM₁₀ μπορεί να συνδέονται είτε με τοπικές πηγές είτε με μεταφορά αφρικανικής σκόνης, ανάλογα με τις συνοπτικές και μετεωρολογικές συνθήκες (Matthaios et al., 2016). Η χρονική μεταβλητότητα των συγκεντρώσεων δεν μπορεί, συνεπώς, να ερμηνεύεται αποκλειστικά με βάση τις τοπικές ανθρωπογενείς εκπομπές.

Το Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης παρέχει δεδομένα για την αξιολόγηση της ποιότητας του αέρα σε διαφορετικές περιοχές της χώρας, συμπεριλαμβανομένης της Στερεάς Ελλάδας. Η διαθεσιμότητα μετρήσεων επιτρέπει τη χωρική και χρονική σύγκριση των συγκεντρώσεων PM₁₀, δεν αρκεί όμως από μόνη της για την ποσοτική απόδοση των τιμών σε συγκεκριμένες πηγές. Για τον σκοπό αυτό απαιτούνται συμπληρωματικά δεδομένα χημικής σύστασης, εκπομπών και μετεωρολογικών παραμέτρων, καθώς και ειδικές μέθοδοι κατανομής πηγών (ΥΠΕΝ, χ.χ.). Στη Στερεά Ελλάδα, οι συγκεντρώσεις PM₁₀ πρέπει να εξετάζονται ανά πόλη, καθώς οι επιμέρους περιοχές διαφοροποιούνται ως προς τη μορφολογία, τις χρήσεις γης και τις συνθήκες διασποράς. Η Άμφισσα αποτελεί μικρό αστικό κέντρο σε ημιορεινό περιβάλλον, όπου ενδέχεται να συνδυάζονται επιδράσεις από την κυκλοφορία, την οικιακή θέρμανση και την επαναιώρηση σκόνης. Η Λαμία, ως μεγαλύτερο αστικό και διοικητικό κέντρο, μπορεί να επηρεάζεται περισσότερο από κυκλοφοριακές και οικιστικές δραστηριότητες. Η Λειβαδιά βρίσκεται σε αστικό-αγροτικό περιβάλλον, στο οποίο είναι πιθανή η

συνδυασμένη επίδραση της κυκλοφορίας, της εδαφικής σκόνης και των γεωργικών δραστηριοτήτων. Η Χαλκίδα, λόγω της παράκτιας θέσης της, μπορεί να δέχεται επιδράσεις από θαλάσσιο αερόλυμα, αστικές εκπομπές και μεταφορά ρύπων. Το Καρπενήσι παρουσιάζει διαφορετικά χαρακτηριστικά λόγω του ορεινού του περιβάλλοντος και της αυξημένης σημασίας της χειμερινής θέρμανσης.

Οι παραπάνω αναφορές περιγράφουν μόνο πιθανούς μηχανισμούς επίδρασης και δεν τεκμηριώνουν την κυριαρχία συγκεκριμένης πηγής σε κάθε πόλη. Στην παρούσα εργασία, οι πόλεις της Στερεάς Ελλάδας αντιμετωπίζονται ως χωρικά διαφοροποιημένες περιπτώσεις, στις οποίες οι συγκεντρώσεις PM_{10} μπορεί να διαμορφώνονται από διαφορετικούς συνδυασμούς τοπικών εκπομπών, φυσικών πηγών, μεταφοράς σκόνης και ατμοσφαιρικών συνθηκών (Kozak, Mihalopoulos & Kubilay, 2009· Diapouli et al., 2017· Grivas et al., 2018).

2.2.4 Μετεωρολογικοί και τοπογραφικοί παράγοντες

Οι συγκεντρώσεις των PM_{10} δεν εξαρτώνται μόνο από το ύψος των εκπομπών, αλλά και από τον τρόπο με τον οποίο οι μετεωρολογικές συνθήκες και το ανάγλυφο επηρεάζουν τη διασπορά, τη μεταφορά, τον μετασχηματισμό και την απομάκρυνση των σωματιδίων (Seinfeld & Pandis, 2016· Karagiannidis et al., 2015· Pateraki et al., 2013). Σε συνθήκες ισχυρού αερισμού και βροχόπτωσης, οι συγκεντρώσεις τείνουν να μειώνονται λόγω αραίωσης και υγρής απόθεσης. Αντίθετα, η άπνοια, οι θερμοκρασιακές αναστροφές και το χαμηλό ύψος ανάμιξης μπορούν να ευνοήσουν τη συσσώρευση των ρύπων κοντά στην επιφάνεια του εδάφους (Pateraki et al., 2013· Grivas, Chaloulakou & Kassomenos, 2008).

Ο άνεμος έχει διττό ρόλο. Οι μέτριοι και ισχυροί άνεμοι μπορούν να ενισχύσουν τον αερισμό της ατμόσφαιρας και να μειώσουν τις τοπικές συγκεντρώσεις, αλλά παράλληλα είναι δυνατόν να προκαλέσουν επαναιώρηση σκόνης από οδοστρώματα και ξηρές επιφάνειες ή να μεταφέρουν σωματίδια από απομακρυσμένες πηγές (Seinfeld & Pandis, 2016· Ntziachristos & Boulter, 2023). Η διεύθυνση του ανέμου καθορίζει επίσης αν μια περιοχή επηρεάζεται κυρίως από τοπικές, περιφερειακές ή υπερτοπικές πηγές, όπως η

μεταφορά σαχαριανής σκόνης προς την Ελλάδα (Mitsakou et al., 2008· Querol et al., 2009· Viana et al., 2011).

Η βροχόπτωση λειτουργεί συνήθως καθαριστικά, απομακρύνοντας τα σωματίδια από την ατμόσφαιρα μέσω υγρής απόθεσης. Η σχετική υγρασία, ωστόσο, μπορεί να επηρεάσει διαφορετικά τη σωματιδιακή ύλη, καθώς ορισμένα σωματίδια είναι υγροσκοπικά και απορροφούν νερό, με αποτέλεσμα να μεταβάλλεται η μάζα και η φυσικοχημική τους συμπεριφορά (Seinfeld & Pandis, 2016· Pateraki et al., 2013).

Η θερμοκρασία συνδέεται τόσο με τις ατμοσφαιρικές διεργασίες όσο και με την εποχικότητα των πηγών. Κατά τους ψυχρούς μήνες, η οικιακή θέρμανση και η καύση βιομάζας μπορούν να αυξήσουν τις συγκεντρώσεις, ιδιαίτερα όταν συνδυάζονται με περιορισμένη διασπορά. Αντίθετα, κατά τις θερμότερες και ξηρότερες περιόδους μπορεί να ενισχύονται η επαναιώρηση σκόνης, η μεταφορά ερημικών σωματιδίων και ο δευτερογενής σχηματισμός αερολυμάτων (Diapouli et al., 2017· Grivas et al., 2018· Tsiaousidis et al., 2023· Querol et al., 2009).

Η τοπογραφία επηρεάζει επίσης τον αερισμό και τη χωρική κατανομή των PM_{10} . Κοιλάδες, λεκανοπέδια και κλειστές μορφολογίες μπορούν να περιορίσουν τη διασπορά, ενώ οι ορεινοί όγκοι είναι δυνατόν να λειτουργήσουν ως φραγμοί στη μεταφορά των αερίων μαζών (Karagiannidis et al., 2015· Seinfeld & Pandis, 2016). Για τον λόγο αυτό, παρόμοια επίπεδα εκπομπών δεν οδηγούν αναγκαστικά σε ίδιες συγκεντρώσεις σε διαφορετικές πόλεις.

Στη Στερεά Ελλάδα, η Χαλκίδα μπορεί να επηρεάζεται από θαλάσσιους ανέμους και θαλάσσιο αερόλυμα, ενώ το Καρπενήσι, λόγω της ορεινής του θέσης, ενδέχεται να παρουσιάζει διαφορετικές συνθήκες αερισμού και στασιμότητας. Αντίστοιχα, η Λαμία και οι υπόλοιπες ηπειρωτικές πόλεις μπορεί να εμφανίζουν διαφορετική δυναμική διασποράς ανάλογα με το ανάγλυφο και τις τοπικές μετεωρολογικές συνθήκες. Οι αναφορές αυτές περιγράφουν πιθανούς μηχανισμούς και δεν αποτελούν απόδειξη της πραγματικής επίδρασής τους στις εξεταζόμενες χρονοσειρές.

Οι μετεωρολογικοί και τοπογραφικοί παράγοντες αποτελούν βασικό πλαίσιο για την ερμηνεία της μεταβλητότητας των συγκεντρώσεων PM_{10} , παρότι δεν εξετάζονται άμεσα στην παρούσα ανάλυση.

2.3 Θεσμικό πλαίσιο και παρακολούθηση PM₁₀

Η αξιολόγηση των συγκεντρώσεων PM₁₀ συνδέεται άμεσα με το θεσμικό και τεχνικό πλαίσιο βάσει του οποίου παράγονται, ελέγχονται και ερμηνεύονται τα δεδομένα ποιότητας του αέρα. Οι μετρήσεις αποκτούν κανονιστική και ερευνητική σημασία όταν εξετάζονται σε σχέση με τις οριακές τιμές, τις υγειονομικές κατευθυντήριες γραμμές, τις πρότυπες μεθόδους μέτρησης και τις απαιτήσεις ποιότητας των δεδομένων. Για τον λόγο αυτό, η ενότητα παρουσιάζει τα ευρωπαϊκά όρια και τις τιμές αναφοράς του Παγκόσμιου Οργανισμού Υγείας, καθώς και τις βασικές προϋποθέσεις αξιοπιστίας και συγκρισιμότητας των μετρήσεων.

2.3.1 Ευρωπαϊκά όρια και κατευθυντήριες τιμές

Η αξιολόγηση της ρύπανσης από PM₁₀ στην Ευρωπαϊκή Ένωση βασίστηκε επί σειρά ετών στην Οδηγία 2008/50/EK, η οποία ενσωματώθηκε στο ελληνικό δίκαιο με την ΚΥΑ Η.Π. 14122/549/Ε.103. Στο πλαίσιο αυτό, θεσπίστηκε ημερήσια οριακή τιμή 50 $\mu\text{g}/\text{m}^3$, η οποία δεν επιτρέπεται να υπερβαίνεται περισσότερες από 35 ημέρες ανά ημερολογιακό έτος, καθώς και ετήσια οριακή τιμή 40 $\mu\text{g}/\text{m}^3$. Οι τιμές αυτές αποτέλεσαν το βασικό θεσμικό σημείο αναφοράς για την αξιολόγηση της συμμόρφωσης των κρατών μελών και για τη διαχείριση της ποιότητας του ατμοσφαιρικού αέρα (ΥΠΕΚΑ, 2011).

Οι κατευθυντήριες οδηγίες του Παγκόσμιου Οργανισμού Υγείας του 2021 προτείνουν αυστηρότερα επίπεδα για την προστασία της ανθρώπινης υγείας. Για τα PM₁₀ ορίζεται ετήσια τιμή-οδηγός 15 $\mu\text{g}/\text{m}^3$ και 24ωρη τιμή 45 $\mu\text{g}/\text{m}^3$, η οποία αντιστοιχεί στο 99ο εκατοστημόριο των ημερήσιων συγκεντρώσεων. Οι τιμές αυτές δεν είναι νομικά δεσμευτικές, αλλά αποτελούν υγειονομικά τεκμηριωμένα επίπεδα αναφοράς, τα οποία βασίζονται σε νεότερα επιστημονικά δεδομένα για δυσμενείς επιδράσεις ακόμη και σε χαμηλότερες συγκεντρώσεις από εκείνες που θεωρούνταν παλαιότερα αποδεκτές (WHO, 2021).

Η Οδηγία (ΕΕ) 2024/2881 αναδιατυπώνει και αυστηροποιεί το ευρωπαϊκό πλαίσιο, στο πλαίσιο του στόχου για μηδενική ρύπανση έως το 2050. Για τα PM₁₀ προβλέπει, με χρονικό ορίζοντα επίτευξης την 1η Ιανουαρίου 2030, ετήσια οριακή τιμή 20 $\mu\text{g}/\text{m}^3$ και ημερήσια τιμή 45 $\mu\text{g}/\text{m}^3$, με έως 18 επιτρεπόμενες υπερβάσεις ανά έτος. Οι νέες τιμές προσεγγίζουν περισσότερο τις συστάσεις του WHO, χωρίς όμως να επιτυγχάνεται πλήρης

σύγκλιση ως προς την ετήσια συγκέντρωση (European Parliament and Council of the European Union, 2024).

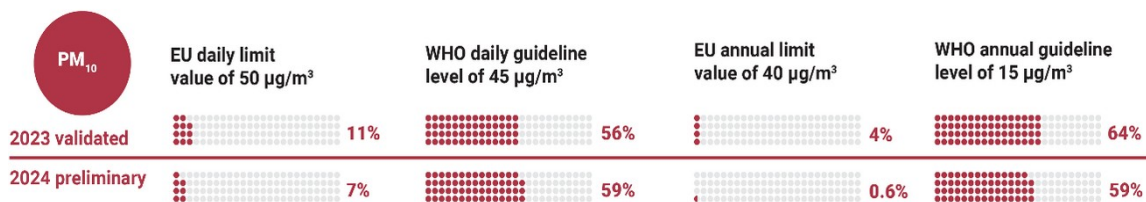
Πλαίσιο αναφοράς	Ετήσια τιμή PM_{10}	Ημερήσια/24 _{ωρη} τιμή PM_{10}	Χαρακτήρας
Οδηγία 2008/50/ΕΚ	40 $\mu\text{g}/\text{m}^3$	50 $\mu\text{g}/\text{m}^3$, έως 35 υπερβάσεις/έτος	Ισχύον ευρωπαϊκό όριο
WHO 2021	15 $\mu\text{g}/\text{m}^3$	45 $\mu\text{g}/\text{m}^3$, 99ο εκατοστημόριο	Υγειονομική κατευθυντήρια τιμή
Οδηγία 2024/2881	20 $\mu\text{g}/\text{m}^3$ από το 2030	45 $\mu\text{g}/\text{m}^3$, έως 18 υπερβάσεις/έτος	Νεότερο ευρωπαϊκό όριο με ορίζοντα το 2030

Πίνακας 2.3 Σύγκριση ευρωπαϊκών οριακών τιμών και κατευθυντήριων τιμών του WHO για τα PM_{10} .

Πηγές: [Οδηγία 2008/50/ΕΚ](#), [WHO \(2021\)](#), [Οδηγία \(European Parliament and Council of the European Union\) 2024/2881](#).

Η σύγκριση δείχνει ότι η αξιολόγηση των PM_{10} διαφοροποιείται ανάλογα με το πλαίσιο αναφοράς. Η συμμόρφωση με τα ευρωπαϊκά όρια δεν συνεπάγεται απαραίτητα ότι οι συγκεντρώσεις βρίσκονται κάτω από τις αυστηρότερες υγειονομικές συστάσεις του WHO. Η διαφορά είναι ιδιαίτερα εμφανής στην ετήσια τιμή, καθώς το όριο των 20 $\mu\text{g}/\text{m}^3$ που προβλέπεται για το 2030 εξακολουθεί να είναι υψηλότερο από την τιμή-οδηγό των 15 $\mu\text{g}/\text{m}^3$. Επομένως, η θεσμική συμμόρφωση και η μέγιστη δυνατή υγειονομική προστασία δεν ταυτίζονται πλήρως.

Η διαφοροποίηση αυτή αποτυπώνεται και στα ευρωπαϊκά δεδομένα παρακολούθησης. Το 2023, το 11% των σταθμών κατέγραψε υπέρβαση της ισχύουσας ημερήσιας οριακής τιμής της ΕΕ για τα PM_{10} , ενώ το ποσοστό των σταθμών που υπερέβησαν την αντίστοιχη ημερήσια τιμή του WHO ήταν 56%. Για την ετήσια συγκέντρωση, τα αντίστοιχα ποσοστά ήταν 4% και 64%. Τα προκαταρκτικά δεδομένα του 2024 παρουσιάζουν παρόμοια εικόνα, γεγονός που αναδεικνύει την απόσταση μεταξύ των νομικά δεσμευτικών ορίων και των αυστηρότερων υγειονομικών κατευθύνσεων (EEA, 2026).



Σχήμα 2.5 Ποσοστό σταθμών παρακολούθησης στην Ευρώπη που κατέγραψαν υπερβάσεις των οριακών τιμών της ΕΕ και των κατευθυντήριων τιμών του WHO για τα PM₁₀, με βάση επικυρωμένα δεδομένα για το 2023 και προκαταρκτικά δεδομένα για το 2024.

Πηγή: [European Environment Agency, Air quality status report 2025](#).

Για την παρούσα εργασία, οι παραπάνω τιμές παρέχουν το αναγκαίο πλαίσιο για την κατανόηση της περιβαλλοντικής και υγειονομικής σημασίας των συγκεντρώσεων. Ωστόσο, η εφαρμοζόμενη στατιστική ανάλυση δεν αποτελεί άμεσο έλεγχο θεσμικής συμμόρφωσης, καθώς βασίζεται σε ωριαίες παρατηρήσεις και αποσκοπεί κυρίως στη διερεύνηση της κατανομικής μορφής, της μεταβλητότητας και της εμφάνισης υψηλών τιμών. Η αξιολόγηση της συμμόρφωσης με ημερήσια ή ετήσια όρια θα απαιτούσε χωριστή επεξεργασία των αντίστοιχων χρονικών μέσων και σύγκρισή τους με τα προβλεπόμενα κανονιστικά κριτήρια.

2.3.2 Μέθοδοι μέτρησης PM₁₀

Από τεχνική άποψη, τα PM₁₀ δεν ορίζονται απλώς ως σωματίδια με γεωμετρική διάμετρο μικρότερη από 10 µm, αλλά ως συγκεκριμένο κλάσμα αιωρούμενων σωματιδίων που επιλέγεται βάσει της αεροδυναμικής του συμπεριφοράς κατά τη δειγματοληψία. Σύμφωνα με την Οδηγία (ΕΕ) 2024/2881, ως PM₁₀ νοούνται τα σωματίδια που διέρχονται από στόμιο επιλογής μεγέθους, όπως αυτό περιγράφεται στη μέθοδο αναφοράς EN 12341:2023, με απόδοση αποκοπής 50% σε αεροδυναμική διάμετρο 10 µm (European Parliament and Council of the European Union, 2024). Ο ορισμός αυτός συνδέει άμεσα τη φυσική συμπεριφορά των σωματιδίων με τη μέθοδο μέτρησής τους, καθώς η κεφαλή

δειγματοληψίας καθορίζει ποιο τμήμα της ατμοσφαιρικής κατανομής μεγέθους καταγράφεται ως PM_{10} .

Η μέθοδος αναφοράς για τη δειγματοληψία και τη μέτρηση των PM_{10} στην Ευρωπαϊκή Ένωση είναι η πρότυπη βαρυμετρική μέθοδος EN 12341:2023 (European Parliament and Council of the European Union, 2024). Κατά την εφαρμογή της, γνωστός όγκος αέρα διέρχεται από κατάλληλο φίλτρο, στο οποίο συλλέγεται το επιλεγμένο σωματιδιακό κλάσμα. Το φίλτρο ζυγίζεται πριν και μετά τη δειγματοληψία υπό ελεγχόμενες συνθήκες θερμοκρασίας και σχετικής υγρασίας. Η συγκέντρωση υπολογίζεται ως ο λόγος της συλλεγείσας μάζας των σωματιδίων προς τον όγκο του αέρα που διήλθε από το σύστημα και εκφράζεται συνήθως σε $\mu g/m^3$. Η τυποποίηση της διαδικασίας συμβάλλει στην ιχνηλασιμότητα και στη συγκρισιμότητα των αποτελεσμάτων μεταξύ διαφορετικών σταθμών και δικτύων παρακολούθησης.

Παράλληλα με τη μέθοδο αναφοράς, χρησιμοποιούνται συχνά αυτόματα συστήματα συνεχούς μέτρησης, τα οποία παρέχουν δεδομένα υψηλότερης χρονικής ανάλυσης. Η χρήση τους επιτρέπει την παραγωγή ωριαίων χρονοσειρών και διευκολύνει την επιχειρησιακή παρακολούθηση της ποιότητας του αέρα, τον υπολογισμό ημερήσιων μέσων τιμών και τη διερεύνηση της βραχυχρόνιας μεταβλητότητας. Η Οδηγία (ΕΕ) 2024/2881 επιτρέπει τη χρήση εναλλακτικών μεθόδων, συμπεριλαμβανομένων αυτόματων συστημάτων που ανταποκρίνονται στις απαιτήσεις του EN 16450:2017, υπό την προϋπόθεση ότι αποδεικνύεται σταθερή σχέση με τη μέθοδο αναφοράς και ότι εφαρμόζονται, όπου χρειάζεται, κατάλληλες διορθώσεις ισοδυναμίας (European Parliament and Council of the European Union, 2024).

Η μέθοδος μέτρησης έχει ιδιαίτερη σημασία για την ανάλυση χρονοσειρών, καθώς οι διαφορές μεταξύ οργάνων, οι διαδικασίες βαθμονόμησης, οι συνθήκες δειγματοληψίας και οι διορθώσεις που εφαρμόζονται μπορούν να επηρεάσουν τη συγκρισιμότητα των τιμών στον χρόνο και μεταξύ σταθμών. Επιπλέον, η σχετική υγρασία και η πτητικότητα ορισμένων συστατικών της σωματιδιακής ύλης ενδέχεται να επηρεάζουν διαφορετικά τις ενδείξεις των επιμέρους συστημάτων (Seinfeld & Pandis, 2016). Για τον λόγο αυτό, η χρήση πρότυπων ή αποδεδειγμένα ισοδύναμων μεθόδων πρέπει να συνοδεύεται από βαθμονόμηση, συντήρηση, επικύρωση και συστηματικό έλεγχο ποιότητας.

Στην Ελλάδα, η παρακολούθηση των PM_{10} πραγματοποιείται στο πλαίσιο του Εθνικού Δικτύου Παρακολούθησης Ατμοσφαιρικής Ρύπανσης, υπό την ευθύνη του Υπουργείου Περιβάλλοντος και Ενέργειας και των αρμόδιων περιφερειακών φορέων. Το δίκτυο

παρέχει τρέχουσες και επικυρωμένες μετρήσεις, ενώ εφαρμόζονται διαδικασίες βαθμονόμησης, συντήρησης και ποιοτικού ελέγχου των οργάνων σύμφωνα με τις απαιτήσεις της ευρωπαϊκής και εθνικής νομοθεσίας (ΥΠΕΝ, χ.χ.). Επιπλέον, η ΚΥΑ 70601/2013 εντάσσει τα PM_{10} σε μηχανισμούς ενημέρωσης και λήψης βραχυπρόθεσμων μέτρων κατά την εμφάνιση αυξημένων συγκεντρώσεων (ΥΠΕΚΑ, 2013).

Οι χρονοσειρές PM_{10} που χρησιμοποιούνται στην παρούσα εργασία προέρχονται από το εθνικό σύστημα παρακολούθησης και αποτελούν τη βάση της στατιστικής ανάλυσης. Ωστόσο, χωρίς αναλυτικά μεταδεδομένα για κάθε σταθμό και περίοδο, δεν είναι ασφαλές να αποδοθούν όλες οι μετρήσεις σε έναν συγκεκριμένο τύπο οργάνου. Η ερμηνεία των αποτελεσμάτων στηρίζεται στην παραδοχή ότι τα δεδομένα έχουν υποβληθεί στις προβλεπόμενες διαδικασίες μέτρησης και ελέγχου ποιότητας. Επομένως, η μέθοδος μέτρησης δεν αποτελεί απλώς τεχνική λεπτομέρεια, αλλά βασική προϋπόθεση για την αξιοπιστία, τη συγκρισιμότητα και την ορθή ερμηνεία των στατιστικών αποτελεσμάτων.

2.3.3 Αβεβαιότητα, πληρότητα και αξιοπιστία δεδομένων

Η ύπαρξη μετρήσεων PM_{10} δεν επαρκεί από μόνη της για την εξαγωγή αξιόπιστων συμπερασμάτων. Οι χρονοσειρές ποιότητας του αέρα μπορεί να περιλαμβάνουν ελλείπουσες τιμές, διακοπές λειτουργίας των οργάνων, περιόδους μη έγκυρης μέτρησης ή ασυνήθιστα υψηλές παρατηρήσεις. Οι υψηλές τιμές δεν πρέπει να απορρίπτονται αυτομάτως, καθώς μπορεί να αντανakλούν πραγματικές ατμοσφαιρικές συνθήκες. Η αφαίρεσή τους δικαιολογείται μόνο όταν υπάρχουν επαρκείς ενδείξεις τεχνικού ή καταγραφικού σφάλματος. Για τον λόγο αυτό, πριν από τη στατιστική ανάλυση απαιτείται έλεγχος της πληρότητας, της συνέχειας και της εγκυρότητας των διαθέσιμων δεδομένων.

Η αβεβαιότητα αποτελεί βασική διάσταση της ποιότητας των μετρήσεων. Η Οδηγία (ΕΕ) 2024/2881 καθορίζει συγκεκριμένους στόχους ποιότητας για τις σταθερές μετρήσεις των PM_{10} . Για τις ετήσιες μέσες συγκεντρώσεις προβλέπεται μέγιστη αβεβαιότητα $4,0 \mu\text{g}/\text{m}^3$ ή 20%, ενώ για τις 24ωρες μέσες συγκεντρώσεις προβλέπεται αβεβαιότητα $11,3 \mu\text{g}/\text{m}^3$ ή 25%, σε επίπεδο εμπιστοσύνης 95% (European Parliament and Council of the European Union, 2024). Οι απαιτήσεις αυτές δείχνουν ότι η αξιολόγηση της ποιότητας του αέρα

εξαρτάται όχι μόνο από το ύψος των καταγραφόμενων συγκεντρώσεων, αλλά και από την ακρίβεια και την τεκμηρίωση της διαδικασίας μέτρησης.

Εξίσου σημαντική είναι η πληρότητα των χρονοσειρών. Η κάλυψη δεδομένων εκφράζει το ποσοστό της εξεταζόμενης περιόδου για το οποίο υπάρχουν έγκυρες μετρήσεις. Για τις σταθερές μετρήσεις PM_{10} και $PM_{2.5}$, η Οδηγία (ΕΕ) 2024/2881 προβλέπει ελάχιστη ετήσια κάλυψη 85%. Παράλληλα, για τον υπολογισμό έγκυρου ημερήσιου μέσου απαιτείται διαθεσιμότητα τουλάχιστον 75% των ωριαίων τιμών, δηλαδή τουλάχιστον 18 έγκυρες παρατηρήσεις σε ένα εικοσιτετράωρο (European Parliament and Council of the European Union, 2024). Τα κριτήρια αυτά περιορίζουν τον κίνδυνο οι ημερήσιοι ή ετήσιοι δείκτες να βασίζονται σε αποσπασματικές παρατηρήσεις.

Στην παρούσα εργασία εφαρμόστηκε αυστηρότερο μεθοδολογικό κριτήριο, σύμφωνα με το οποίο επιλέχθηκαν για στατιστική ανάλυση μόνο τα έτη με πληρότητα τουλάχιστον 90% των αναμενόμενων ωριαίων παρατηρήσεων. Η επιλογή αυτή αποσκοπεί στον περιορισμό της επίδρασης των ελλειπουσών τιμών και στη διατήρηση μιας όσο το δυνατόν πληρέστερης εικόνας της ετήσιας μεταβλητότητας. Για κάθε πόλη επιλέχθηκε το πληρέστερο διαθέσιμο έτος που ικανοποιούσε το συγκεκριμένο όριο. Το Καρπενήσι δεν συμπεριλήφθηκε στη μοντελοποίηση, καθώς κανένα από τα διαθέσιμα έτη δεν παρουσίαζε πληρότητα ίση ή μεγαλύτερη από 90%.

Οι ελλείπουσες παρατηρήσεις δεν αντικαταστάθηκαν με εκτιμώμενες τιμές, αλλά εξαιρέθηκαν από τους υπολογισμούς. Αντίθετα, οι υψηλές καταγεγραμμένες συγκεντρώσεις διατηρήθηκαν, εφόσον δεν υπήρχε διαθέσιμη τεκμηρίωση ότι οφείλονταν σε δυσλειτουργία του οργάνου ή σε σφάλμα καταγραφής. Η προσέγγιση αυτή αποφεύγει την αυθαίρετη αφαίρεση παρατηρήσεων που μπορεί να αποτελούν πραγματικό μέρος της δεξιάς ουράς της εμπειρικής κατανομής.

Η ποιότητα και η χρονική κάλυψη των δεδομένων επηρεάζουν άμεσα τη μορφή της κατανομής των συγκεντρώσεων. Εκτεταμένα ή συστηματικά κενά μπορεί να αλλοιώσουν τον μέσο όρο, τη διασπορά, την ασυμμετρία και τη συχνότητα εμφάνισης υψηλών τιμών. Αντίστοιχα, η απουσία παρατηρήσεων από συγκεκριμένες εποχές ή ώρες μπορεί να οδηγήσει σε χρονοσειρά που δεν αποτυπώνει ομοιόμορφα ολόκληρο το έτος.

Η εφαρμογή του ορίου πληρότητας δεν εγγυάται ότι οι επιλεγμένες χρονοσειρές είναι πλήρως αντιπροσωπευτικές της μακροχρόνιας συμπεριφοράς κάθε πόλης. Διασφαλίζει, όμως, ότι βασίζονται στα πληρέστερα διαθέσιμα δεδομένα και ότι η προσαρμογή των θεωρητικών κατανομών δεν πραγματοποιείται σε έτη με εκτεταμένες ασυνέχειες. Η

αξιοπιστία των αποτελεσμάτων εξαρτάται, επομένως, τόσο από την καταλληλότητα των στατιστικών μεθόδων όσο και από την ποιότητα και την πληρότητα των παρατηρήσεων στις οποίες αυτές εφαρμόζονται.

2.4 Στατιστική μοντελοποίηση περιβαλλοντικών δεδομένων

Η στατιστική μοντελοποίηση αποτελεί βασικό εργαλείο για την περιγραφή της μεταβλητότητας των περιβαλλοντικών δεδομένων. Οι συγκεντρώσεις PM_{10} είναι μη αρνητικές και συχνά παρουσιάζουν θετική ασυμμετρία, αυξημένη διασπορά και περιορισμένο αριθμό υψηλών παρατηρήσεων, λόγω της επίδρασης διαφορετικών πηγών, μετεωρολογικών συνθηκών και μηχανισμών μεταφοράς ή συσσώρευσης.

Για τον λόγο αυτό, η ανάλυση δεν πρέπει να περιορίζεται μόνο σε μέτρα όπως ο μέσος όρος και η διάμεσος, καθώς αυτά δεν αποτυπώνουν πλήρως τη διασπορά, την ασυμμετρία και τη συμπεριφορά της δεξιάς ουράς. Η προσαρμογή θεωρητικών κατανομών επιτρέπει πληρέστερη περιγραφή της εμπειρικής συμπεριφοράς των δεδομένων και σύγκριση μεταξύ διαφορετικών πόλεων ή χρονικών περιόδων.

Στις επόμενες υποενότητες παρουσιάζονται οι βασικές έννοιες των κατανομών πιθανότητας, οι υποψήφιες θεωρητικές κατανομές και τα κριτήρια αξιολόγησης της προσαρμογής τους. Η ανάλυση αφορά την οριακή κατανομή των ωριαίων συγκεντρώσεων PM_{10} και όχι τη χρονική εξάρτηση ή την εποχικότητα των διαδοχικών παρατηρήσεων.

2.4.1 Ρόλος των κατανομών πιθανότητας

Κάθε μέτρηση PM_{10} μπορεί να θεωρηθεί ως παρατήρηση μιας τυχαίας μεταβλητής, της οποίας οι τιμές εκφράζουν συγκεντρώσεις αιωρούμενων σωματιδίων, συνήθως σε $\mu g/m^3$. Η σειρά των ωριαίων ή ημερήσιων συγκεντρώσεων δεν αποτελεί απλώς μια ακολουθία αριθμών, αλλά αποτύπωση ενός στοχαστικού μηχανισμού που επηρεάζεται από εκπομπές, μετεωρολογικές συνθήκες, τοπογραφικά χαρακτηριστικά και διαδικασίες μεταφοράς ή απομάκρυνσης των σωματιδίων. Η κατανομή πιθανότητας περιγράφει τη σχετική συχνότητα εμφάνισης χαμηλών, μέτριων και υψηλών τιμών και επιτρέπει τη συνολική αποτύπωση της συμπεριφοράς των δεδομένων (Wilks, 2019).

Η χρήση μόνο του μέσου όρου δεν επαρκεί για την πλήρη περιγραφή της σωματιδιακής ρύπανσης. Δύο πόλεις ή δύο διαφορετικά έτη μπορεί να εμφανίζουν παρόμοια μέση συγκέντρωση PM_{10} , αλλά να διαφέρουν σημαντικά ως προς τη διασπορά, την ασυμμετρία και τη συχνότητα εμφάνισης υψηλών τιμών. Σε μία περίπτωση οι συγκεντρώσεις μπορεί να είναι σχετικά σταθερές, ενώ σε μια άλλη να παρατηρούνται πολλές χαμηλές τιμές και περιορισμένος αριθμός πολύ υψηλών παρατηρήσεων. Επομένως, η ανάλυση της πλήρους κατανομής αποκαλύπτει πληροφορίες που δεν αποτυπώνονται επαρκώς από έναν μόνο περιγραφικό δείκτη (Wilks, 2019).

Οι κατανομές πιθανότητας παρέχουν ένα συνεπές μαθηματικό πλαίσιο για την περιγραφή της μεταβλητότητας των συγκεντρώσεων, την εκτίμηση της πιθανότητας εμφάνισης υψηλών τιμών και τη σύγκριση μεταξύ πόλεων ή ετών. Μέσω της εμπειρικής κατανομής, των ιστογραμμάτων και παραμετρικών μοντέλων, όπως η κανονική, η λογαριθμοκανονική, η Gamma και η Weibull, εξετάζονται το κέντρο, η διασπορά, η ασυμμετρία και η δεξιά ουρά των δεδομένων.

Η κατανομή πιθανότητας δεν υποκαθιστά την περιβαλλοντική ερμηνεία των συγκεντρώσεων, αλλά λειτουργεί ως εργαλείο συστηματικής περιγραφής και σύγκρισης. Η επιλογή της καταλληλότερης θεωρητικής κατανομής βασίζεται τόσο σε γραφικές διαγνωστικές απεικονίσεις όσο και σε ποσοτικά κριτήρια καλής προσαρμογής, τα οποία παρουσιάζονται στις επόμενες υποενότητες.

2.4.2 Συνάρτηση πυκνότητας πιθανότητας και αθροιστική συνάρτηση κατανομής

Για μια συνεχή τυχαία μεταβλητή, όπως η συγκέντρωση PM_{10} , η πιθανότητα δεν αποδίδεται σε μία μεμονωμένη τιμή, αλλά σε ένα διάστημα τιμών. Επομένως, η πιθανότητα να λάβει η μεταβλητή μία ακριβώς συγκεκριμένη τιμή είναι μηδενική, ενώ η πιθανότητα να βρίσκεται μέσα σε ένα εύρος τιμών μπορεί να υπολογιστεί μέσω της συνάρτησης πυκνότητας πιθανότητας.

Η συνάρτηση πυκνότητας πιθανότητας, γνωστή ως Probability Density Function (PDF), συμβολίζεται συνήθως με $f(x)$. Είναι μη αρνητική και το συνολικό εμβαδόν κάτω από την καμπύλη της ισούται με τη μονάδα:

$$f(x) \geq 0$$

και

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Η πιθανότητα να λάβει η μεταβλητή X τιμή μέσα σε ένα διάστημα $[a, b]$ ισούται με το ολοκλήρωμα της $f(x)$ από το a έως το b :

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

Η PDF δεν εκφράζει άμεσα την πιθανότητα μιας συγκεκριμένης τιμής, αλλά τη σχετική πυκνότητα εμφάνισης τιμών σε διαφορετικές περιοχές του πεδίου ορισμού. Υψηλότερες τιμές της $f(x)$ υποδηλώνουν περιοχές στις οποίες οι παρατηρήσεις αναμένεται να εμφανίζονται συχνότερα, χωρίς όμως η ίδια η τιμή της πυκνότητας να αποτελεί πιθανότητα. Η πιθανότητα προκύπτει μόνο από το εμβαδόν κάτω από την καμπύλη σε συγκεκριμένο διάστημα (Wilks, 2019).

Η αθροιστική συνάρτηση κατανομής, γνωστή ως Cumulative Distribution Function (CDF), συμβολίζεται με $F(x)$ και εκφράζει την πιθανότητα η μεταβλητή X να λάβει τιμή μικρότερη ή ίση από x :

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$

Για μεταβλητές με θετικό ή μη αρνητικό πεδίο ορισμού, όπως οι συγκεντρώσεις PM_{10} , το κάτω όριο του ολοκληρώματος προσαρμόζεται στο αντίστοιχο πεδίο τιμών. Οι τιμές της CDF βρίσκονται πάντοτε στο διάστημα:

$$0 \leq F(x) \leq 1$$

ενώ η συνάρτηση είναι μη φθίνουσα. Καθώς το x αυξάνεται, η $F(x)$ προσεγγίζει τη μονάδα, επειδή περιλαμβάνει σταδιακά μεγαλύτερο μέρος της συνολικής πιθανότητας.

Η πιθανότητα υπέρβασης μιας τιμής x υπολογίζεται ως:

$$P(X > x) = 1 - F(x)$$

Η σχέση αυτή είναι ιδιαίτερα χρήσιμη στην περιβαλλοντική ανάλυση, καθώς επιτρέπει την εκτίμηση της πιθανότητας οι συγκεντρώσεις PM_{10} να υπερβαίνουν ένα συγκεκριμένο επίπεδο αναφοράς. Ωστόσο, στην παρούσα εργασία η χρήση της CDF αφορά κυρίως τον

υπολογισμό θεωρητικών πιθανοτήτων ανά κλάση και όχι άμεσο έλεγχο συμμόρφωσης με ημερήσια ή ετήσια θεσμικά όρια (Wilks, 2019).

Η σύνδεση των θεωρητικών κατανομών με τις εμπειρικές συχνότητες πραγματοποιείται μέσω της CDF. Όταν οι παρατηρήσεις ομαδοποιούνται σε κλάσεις συγκεντρώσεων, για μία κλάση $(a_i, b_i]$ η θεωρητική πιθανότητα υπολογίζεται ως:

$$P(a_i < X \leq b_i) = F(b_i) - F(a_i)$$

Η διαφορά αυτή εκφράζει το ποσοστό της θεωρητικής κατανομής που βρίσκεται μεταξύ του κάτω και του άνω ορίου της κλάσης. Εάν το συνολικό πλήθος των έγκυρων παρατηρήσεων είναι n , τότε η αναμενόμενη συχνότητα της i -στής κλάσης δίνεται από:

$$E_i = n \cdot [F(b_i) - F(a_i)]$$

Όπου E_i είναι ο αναμενόμενος αριθμός παρατηρήσεων στην κλάση σύμφωνα με τη θεωρητική κατανομή. Η τιμή αυτή συγκρίνεται με την αντίστοιχη εμπειρική συχνότητα O_i , δηλαδή με το πλήθος των πραγματικών παρατηρήσεων που ανήκουν στο ίδιο διάστημα.

Για την τελευταία κλάση, όταν δεν υπάρχει πεπερασμένο άνω όριο, η θεωρητική πιθανότητα υπολογίζεται ως:

$$P(X > a_i) = 1 - F(a_i)$$

και η αναμενόμενη συχνότητα ως:

$$E_i = n[1 - F(a_i)]$$

Η διατύπωση αυτή είναι σημαντική όταν η ανώτερη κλάση περιλαμβάνει όλες τις συγκεντρώσεις πάνω από ένα συγκεκριμένο όριο, επειδή επιτρέπει να ενσωματωθεί ολόκληρη η δεξιά ουρά της θεωρητικής κατανομής. Παρότι η ανοικτή κλάση διευκολύνει τον υπολογισμό των συχνοτήτων, περιορίζει τη λεπτομερή απεικόνιση της συμπεριφοράς των ιδιαίτερα υψηλών τιμών. Η διαδικασία αποτελεί τη θεωρητική βάση για τη σύγκριση εμπειρικών και αναμενόμενων συχνοτήτων και για τον υπολογισμό του στατιστικού χ^2 . Με τον τρόπο αυτό, η PDF χρησιμοποιείται κυρίως για την περιγραφή του σχήματος της κατανομής, ενώ η CDF επιτρέπει τον υπολογισμό των πιθανοτήτων ανά διάστημα και τη μετατροπή τους σε αναμενόμενες συχνότητες. Οι δύο συναρτήσεις είναι επομένως συμπληρωματικές και αποτελούν βασικό μέρος της στατιστικής μοντελοποίησης των συγκεντρώσεων PM_{10} .

2.4.3 Κανονική κατανομή

Η κανονική κατανομή αποτελεί μία από τις θεμελιώδεις κατανομές της στατιστικής θεωρίας και χρησιμοποιείται συχνά ως βασικό μοντέλο αναφοράς για τη μελέτη συνεχών μεταβλητών. Η ευρεία εφαρμογή της συνδέεται, μεταξύ άλλων, με το Κεντρικό Οριακό Θεώρημα, σύμφωνα με το οποίο, υπό κατάλληλες προϋποθέσεις, τα αθροίσματα ή οι μέσοι όροι μεγάλου αριθμού τυχαίων μεταβλητών τείνουν να προσεγγίζουν την κανονική κατανομή. Για τον λόγο αυτό, χρησιμοποιείται συχνά στις φυσικές και περιβαλλοντικές επιστήμες, ιδιαίτερα όταν τα δεδομένα εμφανίζουν περίπου συμμετρική μορφή γύρω από μία κεντρική τιμή (Wilks, 2019).

Η κανονική κατανομή καθορίζεται πλήρως από δύο παραμέτρους: τη μέση τιμή μ , η οποία προσδιορίζει το κέντρο της κατανομής, και την τυπική απόκλιση σ , η οποία εκφράζει τη διασπορά των τιμών γύρω από τη μέση τιμή. Η συνάρτηση πυκνότητας πιθανότητας δίνεται από τη σχέση:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < \infty$$

όπου:

- μ είναι η μέση τιμή,
- $\sigma > 0$ είναι η τυπική απόκλιση,
- σ^2 είναι η διακύμανση.

Η κανονική κατανομή είναι συμμετρική ως προς τη μέση τιμή και έχει τη χαρακτηριστική μορφή καμπύλης καμπάνας. Σε αυτήν ισχύει ότι η μέση τιμή, η διάμεσος και η επικρατούσα τιμή συμπίπτουν. Μεγαλύτερες τιμές της σ οδηγούν σε περισσότερο απλωμένη κατανομή, ενώ μικρότερες τιμές συνεπάγονται μεγαλύτερη συγκέντρωση των παρατηρήσεων γύρω από τη μέση τιμή.

Μια κανονικά κατανεμημένη μεταβλητή X μπορεί να μετατραπεί σε τυπική κανονική μεταβλητή μέσω της σχέσης:

$$Z = \frac{X - \mu}{\sigma}$$

Η μεταβλητή Z ακολουθεί την τυπική κανονική κατανομή, με μέση τιμή 0 και τυπική απόκλιση 1. Η τυποποίηση επιτρέπει τον υπολογισμό πιθανοτήτων και ποσοστημορίων με βάση μία κοινή κατανομή, ανεξάρτητα από τις αρχικές τιμές των μ και σ (Wilks, 2019).

Παρά τη μεγάλη θεωρητική και πρακτική σημασία της, η κανονική κατανομή δεν είναι πάντοτε κατάλληλη για την περιγραφή των συγκεντρώσεων PM_{10} . Οι συγκεντρώσεις είναι μη αρνητικές, ενώ η κανονική κατανομή ορίζεται σε ολόκληρο τον πραγματικό άξονα και αποδίδει θετική πιθανότητα και σε αρνητικές τιμές. Επιπλέον, οι χρονοσειρές PM_{10} παρουσιάζουν συχνά θετική ασυμμετρία, με μεγάλο αριθμό χαμηλών ή μέτριων τιμών και μικρότερο αριθμό ιδιαίτερα υψηλών παρατηρήσεων. Η μορφή αυτή αποκλίνει από τη συμμετρική δομή της κανονικής κατανομής.

Στην παρούσα εργασία, η κανονική κατανομή εξετάζεται κυρίως ως μοντέλο αναφοράς για τη σύγκριση με κατανομές που ορίζονται για θετικές τιμές και μπορούν να αποδώσουν μεγαλύτερη δεξιά ασυμμετρία, όπως η λογαριθμοκανονική, η Gamma και η Weibull. Η καταλληλότητά της δεν προεξοφλείται, αλλά αξιολογείται με τα ίδια αριθμητικά και γραφικά κριτήρια που εφαρμόζονται και στις υπόλοιπες υποψήφιες κατανομές.

2.4.4 Λογαριθμοκανονική κατανομή

Η λογαριθμοκανονική κατανομή αποτελεί συνεχή παραμετρική κατανομή, η οποία εφαρμόζεται σε τυχαίες μεταβλητές που λαμβάνουν μόνο θετικές τιμές και παρουσιάζουν δεξιά ασυμμετρία. Μια τυχαία μεταβλητή X θεωρείται λογαριθμοκανονικά κατανομημένη όταν ο φυσικός λογάριθμός της ακολουθεί κανονική κατανομή.

Δηλαδή:

$$Y = \ln(X) \sim N(\mu_y, \sigma_y^2)$$

όπου:

- μ_y είναι η μέση τιμή
- σ_y^2 η διακύμανση των λογαριθμισμένων τιμών.

Η αντίστοιχη συνάρτηση πυκνότητας πιθανότητας της λογαριθμοκανονικής κατανομής δίνεται από τη σχέση:

$$f(x) = \frac{1}{x \sigma_y \sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu_y)^2}{2\sigma_y^2}\right], x > 0$$

Η κατανομή αυτή ορίζεται μόνο για θετικές τιμές και η συνάρτηση πυκνότητας πιθανότητας τείνει στο μηδέν τόσο όταν οι τιμές προσεγγίζουν το μηδέν όσο και όταν

γίνονται πολύ μεγάλες. Οι παράμετροι μ^y και σ_y^2 εκφράζουν αντίστοιχα τη μέση τιμή και τη διακύμανση των λογαριθμισμένων παρατηρήσεων. Η μ_y επηρεάζει κυρίως τη θέση της κατανομής, ενώ η σ_y καθορίζει τη διασπορά και τον βαθμό της θετικής ασυμμετρίας. Οι παράμετροι αυτές συνδέονται επίσης με τη γεωμετρική μέση τιμή και τη γεωμετρική τυπική απόκλιση της αρχικής μεταβλητής. Για τον λόγο αυτό, η λογαριθμοκανονική κατανομή χρησιμοποιείται συχνά για την περιγραφή θετικών περιβαλλοντικών μεγεθών που παρουσιάζουν δεξιά ασυμμετρία και περιορισμένο αριθμό υψηλών παρατηρήσεων (Seinfeld & Pandis, 2016· Wilks, 2019). Στις συγκεντρώσεις ατμοσφαιρικών ρύπων, η λογαριθμοκανονική κατανομή, μαζί με τη Weibull, συγκαταλέγεται στις κατανομές που έχουν χρησιμοποιηθεί ευρέως για την περιγραφή εμπειρικών κατανομών συγκέντρωσης (Seinfeld & Pandis, 2016). Οι συγκεντρώσεις PM_{10} σε αστικά και περιαστικά περιβάλλοντα εμφανίζουν συχνά χαρακτηριστικά συμβατά με λογαριθμοκανονική μορφή. Σε πολλές χρονοσειρές παρατηρείται μεγάλος αριθμός χαμηλών ή μέτριων τιμών και μικρότερος αριθμός ιδιαίτερα υψηλών παρατηρήσεων. Η μορφή αυτή δημιουργεί επιμήκη δεξιά ουρά στην εμπειρική κατανομή και καθιστά τη λογαριθμοκανονική κατανομή εύλογη υποψήφια για την περιγραφή των δεδομένων. Ωστόσο, η θεωρητική συμβατότητα δεν συνεπάγεται ότι η κατανομή είναι πάντοτε κατάλληλη, καθώς η ποιότητα της προσαρμογής πρέπει να αξιολογείται χωριστά για κάθε χρονοσειρά με αριθμητικά και γραφικά κριτήρια.

Η εφαρμογή της λογαριθμοκανονικής κατανομής σε δεδομένα PM_{10} προϋποθέτει κατάλληλο έλεγχο ποιότητας και προεπεξεργασία, καθώς ο φυσικός λογάριθμος ορίζεται μόνο για θετικές τιμές. Πριν από τη λογαριθμική μετατροπή πρέπει να απομακρύνονται οι ελλείπουσες ή μη έγκυρες παρατηρήσεις, όπως οι ειδικοί κωδικοί απουσίας μέτρησης τύπου -9999. Οι αρνητικές τιμές δεν έχουν φυσικό νόημα για συγκεντρώσεις ατμοσφαιρικών ρύπων, ενώ οι μηδενικές τιμές δεν μπορούν να χρησιμοποιηθούν απευθείας σε λογαριθμικό μετασχηματισμό. Επομένως, η διατήρηση μόνο έγκυρων θετικών συγκεντρώσεων αποτελεί αναγκαία προϋπόθεση για την ορθή εφαρμογή της κατανομής.

Στην παρούσα εργασία, η λογαριθμοκανονική κατανομή εξετάζεται ως μία από τις βασικές υποψήφιες κατανομές για τη μοντελοποίηση των ωριαίων συγκεντρώσεων PM_{10} . Η τελική αξιολόγηση της καταλληλότητάς της βασίζεται στη σύγκριση εμπειρικών και αναμενόμενων συχνοτήτων, στο στατιστικό χ^2 και στις αντίστοιχες γραφικές διαγνωστικές απεικονίσεις.

2.4.5 Κατανομή Gamma

Η κατανομή Gamma είναι μία συνεχής παραμετρική κατανομή, ορισμένη για θετικές τιμές της τυχαίας μεταβλητής, δηλαδή για $x > 0$.

Η συνάρτηση πυκνότητας πιθανότητας δίνεται από τη σχέση:

$$f(x) = \frac{(x/b)^{a-1} e^{-x/b}}{b\Gamma(a)}, x > 0, a > 0, b > 0$$

όπου:

- a είναι η παράμετρος μορφής (shape parameter),
- b η παράμετρος κλίμακας (scale parameter),
- $\Gamma(a)$ η συνάρτηση Gamma.

Οι δύο αυτές παράμετροι καθορίζουν τόσο τη μορφή όσο και την οριζόντια κλίμακα της κατανομής. Η μέση τιμή και η διακύμανση της κατανομής δίνονται αντίστοιχα από τις σχέσεις:

$$E(X) = ab$$

και

$$\text{Var}(X) = ab^2$$

Μικρές τιμές της παραμέτρου μορφής a οδηγούν σε έντονα δεξιά ασύμμετρες κατανομές, ενώ όσο η τιμή της αυξάνεται η κατανομή γίνεται περισσότερο συμμετρική. Για αρκετά μεγάλες τιμές της a , η μορφή της μπορεί να προσεγγίζει την κανονική κατανομή. Αντίστοιχα, η παράμετρος κλίμακας b επηρεάζει την έκταση της κατανομής στον οριζόντιο άξονα, χωρίς να μεταβάλλει τη βασική μορφή που καθορίζεται από την a (Wilks, 2019).

Η ευελιξία αυτή καθιστά την κατανομή Gamma κατάλληλη για τη μοντελοποίηση θετικών περιβαλλοντικών μεταβλητών που παρουσιάζουν θετική ασυμμετρία. Σε τέτοια δεδομένα, οι περισσότερες παρατηρήσεις συγκεντρώνονται συνήθως σε χαμηλές ή μέτριες τιμές, ενώ μικρότερος αριθμός παρατηρήσεων εμφανίζεται στη δεξιά ουρά. Για τον λόγο αυτό, η κατανομή Gamma χρησιμοποιείται ευρέως στην περιβαλλοντική και ατμοσφαιρική στατιστική, ιδιαίτερα για μεταβλητές όπως η βροχόπτωση και άλλα υδρομετεωρολογικά μεγέθη (Wilks, 2019).

Η κατανομή Gamma μπορεί να εξεταστεί και ως υποψήφιο μοντέλο για τις συγκεντρώσεις PM_{10} , καθώς οι έγκυρες παρατηρήσεις είναι θετικές και συχνά παρουσιάζουν δεξιά

ασυμμετρία. Σε πολλές χρονοσειρές καταγράφεται μεγάλος αριθμός χαμηλών ή μέτριων συγκεντρώσεων και μικρότερος αριθμός υψηλών τιμών. Σε αντίθεση με την κανονική κατανομή, η Gamma δεν αποδίδει πιθανότητα σε αρνητικές τιμές και, επομένως, είναι θεωρητικά περισσότερο συμβατή με δεδομένα συγκεντρώσεως. Η καταλληλότητά της, ωστόσο, δεν θεωρείται δεδομένη, αλλά πρέπει να αξιολογείται εμπειρικά για κάθε χρονοσειρά.

Η Gamma και η λογαριθμοκανονική κατανομή αποτελούν δύο βασικές επιλογές για τη μοντελοποίηση θετικών και δεξιά ασύμμετρων δεδομένων, αλλά διαφέρουν ως προς τη μαθηματική τους δομή. Η λογαριθμοκανονική προκύπτει όταν ο φυσικός λογάριθμος της μεταβλητής ακολουθεί κανονική κατανομή, ενώ η Gamma εφαρμόζεται απευθείας στις αρχικές τιμές. Επιπλέον, οι δύο κατανομές μπορεί να αποδίδουν διαφορετικά το κεντρικό τμήμα και τη δεξιά ουρά των δεδομένων, γεγονός που καθιστά αναγκαία τη μεταξύ τους σύγκριση με κοινά κριτήρια προσαρμογής (Wilks, 2019).

Στην παρούσα εργασία, η κατανομή Gamma εξετάζεται ως μία από τις υποψήφιες κατανομές για τις ωριαίες συγκεντρώσεις PM₁₀. Η τελική αξιολόγηση της προσαρμογής της βασίζεται στη σύγκριση εμπειρικών και αναμενόμενων συχνοτήτων, στο στατιστικό χ^2 και στις αντίστοιχες γραφικές διαγνωστικές απεικονίσεις.

2.4.6 Κατανομή Weibull

Η κατανομή Weibull αποτελεί συνεχή διπαραμετρική κατανομή πιθανότητας, η οποία ορίζεται για μη αρνητικές τιμές και χρησιμοποιείται ως ευέλικτο μοντέλο για θετικά συνεχή μεγέθη. Στην ατμοσφαιρική ρύπανση, οι συγκεντρώσεις αντιμετωπίζονται ως τυχαίες μεταβλητές, καθώς παρουσιάζουν μεταβλητότητα που συνδέεται με διαφορετικές πηγές, μετεωρολογικές συνθήκες και διεργασίες μεταφοράς και διασποράς. Η Weibull, μαζί με τη λογαριθμοκανονική κατανομή, έχει χρησιμοποιηθεί ευρέως για την περιγραφή εμπειρικών κατανομών συγκεντρώσεων ατμοσφαιρικών ρύπων (Seinfeld & Pandis, 2016).

Η συνάρτηση πυκνότητας πιθανότητας της κατανομής Weibull δίνεται από τη σχέση:

$$f(c) = \frac{\lambda}{\sigma} \left(\frac{c}{\sigma} \right)^{\lambda-1} \exp \left[- \left(\frac{c}{\sigma} \right)^{\lambda} \right], \lambda, \sigma > 0$$

όπου:

- λ είναι η παράμετρος μορφής,
- σ η παράμετρος κλίμακας.

Η παράμετρος μορφής καθορίζει τη βασική μορφή και τον βαθμό ασυμμετρίας της κατανομής, ενώ η παράμετρος κλίμακας επηρεάζει την έκταση των τιμών στον οριζόντιο άξονα. Η αντίστοιχη αθροιστική συνάρτηση κατανομής είναι:

$$F(c) = 1 - \exp \left[- \left(\frac{c}{\sigma} \right)^\lambda \right]$$

ενώ η πιθανότητα υπέρβασης ενός επιπέδου συγκέντρωσης x δίνεται από τη συμπληρωματική συνάρτηση κατανομής:

$$P(C > x) = \exp \left[- \left(\frac{x}{\sigma} \right)^\lambda \right]$$

Η τελευταία σχέση είναι ιδιαίτερα χρήσιμη σε περιβαλλοντικές εφαρμογές, καθώς επιτρέπει την εκτίμηση της πιθανότητας υπέρβασης συγκεκριμένων ορίων συγκέντρωσης. Η κατανομή Weibull έχει χρησιμοποιηθεί για την προσαρμογή δεδομένων ωριαίων μέσων και ημερήσιων μέγιστων ωριαίων μέσων συγκεντρώσεων όζοντος, δείχνοντας ότι μπορεί να αποδώσει ικανοποιητικά τόσο τη γενική μορφή όσο και τα υψηλά ποσοστημόρια της εμπειρικής κατανομής (Seinfeld & Pandis, 2016).

Σε σχέση με τις συγκεντρώσεις PM_{10} , η κατανομή Weibull μπορεί να λειτουργήσει ως υποψήφιο πρότυπο όταν οι παρατηρούμενες τιμές εμφανίζουν μη αρνητικό χαρακτήρα, δεξιά ασυμμετρία, αυξημένη μεταβλητότητα και παρουσία υψηλών τιμών, χαρακτηριστικά που δεν αποδίδονται πάντοτε ικανοποιητικά από την κανονική κατανομή. Ωστόσο, η καταλληλότητά της δεν πρέπει να θεωρείται δεδομένη εκ των προτέρων. Η κατανομή που προσαρμόζεται καλύτερα σε ένα συγκεκριμένο σύνολο ατμοσφαιρικών δεδομένων εξαρτάται από τα εμπειρικά χαρακτηριστικά του ίδιου του συνόλου δεδομένων (Seinfeld & Pandis, 2016).

Στην παρούσα εργασία, η κατανομή Weibull εξετάζεται ως μία από τις υποψήφιες κατανομές προσαρμογής, λόγω της ικανότητάς της να περιγράφει θετικές και ασύμμετρες μεταβλητές με διαφορετικά σχήματα κατανομής. Ο ρόλος της δεν είναι προεξοφλημένα προνομιούχος έναντι άλλων κατανομών, όπως η λογαριθμοκανονική και η Gamma, αλλά εντάσσεται σε συγκριτική αξιολόγηση εναλλακτικών μοντέλων. Η τελική αποτίμηση της καταλληλότητάς της βασίζεται σε κριτήρια προσαρμογής, στην αναπαράσταση των

υψηλών συγκεντρώσεων, στην εκτίμηση πιθανοτήτων υπέρβασης και στη συμφωνία της θεωρητικής κατανομής με τα εμπειρικά χαρακτηριστικά των δεδομένων PM_{10} .

2.4.7 Αξιολόγηση προσαρμογής και επιλογή βέλτιστης κατανομής

Η αξιολόγηση της προσαρμογής μιας θεωρητικής κατανομής σε δεδομένα ατμοσφαιρικής ρύπανσης βασίζεται στη σύγκριση της εμπειρικής συμπεριφοράς των παρατηρήσεων με τις πιθανότητες και τις συχνότητες που προβλέπει το αντίστοιχο θεωρητικό μοντέλο. Μια κατανομή θεωρείται περισσότερο κατάλληλη όταν αποδίδει ικανοποιητικά τη θέση, τη διασπορά, την ασυμμετρία και τη δεξιά ουρά των δεδομένων, χωρίς έντονες ή συστηματικές αποκλίσεις (Wilks, 2019). Η αξιολόγηση αυτή είναι ιδιαίτερα σημαντική για τις συγκεντρώσεις PM_{10} , οι οποίες είναι μη αρνητικές και συχνά εμφανίζουν θετική ασυμμετρία και περιορισμένο αριθμό υψηλών παρατηρήσεων.

Όταν οι παρατηρήσεις ομαδοποιούνται σε κλάσεις, η θεωρητική πιθανότητα της i -στής κλάσης $(a_i, b_i]$ υπολογίζεται από την αθροιστική συνάρτηση κατανομής ως:

$$P(a_i < X \leq b_i) = F(b_i) - F(a_i)$$

Εάν n είναι το συνολικό πλήθος των έγκυρων παρατηρήσεων, η αναμενόμενη συχνότητα της κλάσης δίνεται από τη σχέση:

$$E_i = n \cdot [F(b_i) - F(a_i)]$$

όπου:

- E_i είναι η αναμενόμενη συχνότητα της i -στής κλάσης,
- n το πλήθος των έγκυρων παρατηρήσεων,
- a_i και b_i το κάτω και άνω όριο της κλάσης,
- F η θεωρητική CDF.

Η χρήση της CDF είναι απαραίτητη, καθώς η πιθανότητα ενός διαστήματος δεν προκύπτει από μία μεμονωμένη τιμή της συνάρτησης πυκνότητας πιθανότητας, αλλά από τη διαφορά των αθροιστικών πιθανοτήτων στα δύο όριά του.

Για την τελευταία ανοικτή κλάση, όταν δεν υπάρχει πεπερασμένο άνω όριο, η πιθανότητα υπολογίζεται ως:

$$P(X > a_i) = 1 - F(a_i)$$

και η αντίστοιχη αναμενόμενη συχνότητα ως:

$$E_i = n[1 - F(a_i)]$$

Η χρήση ανοικτής ανώτερης κλάσης επιτρέπει την ενσωμάτωση όλων των υψηλών παρατηρήσεων, αλλά περιορίζει τη λεπτομερή αξιολόγηση της ακραίας δεξιάς ουράς, επειδή διαφορετικές υψηλές τιμές ομαδοποιούνται στο ίδιο διάστημα.

Η αριθμητική σύγκριση μεταξύ εμπειρικών και αναμενόμενων συχνοτήτων πραγματοποιείται μέσω του στατιστικού:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

όπου:

- O_i είναι η παρατηρούμενη συχνότητα της i -στής κλάσης,
- E_i η αντίστοιχη αναμενόμενη συχνότητα
- k ο συνολικός αριθμός των κλάσεων.

Μικρότερη τιμή του χ^2 υποδηλώνει μικρότερη συνολική απόκλιση μεταξύ εμπειρικών και θεωρητικών συχνοτήτων.

Στην παρούσα εργασία, το χ^2 χρησιμοποιείται κυρίως ως συγκριτικό μέτρο μεταξύ των υποψήφιων κατανομών για την ίδια χρονοσειρά και με κοινή ομαδοποίηση των παρατηρήσεων. Δεν ερμηνεύεται ως αυτοτελής τυπικός έλεγχος αποδοχής ή απόρριψης μιας κατανομής, καθώς δεν υπολογίζονται κρίσιμες τιμές ή p-values. Η σύγκριση είναι ουσιαστική μόνο όταν αφορά το ίδιο δείγμα, τις ίδιες κλάσεις και τον ίδιο τρόπο υπολογισμού των αναμενόμενων συχνοτήτων.

Η τιμή του χ^2 επηρεάζεται από τον αριθμό και τα όρια των κλάσεων. Πολύ ευρείες κλάσεις μπορεί να αποκρύψουν τοπικές αποκλίσεις, ενώ υπερβολικά στενές κλάσεις μπορεί να οδηγήσουν σε μικρές αναμενόμενες συχνότητες και μεγαλύτερη αστάθεια του μέτρου. Για τον λόγο αυτό, η αριθμητική αξιολόγηση δεν εξετάζεται ανεξάρτητα από τις γραφικές διαγνωστικές απεικονίσεις (Wilks, 2019).

Τα Q-Q διαγράμματα συγκρίνουν τα εμπειρικά με τα θεωρητικά ποσοστημόρια. Όταν τα σημεία βρίσκονται κοντά στην ευθεία αναφοράς, η προσαρμογή θεωρείται ικανοποιητική, ενώ συστηματικές αποκλίσεις μπορεί να υποδηλώνουν διαφορές στο κεντρικό τμήμα ή στις ουρές της κατανομής. Τα Q-Q διαγράμματα είναι ιδιαίτερα χρήσιμα, επειδή

επιτρέπουν να διαπιστωθεί αν μια κατανομή αποδίδει εξίσου καλά τις χαμηλές, τις μέσες και τις υψηλές παρατηρήσεις.

Στα διαγράμματα παρατηρούμενων-αναμενόμενων συχνοτήτων συγκρίνονται οι εμπειρικές με τις θεωρητικές συχνότητες ανά κλάση. Ο οριζόντιος άξονας αντιστοιχεί στις αναμενόμενες συχνότητες και ο κατακόρυφος στις εμπειρικές. Ικανοποιητική προσαρμογή αναμένεται όταν τα σημεία βρίσκονται κοντά στη γραμμή ταύτισης, ενώ συστηματικές αποκλίσεις από αυτήν υποδηλώνουν περιοχές στις οποίες το θεωρητικό μοντέλο υπερεκτιμά ή υποεκτιμά τις παρατηρούμενες συχνότητες.

Τα διαγράμματα υπολοίπων βασίζονται στις διαφορές:

$$r_i = O_i - E_i$$

Σε περίπτωση ικανοποιητικής προσαρμογής, τα υπόλοιπα αναμένεται να κυμαίνονται γύρω από το μηδέν χωρίς έντονο ή συστηματικό μοτίβο. Στην παρούσα εφαρμογή, ο οριζόντιος άξονας αντιστοιχεί σε διαδοχικά ταξινομημένες συγκεντρώσεις ή κλάσεις και όχι στη χρονική ακολουθία των παρατηρήσεων. Επομένως, τα συγκεκριμένα διαγράμματα δεν αξιολογούν χρονική εξάρτηση ή αυτοσυσχέτιση.

Κάθε διαγνωστικό εργαλείο εξετάζει διαφορετική πτυχή της προσαρμογής. Το χ^2 συνοψίζει τις αποκλίσεις των ομαδοποιημένων συχνοτήτων, τα Q-Q διαγράμματα εξετάζουν τη συμφωνία των ποσοστημορίων, ενώ τα διαγράμματα παρατηρούμενων-αναμενόμενων συχνοτήτων και υπολοίπων αναδεικνύουν τοπικές αποκλίσεις ανά κλάση. Για τον λόγο αυτό, οι επιμέρους μέθοδοι δεν είναι απαραίτητο να κατατάσσουν πάντοτε τις υποψήφιες κατανομές με τον ίδιο τρόπο. Η τελική επιλογή πρέπει να βασίζεται στη συνδυαστική αξιολόγηση όλων των διαθέσιμων ενδείξεων.

Η επιλογή της «βέλτιστης» κατανομής έχει σχετικό και όχι απόλυτο χαρακτήρα. Ως βέλτιστη ορίζεται η κατανομή που, μεταξύ της κανονικής, της λογαριθμοκανονικής, της Gamma και της Weibull, παρουσιάζει τη μικρότερη συνολική απόκλιση και την περισσότερο ικανοποιητική εικόνα στα αριθμητικά και γραφικά κριτήρια. Η επιλογή αυτή δεν σημαίνει ότι το συγκεκριμένο μοντέλο αποτελεί τη μοναδική ή πραγματική κατανομή της διαδικασίας που παράγει τις συγκεντρώσεις PM₁₀, αλλά ότι περιγράφει καλύτερα τα διαθέσιμα δεδομένα από τα υπόλοιπα μοντέλα που εξετάστηκαν.

Η αξιολόγηση αφορά την οριακή κατανομή των ωριαίων συγκεντρώσεων PM₁₀ και όχι τη χρονική εξάρτηση, την αυτοσυσχέτιση ή την εποχικότητα των διαδοχικών παρατηρήσεων. Επιπλέον, τα αποτελέσματα αφορούν τα συγκεκριμένα έτη που επιλέχθηκαν με βάση την

πληρότητα των διαθέσιμων δεδομένων και δεν πρέπει να θεωρούνται αυτομάτως αντιπροσωπευτικά της μακροχρόνιας συμπεριφοράς κάθε πόλης.

2.5 Προηγούμενες μελέτες για κατανομές

Οι συγκεντρώσεις PM_{10} συχνά αποκλίνουν από τη συμμετρική κανονική κατανομή, καθώς διαμορφώνονται από την αλληλεπίδραση εκπομπών, μετεωρολογικών συνθηκών, διεργασιών μεταφοράς, χημικού μετασχηματισμού και απομάκρυνσης. Ως αποτέλεσμα, τα δεδομένα είναι μη αρνητικά και παρουσιάζουν συχνά θετική ασυμμετρία, με μεγαλύτερη συχνότητα χαμηλών ή μέτριων τιμών και μικρότερο αριθμό υψηλών συγκεντρώσεων. Για τον λόγο αυτό, η κανονική κατανομή χρησιμοποιείται συνήθως ως μοντέλο αναφοράς, ενώ η βιβλιογραφία εξετάζει συστηματικά κατανομές ορισμένες σε θετικές τιμές, όπως η λογαριθμοκανονική, η Gamma και η Weibull (Lu & Fang, 2002· Diapouli et al., 2017· Plocoste et al., 2020).

Η λογαριθμοκανονική κατανομή αποτελεί μία από τις συχνότερα εφαρμοζόμενες προσεγγίσεις στη στατιστική περιγραφή των PM_{10} . Η θεωρητική της καταλληλότητα συνδέεται με το γεγονός ότι οι συγκεντρώσεις ατμοσφαιρικών ρύπων μπορεί να προκύπτουν από διαδοχικές διεργασίες αραίωσης, ανάμιξης και μεταφοράς, οι οποίες οδηγούν συχνά σε δεξιά ασύμμετρη μορφή. Σε μελέτες για την Ταϊβάν, οι Lu και Fang συνέκριναν τη λογαριθμοκανονική, τη Weibull και την Pearson τύπου V και έδειξαν ότι η λογαριθμοκανονική μπορεί να περιγράψει ικανοποιητικά τη συνολική κατανομή των ημερήσιων συγκεντρώσεων PM_{10} , ενώ για την ανάλυση πολύ υψηλών τιμών χρησιμοποίησαν ειδικότερες κατανομές ακραίων τιμών (Lu, 2002· Lu & Fang, 2003).

Αντίστοιχα, στον ευρωπαϊκό και ελληνικό χώρο έχουν εφαρμοστεί κατανομές θετικών και δεξιά ασύμμετρων μεταβλητών. Οι Giavis, Kambezidis και Lykoudis εξέτασαν ωριαίες συγκεντρώσεις PM_{10} στην Αθήνα και στο Manchester και συνέκριναν τη λογαριθμοκανονική, τη Gamma και τη Weibull (Giavis, Kambezidis & Lykoudis, 2008). Τα αποτελέσματα έδειξαν καλύτερη συνολική προσαρμογή της λογαριθμοκανονικής κατανομής με βάση δείκτες όπως το MBE, το RMSE και ο δείκτης συμφωνίας d, ενώ οι Gamma και Weibull παρουσίασαν μεγαλύτερες αποκλίσεις, ιδιαίτερα στις υψηλότερες τιμές.

Στην Κεντρική Ελλάδα, οι Papanastasiou και Melas μελέτησαν ημερήσιες συγκεντρώσεις PM_{10} στον Βόλο και έδειξαν ότι η λογαριθμοκανονική κατανομή μπορούσε να περιγράψει ικανοποιητικά τα δεδομένα. Η προσαρμογή χρησιμοποιήθηκε και για την εκτίμηση ημερών υπέρβασης του ημερήσιου ορίου των $50 \mu g/m^3$, καθώς και για τη διερεύνηση της απαιτούμενης μείωσης εκπομπών (Papanastasiou & Melas, 2010). Η μελέτη αυτή είναι ιδιαίτερα σχετική με την παρούσα εργασία, επειδή αφορά ελληνική αστική περιοχή και συνδέει τη στατιστική μοντελοποίηση με την αξιολόγηση της ποιότητας του αέρα.

Συμπληρωματικά, οι Moustris, Petraki, Ntourou, Priniotakis και Nikolopoulos ανέλυσαν ημερήσιες συγκεντρώσεις PM_{10} στην ευρύτερη περιοχή της Αθήνας για την περίοδο 2001-2018. Η μελέτη δεν επικεντρώθηκε στην προσαρμογή θεωρητικών κατανομών, αλλά ανέδειξε έντονη χωρική, εποχική και μακροχρόνια μεταβλητότητα, καθώς και διαφοροποιήσεις στις υπερβάσεις του ημερήσιου ορίου (Moustris et al., 2020). Τα ευρήματα αυτά ενισχύουν την ανάγκη στατιστικής διερεύνησης των PM_{10} , καθώς δείχνουν ότι οι χρονοσειρές τους διαμορφώνονται από πολλαπλούς και μεταβαλλόμενους παράγοντες.

Η ελληνική βιβλιογραφία έχει εξετάσει και πιο γενικευμένες οικογένειες κατανομών. Οι Gavriil et al. (2006) εφάρμοσαν οκτώ θεωρητικές κατανομές σε ημερήσιες συγκεντρώσεις PM_{10} και $PM_{2.5}$ στην Αθήνα, μεταξύ των οποίων οι λογαριθμοκανονική, Gamma, Weibull, beta, log-logistic, inverse Gaussian και Pearson τύπου V και VI. Οι Pearson VI και V παρουσίασαν την καλύτερη συνολική προσαρμογή, ενώ η inverse Gaussian και η λογαριθμοκανονική ακολούθησαν. Οι Gamma και Weibull εμφάνισαν ασθενέστερη προσαρμογή, κυρίως στις υψηλές συγκεντρώσεις. Το αποτέλεσμα αυτό δείχνει ότι, σε ορισμένα σύνολα δεδομένων, μπορεί να απαιτούνται κατανομές με μεγαλύτερη ευελιξία ή βαρύτερες ουρές.

Διεθνείς μελέτες επιβεβαιώνουν ότι η κατάταξη των υποψήφιων κατανομών δεν είναι σταθερή. Οι Wang, Chen, Chen και Kan, εξετάζοντας ημερήσιες συγκεντρώσεις PM_{10} σε πέντε μεγάλες πόλεις της Κίνας, συνέκριναν τη λογαριθμοκανονική, τη Weibull και τη Gamma και διαπίστωσαν ότι η λογαριθμοκανονική περιέγραφε ικανοποιητικά τα δεδομένα. Παράλληλα, έδειξαν ότι διαφορετικά κριτήρια καλής προσαρμογής μπορεί να οδηγούν σε διαφορετική κατάταξη, επειδή δεν αποδίδουν το ίδιο βάρος στο κεντρικό τμήμα και στις ουρές της κατανομής (Wang et al., 2013).

Η χωρική περιοχή, η χρονική ανάλυση και το τμήμα της κατανομής που ενδιαφέρει κάθε μελέτη επηρεάζουν σημαντικά την επιλογή μοντέλου. Οι ωριαίες τιμές διατηρούν εντονότερη βραχυχρόνια μεταβλητότητα και συχνά μεγαλύτερη δεξιά ασυμμετρία, ενώ οι ημερήσιες μέσες τιμές εξομαλύνουν μέρος της μεταβλητότητας. Παράλληλα, μια κατανομή που περιγράφει ικανοποιητικά το κύριο σώμα των δεδομένων δεν είναι απαραίτητα η καταλληλότερη για τις υψηλότερες τιμές. Για τον λόγο αυτό, αρκετές μελέτες εξετάζουν χωριστά την άνω ουρά ή χρησιμοποιούν πιο ευέλικτες κατανομές, όπως Pearson, Burr ή κατανομές ακραίων τιμών (Lu & Fang, 2003· Gavriil et al., 2006· He, Schäfer & Beck, 2022).

Η σημασία της περιβαλλοντικής ιδιαιτερότητας φαίνεται και σε νεότερες εφαρμογές. Οι Plocoste et al. (2020), χρησιμοποιώντας ενδεκαετή σειρά ημερήσιων συγκεντρώσεων PM_{10} στη Γουαδελούπη, συνέκριναν λογαριθμοκανονική, Weibull, Burr και stable κατανομή, καθώς και μίγματα κατανομών. Η Burr απέδωσε καλύτερα τη γενική μορφή των δεδομένων, αλλά η εποχική διμορφία που συνδεόταν με τη μεταφορά αφρικανικής σκόνης περιγράφηκε πληρέστερα από μίγμα Burr-Weibull. Το εύρημα αυτό δείχνει ότι διαφορετικές περιβαλλοντικές συνθήκες μέσα στο ίδιο σύνολο δεδομένων μπορούν να παράγουν διαφορετικά στατιστικά καθεστώτα.

Σε ευρωπαϊκή κλίμακα, οι He, Schäfer & Beck, ανέλυσαν πενταετή ωριαία δεδομένα ατμοσφαιρικών ρύπων από μεγάλο αριθμό σταθμών και έδειξαν ότι οι ουρές των εμπειρικών κατανομών παρουσιάζουν συχνά βαριά συμπεριφορά και έντονη χωρική ετερογένεια. Η μελέτη τους υποστηρίζει ότι οι κλασικές κατανομές μπορούν να είναι χρήσιμες για το κύριο σώμα των δεδομένων, αλλά οι υψηλές συγκεντρώσεις μπορεί να απαιτούν διαφορετική στατιστική προσέγγιση (He, Schäfer & Beck, 2022).

Πρόσφατες εφαρμογές επιβεβαιώνουν ότι ακόμη και όταν εξετάζονται οι ίδιες κατανομές, η τελική επιλογή μπορεί να διαφοροποιείται ανά περιοχή. Οι Meirelles και Vasconcelos, σε δεδομένα από τις Αζόρες, διαπίστωσαν ότι τα $PM_{2.5}$ περιγράφονταν καλύτερα από λογαριθμοκανονική κατανομή, ενώ τα PM_{10} από Gamma (Meirelles & Vasconcelos, 2025). Αντίστοιχα, ο Bagci, μελετώντας χειμερινές ημερήσιες συγκεντρώσεις PM_{10} στην πόλη Van της Τουρκίας, βρήκε ελαφρώς καλύτερη προσαρμογή της Gamma σε σχέση με τη λογαριθμοκανονική και τη Weibull (Bagci, 2023).

Μελέτη	Περιοχή	Δεδομένα	Κατανομές	Βασικό εύρημα
Lu & Fang (2003)	Ταϊβάν	Ημερήσια	Λογαριθμοκανονική, Weibull, Pearson V, ακραίων τιμών	Η λογαριθμοκανονική περιέγραψε ικανοποιητικά τη συνολική κατανομή, ενώ οι ακραίες τιμές απαιτούσαν ειδικότερη προσέγγιση
Giavis, Kambezidis & Lykoudis (2008)	Αθήνα, Manchester	Ωριαία PM ₁₀	Λογαριθμοκανονική, Gamma, Weibull	Η λογαριθμοκανονική εμφάνισε καλύτερη προσαρμογή
Papanastasiou & Melas (2010)	Βόλος	Ημερήσια PM ₁₀	Λογαριθμοκανονική	Χρησιμοποιήθηκε για πρόβλεψη υπερβάσεων του ορίου 50 µg/m ³
Gavriil et al. (2006)	Αθήνα	Ημερήσια PM ₁₀ /PM _{2.5}	Pearson V/VI, lognormal, Gamma, Weibull κ.ά.	Οι Pearson V/VI είχαν καλύτερη συνολική προσαρμογή
Wang et al. (2013)	Κίνα	Ημερήσια PM ₁₀	Λογαριθμοκανονική, Weibull, Gamma	Η λογαριθμοκανονική περιέγραψε ικανοποιητικά τα δεδομένα
Plocoste et al. (2020)	Γουαδελούπη	Ημερήσια PM ₁₀	Lognormal, Weibull, Burr, stable, μίγματα	Το μίγμα Burr-Weibull απέδωσε καλύτερα την εποχική διμορφία
Meirelles & Vasconcelos (2025)	Αζόρες	PM ₁₀ /PM _{2.5}	Κανονική, lognormal, Weibull, Gamma	Τα PM ₁₀ περιγράφηκαν καλύτερα από Gamma
Bagcı (2023)	Van, Τουρκία	Χειμερινές ημερήσιες PM ₁₀	Lognormal, Weibull, Gamma	Η Gamma απέδωσε ελαφρώς καλύτερα

Πίνακας 2.4 Ενδεικτικές προηγούμενες μελέτες προσαρμογής κατανομών σε δεδομένα PM₁₀.

Η βιβλιογραφία δείχνει ότι δεν υπάρχει μία καθολικά βέλτιστη κατανομή για την περιγραφή των συγκεντρώσεων PM₁₀. Η καταλληλότητα κάθε μοντέλου εξαρτάται από την περιοχή μελέτης, τη χρονική ανάλυση των δεδομένων, τις επικρατούσες

περιβαλλοντικές συνθήκες, το μέγεθος και την ποιότητα του δείγματος, καθώς και από το κριτήριο που χρησιμοποιείται για την αξιολόγηση της προσαρμογής. Η λογαριθμοκανονική, η Gamma και η Weibull χρησιμοποιούνται συστηματικά ως βασικά συγκριτικά μοντέλα, επειδή μπορούν να περιγράψουν θετικές και δεξιά ασύμμετρες μεταβλητές. Ωστόσο, σε δεδομένα με εντονότερη ασυμμετρία, βαρύτερες ουρές ή σύνθετη στατιστική δομή, πιο ευέλικτες κατανομές, όπως οι Pearson ή Burr, ενδέχεται να προσφέρουν καλύτερη προσαρμογή, ιδιαίτερα στην περιοχή των υψηλών συγκεντρώσεων. Στο πλαίσιο αυτό, η παρούσα εργασία εντάσσεται σε μια ήδη τεκμηριωμένη ερευνητική παράδοση, καθώς συγκρίνει την κανονική, τη λογαριθμοκανονική, τη Gamma και τη Weibull σε ωριαίες χρονοσειρές PM_{10} από πόλεις της Στερεάς Ελλάδας. Καμία από τις υποψήφιες κατανομές δεν θεωρείται εκ των προτέρων καταλληλότερη ή καθολικά αποδεκτή. Η επιλογή της βέλτιστης κατανομής πραγματοποιείται ξεχωριστά για κάθε χρονοσειρά και βασίζεται στη συνολική συμφωνία με την εμπειρική κατανομή, στη συγκριτική τιμή του στατιστικού χ^2 και στις γραφικές διαγνωστικές απεικονίσεις. Ιδιαίτερη προσοχή δίνεται στις συστηματικές αποκλίσεις που εμφανίζονται στο κεντρικό τμήμα και στις υψηλότερες τιμές, ώστε η τελική επιλογή να στηρίζεται σε συνδυαστική και όχι μονοδιάστατη αξιολόγηση.

3. Μεθοδολογία

Η παρούσα εργασία έχει ως αντικείμενο τη στατιστική διερεύνηση της κατανομικής συμπεριφοράς των ωριαίων συγκεντρώσεων PM_{10} σε επιλεγμένες πόλεις της Περιφέρειας Στερεάς Ελλάδας. Τα δεδομένα προέρχονται από το Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης και αφορούν μετρήσεις από τους σταθμούς της Λαμίας, της Χαλκίδας, της Λειβαδιάς, της Άμφισσας και του Καρπενησίου. Οι συγκεντρώσεις εκφράζονται σε $\mu g/m^3$ και εξετάζονται σε ωριαία χρονική κλίμακα.

Η ανάλυση οργανώθηκε σε διαδοχικά στάδια, ώστε η εφαρμογή των θεωρητικών κατανομών να βασίζεται σε επαρκή και συγκρίσιμα δεδομένα. Αρχικά πραγματοποιήθηκαν η συλλογή, η οργάνωση και η προεπεξεργασία των διαθέσιμων χρονοσειρών. Στη συνέχεια ελέγχθηκαν η πληρότητα και η χρονική συνέχεια κάθε ετήσιου συνόλου δεδομένων, προκειμένου να επιλεγούν μόνο οι περιπτώσεις που μπορούσαν να αξιοποιηθούν με επαρκή μεθοδολογική ασφάλεια.

Μετά τον έλεγχο ποιότητας υπολογίστηκαν τα βασικά περιγραφικά στατιστικά και δημιουργήθηκαν εμπειρικές συχνότητες των συγκεντρώσεων PM_{10} σε προκαθορισμένες κλάσεις. Οι παρατηρούμενες αυτές συχνότητες αποτέλεσαν τη βάση για τον υπολογισμό των θεωρητικά αναμενόμενων συχνοτήτων και τη σύγκριση των υποψήφιων μοντέλων.

Στις επιλεγμένες χρονοσειρές εφαρμόστηκαν τέσσερις παραμετρικές κατανομές πιθανότητας: η Normal, η Gamma, η Weibull και η Lognormal. Η Normal χρησιμοποιήθηκε ως συμμετρικό μοντέλο αναφοράς, ενώ οι Gamma, Weibull και Lognormal επιλέχθηκαν επειδή μπορούν να περιγράψουν μη αρνητικά δεδομένα με διαφορετικό βαθμό θετικής ασυμμετρίας και δεξιά ουρά.

Η ποιότητα προσαρμογής αξιολογήθηκε με συνδυασμό αριθμητικών και γραφικών κριτηρίων. Συγκεκριμένα, εξετάστηκαν οι παρατηρούμενες και οι αναμενόμενες συχνότητες ανά κλάση, τα observed-expected διαγράμματα, τα διαγράμματα υπολοίπων και τα Q-Q διαγράμματα. Παράλληλα, υπολογίστηκε το στατιστικό χ^2 , το οποίο χρησιμοποιήθηκε ως συγκριτικός δείκτης της συνολικής απόκλισης μεταξύ παρατηρούμενων και θεωρητικά αναμενόμενων συχνοτήτων.

Η τελική επιλογή της καταλληλότερης κατανομής δεν βασίστηκε αποκλειστικά στη χαμηλότερη τιμή χ^2 . Προέκυψε από τη συνδυαστική ερμηνεία της αριθμητικής κατάταξης και των γραφικών διαγνωστικών, ώστε να λαμβάνονται υπόψη τόσο η γενική συμφωνία

των συχνοτήτων όσο και η συμπεριφορά των μοντέλων σε επιμέρους τμήματα της εμπειρικής κατανομής. Η αξιολόγηση πραγματοποιήθηκε χωριστά για κάθε πόλη και έτος.

3.1 Πηγή και χαρακτηριστικά των δεδομένων

Τα δεδομένα της παρούσας εργασίας προέρχονται από το Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης, το οποίο λειτουργεί υπό την εποπτεία του Υπουργείου Περιβάλλοντος και Ενέργειας. Το δίκτυο παρέχει μετρήσεις ατμοσφαιρικών ρύπων από σταθερούς σταθμούς παρακολούθησης της ποιότητας του αέρα.

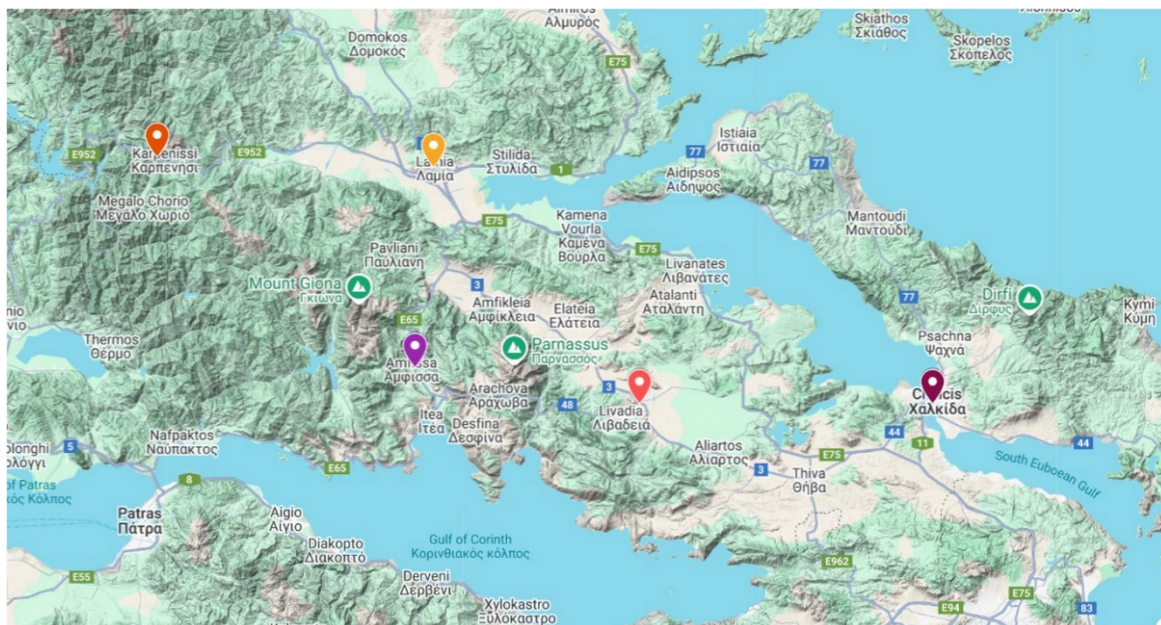
Η μελέτη επικεντρώνεται στις ωριαίες συγκεντρώσεις PM_{10} , δηλαδή των αιωρούμενων σωματιδίων με αεροδυναμική διάμετρο έως $10 \mu m$. Οι τιμές εκφράζονται σε $\mu g/m^3$. Η επιλογή των PM_{10} συνδέεται τόσο με τη σημασία τους ως βασικού δείκτη ποιότητας του αέρα όσο και με τη διαθεσιμότητα μετρήσεων στους σταθμούς της Περιφέρειας Στερεάς Ελλάδας.

Τα βασικά μεταδεδομένα των σταθμών που εξετάστηκαν, όπως η γεωγραφική θέση, το υψόμετρο, ο χαρακτηρισμός και οι καταγραφόμενοι ρύποι, παρουσιάζονται στον Πίνακα 3.1. Όλοι οι σταθμοί χαρακτηρίζονται ως σταθμοί αστικού υποβάθρου και διαθέτουν μετρήσεις αιωρούμενων σωματιδίων PM_{10} .

Περιφέρεια ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ													
Σταθμός							Μετρούμενοι ρύποι						
Όνομα	Θέση					Χαρακτηρισμός	SO ₂	NO _x	CO	O ₃	ΑΣ ₁₀	ΑΣ _{2,5}	C ₆ H ₆
	WGS84		ΕΓΣΑ87		Υψόμετρο (m –asl)								
	Γεωγρ. Μήκος	Γεωγρ. Πλάτος	X (m)	Y(m)									
ΛΑΜΙΑ (LAM)	22,431707	38,902170	363858	4306802	105	Αστικός-Υποβάθρου					v		
ΧΑΛΚΙΔΑ (HAL)	23,612019	38,469880	466005	4257735	12	Αστικός-Υποβάθρου					v		
ΛΕΙΒΑΔΙΑ (LEI)	22,877516	38,440615	401890	4255013	140	Αστικός-Υποβάθρου					v		
ΑΜΦΙΣΣΑ (AMF)	22,380409	38,528371	358676	4265397	180	Αστικός-Υποβάθρου					v		
ΚΑΡΠΕΝΗΣΙ (KAR)	21,791822	38,916588	308405	4309551	960	Αστικός-Υποβάθρου					v		

Πίνακας 3.1. Μεταδεδομένα σταθμών μέτρησης ατμοσφαιρικής ρύπανσης του ΕΔΠΑΡ που την ευθύνη λειτουργίας των οποίων έχει η Περιφέρεια Στερεάς Ελλάδας. Τα στοιχεία προέρχονται από τον αντίστοιχο πίνακα της ιστοσελίδας του ΥΠΕΝ <https://ypen.gov.gr/perivallon/poiotita-tis-atmosfairas/dedomena-metriseonatosfairikis-rypansis/>

Η χωρική θέση των σταθμών παρακολούθησης παρουσιάζεται στο Σχήμα 3.1. Οι σταθμοί βρίσκονται στη Λαμία, στη Χαλκίδα, στη Λειβαδιά, στην Άμφισσα και στο Καρπενήσι.



Σχήμα 3.1. Χωρική θέση των σταθμών παρακολούθησης PM_{10} στην Περιφέρεια Στερεάς Ελλάδας.

Πηγή: Ιδία επεξεργασία με βάση τις γεωγραφικές συντεταγμένες του ΥΠΕΝ.

Η χωρική κάλυψη περιλαμβάνει τους σταθμούς της Λαμίας, της Χαλκίδας, της Λειβαδιάς, της Άμφισσας και του Καρπενησίου. Τα δεδομένα κάθε σταθμού οργανώθηκαν ανά ημερολογιακό έτος, ώστε κάθε ετήσια χρονοσειρά να αποτελεί αυτοτελή μονάδα ελέγχου και ανάλυσης. Με τον τρόπο αυτό αποφεύχθηκε η συγχώνευση δεδομένων με διαφορετική πληρότητα ή χρονική κάλυψη.

Το αρχικό διαθέσιμο υλικό κάλυπτε την περίοδο 2016-2024. Ωστόσο, τα δεδομένα του 2016 παρέχονταν σε ημερήσια και όχι σε ωριαία χρονική ανάλυση και, επομένως, δεν ήταν άμεσα συγκρίσιμα με τις υπόλοιπες χρονοσειρές. Για τον λόγο αυτό, η τελική διερεύνηση περιορίστηκε στα ωριαία δεδομένα της περιόδου 2017-2024.

Για κάθε πόλη και έτος υπολογίστηκαν ο αριθμός των έγκυρων ωριαίων παρατηρήσεων και το ποσοστό πληρότητας ως προς τον θεωρητικά αναμενόμενο αριθμό ωρών του έτους. Παράλληλα, εξετάστηκε η χρονική συνέχεια των διαθέσιμων μετρήσεων, ώστε να εντοπιστούν πιθανές εκτεταμένες περιόδους απουσίας δεδομένων. Ως εσωτερικό κριτήριο

ένταξης στην τελική ανάλυση χρησιμοποιήθηκε πληρότητα τουλάχιστον 90%, σε συνδυασμό με ικανοποιητική χρονική κάλυψη.

Με βάση τον έλεγχο αυτό, στην τελική κατανομική ανάλυση εντάχθηκαν η Άμφισσα για το 2018, η Λαμία για το 2017, η Λειβαδιά για το 2019 και η Χαλκίδα για το 2017. Οι συγκεκριμένες χρονοσειρές πληρούσαν τα κριτήρια που είχαν τεθεί και κρίθηκαν κατάλληλες για περιγραφική και κατανομική ανάλυση.

Αντίθετα, το Καρπενήσι δεν εντάχθηκε στην προσαρμογή των θεωρητικών κατανομών, καθώς κανένα διαθέσιμο έτος δεν πληρούσε το απαιτούμενο επίπεδο πληρότητας. Η μη συμπερίληψή του αποτέλεσε εφαρμογή των ίδιων κριτηρίων ποιότητας που χρησιμοποιήθηκαν για όλους τους σταθμούς και όχι αυθαίρετο αποκλεισμό.

Το τελικό αναλυτικό σώμα περιλάμβανε, επομένως, τέσσερις ετήσιες χρονοσειρές ωριαίων συγκεντρώσεων PM_{10} . Οι χρονοσειρές αυτές χρησιμοποιήθηκαν για την περιγραφική στατιστική, τη δημιουργία των εμπειρικών συχνοτήτων και την προσαρμογή των θεωρητικών κατανομών.

Πόλη / Σταθμός	Ρύπος	Μονάδα μέτρησης	Χρονική κλίμακα	Τελικό αναλύσιμο έτος
Λαμία	PM_{10}	$\mu g/m^3$	Ωριαία	2017
Χαλκίδα	PM_{10}	$\mu g/m^3$	Ωριαία	2017
Λειβαδιά	PM_{10}	$\mu g/m^3$	Ωριαία	2019
Άμφισσα	PM_{10}	$\mu g/m^3$	Ωριαία	2018
Καρπενήσι	PM_{10}	$\mu g/m^3$	Ωριαία	Δεν εντάχθηκε

Πίνακας 3.2. Τελικό σώμα δεδομένων που χρησιμοποιήθηκε στην ανάλυση.

Σημείωση: Η επιλογή των ετών βασίστηκε σε κριτήριο πληρότητας τουλάχιστον 90% και στον έλεγχο της χρονικής συνέχειας των ωριαίων μετρήσεων. Τα έτη με ανεπαρκή κάλυψη δεν εντάχθηκαν στην προσαρμογή των θεωρητικών κατανομών.

3.2 Προεπεξεργασία και καθαρισμός δεδομένων PM_{10}

Πριν από τη στατιστική ανάλυση πραγματοποιήθηκε προεπεξεργασία και καθαρισμός των αρχικών δεδομένων PM_{10} , με σκοπό τη δημιουργία ομοιογενών και κατάλληλων χρονοσειρών για την εφαρμογή των θεωρητικών κατανομών. Η διαδικασία αυτή ήταν αναγκαία, καθώς τα δεδομένα περιβαλλοντικής παρακολούθησης μπορεί να περιλαμβάνουν ελλείπουσες τιμές, ασυνέχειες και διαφοροποιήσεις στη χρονική ανάλυση. Αρχικά, τα αρχεία οργανώθηκαν ανά πόλη και έτος. Επειδή οι καταγραφές του 2016 παρέχονταν σε ημερήσια χρονική ανάλυση, ενώ για την περίοδο 2017-2024 ήταν διαθέσιμες ωριαίες μετρήσεις, τα δεδομένα του 2016 δεν εντάχθηκαν στην τελική διερεύνηση. Η ανάλυση περιορίστηκε στις ωριαίες χρονοσειρές, ώστε να διατηρηθεί κοινή χρονική κλίμακα σε όλες τις εξεταζόμενες περιπτώσεις. Στη συνέχεια πραγματοποιήθηκε έλεγχος για ελλείπουσες και μη έγκυρες καταγραφές. Στα πρωτογενή αρχεία οι μη διαθέσιμες τιμές σημειώνονταν με τον ειδικό αριθμητικό κωδικό -9999¹, ο οποίος αντικαταστάθηκε με NA, ώστε να μην αντιμετωπίζεται ως πραγματική συγκέντρωση PM_{10} . Δεν εφαρμόστηκε συμπλήρωση ή παρεμβολή των ελλειπουσών τιμών, καθώς στόχος της εργασίας ήταν η προσαρμογή των κατανομών στις πραγματικά διαθέσιμες παρατηρήσεις. Για τις κατανομές Lognormal, Gamma και Weibull χρησιμοποιήθηκαν μόνο οι έγκυρες θετικές συγκεντρώσεις, λόγω του πεδίου ορισμού τους. Ειδικά για τη Lognormal, υπολογίστηκε ο φυσικός λογάριθμος των θετικών τιμών PM_{10} , ο οποίος χρησιμοποιήθηκε για την εκτίμηση των αντίστοιχων παραμέτρων.

Τα ωριαία δεδομένα μετασχηματίστηκαν από οριζόντια μορφή, στην οποία οι ώρες εμφανίζονταν σε διαφορετικές στήλες, σε κάθετη μορφή, όπου κάθε γραμμή αντιστοιχούσε σε μία ωριαία παρατήρηση. Δημιουργήθηκαν διακριτά πεδία για την ημερομηνία, την ώρα και τη συγκέντρωση PM_{10} . Η μετατροπή αυτή διευκόλυνε τον έλεγχο της χρονικής σειράς, τον υπολογισμό της πληρότητας και την κατασκευή των εμπειρικών συχνοτήτων. Η προεπεξεργασία και η αναδιοργάνωση των δεδομένων πραγματοποιήθηκαν με τη γλώσσα προγραμματισμού R και εργαλεία του οικοσυστήματος tidyverse. Μετά την ολοκλήρωση του καθαρισμού, κάθε ετήσια χρονοσειρά εξετάστηκε ως προς την πληρότητα και τη χρονική της συνέχεια, σύμφωνα με τα κριτήρια που παρουσιάζονται στην επόμενη ενότητα.

¹Σύμφωνα με την τεκμηρίωση του ΕΔΠΑΡ για την ονοματολογία και τη δομή των αρχείων, ο κωδικός -9999 χρησιμοποιείται για τη σήμανση ελλειπουσών τιμών συγκέντρωσης και, συνεπώς, δεν αντιστοιχεί σε πραγματική μέτρηση ατμοσφαιρικού ρύπου.

3.3 Κριτήρια επιλογής αναλύσιμων χρονοσειρών

Μετά την προεπεξεργασία πραγματοποιήθηκε έλεγχος πληρότητας και χρονικής συνέχειας των ωριαίων χρονοσειρών PM_{10} ανά πόλη και έτος. Ο έλεγχος αυτός προηγήθηκε της κατανομικής ανάλυσης, επειδή εκτεταμένες ελλείψεις ή μεγάλα χρονικά κενά μπορούν να μεταβάλουν την εμπειρική μορφή των συγκεντρώσεων και να επηρεάσουν την αξιοπιστία της προσαρμογής.

Η πληρότητα κάθε χρονοσειράς υπολογίστηκε ως το ποσοστό των έγκυρων ωριαίων παρατηρήσεων προς τον συνολικό αριθμό ωρών του αντίστοιχου έτους. Ως εσωτερικό κριτήριο ένταξης στην τελική ανάλυση ορίστηκε πληρότητα τουλάχιστον 90%. Παράλληλα, εξετάστηκε η χρονική κατανομή των διαθέσιμων μετρήσεων, ώστε να διαπιστωθεί αν υπήρχαν εκτεταμένα συνεχόμενα κενά που θα μπορούσαν να οδηγήσουν σε υποεκπροσώπηση συγκεκριμένων περιόδων του έτους.

Η επιλογή δεν βασίστηκε στο ύψος των συγκεντρώσεων ή στην περιβαλλοντική σημασία κάθε σταθμού, αλλά αποκλειστικά στην καταλληλότητα των δεδομένων για στατιστική επεξεργασία. Με αυτόν τον τρόπο εφαρμόστηκαν κοινά κριτήρια σε όλες τις πόλεις και περιορίστηκε ο κίνδυνος εξαγωγής συμπερασμάτων από αποσπασματικές χρονοσειρές.

Με βάση τον έλεγχο, στην τελική κατανομική ανάλυση εντάχθηκαν η Άμφισσα για το 2018, η Λαμία για το 2017, η Λειβαδιά για το 2019 και η Χαλκίδα για το 2017. Οι συγκεκριμένες χρονοσειρές παρουσίασαν πληρότητα μεγαλύτερη από 90% και κρίθηκαν επαρκείς για την περιγραφική ανάλυση, την κατασκευή εμπειρικών συχνοτήτων και την εφαρμογή θεωρητικών κατανομών.

Το Καρπενήσι δεν εντάχθηκε στην τελική προσαρμογή, καθώς κανένα από τα διαθέσιμα έτη δεν πληρούσε το κριτήριο πληρότητας. Η περίπτωση αυτή παρουσιάστηκε χωριστά στην ανάλυση, ώστε να τεκμηριωθεί η εφαρμογή των ίδιων μεθοδολογικών προϋποθέσεων σε όλους τους σταθμούς.

Η διαδικασία επιλογής διασφάλισε ότι οι θεωρητικές κατανομές εφαρμόστηκαν μόνο σε ετήσιες χρονοσειρές με επαρκή αριθμό παρατηρήσεων και ικανοποιητική χρονική κάλυψη. Τα επιλεγμένα έτη αποτελούν τις πληρέστερες διαθέσιμες περιπτώσεις για κάθε πόλη, χωρίς να θεωρούνται κατ' ανάγκη αντιπροσωπευτικά της μακροχρόνιας συμπεριφοράς των συγκεντρώσεων PM_{10} .

3.4 Δημιουργία εμπειρικών συχνοτήτων

Για την εφαρμογή και την αξιολόγηση των θεωρητικών κατανομών, οι έγκυρες ωριαίες συγκεντρώσεις PM_{10} κάθε επιλεγμένης χρονοσειράς ομαδοποιήθηκαν σε κλάσεις συγκέντρωσης. Η διαδικασία αυτή επέτρεψε τη μετατροπή των αρχικών παρατηρήσεων σε εμπειρική κατανομή συχνοτήτων και τη σύγκρισή της με τις αντίστοιχες θεωρητικά αναμενόμενες συχνότητες.

Οι συγκεντρώσεις, εκφρασμένες σε $\mu g/m^3$, ταξινομήθηκαν σε διαδοχικές κλάσεις με προκαθορισμένα όρια. Για κάθε χρονοσειρά χρησιμοποιήθηκαν τα ίδια όρια κλάσεων σε όλες τις υπό εξέταση κατανομές, ώστε η σύγκριση μεταξύ των μοντέλων να πραγματοποιείται υπό κοινές συνθήκες. Η επιλογή του εύρους των κλάσεων έγινε με τρόπο που να αποτυπώνει ικανοποιητικά την κύρια μορφή της εμπειρικής κατανομής, χωρίς να δημιουργούνται υπερβολικά πολλές κατηγορίες με πολύ μικρές ή μηδενικές συχνότητες.

Η ομαδοποίηση των δεδομένων αποτελεί κρίσιμο στάδιο, καθώς επηρεάζει τον τρόπο με τον οποίο αποδίδεται η κατανομή των συγκεντρώσεων. Πολύ μεγάλο εύρος κλάσης μπορεί να αποκρύψει σημαντικές διαφοροποιήσεις, ενώ πολύ μικρό εύρος μπορεί να οδηγήσει σε ασταθή εικόνα με πολλές αραιές κατηγορίες. Για τον λόγο αυτό, επιδιώχθηκε ισορροπία μεταξύ επαρκούς λεπτομέρειας και στατιστικής σταθερότητας.

Για κάθε κλάση i , με κατώτερο όριο a_i και ανώτερο όριο b_i , η παρατηρούμενη συχνότητα υπολογίστηκε ως το πλήθος των έγκυρων ωριαίων τιμών που ανήκαν στο αντίστοιχο διάστημα:

$$O_i = \#\{x_j : a_i < x_j \leq b_i\}$$

όπου:

- O_i είναι η παρατηρούμενη συχνότητα
- x_j η j -στή έγκυρη συγκέντρωση PM_{10} .

Με τον τρόπο αυτό, το αρχικό σύνολο των ωριαίων παρατηρήσεων μετατράπηκε σε εμπειρική κατανομή συχνοτήτων. Η κατανομή αυτή αποτυπώνει πόσο συχνά εμφανίζονται χαμηλές, μέσες και υψηλές συγκεντρώσεις και επιτρέπει την άμεση σύγκριση της πραγματικής συμπεριφοράς των δεδομένων με εκείνη που προβλέπει κάθε θεωρητικό μοντέλο.

Οι θεωρητικά αναμενόμενες συχνότητες υπολογίστηκαν για τις ίδιες ακριβώς κλάσεις. Για κάθε κατανομή, η πιθανότητα εμφάνισης μιας τιμής στο διάστημα $(a_i, b_i]$ προέκυψε από τη διαφορά της αθροιστικής συνάρτησης κατανομής στα όρια της κλάσης:

$$p_i = P(a_i < X \leq b_i) = F(b_i) - F(a_i)$$

όπου:

- p_i είναι η θεωρητική πιθανότητα της κλάσης,
- X η τυχαία μεταβλητή που εκφράζει τη συγκέντρωση PM_{10}
- F η αθροιστική συνάρτηση της εκάστοτε θεωρητικής κατανομής.

Η αναμενόμενη συχνότητα υπολογίστηκε στη συνέχεια ως:

$$E_i = N \cdot p_i = N \cdot [F(b_i) - F(a_i)]$$

όπου:

- E_i είναι η αναμενόμενη συχνότητα της i -στής κλάσης,
- N το συνολικό πλήθος των έγκυρων ωριαίων παρατηρήσεων.

Η χρήση της αθροιστικής συνάρτησης ήταν απαραίτητη, επειδή η πιθανότητα μιας ολόκληρης κλάσης δεν προκύπτει από μία μεμονωμένη τιμή της συνάρτησης πυκνότητας, αλλά από τη διαφορά των αθροιστικών πιθανοτήτων στα δύο όριά της. Με τον τρόπο αυτό, η θεωρητική πιθανότητα κάθε κλάσης μετατράπηκε σε αναμενόμενο αριθμό παρατηρήσεων άμεσα συγκρίσιμο με την αντίστοιχη εμπειρική συχνότητα.

Οι παρατηρούμενες και οι αναμενόμενες συχνότητες χρησιμοποιήθηκαν τόσο στη γραφική σύγκριση των μοντέλων όσο και στον υπολογισμό του στατιστικού χ^2 . Ιδιαίτερη προσοχή δόθηκε στις κλάσεις με πολύ χαμηλές αναμενόμενες συχνότητες, επειδή μπορούν να επηρεάσουν δυσανάλογα το αποτέλεσμα. Όπου ήταν αναγκαίο, η αντιμετώπισή τους έγινε με ενιαίο τρόπο για όλες τις κατανομές της ίδιας χρονοσειράς, ώστε να διατηρείται η συγκρισιμότητα της αξιολόγησης.

Η ανώτερη κλάση διατηρήθηκε ανοικτή, ώστε να περιλαμβάνει όλες τις τιμές που υπερέβαιναν το αντίστοιχο όριο συγκέντρωσης. Η επιλογή αυτή εξασφάλισε ότι καμία έγκυρη παρατήρηση δεν αποκλείστηκε από την ανάλυση, αν και περιορίζει τη λεπτομερή διάκριση μεταξύ των υψηλότερων τιμών της δεξιάς ουράς. Για τον λόγο αυτό, η αξιολόγηση των ακραίων παρατηρήσεων συμπληρώθηκε από τα Q-Q διαγράμματα.

Η δημιουργία των εμπειρικών συχνοτήτων αποτέλεσε το ενδιάμεσο στάδιο μεταξύ των αρχικών ωριαίων μετρήσεων και της τελικής αξιολόγησης των θεωρητικών κατανομών.

Μέσω αυτής της διαδικασίας διαμορφώθηκε ένα κοινό πλαίσιο σύγκρισης, το οποίο χρησιμοποιήθηκε στις επόμενες ενότητες για την αριθμητική και γραφική αποτίμηση της προσαρμογής.

3.5 Εφαρμογή θεωρητικών κατανομών

Η εφαρμογή των θεωρητικών κατανομών αποτέλεσε το βασικό στάδιο της στατιστικής μοντελοποίησης των ωριαίων συγκεντρώσεων PM_{10} . Οι εμπειρικές κατανομές των επιλεγμένων χρονοσειρών συγκρίθηκαν με παραμετρικά μοντέλα, με σκοπό να αξιολογηθεί η ικανότητά τους να περιγράφουν τη γενική μορφή, τη διασπορά, τη θετική ασυμμετρία και την άνω ουρά των δεδομένων.

Στην ανάλυση εξετάστηκαν οι κατανομές Lognormal, Gamma και Weibull, οι οποίες είναι κατάλληλες για θετικές ή μη αρνητικές συνεχείς μεταβλητές, ενώ η Normal χρησιμοποιήθηκε ως συμμετρικό μοντέλο αναφοράς. Η εφαρμογή πραγματοποιήθηκε χωριστά για κάθε πόλη και έτος, ώστε να διατηρηθούν τα ιδιαίτερα χαρακτηριστικά κάθε χρονοσειράς και να αποφευχθεί η συνένωση δεδομένων με διαφορετικά επίπεδα συγκέντρωσης και μεταβλητότητας.

Για κάθε κατανομή εκτιμήθηκαν οι αντίστοιχες παράμετροι από τις έγκυρες παρατηρήσεις. Οι εκτιμημένες παράμετροι χρησιμοποιήθηκαν για τον υπολογισμό των θεωρητικών πιθανοτήτων στις κλάσεις συγκέντρωσης και, στη συνέχεια, των αναμενόμενων συχνοτήτων. Με τον τρόπο αυτό, κάθε θεωρητικό μοντέλο εκφράστηκε στην ίδια κλίμακα με την εμπειρική κατανομή και κατέστη δυνατή η άμεση σύγκριση των αποτελεσμάτων.

Η ποιότητα προσαρμογής αξιολογήθηκε με κοινή διαδικασία για όλα τα μοντέλα. Εξετάστηκαν οι παρατηρούμενες και αναμενόμενες συχνότητες, τα observed-expected διαγράμματα, τα διαγράμματα υπολοίπων, τα Q-Q διαγράμματα και το στατιστικό χ^2 . Η χρήση πολλαπλών κριτηρίων κρίθηκε αναγκαία, επειδή κάθε διαγνωστικό εργαλείο αποτυπώνει διαφορετική πλευρά της προσαρμογής και ένα μοντέλο μπορεί να αποδίδει ικανοποιητικά την κύρια μάζα των δεδομένων, αλλά λιγότερο καλά τις υψηλότερες τιμές.

3.5.1 Γενική λογική εφαρμογής των θεωρητικών κατανομών

Μετά τον καθαρισμό και την επιλογή των αναλύσιμων χρονοσειρών, οι θεωρητικές κατανομές εφαρμόστηκαν στις έγκυρες ωριαίες συγκεντρώσεις PM_{10} . Η ανάλυση πραγματοποιήθηκε αυτοτελώς για κάθε πόλη και έτος, καθώς οι χρονοσειρές παρουσίαζαν διαφορετικά χαρακτηριστικά ως προς την κεντρική τάση, τη διασπορά, την ασυμμετρία και την παρουσία υψηλών παρατηρήσεων.

Για κάθε μοντέλο εκτιμήθηκαν οι παράμετροι από τα δεδομένα της αντίστοιχης χρονοσειράς. Στη συνέχεια, με χρήση της αθροιστικής συνάρτησης κατανομής, υπολογίστηκε η θεωρητική πιθανότητα κάθε κλάσης συγκέντρωσης. Οι πιθανότητες μετατράπηκαν σε αναμενόμενες συχνότητες μέσω πολλαπλασιασμού με το συνολικό πλήθος των έγκυρων ωριαίων παρατηρήσεων.

Οι αναμενόμενες συχνότητες υπολογίστηκαν για τα ίδια ακριβώς όρια κλάσεων που χρησιμοποιήθηκαν στην εμπειρική κατανομή. Η επιλογή αυτή διασφάλισε ότι οι αποκλίσεις μεταξύ παρατηρούμενων και θεωρητικών τιμών δεν οφείλονταν σε διαφορετική ομαδοποίηση των δεδομένων.

Η αξιολόγηση της προσαρμογής περιέλαβε τόσο τη συνολική αριθμητική απόκλιση όσο και τη θέση και τη φορά των επιμέρους σφαλμάτων. Το χ^2 χρησιμοποιήθηκε ως συγκριτικός δείκτης κατάταξης των μοντέλων μέσα στην ίδια χρονοσειρά. Παράλληλα, τα γραφικά διαγνωστικά επέτρεψαν να εξεταστεί αν οι αποκλίσεις εμφανίζονταν συστηματικά στις χαμηλές, μεσαίες ή υψηλές συγκεντρώσεις.

Η έννοια της «καταλληλότερης κατανομής» χρησιμοποιείται με σχετικό χαρακτήρα. Αναφέρεται στο μοντέλο που παρουσίασε καλύτερη αριθμητική και γραφική συμφωνία μεταξύ των τεσσάρων κατανομών που εξετάστηκαν και όχι σε απόλυτη ή καθολική περιγραφή όλων των χαρακτηριστικών της χρονοσειράς.

3.5.2 Λογαριθμοκανονική κατανομή

Η λογαριθμοκανονική κατανομή χρησιμοποιήθηκε ως ένα από τα κύρια θεωρητικά μοντέλα προσαρμογής των ωριαίων συγκεντρώσεων PM_{10} . Η επιλογή της κρίθηκε μεθοδολογικά εύλογη, καθώς εφαρμόζεται σε θετικές συνεχείς μεταβλητές και μπορεί να αποδώσει δεδομένα με δεξιά ασυμμετρία, όπου παρατηρείται μεγάλος αριθμός χαμηλών ή

μέτριων τιμών και μικρότερος αριθμός υψηλών συγκεντρώσεων. Σε τέτοιες περιπτώσεις, η λογαριθμοκανονική κατανομή αποτελεί κατάλληλο πιθανοτικό μοντέλο, εφόσον ο φυσικός λογάριθμος των αρχικών τιμών προσεγγίζει την κανονική κατανομή.

Για την εφαρμογή της χρησιμοποιήθηκαν μόνο οι έγκυρες θετικές τιμές PM_{10} κάθε επιλεγμένης χρονοσειράς. Οι μη διαθέσιμες, μηδενικές και μη έγκυρες τιμές δεν χρησιμοποιήθηκαν στον υπολογισμό των παραμέτρων, επειδή ο φυσικός λογάριθμος δεν ορίζεται για μηδενικές ή αρνητικές τιμές. Για κάθε θετική συγκέντρωση υπολογίστηκε η μεταβλητή:

$$y_j = \ln(x_j)$$

όπου:

- x_j είναι η j-στή θετική παρατήρηση PM_{10} .

Οι παράμετροι της λογαριθμοκανονικής κατανομής εκτιμήθηκαν από τις λογαριθμισμένες τιμές. Η παράμετρος μ αντιστοιχεί στον μέσο όρο των τιμών $\ln(PM_{10})$, ενώ η παράμετρος σ αντιστοιχεί στην τυπική απόκλιση των λογαριθμισμένων τιμών. Η διάκριση αυτή είναι ουσιώδης, καθώς η προσαρμογή δεν πραγματοποιείται αντιμετωπίζοντας τις αρχικές συγκεντρώσεις ως κανονικά κατανεμημένες, αλλά μέσω του λογαριθμικού τους μετασχηματισμού.

Με βάση τις εκτιμημένες παραμέτρους υπολογίστηκε, για κάθε κλάση συγκέντρωσης, η θεωρητική πιθανότητα εμφάνισης τιμών εντός των ορίων της. Η πιθανότητα αυτή προέκυψε από την αθροιστική συνάρτηση της λογαριθμοκανονικής κατανομής και μετατράπηκε σε αναμενόμενη συχνότητα με πολλαπλασιασμό επί το συνολικό πλήθος των έγκυρων παρατηρήσεων.

Η ποιότητα προσαρμογής αξιολογήθηκε με βάση τη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων, τα observed-expected διαγράμματα, τα διαγράμματα υπολοίπων, τα Q-Q διαγράμματα και το στατιστικό χ^2 . Ιδιαίτερη προσοχή δόθηκε στην άνω ουρά, όπου οι αποκλίσεις μεταξύ εμπειρικών και θεωρητικών ποσοστημορίων μπορεί να είναι εντονότερες.

3.5.3 Κατανομή Gamma

Η κατανομή Gamma χρησιμοποιήθηκε ως ένα από τα βασικά θεωρητικά μοντέλα για την προσαρμογή των ωριαίων συγκεντρώσεων PM_{10} . Η επιλογή της κρίθηκε κατάλληλη, επειδή ορίζεται για θετικές συνεχείς μεταβλητές και μπορεί να αποδώσει κατανομές με διαφορετικό βαθμό δεξιάς ασυμμετρίας.

Σε αντίθεση με τη λογαριθμοκανονική, η Gamma εφαρμόζεται απευθείας στις αρχικές θετικές συγκεντρώσεις, χωρίς να απαιτείται λογαριθμικός μετασχηματισμός. Η μορφή της καθορίζεται από την παράμετρο σχήματος α και την παράμετρο κλίμακας β .

Στην παρούσα εφαρμογή, οι παράμετροι εκτιμήθηκαν από τα εμπειρικά χαρακτηριστικά κάθε χρονοσειράς.

Οι παράμετροι εκτιμήθηκαν με τη μέθοδο των ροπών και δίνονται από τις σχέσεις:

$$\alpha = \frac{\bar{X}^2}{S^2}$$

Ενώ η παράμετρος κλίμακας ως:

$$\beta = \frac{S^2}{\bar{X}}$$

όπου:

- $\bar{X}$ είναι ο μέσος όρος,
- S^2 η διακύμανση των συγκεντρώσεων PM_{10} .

Η παράμετρος α επηρεάζει κυρίως τη μορφή και τον βαθμό ασυμμετρίας της κατανομής, ενώ η β καθορίζει την κλίμακα και τη διασπορά των τιμών.

Για κάθε κλάση συγκέντρωσης υπολογίστηκε η θεωρητική πιθανότητα με βάση την αθροιστική συνάρτηση της Gamma και στη συνέχεια η αντίστοιχη αναμενόμενη συχνότητα. Η κατανομή εκφράστηκε έτσι στην ίδια κλίμακα με τις εμπειρικές συχνότητες και κατέστη δυνατή η άμεση σύγκριση μεταξύ παρατηρούμενων και θεωρητικών τιμών.

Η αξιολόγηση της προσαρμογής πραγματοποιήθηκε με βάση τα ίδια αριθμητικά και γραφικά κριτήρια που χρησιμοποιήθηκαν και στις υπόλοιπες κατανομές. Το χ^2 χρησιμοποιήθηκε κυρίως ως συγκριτικός δείκτης κατάταξης των μοντέλων μέσα στην ίδια χρονοσειρά και όχι ως μοναδικό κριτήριο αποδοχής ή απόρριψης της Gamma.

3.5.4 Κατανομή Weibull

Η κατανομή Weibull χρησιμοποιήθηκε ως τρίτο κύριο θεωρητικό μοντέλο προσαρμογής των ωριαίων συγκεντρώσεων PM_{10} . Η επιλογή της βασίστηκε στην ευελιξία της ως προς τη μορφή της κατανομής και στην ικανότητά της να περιγράφει θετικές συνεχείς μεταβλητές με διαφορετικό βαθμό ασυμμετρίας και μεταβλητότητας.

Η κατανομή Weibull χαρακτηρίζεται από την παράμετρο σχήματος k και την παράμετρο κλίμακας λ . Η παράμετρος k επηρεάζει τη μορφή και τον βαθμό ασυμμετρίας της κατανομής, ενώ η παράμετρος λ συνδέεται με το εύρος και το επίπεδο των τιμών. Η αθροιστική συνάρτηση κατανομής της Weibull δίνεται από τη σχέση:

$$F(x) = 1 - \exp\left[-\left(\frac{x}{\lambda}\right)^k\right], \quad x \geq 0$$

Οι παράμετροι k και λ εκτιμήθηκαν χωριστά για κάθε επιλεγμένη χρονοσειρά, με βάση τις έγκυρες θετικές συγκεντρώσεις PM_{10} .

Για κάθε κλάση συγκέντρωσης υπολογίστηκε η θεωρητική πιθανότητα μέσω της αθροιστικής συνάρτησης κατανομής και, στη συνέχεια, η αντίστοιχη αναμενόμενη συχνότητα, μέσω του πολλαπλασιασμού της πιθανότητας με το συνολικό πλήθος των παρατηρήσεων. Η σύγκριση πραγματοποιήθηκε με το ίδιο σύστημα κλάσεων που χρησιμοποιήθηκε για τις εμπειρικές συχνότητες.

Η προσαρμογή της κατανομής Weibull αξιολογήθηκε μέσω της σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων, των διαγνωστικών διαγραμμάτων παρατηρούμενων έναντι αναμενόμενων τιμών, των διαγραμμάτων υπολοίπων, των διαγραμμάτων Q-Q όπου εφαρμόστηκαν και του συγκριτικού στατιστικού χ^2 .

Με τον τρόπο αυτό, η Weibull εξετάστηκε ως ευέλικτο εναλλακτικό μοντέλο, χωρίς να θεωρείται εκ των προτέρων καταλληλότερη από τη λογαριθμοκανονική κατανομή ή την κατανομή Gamma.

3.5.5 Κανονική κατανομή ως συγκριτική βάση

Η κανονική κατανομή εξετάστηκε ως συμμετρικό μοντέλο αναφοράς και όχι ως εκ των προτέρων καταλληλότερη επιλογή για τις συγκεντρώσεις PM_{10} . Η χρήση της κρίθηκε χρήσιμη, επειδή παρέχει ένα απλό και ευρέως χρησιμοποιούμενο πλαίσιο σύγκρισης με τις θετικά ορισμένες και ασύμμετρες κατανομές.

Η Normal χαρακτηρίζεται από τον μέσο όρο μ και την τυπική απόκλιση σ , οι οποίοι εκτιμήθηκαν από τις έγκυρες ωριαίες τιμές κάθε χρονοσειράς. Σε αντίθεση με τη Lognormal, τη Gamma και τη Weibull, είναι συμμετρική και ορίζεται σε ολόκληρο το πραγματικό πεδίο τιμών. Κατά συνέπεια, αποδίδει θεωρητική πιθανότητα και σε αρνητικές τιμές, οι οποίες δεν έχουν φυσική σημασία για συγκεντρώσεις PM_{10} .

Για κάθε κλάση συγκέντρωσης υπολογίστηκε η θεωρητική πιθανότητα με βάση την αθροιστική συνάρτηση της κανονικής κατανομής και στη συνέχεια η αντίστοιχη αναμενόμενη συχνότητα. Η αξιολόγηση πραγματοποιήθηκε με την ίδια μεθοδολογική διαδικασία που εφαρμόστηκε και στις υπόλοιπες κατανομές.

Η Normal δεν χρησιμοποιήθηκε για να υποστηριχθεί ότι οι συγκεντρώσεις PM_{10} ακολουθούν συμμετρική συμπεριφορά. Ο ρόλος της ήταν να λειτουργήσει ως συγκριτική βάση και να αναδείξει κατά πόσο οι Lognormal, Gamma και Weibull, οι οποίες είναι περισσότερο συμβατές με τη μη αρνητικότητα και τη θετική ασυμμετρία των δεδομένων, παρέχουν καλύτερη περιγραφή της εμπειρικής κατανομής.

3.6 Κριτήριο και αξιολόγηση προσαρμογής

Για την αριθμητική αξιολόγηση της προσαρμογής των θεωρητικών κατανομών στις εμπειρικές συγκεντρώσεις PM_{10} χρησιμοποιήθηκε το στατιστικό χ^2 . Η αξιολόγηση βασίστηκε στη σύγκριση των παρατηρούμενων συχνοτήτων με τις θεωρητικά αναμενόμενες συχνότητες που προέκυψαν από κάθε κατανομή. Η σύγκριση πραγματοποιήθηκε σε επίπεδο κλάσεων συγκέντρωσης και όχι σε επίπεδο μεμονωμένων ωριαίων παρατηρήσεων.

Για κάθε κλάση i , υπολογίστηκε η παρατηρούμενη συχνότητα O_i , δηλαδή ο αριθμός των πραγματικών ωριαίων τιμών PM_{10} που ανήκαν στην αντίστοιχη κλάση, καθώς και η

αναμενόμενη συχνότητα E_i , δηλαδή ο αριθμός των τιμών που θα αναμένονταν στην ίδια κλάση σύμφωνα με την εκάστοτε θεωρητική κατανομή. Η επιμέρους συνεισφορά κάθε κλάσης στο χ^2 υπολογίστηκε ως:

$$\chi_i^2 = \frac{(O_i - E_i)^2}{E_i}$$

όπου:

- O_i είναι η παρατηρούμενη συχνότητα της i -στής κλάσης,
- E_i η αναμενόμενη συχνότητα της ίδιας κλάσης σύμφωνα με τη θεωρητική κατανομή
- χ_i^2 η επιμέρους συνεισφορά της κλάσης στη συνολική τιμή του δείκτη.

Η συνολική τιμή του χ^2 για κάθε κατανομή προέκυψε από το άθροισμα των επιμέρους αποκλίσεων μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

όπου k είναι ο αριθμός των κλάσεων που χρησιμοποιήθηκαν στην αξιολόγηση.

Μικρότερες τιμές του χ^2 υποδηλώνουν μικρότερη συνολική απόκλιση μεταξύ εμπειρικής και θεωρητικής κατανομής, ενώ μεγαλύτερες τιμές υποδηλώνουν εντονότερη απόκλιση.

Στην παρούσα εργασία, το στατιστικό χ^2 χρησιμοποιήθηκε κυρίως ως συγκριτικός διαγνωστικός δείκτης μεταξύ των εξεταζόμενων θεωρητικών κατανομών και όχι ως αυτοτελές κριτήριο αποδοχής ή απόρριψης ενός μοντέλου. Η επιλογή αυτή κρίθηκε κατάλληλη, επειδή στις υψηλότερες κλάσεις συγκέντρωσης, δηλαδή στην άνω ουρά της κατανομής, οι αναμενόμενες συχνότητες μπορεί να είναι χαμηλές, γεγονός που μπορεί να επηρεάσει δυσανάλογα τη συνολική τιμή του δείκτη. Για τον λόγο αυτό, οι τιμές του χ^2 ερμηνεύτηκαν σε συνδυασμό με τα γραφικά διαγνωστικά και με βάση τη σχετική κατάταξη των μοντέλων.

Για τον λόγο αυτό, η αξιολόγηση της προσαρμογής δεν βασίστηκε αποκλειστικά στο χ^2 , αλλά συνδυάστηκε με οπτική και διαγνωστική διερεύνηση. Συγκεκριμένα, για κάθε πόλη, έτος και θεωρητική κατανομή εξετάστηκαν τα γραφήματα σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων, τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων τιμών και τα διαγράμματα υπολοίπων.

Τα υπόλοιπα υπολογίστηκαν ως:

$$R_i = O_i - E_i$$

όπου:

- R_i είναι το υπόλοιπο της i -στής κλάσης,
- O_i η παρατηρούμενη συχνότητα και
- E_i η αναμενόμενη συχνότητα.

Θετικές τιμές υποδηλώνουν ότι η παρατηρούμενη συχνότητα είναι μεγαλύτερη από την αναμενόμενη, ενώ αρνητικές τιμές ότι το θεωρητικό μοντέλο υπερεκτιμά τη συχνότητα της συγκεκριμένης κλάσης.

Η μικρότερη τιμή χ^2 δεν ερμηνεύθηκε ως απόδειξη ότι η αντίστοιχη κατανομή αποτελεί τη μοναδική ή πραγματική μορφή των συγκεντρώσεων PM_{10} . Αντίθετα, χρησιμοποιήθηκε ως σχετικό μέτρο σύγκρισης μεταξύ των τεσσάρων υποψήφιων μοντέλων. Η τελική επιλογή της καταλληλότερης κατανομής βασίστηκε στη συνδυαστική αξιολόγηση του χ^2 και των γραφικών διαγνωστικών, ώστε να λαμβάνονται υπόψη τόσο η συνολική απόκλιση όσο και η θέση και η μορφή των επιμέρους αποκλίσεων.

3.7 Γραφική αξιολόγηση της προσαρμογής των θεωρητικών κατανομών

Η αξιολόγηση της προσαρμογής δεν βασίστηκε αποκλειστικά στο στατιστικό χ^2 , αλλά συμπληρώθηκε με γραφική διερεύνηση. Η προσέγγιση αυτή επέτρεψε τον εντοπισμό αποκλίσεων που δεν αποτυπώνονται πλήρως σε έναν συνολικό αριθμητικό δείκτη, ιδίως στην κύρια περιοχή των δεδομένων και στην άνω ουρά της κατανομής.

Για κάθε πόλη, έτος και θεωρητικό μοντέλο εξετάστηκαν αρχικά τα γραφήματα σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων ανά κλάση συγκέντρωσης. Μέσω αυτών αξιολογήθηκε κατά πόσο κάθε κατανομή αναπαρήγαγε τη γενική μορφή της εμπειρικής κατανομής και σε ποιες κλάσεις παρουσίαζε υπερεκτίμηση ή υποεκτίμηση.

Παράλληλα, χρησιμοποιήθηκαν observed-expected διαγράμματα. Στα διαγράμματα αυτά, η εγγύτητα των σημείων προς τη διαγώνια γραμμή $y=x$ θεωρήθηκε ένδειξη καλύτερης συμφωνίας μεταξύ θεωρητικών και εμπειρικών συχνοτήτων. Συστηματικές αποκλίσεις

από τη γραμμή αναφοράς υποδήλωναν ότι το μοντέλο δεν απέδιδε ομοιόμορφα τις επιμέρους κλάσεις.

Τα διαγράμματα υπολοίπων χρησιμοποιήθηκαν για την εξέταση της φοράς και της συστηματικότητας των αποκλίσεων. Τα υπόλοιπα ορίστηκαν ως η διαφορά μεταξύ παρατηρούμενης και αναμενόμενης συχνότητας. Η τυχαία διασπορά τους γύρω από το μηδέν θεωρήθηκε ένδειξη ικανοποιητικής προσαρμογής, ενώ οργανωμένα ή επαναλαμβανόμενα πρότυπα έδειχναν συστηματική αστοχία του μοντέλου σε συγκεκριμένες περιοχές της κατανομής.

Συμπληρωματικά, χρησιμοποιήθηκαν Q-Q διαγράμματα για τη σύγκριση των εμπειρικών με τα θεωρητικά ποσοστημόρια. Η προσέγγιση των σημείων στη διαγώνια γραμμή αναφοράς υποδήλωνε καλύτερη συμφωνία, ενώ αποκλίσεις στα άκρα αποκάλυπταν αδυναμία του μοντέλου να αποδώσει τις χαμηλότερες ή, κυρίως, τις υψηλότερες συγκεντρώσεις.

Η γραφική αξιολόγηση χρησιμοποιήθηκε ως συμπληρωματικό διαγνωστικό εργαλείο και όχι ως ανεξάρτητο ή αποκλειστικό κριτήριο. Σε συνδυασμό με το χ^2 , επέτρεψε πιο ολοκληρωμένη αποτίμηση της προσαρμογής και καλύτερη ερμηνεία των διαφορών μεταξύ των υποψήφιων κατανομών.

3.8 Επιλογή καταλληλότερης κατανομής ανά πόλη και έτος

Μετά την εφαρμογή των θεωρητικών κατανομών ακολούθησε η συγκριτική αξιολόγησή τους για κάθε επιλεγμένη χρονοσειρά. Η τελική επιλογή δεν βασίστηκε σε ένα μόνο κριτήριο, αλλά στη συνδυαστική εξέταση της αριθμητικής και της γραφικής απόδοσης κάθε μοντέλου.

Ως βασικό αριθμητικό κριτήριο χρησιμοποιήθηκε η συνολική τιμή του χ^2 . Οι τιμές συγκρίθηκαν μόνο μεταξύ κατανομών που εφαρμόστηκαν στην ίδια χρονοσειρά, με κοινό πλήθος παρατηρήσεων και ίδια όρια κλάσεων. Μικρότερη τιμή χ^2 θεωρήθηκε ένδειξη καλύτερης συνολικής προσαρμογής, χωρίς όμως να αποτελεί από μόνη της επαρκή προϋπόθεση για την τελική επιλογή.

Η αριθμητική κατάταξη εξετάστηκε σε συνδυασμό με τα γραφήματα παρατηρούμενων και αναμενόμενων συχνοτήτων, τα διαγράμματα observed-expected, τα διαγράμματα υπολοίπων και τα Q-Q διαγράμματα. Ιδιαίτερη σημασία δόθηκε στην ύπαρξη συστηματικών αποκλίσεων, στην κατανομή των υπολοίπων γύρω από το μηδέν και στη συμπεριφορά κάθε μοντέλου στην άνω ουρά.

Ως καταλληλότερη θεωρήθηκε η κατανομή που συνδύαζε χαμηλή τιμή χ^2 , ικανοποιητική συμφωνία των συχνοτήτων, περιορισμένα και μη συστηματικά υπόλοιπα και αποδεκτή συμφωνία μεταξύ θεωρητικών και εμπειρικών ποσοστημορίων. Ισχυρότερη ένδειξη καταλληλότητας θεωρήθηκε η σύγκλιση περισσότερων ανεξάρτητων διαγνωστικών προς την ίδια επιλογή, καθώς αυτή περιορίζει τον κίνδυνο μιας μηχανιστικής κατάταξης που θα βασιζόταν αποκλειστικά σε ένα αριθμητικό μέτρο.

Η αξιολόγηση πραγματοποιήθηκε χωριστά για κάθε πόλη και έτος, καθώς οι χρονοσειρές διέφεραν ως προς την κεντρική τάση, τη διασπορά, την ασυμμετρία και την παρουσία υψηλών τιμών. Με τον τρόπο αυτό, η τελική επιλογή αντανάκλα τα ειδικά χαρακτηριστικά κάθε συνόλου δεδομένων και όχι μία προκαθορισμένη παραδοχή κοινής κατανομικής συμπεριφοράς για όλες τις περιοχές.

Η επιλογή ερμηνεύεται με σχετικό χαρακτήρα. Η κατανομή που αναδείχθηκε ως καταλληλότερη θεωρείται η καλύτερη μόνο μεταξύ των τεσσάρων μοντέλων που εξετάστηκαν και για το συγκεκριμένο σύνολο δεδομένων. Δεν υποδηλώνεται ότι αποτελεί τη μοναδική ή οριστική περιγραφή της κατανομικής συμπεριφοράς των συγκεντρώσεων PM_{10} . Η επανάληψη της διαδικασίας σε διαφορετικές περιόδους ή με πρόσθετα μοντέλα ενδέχεται να οδηγήσει σε διαφορετική κατάταξη.

4. Αποτελέσματα

Στο παρόν κεφάλαιο παρουσιάζονται τα αποτελέσματα της στατιστικής ανάλυσης των ωριαίων συγκεντρώσεων PM_{10} στις επιλεγμένες πόλεις της Στερεάς Ελλάδας. Η ανάλυση περιορίζεται στις ετήσιες χρονοσειρές που ικανοποίησαν τα προκαθορισμένα κριτήρια πληρότητας και χρονικής συνέχειας, ώστε η αξιολόγηση των κατανομικών χαρακτηριστικών να βασίζεται σε επαρκές και συγκρίσιμο σύνολο παρατηρήσεων.

Η παρουσίαση των αποτελεσμάτων οργανώνεται σε διαδοχικά επίπεδα. Αρχικά αποτυπώνεται το τελικό αναλυτικό σύνολο και περιγράφονται τα βασικά στατιστικά χαρακτηριστικά των χρονοσειρών. Στη συνέχεια εξετάζεται η εμπειρική μορφή των συγκεντρώσεων και αξιολογείται η προσαρμογή της κανονικής, της λογαριθμοκανονικής, της Gamma και της Weibull κατανομής. Η αξιολόγηση βασίζεται στη συνδυαστική ερμηνεία της συμφωνίας παρατηρούμενων και αναμενόμενων συχνοτήτων, των διαγνωστικών διαγραμμάτων, των υπολοίπων, των Q-Q διαγραμμάτων και του στατιστικού χ^2 .

Η επιλογή της καταλληλότερης κατανομής δεν αντιμετωπίζεται ως μηχανική κατάταξη βάσει ενός μόνο δείκτη, αλλά ως διαδικασία σύνθεσης αριθμητικών και γραφικών ενδείξεων. Με αυτή την προσέγγιση διερευνάται τόσο η συνολική ικανότητα κάθε θεωρητικού μοντέλου να αναπαράγει την εμπειρική κατανομή όσο και η συμπεριφορά του στο κεντρικό τμήμα και στις ουρές των χρονοσειρών. Η συγκριτική αποτίμηση επιτρέπει, τέλος, την ανάδειξη κοινών κατανομικών προτύπων και διαφοροποιήσεων μεταξύ των εξεταζόμενων πόλεων, χωρίς να αποδίδονται οι διαφορές αυτές σε συγκεκριμένους μηχανισμούς εκπομπής ή διασποράς πέρα από όσα μπορούν να τεκμηριωθούν από τα διαθέσιμα δεδομένα.

4.1 Τελικό αναλυτικό σύνολο δεδομένων

Πριν από την εφαρμογή των στατιστικών αναλύσεων και την προσαρμογή των θεωρητικών κατανομών πιθανότητας, αξιολογήθηκαν η πληρότητα και η χρονική συνέχεια των διαθέσιμων ωριαίων χρονοσειρών συγκεντρώσεων PM_{10} για τις πέντε υπό μελέτη πόλεις κατά την περίοδο 2017-2024. Η διαδικασία αυτή αποσκοπούσε στον εντοπισμό των χρονοσειρών που παρείχαν επαρκή αριθμό παρατηρήσεων και

ικανοποιητική χρονική κάλυψη, ώστε η κατανομική ανάλυση να βασίζεται σε δεδομένα κατάλληλης ποιότητας.

Σύμφωνα με τα κριτήρια που καθορίστηκαν στο Κεφάλαιο 3, στη στατιστική ανάλυση εντάχθηκαν μόνο οι ετήσιες χρονοσειρές που υπερέβαιναν το προκαθορισμένο κατώφλι πληρότητας και παρουσίαζαν επαρκή συνέχεια στον χρόνο. Με τον τρόπο αυτό περιορίστηκε η πιθανότητα στρέβλωσης των εμπειρικών κατανομών από εκτεταμένα ή συστηματικά κενά στις μετρήσεις. Από κάθε πόλη επιλέχθηκε η πληρέστερη διαθέσιμη ετήσια χρονοσειρά που ικανοποιούσε τα κριτήρια ποιότητας της παρούσας εργασίας. Οι αναλυτικοί έλεγχοι πληρότητας και χρονικής συνέχειας για όλες τις πόλεις και όλα τα διαθέσιμα έτη παρατίθενται στο Παράρτημα Α.

Πόλη	Έτος	Έγκυρες ώρες	Πληρότητα (%)
Αμφισσα	2018	8750	99,89
Λαμία	2017	8059	92,00
Λειβαδιά	2019	8745	99,83
Χαλκίδα	2017	8702	99,34
Καρπενήσι	-	-	Μη αναλύσιμο

Πίνακας 4.1 Τελικό αναλυτικό σύνολο δεδομένων PM₁₀ που χρησιμοποιήθηκε στις στατιστικές αναλύσεις.

Σημείωση: Για το Καρπενήσι δεν εντοπίστηκε ετήσια χρονοσειρά που να ικανοποιεί τα κριτήρια πληρότητας και χρονικής συνέχειας της παρούσας εργασίας.

Όπως προκύπτει από τον Πίνακα 4.1, τέσσερις από τις πέντε εξεταζόμενες πόλεις παρείχαν τουλάχιστον μία ετήσια χρονοσειρά κατάλληλη για περαιτέρω ανάλυση. Η υψηλότερη πληρότητα καταγράφηκε στην Αμφισσα το 2018, με ποσοστό 99,89%, και

ακολούθησαν η Λειβαδιά το 2019 με 99,83% και η Χαλκίδα το 2017 με 99,34%. Η χρονοσειρά της Λαμίας για το 2017 κάλυψε το 92,00% των αναμενόμενων ωριαίων παρατηρήσεων και διατήρησε επαρκή χρονική συνέχεια, με αποτέλεσμα να κριθεί κατάλληλη για κατανομική ανάλυση. Αντίθετα, καμία από τις διαθέσιμες χρονοσειρές του Καρπενησίου δεν ικανοποίησε τα προκαθορισμένα κριτήρια και, ως εκ τούτου, η συγκεκριμένη πόλη δεν εντάχθηκε στην κύρια στατιστική ανάλυση.

Η διαφοροποίηση της πληρότητας μεταξύ των τεσσάρων επιλεγμένων χρονοσειρών δεν αναιρεί τη δυνατότητα συγκριτικής αξιολόγησής τους, καθώς όλες υπερβαίνουν το κατώφλι αποδοχής και διαθέτουν επαρκή χρονική κάλυψη. Ωστόσο, η χρονοσειρά της Λαμίας, λόγω του χαμηλότερου ποσοστού πληρότητας σε σχέση με τις υπόλοιπες, ερμηνεύεται με μεγαλύτερη προσοχή, ιδιαίτερα ως προς την πιθανή επίδραση των ελλειπυσών παρατηρήσεων στη μορφή της εμπειρικής κατανομής. Η ένταξη της στο τελικό σύνολο δεν σημαίνει ότι είναι απολύτως ισοδύναμη με τις σχεδόν πλήρεις χρονοσειρές των άλλων πόλεων, αλλά ότι πληροί το ελάχιστο επίπεδο ποιότητας που τέθηκε για την παρούσα ανάλυση.

Η επιλογή μίας ετήσιας χρονοσειράς ανά πόλη εξασφαλίζει κοινή χρονική κλίμακα ανάλυσης και διευκολύνει τη συγκριτική αποτίμηση των κατανομικών χαρακτηριστικών. Παράλληλα, επειδή οι επιλεγμένες χρονοσειρές αντιστοιχούν σε διαφορετικά ημερολογιακά έτη, οι μεταξύ τους συγκρίσεις αφορούν κυρίως τη στατιστική μορφή των συγκεντρώσεων και όχι την άμεση σύγκριση ταυτόχρονων περιβαλλοντικών συνθηκών. Συνεπώς, οι παρατηρούμενες διαφορές δεν μπορούν να αποδοθούν αποκλειστικά σε χωρικά χαρακτηριστικά των πόλεων, καθώς ενδέχεται να επηρεάζονται και από τη διαφορετική χρονική περίοδο αναφοράς.

Το τελικό αναλυτικό σύνολο περιλαμβάνει, επομένως, τις χρονοσειρές της Άμφισσας για το 2018, της Λαμίας για το 2017, της Λειβαδιάς για το 2019 και της Χαλκίδας για το 2017. Οι χρονοσειρές αυτές αποτελούν τη βάση για τον υπολογισμό των περιγραφικών στατιστικών, την αποτύπωση των εμπειρικών κατανομών και την αξιολόγηση της προσαρμογής των θεωρητικών κατανομών πιθανότητας στις επόμενες ενότητες. Τα αποτελέσματα που προκύπτουν αφορούν τις συγκεκριμένες ετήσιες περιόδους και δεν πρέπει να ερμηνεύονται ως πλήρης αποτύπωση της μακροχρόνιας συμπεριφοράς των συγκεντρώσεων PM_{10} σε κάθε πόλη.

4.2 Περιγραφική στατιστική των αναλύσιμων χρονοσειρών PM₁₀

Μετά τον καθορισμό του τελικού αναλυτικού συνόλου δεδομένων, πραγματοποιήθηκε περιγραφική στατιστική ανάλυση των ωριαίων συγκεντρώσεων PM₁₀ για καθεμία από τις τέσσερις επιλεγμένες ετήσιες χρονοσειρές. Υπολογίστηκαν το πλήθος των έγκυρων παρατηρήσεων N, η αριθμητική μέση τιμή, η διάμεσος, η τυπική απόκλιση, η ελάχιστη και η μέγιστη τιμή, καθώς και οι συντελεστές ασυμμετρίας και κύρτωσης.

Οι δείκτες αυτοί παρέχουν μια πρώτη ποσοτική αποτύπωση της θέσης, της διασποράς και της μορφής των εμπειρικών κατανομών. Ιδιαίτερη σημασία έχουν η σχέση μεταξύ μέσης τιμής και διαμέσου, καθώς και οι τιμές ασυμμετρίας και κύρτωσης, επειδή επιτρέπουν την προκαταρκτική διερεύνηση της συμμετρίας, της έκτασης της δεξιάς ουράς και της επίδρασης ακραίων παρατηρήσεων. Η περιγραφική ανάλυση δεν υποκαθιστά τους ελέγχους προσαρμογής, αλλά λειτουργεί ως αρχικό διαγνωστικό στάδιο για την επιλογή και αξιολόγηση κατάλληλων θεωρητικών κατανομών πιθανότητας.

Πόλη	Έτος	N	Μέση τιμή (μg/m ³)	Διάμεσος (μg/m ³)	Τυπική Απόκλιση (μg/m ³)	Ελάχιστη τιμή (μg/m ³)	Μέγιστη τιμή (μg/m ³)	Ασυμμετρία	Κύρτωση
Αμφισσα	2018	8750	22,21	18,00	18,78	1	260	3,051	16,955
Λαμία	2017	8059	23,86	20,00	17,57	1	250	3,348	22,643
Λειβαδιά	2019	8745	23,57	21,00	15,62	1	438	6,528	117,861
Χαλκίδα	2017	8702	36,48	31,00	30,01	1	891	7,946	161,189

Πίνακας 4.2 Περιγραφικά στατιστικά χαρακτηριστικά των αναλύσιμων χρονοσειρών PM₁₀.

Σημείωση: Οι περιγραφικοί δείκτες υπολογίστηκαν μετά την εφαρμογή των κριτηρίων πληρότητας και ποιοτικού ελέγχου που περιγράφηκαν στο Κεφάλαιο 3.

Όπως προκύπτει από τον Πίνακα 4.2, οι μέσες ωριαίες συγκεντρώσεις κυμάνθηκαν από 22,21 $\mu\text{g}/\text{m}^3$ στην Άμφισσα έως 36,48 $\mu\text{g}/\text{m}^3$ στη Χαλκίδα. Η Άμφισσα, η Λαμία και η Λειβαδιά παρουσίασαν σχετικά κοντινές μέσες τιμές, από 22,21 έως 23,86 $\mu\text{g}/\text{m}^3$, ενώ η χρονοσειρά της Χαλκίδας εμφάνισε υψηλότερη μέση τιμή και διάμεσο. Η διαπίστωση αυτή αφορά τις συγκεκριμένες ετήσιες χρονοσειρές και δεν πρέπει να ερμηνεύεται ως άμεση κατάταξη της μακροχρόνιας ατμοσφαιρικής επιβάρυνσης των πόλεων, καθώς τα δεδομένα αντιστοιχούν σε διαφορετικά ημερολογιακά έτη.

Σε όλες τις περιπτώσεις, η μέση τιμή υπερέβη τη διάμεσο. Η συστηματική αυτή διαφορά αποτελεί πρώτη ένδειξη θετικής ασυμμετρίας, καθώς οι υψηλότερες παρατηρήσεις μετατοπίζουν τη μέση τιμή προς τα δεξιά περισσότερο από τη διάμεσο. Η ένδειξη επιβεβαιώνεται από τους θετικούς συντελεστές ασυμμετρίας, οι οποίοι κυμαίνονται από 3,051 στην Άμφισσα έως 7,946 στη Χαλκίδα. Επομένως, καμία από τις εξεταζόμενες χρονοσειρές δεν εμφανίζει μορφή προσεγγιστικά συμμετρικής κατανομής.

Η μικρότερη ασυμμετρία καταγράφηκε στην Άμφισσα και στη Λαμία, χωρίς αυτό να σημαίνει ότι οι κατανομές τους είναι ήπιες ή σχεδόν συμμετρικές, καθώς και στις δύο περιπτώσεις οι τιμές υπερβαίνουν σημαντικά το μηδέν. Εντονότερη δεξιά ασυμμετρία παρατηρήθηκε στη Λειβαδιά και κυρίως στη Χαλκίδα. Στις δύο αυτές χρονοσειρές, η μορφή της κατανομής φαίνεται να επηρεάζεται ισχυρότερα από περιορισμένο αριθμό πολύ υψηλών ωριαίων παρατηρήσεων.

Οι μέγιστες τιμές ήταν πολλαπλάσιες των αντίστοιχων μέσων τιμών. Στην Άμφισσα και στη Λαμία καταγράφηκαν μέγιστα 260 και 250 $\mu\text{g}/\text{m}^3$ αντίστοιχα, ενώ στη Λειβαδιά και στη Χαλκίδα οι μέγιστες τιμές ανήλθαν σε 438 και 891 $\mu\text{g}/\text{m}^3$. Οι παρατηρήσεις αυτές υποδηλώνουν εκτεταμένη δεξιά ουρά και παρουσία ακραίων τιμών. Ωστόσο, από την περιγραφική στατιστική και μόνο δεν είναι δυνατόν να προσδιοριστεί αν οι συγκεκριμένες τιμές αντιστοιχούν σε πραγματικά επεισόδια αυξημένων συγκεντρώσεων, σε ειδικές τοπικές ή περιφερειακές συνθήκες ή σε άλλους μηχανισμούς. Η αιτιολογική τους ερμηνεία θα απαιτούσε πρόσθετα δεδομένα, όπως μετεωρολογικές παρατηρήσεις, πληροφορίες για πηγές εκπομπής και ανεξάρτητη επιβεβαίωση των επεισοδίων.

Ως προς τη διασπορά, η χαμηλότερη τυπική απόκλιση καταγράφηκε στη Λειβαδιά, με 15,62 $\mu\text{g}/\text{m}^3$, και η υψηλότερη στη Χαλκίδα, με 30,01 $\mu\text{g}/\text{m}^3$. Η αυξημένη τυπική απόκλιση της Χαλκίδας συμφωνεί με το μεγαλύτερο εύρος τιμών και την παρουσία εξαιρετικά υψηλών παρατηρήσεων. Αντίθετα, η χαμηλότερη τυπική απόκλιση της Λειβαδιάς δεν αναιρεί την ύπαρξη ακραίων τιμών. Το εύρημα αυτό δείχνει ότι η

πλειονότητα των παρατηρήσεων μπορεί να συγκεντρώνεται σχετικά κοντά στο κεντρικό τμήμα της κατανομής, ενώ ένας μικρός αριθμός πολύ υψηλών τιμών επηρεάζει έντονα τα χαρακτηριστικά της ουράς.

Οι τιμές κύρτωσης ήταν ιδιαίτερα υψηλές σε όλες τις χρονοσειρές και αυξήθηκαν αισθητά στη Λειβαδιά και στη Χαλκίδα. Ανεξάρτητα από το αν οι τιμές έχουν υπολογιστεί με τη σύμβαση Pearson ή ως υπερβάλλουσα κύρτωση, η μεγάλη απόστασή τους από την αντίστοιχη τιμή αναφοράς της κανονικής κατανομής καταδεικνύει ισχυρή επίδραση ακραίων παρατηρήσεων και ουρών που δεν είναι συμβατές με απλή κανονική συμπεριφορά. Επιστημονικά, είναι ακριβέστερο οι τιμές αυτές να ερμηνεύονται πρωτίστως ως ένδειξη βαριών ουρών και αυξημένης συχνότητας ακραίων αποκλίσεων και όχι μόνο ως ένδειξη «έντονης αιχμής».

Η συνδυαστική εξέταση της μέσης τιμής, της διαμέσου, της ασυμμετρίας, της κύρτωσης και των ακραίων τιμών καταδεικνύει ότι οι τέσσερις εμπειρικές κατανομές είναι μη αρνητικές, έντονα δεξιά ασύμμετρες και επηρεάζονται σε διαφορετικό βαθμό από υψηλές παρατηρήσεις. Η κανονική κατανομή, λόγω της συμμετρίας της και του απεριόριστου θεωρητικού πεδίου τιμών της, δεν αναμένεται να αποδώσει ικανοποιητικά όλα τα χαρακτηριστικά των συγκεκριμένων δεδομένων. Η διαπίστωση αυτή δεν προδικάζει την τελική κατάταξη των μοντέλων, αλλά τεκμηριώνει την ανάγκη διερεύνησης κατανομών ορισμένων στο θετικό ημιάξονα, όπως η λογαριθμοκανονική, η Gamma και η Weibull.

Στην επόμενη ενότητα εξετάζεται γραφικά η εμπειρική μορφή των συγκεντρώσεων PM_{10} , με έμφαση στη θέση της κύριας μάζας των παρατηρήσεων, στην έκταση της δεξιάς ουράς και στην παρουσία ακραίων τιμών. Η ανάλυση αυτή παρέχει το αναγκαίο υπόβαθρο για τη μετέπειτα προσαρμογή και συγκριτική αξιολόγηση των θεωρητικών κατανομών πιθανότητας.

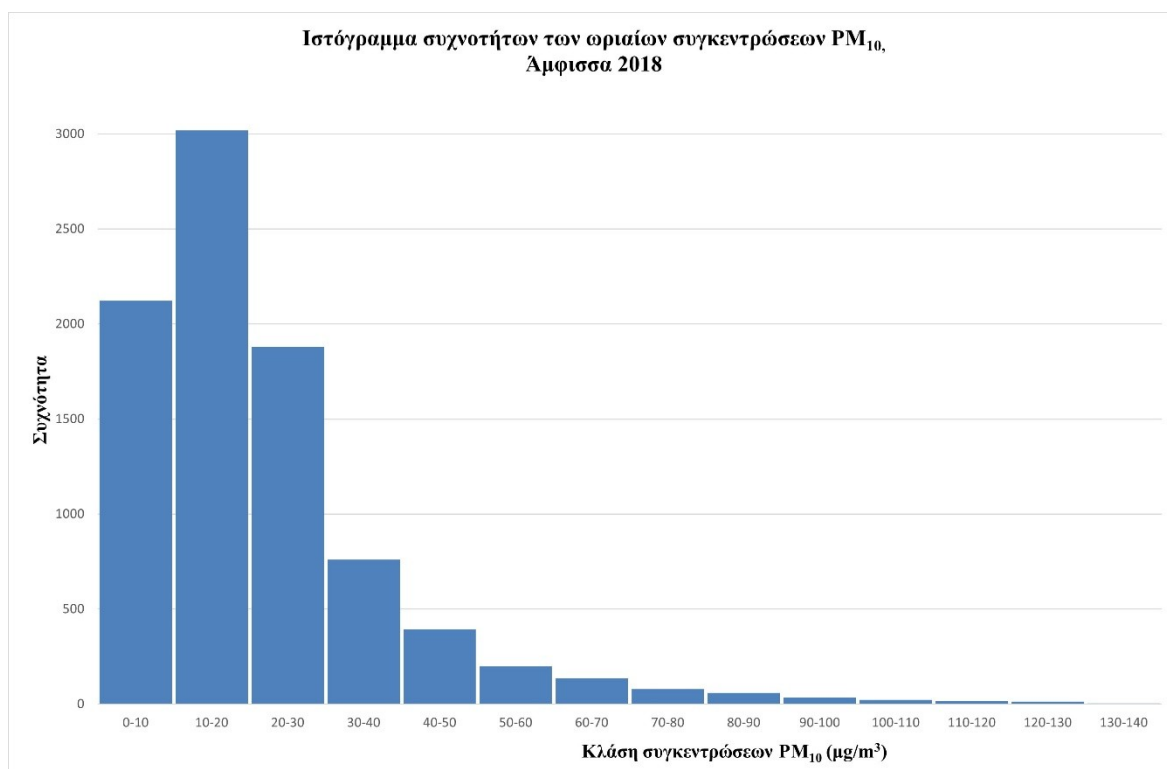
4.3 Εμπειρικές κατανομές συχνότητων των PM_{10}

Για τη γραφική διερεύνηση της εμπειρικής μορφής των ωριαίων συγκεντρώσεων PM_{10} κατασκευάστηκαν ιστογράμματα συχνότητων για τις τέσσερις επιλεγμένες ετήσιες χρονοσειρές. Οι συγκεντρώσεις ομαδοποιήθηκαν σε διαδοχικές κλάσεις ίσου εύρους, ώστε να αποτυπωθούν η θέση της κύριας μάζας των παρατηρήσεων, η διασπορά των τιμών και η έκταση της δεξιάς ουράς.

Τα ιστογράμματα χρησιμοποιούνται στο παρόν στάδιο ως διερευνητικά γραφικά εργαλεία. Η μορφή τους επηρεάζεται σε έναν βαθμό από το εύρος και τα όρια των επιλεγμένων κλάσεων και, επομένως, δεν επαρκεί από μόνη της για την επιλογή της καταλληλότερης θεωρητικής κατανομής. Η τελική αξιολόγηση πραγματοποιείται στις επόμενες ενότητες μέσω της συνδυαστικής εξέτασης των παρατηρούμενων και αναμενόμενων συχνοτήτων, των διαγνωστικών διαγραμμάτων, των υπολοίπων, των Q-Q διαγραμμάτων και του στατιστικού χ^2 .

Για λόγους ευκρίνειας, τα ιστογράμματα εστιάζουν στο εύρος συγκεντρώσεων μέσα στο οποίο εντοπίζεται η κύρια μάζα των παρατηρήσεων. Οι τιμές που υπερβαίνουν το ανώτερο όριο της γραφικής απεικόνισης έχουν συμπεριληφθεί στους περιγραφικούς και κατανομικούς υπολογισμούς, αλλά δεν αποτυπώνονται αναλυτικά στις επιμέρους κλάσεις των σχημάτων.

Άμφισσα, 2018

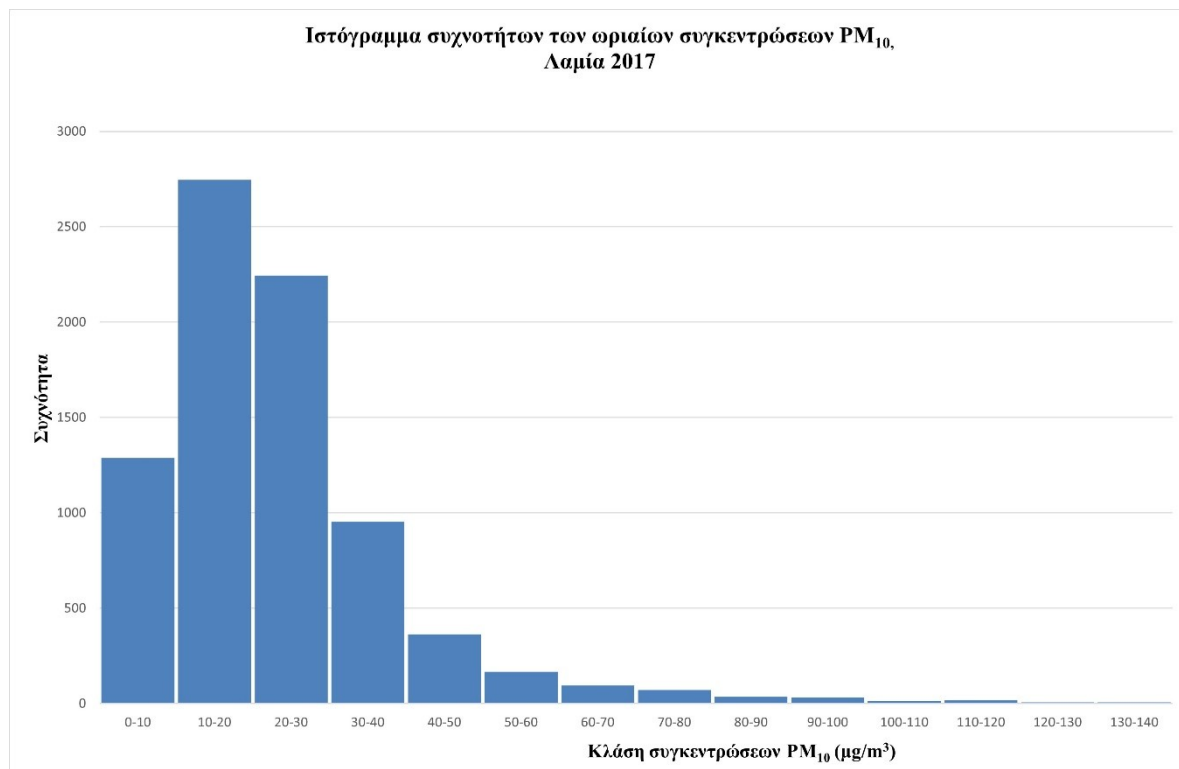


Σχήμα 4.3.α. Ιστόγραμμα συχνοτήτων των ωριαίων συγκεντρώσεων PM_{10} στην Άμφισσα για το έτος 2018.

Η εμπειρική κατανομή της Αμφισσας εμφανίζει σαφή επικράτηση των χαμηλών και μεσαίων συγκεντρώσεων. Η συχνότερη κλάση είναι η 10-20 $\mu\text{g}/\text{m}^3$, ενώ μεγάλο μέρος των παρατηρήσεων εντοπίζεται συνολικά στο εύρος 0-30 $\mu\text{g}/\text{m}^3$. Πέρα από το εύρος αυτό, οι συχνότητες μειώνονται προοδευτικά, χωρίς όμως να μηδενίζονται άμεσα, σχηματίζοντας εκτεταμένη δεξιά ουρά.

Η μορφή αυτή είναι συμβατή με θετική ασυμμετρία και συμφωνεί με τη διαφορά μεταξύ μέσης τιμής και διαμέσου, καθώς και με τον θετικό συντελεστή ασυμμετρίας που παρουσιάστηκε στον Πίνακα 4.2. Η κύρια μάζα των δεδομένων συγκεντρώνεται, επομένως, σε σχετικά χαμηλές τιμές, ενώ περιορισμένος αριθμός υψηλών παρατηρήσεων επεκτείνει την κατανομή προς τα δεξιά.

Λαμία, 2017

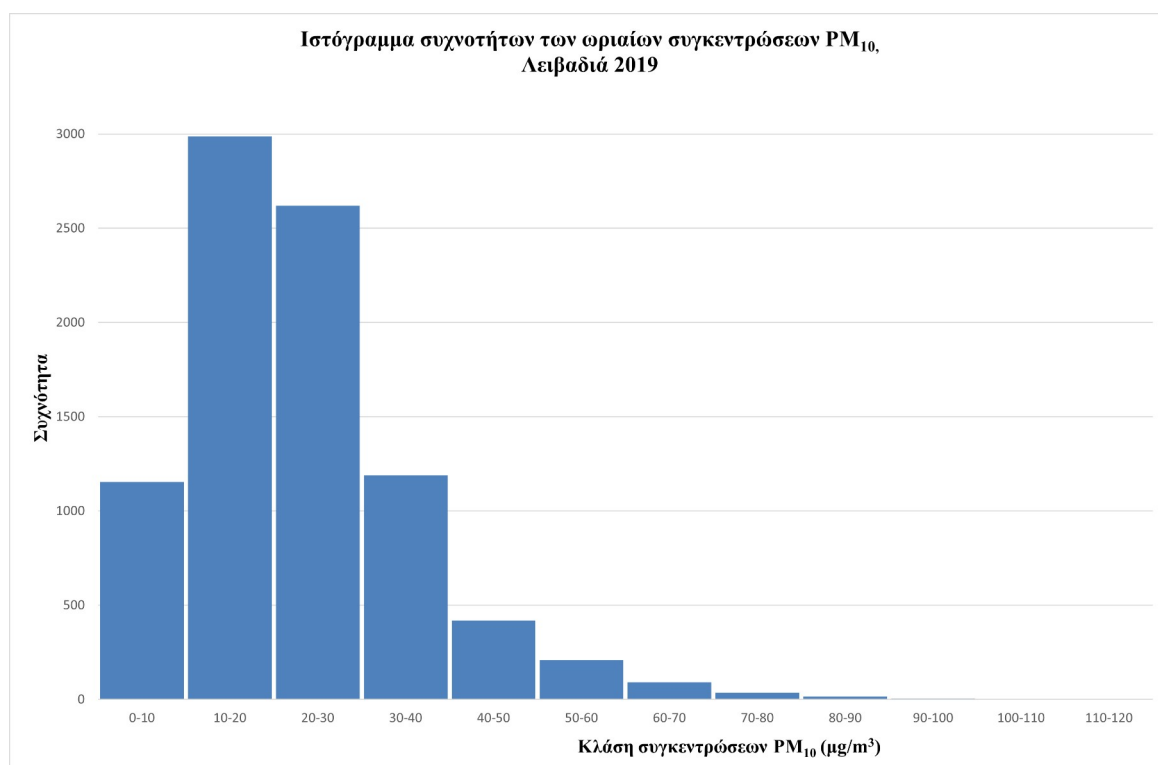


Σχήμα 4.3.β. Ιστόγραμμα συχνότητων των ωριαίων συγκεντρώσεων PM_{10} στη Λαμία για το έτος 2017.

Η χρονοσειρά της Λαμίας παρουσιάζει επίσης μονοκορυφική και δεξιά ασύμμετρη μορφή. Η υψηλότερη συχνότητα καταγράφεται στην κλάση 10-20 $\mu\text{g}/\text{m}^3$, ενώ σημαντικός αριθμός παρατηρήσεων εντοπίζεται και στην κλάση 20-30 $\mu\text{g}/\text{m}^3$. Από την κλάση 30-40 $\mu\text{g}/\text{m}^3$ και μετά παρατηρείται σταδιακή μείωση των συχνοτήτων.

Η κατανομή παρουσιάζει παρόμοια γενική μορφή με εκείνη της Αμφισσας, με τη βασική μάζα των δεδομένων να συγκεντρώνεται στις χαμηλότερες κλάσεις και με μικρότερο αριθμό παρατηρήσεων να εκτείνεται προς υψηλότερες συγκεντρώσεις. Η γραφική εικόνα επιβεβαιώνει τη θετική ασυμμετρία που καταγράφηκε στην περιγραφική ανάλυση.

Λειβαδιά, 2019



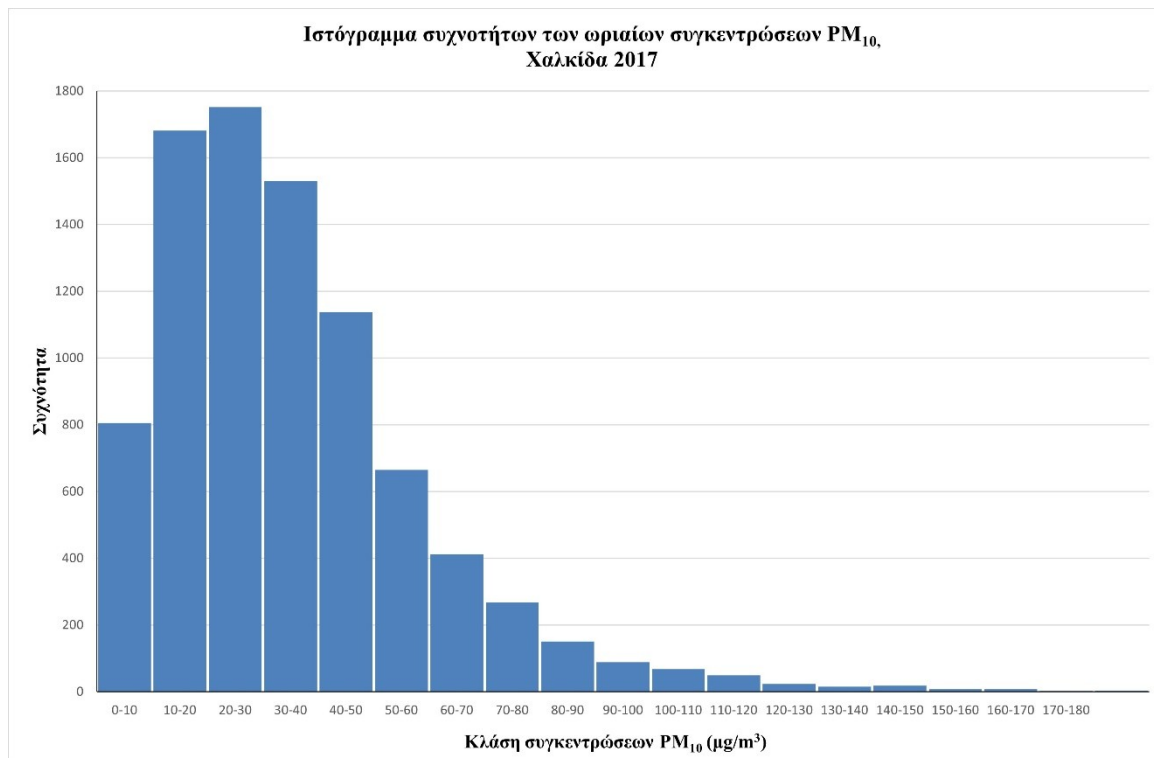
Σχήμα 4.3.γ. Ιστόγραμμα συχνότητων των ωριαίων συγκεντρώσεων PM_{10} στη Λειβαδιά για το έτος 2019.

Στη Λειβαδιά, η μεγαλύτερη συχνότητα εμφανίζεται στην κλάση $10-20 \mu g/m^3$ και ακολουθεί η κλάση $20-30 \mu g/m^3$. Η πλειονότητα των παρατηρήσεων συγκεντρώνεται συνολικά στο εύρος $10-40 \mu g/m^3$, ενώ μετά την κλάση $30-40 \mu g/m^3$ οι συχνότητες μειώνονται σχετικά γρήγορα.

Σε σύγκριση με τις υπόλοιπες χρονοσειρές, η κύρια μάζα της κατανομής εμφανίζεται περισσότερο συγκεντρωμένη, στοιχείο συμβατό με τη χαμηλότερη τυπική απόκλιση της Λειβαδιάς. Παράλληλα, η πολύ υψηλή ασυμμετρία και κύρτωση του Πίνακα 4.2 δείχνουν ότι, πέρα από το εύρος στο οποίο συγκεντρώνεται η πλειονότητα των παρατηρήσεων,

υπάρχουν περιορισμένες αλλά πολύ υψηλές τιμές που επηρεάζουν έντονα τη δεξιά ουρά. Για τον λόγο αυτό, η φαινομενικά περιορισμένη διασπορά του κεντρικού τμήματος δεν πρέπει να συγχέεται με απουσία ακραίων παρατηρήσεων.

Χαλκίδα, 2017



Σχήμα 4.3.δ. Ιστόγραμμα συχνοτήτων των ωριαίων συγκεντρώσεων PM₁₀ στη Χαλκίδα για το έτος 2017.

Η εμπειρική κατανομή της Χαλκίδας είναι εμφανώς ευρύτερη από εκείνες των άλλων τριών χρονοσειρών. Η μέγιστη συχνότητα καταγράφεται στην κλάση 20-30 µg/m³, ενώ υψηλές συχνότητες διατηρούνται σε μεγαλύτερο εύρος τιμών, περίπου από 10 έως 50 µg/m³. Η μείωση των συχνοτήτων προς τις υψηλότερες κλάσεις είναι περισσότερο σταδιακή και οι παρατηρήσεις κατανέμονται σε ευρύτερο φάσμα συγκεντρώσεων.

Η μορφή αυτή συμφωνεί με την υψηλότερη μέση τιμή, τη μεγαλύτερη τυπική απόκλιση και τον υψηλότερο συντελεστή ασυμμετρίας του Πίνακα 4.2. Η Χαλκίδα παρουσιάζει, επομένως, μεγαλύτερη διασπορά τόσο στο κεντρικό τμήμα της κατανομής όσο και προς τις υψηλότερες συγκεντρώσεις. Ωστόσο, επειδή η μέγιστη καταγεγραμμένη τιμή υπερβαίνει κατά πολύ το εύρος των βασικών κλάσεων του σχήματος, η ουρά πρέπει να

ερμηνεύεται σε συνδυασμό με τα αριθμητικά περιγραφικά μέτρα και όχι αποκλειστικά από την οπτική απεικόνιση.

Συγκριτική αποτίμηση των εμπειρικών κατανομών

Η συγκριτική εξέταση των τεσσάρων ιστογραμμάτων αναδεικνύει ένα κοινό κατανομικό πρότυπο: οι περισσότερες ωριαίες συγκεντρώσεις PM_{10} εντοπίζονται στις χαμηλές και μεσαίες κλάσεις, ενώ η συχνότητα μειώνεται προς τις υψηλότερες τιμές. Και στις τέσσερις περιπτώσεις σχηματίζεται δεξιά ουρά, γεγονός που επιβεβαιώνει ότι οι εμπειρικές κατανομές αποκλίνουν από μια συμμετρική κανονική μορφή.

Παρά το κοινό αυτό χαρακτηριστικό, η ένταση της διασποράς και η έκταση της δεξιάς ουράς διαφοροποιούνται μεταξύ των χρονοσειρών. Η Άμφισσα και η Λαμία παρουσιάζουν παρόμοια συγκέντρωση παρατηρήσεων κυρίως στο εύρος $10-30 \mu g/m^3$. Η Λειβαδιά εμφανίζει περισσότερο συμπαγές κεντρικό τμήμα, αλλά τα αριθμητικά χαρακτηριστικά της δείχνουν ισχυρή επίδραση περιορισμένων ακραίων παρατηρήσεων. Η Χαλκίδα διαφοροποιείται εντονότερα, καθώς η κύρια μάζα των δεδομένων εκτείνεται σε μεγαλύτερο εύρος και η μείωση των συχνοτήτων προς τα δεξιά είναι περισσότερο σταδιακή.

Τα ιστογράμματα επιβεβαιώνουν, επομένως, τα βασικά ευρήματα της περιγραφικής στατιστικής ως προς τη μη αρνητικότητα, τη θετική ασυμμετρία και την παρουσία δεξιών ουρών. Δεν αρκούν, ωστόσο, για να προσδιοριστεί εκ των προτέρων ποια θεωρητική κατανομή παρέχει την καλύτερη προσαρμογή. Η κανονική, η Gamma, η Weibull και η λογαριθμοκανονική κατανομή αξιολογούνται στις επόμενες ενότητες με ενιαία γραφικά και αριθμητικά κριτήρια, ώστε η τελική επιλογή να στηρίζεται στη συνολική τους απόδοση και όχι μόνο στην οπτική ομοιότητα με τα ιστογράμματα.

4.4 Προσαρμογή θεωρητικών κατανομών στην Άμφισσα για το έτος 2018

Στην παρούσα ενότητα αξιολογείται η προσαρμογή τεσσάρων θεωρητικών κατανομών πιθανότητας στις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας για το έτος 2018. Η συγκεκριμένη χρονοσειρά εξετάζεται αναλυτικά με ενιαία ακολουθία γραφικών και

αριθμητικών διαγνωστικών, ώστε να αποτιμηθεί η ικανότητα κάθε κατανομής να αναπαράγει τόσο το κεντρικό τμήμα όσο και τη δεξιά ουρά της εμπειρικής κατανομής.

Όπως προέκυψε από την περιγραφική ανάλυση και το αντίστοιχο ιστόγραμμα, η χρονοσειρά χαρακτηρίζεται από μη αρνητικές τιμές, θετική ασυμμετρία και περιορισμένο αριθμό υψηλών παρατηρήσεων. Τα χαρακτηριστικά αυτά καθιστούν αναγκαία τη σύγκριση ενός συμμετρικού μοντέλου αναφοράς με κατανομές που ορίζονται στο θετικό ημιάξονα και μπορούν να αποδώσουν δεξιά ασύμμετρη συμπεριφορά.

Για τον σκοπό αυτό εξετάστηκαν η κανονική κατανομή ως συμμετρικό πρότυπο αναφοράς και οι κατανομές Gamma, Weibull και λογαριθμοκανονική ως εναλλακτικά μοντέλα για μη αρνητικά δεδομένα. Η αξιολόγηση βασίστηκε στη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων, στα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων τιμών, στα διαγράμματα υπολοίπων, στα Q-Q διαγράμματα και στο στατιστικό χ^2 .

Η τελική επιλογή της καταλληλότερης κατανομής δεν βασίζεται σε ένα μόνο κριτήριο, αλλά στη συνδυαστική αποτίμηση της συνολικής συμφωνίας, της συμπεριφοράς στις συχνότερες κλάσεις και της απόδοσης στις υψηλές συγκεντρώσεις. Με τον τρόπο αυτό αποφεύγεται η μηχανική επιλογή ενός μοντέλου αποκλειστικά βάσει μίας αριθμητικής ή οπτικής ένδειξης.

4.4.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων

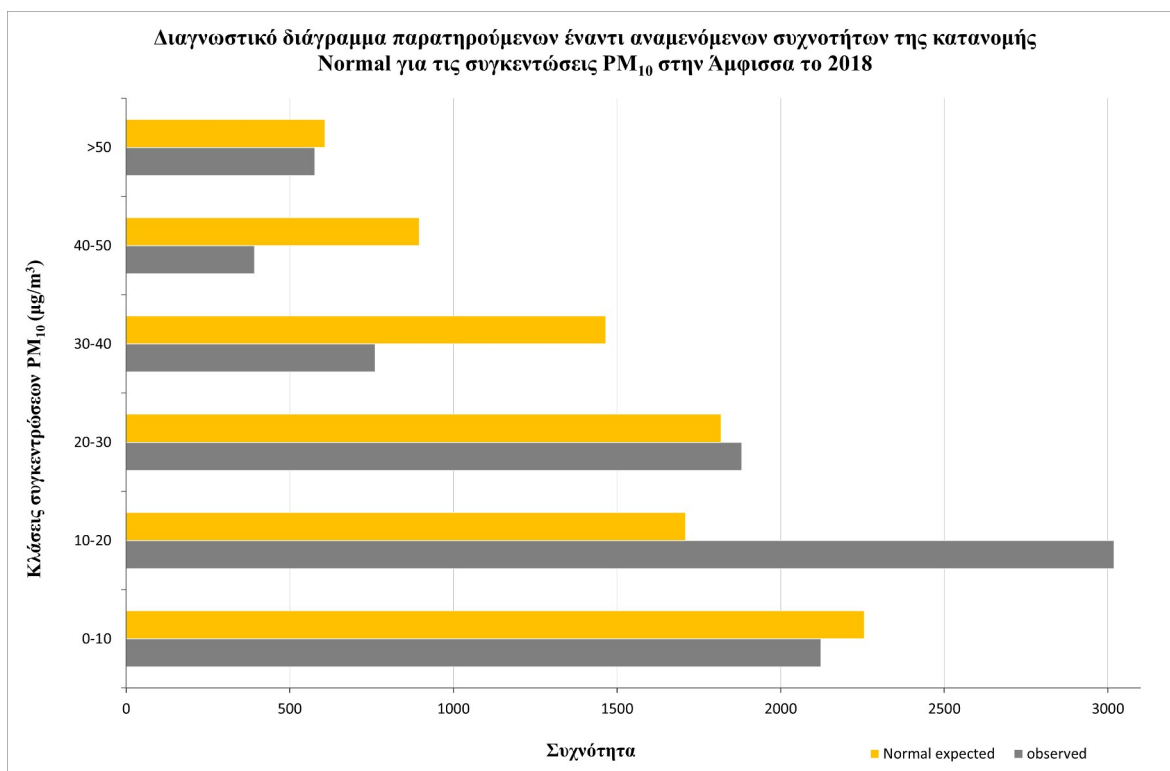
Στην παρούσα υποενότητα εξετάζεται η προσαρμογή των τεσσάρων θεωρητικών κατανομών πιθανότητας στις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας για το έτος 2018. Συγκεκριμένα, συγκρίνονται οι παρατηρούμενες συχνότητες των δεδομένων με τις αναμενόμενες συχνότητες που προκύπτουν από την κανονική, την Gamma, τη Weibull και τη λογαριθμοκανονική κατανομή. Η σύγκριση πραγματοποιείται ανά κλάση συγκέντρωσης, με χρήση των κλάσεων 0-10, 10-20, 20-30, 30-40, 40-50 και $>50 \mu g/m^3$.

Η ανάλυση αυτή αποτελεί το πρώτο στάδιο αξιολόγησης της προσαρμογής των θεωρητικών κατανομών στα εμπειρικά δεδομένα. Μέσω της άμεσης σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων μπορούν να εντοπιστούν οι κλάσεις στις οποίες κάθε μοντέλο παρουσιάζει ικανοποιητική συμφωνία, καθώς και εκείνες στις οποίες εμφανίζονται συστηματικές υποεκτιμήσεις ή υπερεκτιμήσεις. Με τον τρόπο αυτό, η

αξιολόγηση δεν περιορίζεται σε έναν συνολικό αριθμητικό δείκτη, αλλά εξετάζει τη συμπεριφορά των μοντέλων σε διαφορετικά τμήματα της εμπειρικής κατανομής.

Ιδιαίτερη σημασία έχει η απόδοση στις κλάσεις 0-30 $\mu\text{g}/\text{m}^3$, όπου συγκεντρώνεται το μεγαλύτερο μέρος των παρατηρήσεων, χωρίς να παραβλέπεται η συμπεριφορά στις υψηλότερες κλάσεις, οι οποίες συγκροτούν τη δεξιά ουρά της κατανομής. Η αξιολόγηση που ακολουθεί είναι αρχικά γραφική και δεν οδηγεί από μόνη της στην τελική επιλογή μοντέλου. Η τελική αποτίμηση βασίζεται στη συνδυαστική εξέταση των επιμέρους γραφικών διαγνωστικών και της συγκριτικής τιμής του στατιστικού χ^2 .

Κανονική κατανομή



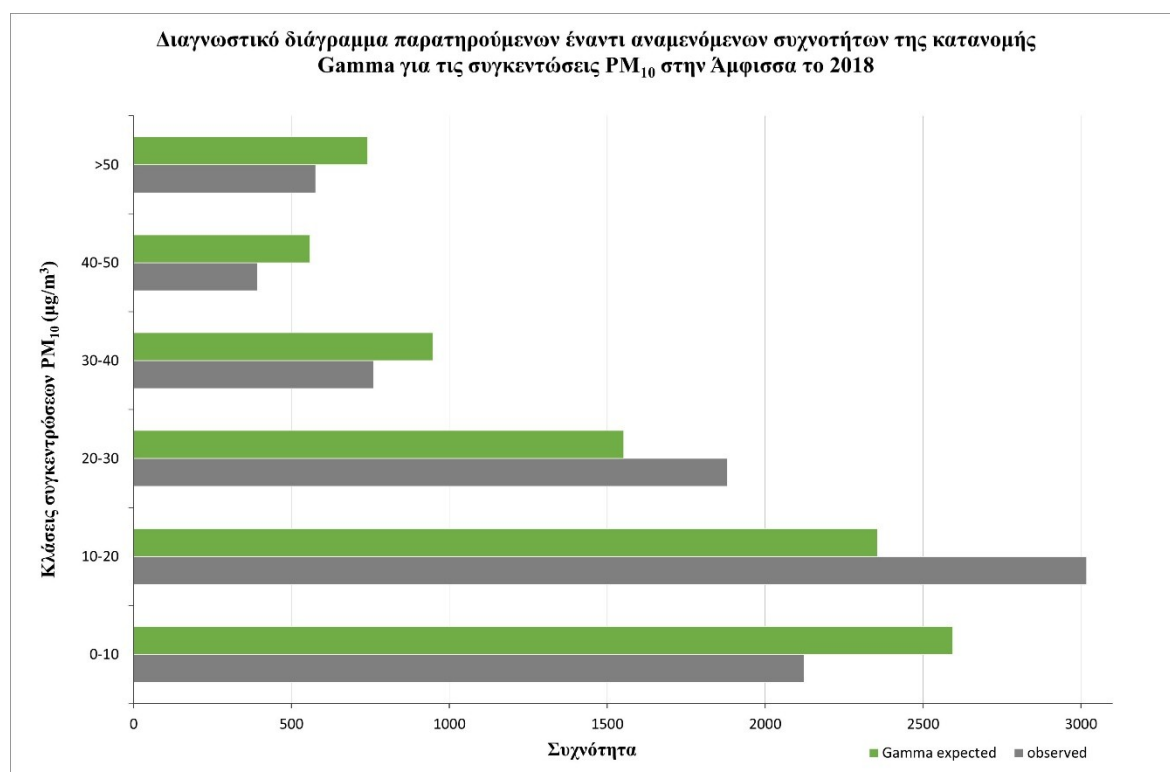
Σχήμα 4.4.1α. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κανονικής κατανομής στην Άμφισσα για το έτος 2018.

Η κανονική κατανομή παρουσιάζει σημαντικές αποκλίσεις από τις παρατηρούμενες συχνότητες. Η μεγαλύτερη διαφορά εντοπίζεται στην κλάση 10-20 $\mu\text{g}/\text{m}^3$, όπου οι αναμενόμενες συχνότητες είναι αισθητά χαμηλότερες από τις παρατηρούμενες. Αυτό

δείχνει ότι το μοντέλο υποεκτιμά την κύρια εμπειρική κορυφή, δηλαδή την περιοχή στην οποία συγκεντρώνεται σημαντικό μέρος των ωριαίων τιμών.

Αντίθετα, στις κλάσεις 30-40 και 40-50 $\mu\text{g}/\text{m}^3$ οι αναμενόμενες συχνότητες υπερβαίνουν τις παρατηρούμενες. Η εναλλαγή υποεκτίμησης και υπερεκτίμησης δείχνει ότι η κανονική κατανομή δεν αναπαράγει ικανοποιητικά ούτε τη συγκέντρωση των δεδομένων στις χαμηλότερες κλάσεις ούτε τη σταδιακή μείωση των συχνοτήτων προς τις υψηλότερες τιμές. Η συμπεριφορά αυτή είναι συμβατή με τη συμμετρική μορφή της κανονικής κατανομής, η οποία δεν ανταποκρίνεται επαρκώς στη δεξιά ασυμμετρία της συγκεκριμένης χρονοσειράς.

Αμφισσα 2018



Σχήμα 4.4.1β. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κατανομής Gamma στην Αμφισσα για το έτος 2018.

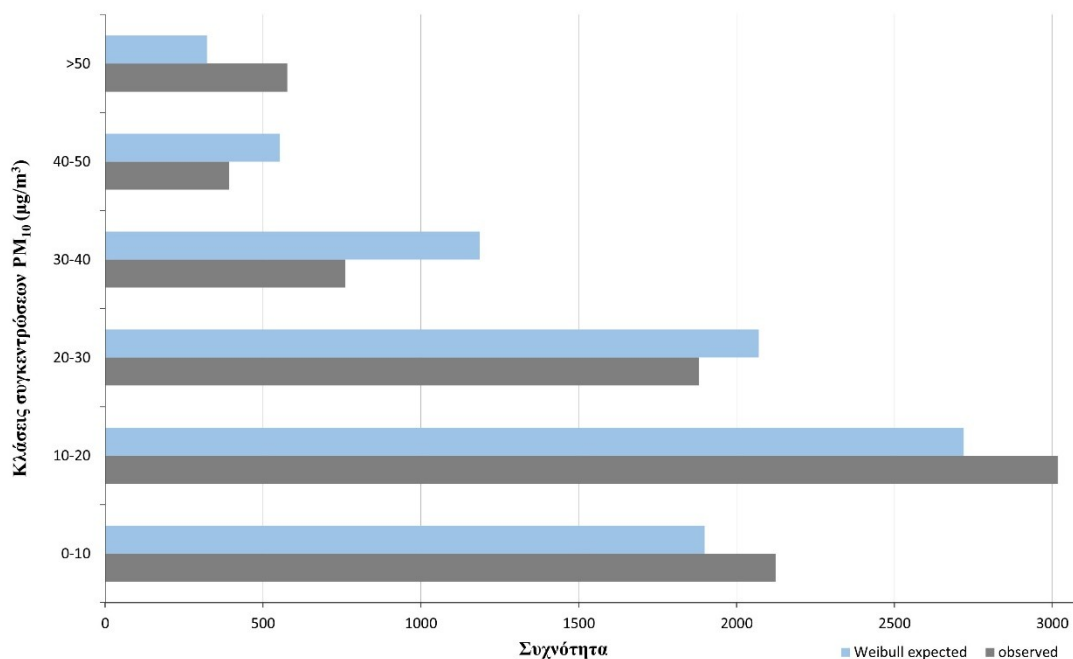
Η κατανομή Gamma παρουσιάζει καλύτερη προσαρμογή σε σχέση με την κανονική, καθώς αποδίδει αποτελεσματικότερα τη γενική ασύμμετρη μορφή των δεδομένων. Οι αναμενόμενες συχνότητες ακολουθούν σε μεγαλύτερο βαθμό τη συγκέντρωση των

παρατηρήσεων στις χαμηλές και μεσαίες κλάσεις, καθώς και τη φθίνουσα πορεία προς τις υψηλότερες συγκεντρώσεις.

Παρά τη βελτιωμένη συνολική εικόνα, εξακολουθούν να εμφανίζονται αξιοσημείωτες αποκλίσεις. Η κλάση 10-20 $\mu\text{g}/\text{m}^3$ υποεκτιμάται, ενώ στις κλάσεις 30-40, 40-50 και >50 $\mu\text{g}/\text{m}^3$ καταγράφεται τάση υπερεκτίμησης. Επομένως, η Gamma αποτελεί σαφή βελτίωση έναντι της κανονικής κατανομής, αλλά δεν αναπαράγει με την ίδια ακρίβεια την εμπειρική κορυφή και τη συμπεριφορά της δεξιάς ουράς.

Κατανομή Weibull

Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις συγκεντρώσεις PM_{10} στην Αμφισσα το 2018

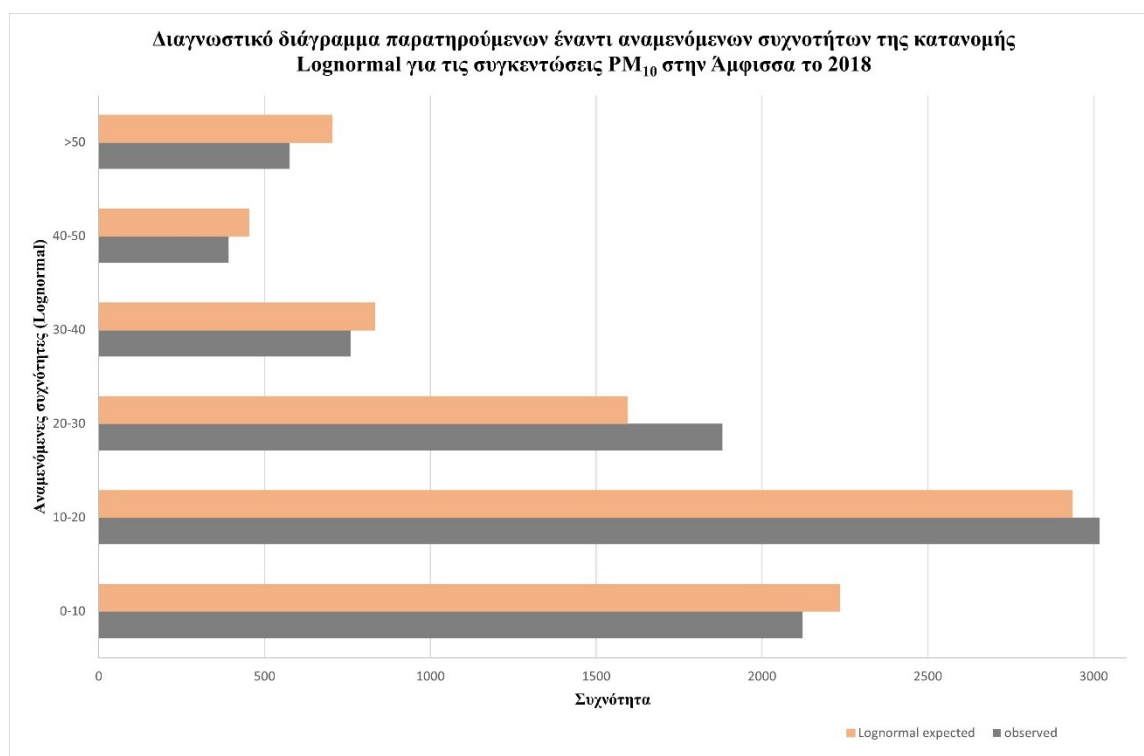


Σχήμα 4.4.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κατανομής Weibull στην Αμφισσα για το έτος 2018.

Η κατανομή Weibull παρουσιάζει σχετικά ικανοποιητική συμφωνία σε μέρος του κεντρικού εύρους των δεδομένων. Στις κλάσεις 10-20 και 20-30 $\mu\text{g}/\text{m}^3$ οι αναμενόμενες συχνότητες βρίσκονται αρκετά κοντά στις παρατηρούμενες, γεγονός που δείχνει ότι το μοντέλο προσεγγίζει σε ικανοποιητικό βαθμό την κύρια μάζα της εμπειρικής κατανομής.

Η απόδοσή της είναι ασθενέστερη στα άκρα. Στην κλάση 0-10 $\mu\text{g}/\text{m}^3$ οι αναμενόμενες συχνότητες είναι χαμηλότερες από τις παρατηρούμενες, ενώ υποεκτίμηση εμφανίζεται και στην ανοικτή ανώτερη κλάση $>50 \mu\text{g}/\text{m}^3$. Παράλληλα, στην κλάση 30-40 $\mu\text{g}/\text{m}^3$ παρατηρείται υπερεκτίμηση. Συνεπώς, η Weibull αποδίδει καλύτερα το κεντρικό τμήμα της εμπειρικής κατανομής από ό,τι τις πολύ χαμηλές και τις υψηλότερες συγκεντρώσεις.

Λογαριθμοκανονική κατανομή



Σχήμα 4.4.1δ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων των PM_{10} βάσει της κατανομής Lognormal στην Αμφισσα για το έτος 2018.

Η λογαριθμοκανονική κατανομή παρουσιάζει την πιο ισορροπημένη γραφική συμφωνία με τις παρατηρούμενες συχνότητες. Στις κλάσεις 0-10 και 10-20 $\mu\text{g}/\text{m}^3$ οι αποκλίσεις είναι σχετικά περιορισμένες, ενώ και στην κλάση 20-30 $\mu\text{g}/\text{m}^3$ η θεωρητική συχνότητα διατηρεί ικανοποιητική εγγύτητα προς την παρατηρούμενη. Η συμφωνία αυτή είναι σημαντική, καθώς οι συγκεκριμένες κλάσεις περιλαμβάνουν μεγάλο μέρος των ωριαίων παρατηρήσεων και καθορίζουν τη βασική μορφή της εμπειρικής κατανομής.

Στις κλάσεις 30-40, 40-50 και $>50 \mu\text{g}/\text{m}^3$ εξακολουθούν να εμφανίζονται αποκλίσεις, κυρίως με τάση υπερεκτίμησης. Ωστόσο, η ένταση και η κατανομή τους είναι περισσότερο ισορροπημένες σε σχέση με τη συμπεριφορά των υπόλοιπων μοντέλων. Η λογαριθμοκανονική κατανομή φαίνεται, επομένως, να αποδίδει αποτελεσματικότερα τον συνδυασμό υψηλής συγκέντρωσης παρατηρήσεων στις χαμηλές κλάσεις και σταδιακής εξασθένησης των συχνοτήτων προς τη δεξιά ουρά.

Συγκριτική αξιολόγηση

Η σύγκριση των τεσσάρων θεωρητικών μοντέλων αναδεικνύει ουσιαστικές διαφορές στην κατανομή των αποκλίσεων ανά κλάση. Η κανονική κατανομή εμφανίζει τη λιγότερο ικανοποιητική συνολική απόδοση, καθώς υποεκτιμά έντονα την κύρια εμπειρική κορυφή και υπερεκτιμά τμήματα των μεσαίων και υψηλότερων συγκεντρώσεων. Η εικόνα αυτή δείχνει ότι το συμμετρικό πρότυπο της κανονικής κατανομής δεν αποδίδει επαρκώς τη δεξιά ασύμμετρη μορφή των συγκεκριμένων δεδομένων.

Οι κατανομές Gamma και Weibull βελτιώνουν αισθητά την προσαρμογή, αλλά παρουσιάζουν διαφορετικό πρότυπο αποκλίσεων. Η Gamma αποδίδει αποτελεσματικότερα τη γενική ασύμμετρη δομή, εμφανίζει όμως υπερεκτίμηση σε αρκετές υψηλότερες κλάσεις. Η Weibull προσεγγίζει σχετικά καλά το κεντρικό εύρος των δεδομένων, αλλά υποεκτιμά τόσο τις πολύ χαμηλές όσο και τις υψηλότερες συγκεντρώσεις.

Η λογαριθμοκανονική κατανομή παρουσιάζει την πιο ισορροπημένη συμφωνία μεταξύ της κύριας μάζας των παρατηρήσεων και της δεξιάς ουράς. Δεν εξαλείφει πλήρως τις αποκλίσεις, αλλά αποφεύγει τις εντονότερες συστηματικές αστοχίες που εμφανίζονται στα υπόλοιπα μοντέλα. Με βάση αποκλειστικά τη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων, παρέχει συνεπώς την ισχυρότερη αρχική ένδειξη καταλληλότητας για τη χρονοσειρά της Άμφισσας το 2018.

Το συμπέρασμα αυτό παραμένει προκαταρκτικό και αφορά μόνο το συγκεκριμένο στάδιο αξιολόγησης. Η οπτική εγγύτητα των ράβδων δεν ποσοτικοποιεί πλήρως το συνολικό μέγεθος των αποκλίσεων ούτε αξιολογεί άμεσα τη συμπεριφορά των ποσοστημορίων και των υπολοίπων. Για τον λόγο αυτό, η κατάταξη των μοντέλων εξετάζεται περαιτέρω μέσω των διαγνωστικών διαγραμμάτων παρατηρούμενων-αναμενόμενων συχνοτήτων, των διαγραμμάτων υπολοίπων και της συγκριτικής τιμής του στατιστικού χ^2 .

4.4.2 Διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων

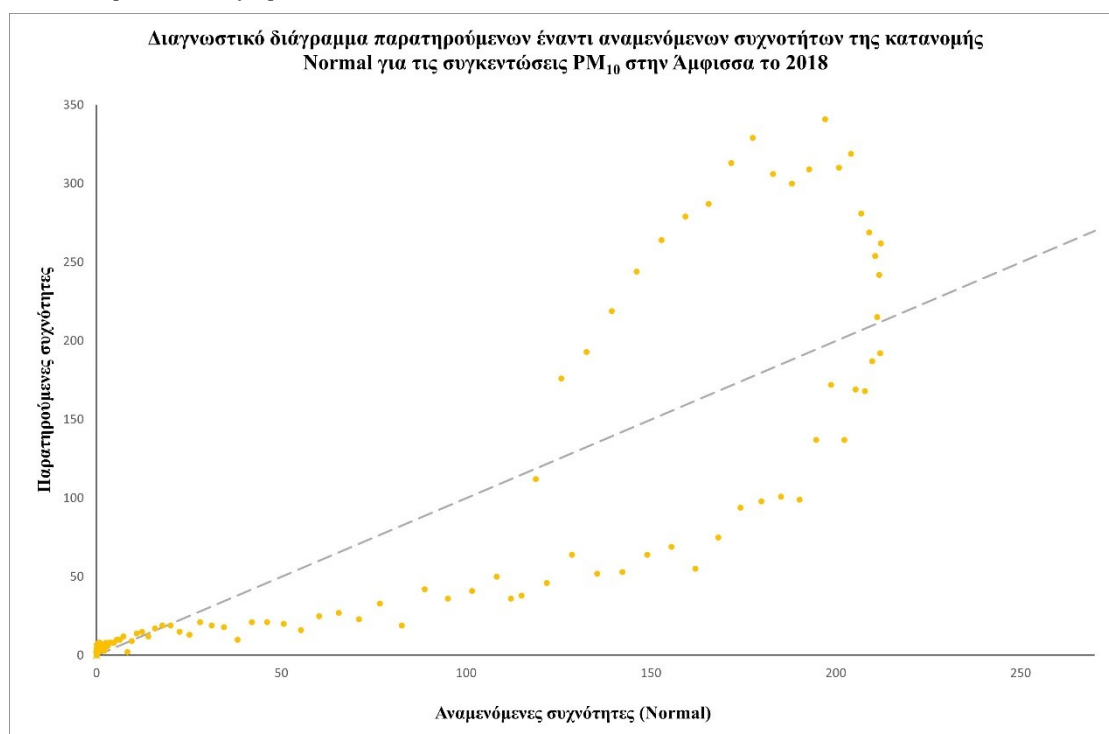
Τα Σχήματα 4.4.2α-δ παρουσιάζουν τη σχέση μεταξύ των παρατηρούμενων και των αναμενόμενων συχνοτήτων για την κανονική, την Gamma, τη Weibull και τη λογαριθμοκανονική κατανομή, όπως αυτές προσαρμόστηκαν στις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας για το έτος 2018. Στον οριζόντιο άξονα απεικονίζονται οι αναμενόμενες συχνότητες κάθε θεωρητικού μοντέλου και στον κατακόρυφο άξονα οι αντίστοιχες παρατηρούμενες συχνότητες.

Η διακεκομμένη διαγώνια γραμμή $y=x$ εκφράζει την ιδανική περίπτωση πλήρους συμφωνίας μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Όσο πλησιέστερα βρίσκονται τα σημεία στη γραμμή αναφοράς, τόσο καλύτερη θεωρείται η προσαρμογή του μοντέλου. Σημεία πάνω από τη διαγώνιο υποδηλώνουν ότι η παρατηρούμενη συχνότητα είναι μεγαλύτερη από την αναμενόμενη και, συνεπώς, ότι το μοντέλο την υποεκτιμά. Αντίθετα, σημεία κάτω από τη διαγώνιο υποδηλώνουν υπερεκτίμηση της παρατηρούμενης συχνότητας από το θεωρητικό μοντέλο.

Τα διαγράμματα αυτά συμπληρώνουν τη σύγκριση ανά κλάση της προηγούμενης υποενότητας, καθώς παρέχουν συνοπτική εικόνα της συνολικής σχέσης μεταξύ εμπειρικών και θεωρητικών συχνοτήτων. Η αξιολόγηση βασίστηκε στην εγγύτητα των σημείων προς τη διαγώνιο, στην έκταση της διασποράς τους και στην παρουσία συστηματικών μοτίβων απόκλισης σε διαφορετικά επίπεδα συχνοτήτων.

Ιδιαίτερη σημασία αποδόθηκε όχι μόνο στο μέγεθος των αποκλίσεων, αλλά και στο αν αυτές εμφάνιζαν οργανωμένη δομή, καθώς ένα συστηματικό μοτίβο υποδηλώνει ότι το μοντέλο αποτυγχάνει να αναπαραγάγει συγκεκριμένα χαρακτηριστικά της εμπειρικής κατανομής.

Κανονική κατανομή



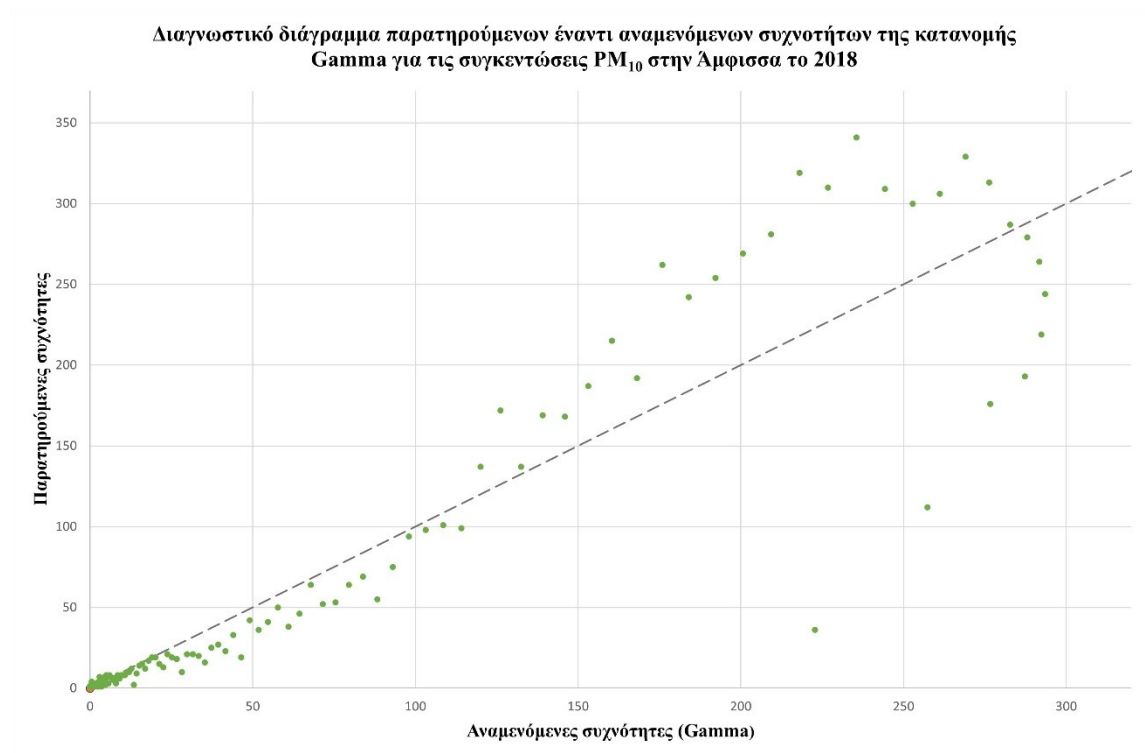
Σχήμα 4.4.2α .Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Αμφισσα το 2018.

Η κανονική κατανομή παρουσιάζει τη λιγότερο ικανοποιητική συνολική συμφωνία με τις παρατηρούμενες συχνότητες. Παρότι αρκετά από τα σημεία που αντιστοιχούν σε μικρές συχνότητες βρίσκονται σχετικά κοντά στην αρχή των αξόνων και στη γραμμή αναφοράς, η εικόνα μεταβάλλεται σημαντικά καθώς αυξάνονται οι αναμενόμενες συχνότητες.

Στο μεσαίο και υψηλότερο εύρος παρατηρείται έντονη διασπορά και σαφής μη γραμμική διάταξη των σημείων. Για παρόμοιες αναμενόμενες συχνότητες εμφανίζονται σημαντικά διαφορετικές παρατηρούμενες τιμές, ενώ τα σημεία βρίσκονται εναλλάξ πάνω και κάτω από τη διαγώνιο. Το μοτίβο αυτό δείχνει ότι η κανονική κατανομή δεν αναπαράγει με συνέπεια τη δομή των εμπειρικών συχνοτήτων και παρουσιάζει τόσο υποεκτιμήσεις όσο και υπερεκτιμήσεις σε διαφορετικά τμήματα του εύρους.

Η εικόνα συμφωνεί με τα αποτελέσματα της προηγούμενης υποενότητας, όπου η κανονική κατανομή εμφάνισε σημαντικές αποκλίσεις από τις παρατηρούμενες συχνότητες. Η συμμετρική μορφή της δεν φαίνεται να ανταποκρίνεται επαρκώς στη δεξιά ασυμμετρία των ωριαίων συγκεντρώσεων PM_{10} στην Αμφισσα.

Κατανομή Gamma



Σχήμα 4.4.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στην Αμφισσα το 2018.

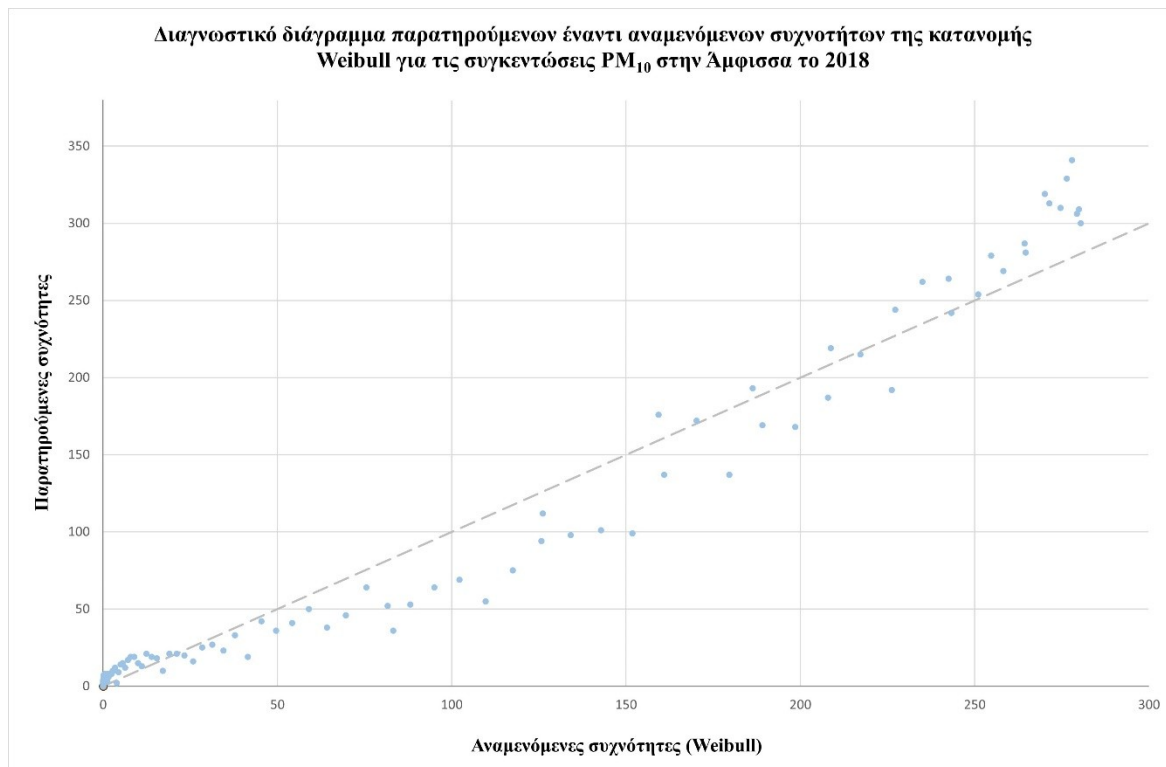
Η κατανομή Gamma παρουσιάζει βελτιωμένη εικόνα σε σχέση με την κανονική. Στο χαμηλό εύρος συχνοτήτων αρκετά σημεία βρίσκονται κοντά στη διαγώνια γραμμή, γεγονός που υποδηλώνει καλύτερη αντιστοιχία μεταξύ παρατηρούμενων και αναμενόμενων τιμών. Σχετικά ικανοποιητική είναι επίσης η συμφωνία σε μέρος του μεσαίου εύρους.

Ωστόσο, η εγγύτητα προς τη διαγώνιο δεν διατηρείται με την ίδια σταθερότητα σε ολόκληρο το διάγραμμα. Καθώς αυξάνονται οι αναμενόμενες συχνότητες, η διασπορά των σημείων γίνεται εντονότερη και εμφανίζονται σημαντικές αποκλίσεις και προς τις δύο πλευρές της γραμμής αναφοράς. Σε ορισμένες περιπτώσεις, παρατηρούνται πολύ διαφορετικές εμπειρικές συχνότητες για σχετικά παρόμοιες θεωρητικές τιμές.

Επομένως, η Gamma αποδίδει καλύτερα από την κανονική τη γενική σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων, αλλά η προσαρμογή της δεν είναι εξίσου σταθερή σε όλο το εύρος. Οι αποκλίσεις στις υψηλότερες συχνότητες δείχνουν ότι

το μοντέλο δεν αναπαράγει με την ίδια ακρίβεια όλα τα τμήματα της εμπειρικής κατανομής.

Κατανομή Weibull



Σχήμα 4.4.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στην Αμφισσα το 2018.

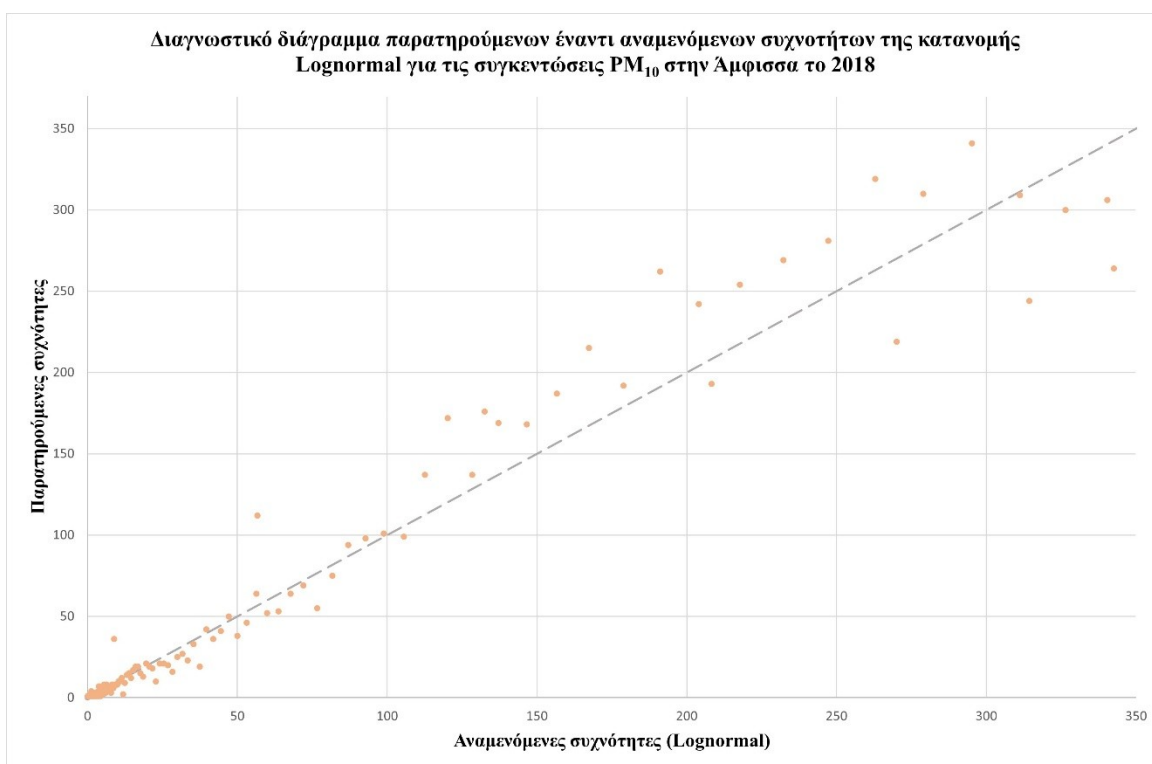
Η κατανομή Weibull εμφανίζει αρκετά συνεπή σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Στο χαμηλό εύρος τα περισσότερα σημεία βρίσκονται σχετικά κοντά στη διαγώνιο, ενώ και στις μεγαλύτερες συχνότητες η γενική τάση τους ακολουθεί σε σημαντικό βαθμό τη γραμμή αναφοράς. Συγκριτικά με την κανονική και την Gamma, η διάταξη των σημείων είναι περισσότερο οργανωμένη και προσεγγίζει καλύτερα μια γραμμική σχέση.

Παραμένουν, ωστόσο, ορισμένες συστηματικές αποκλίσεις. Σε μέρος του μεσαίου εύρους αρκετά σημεία βρίσκονται κάτω από τη διαγώνιο, γεγονός που αντιστοιχεί σε υπερεκτίμηση των συχνοτήτων από το μοντέλο. Αντίθετα, σε τμήμα των υψηλότερων

συχνοτήτων εμφανίζονται σημεία πάνω από τη γραμμή, υποδηλώνοντας υποεκτίμηση των παρατηρούμενων τιμών.

Παρά τις αποκλίσεις αυτές, η Weibull παρουσιάζει καλή συνολική εικόνα και αποδίδει με σχετική συνέπεια τη βασική τάση των δεδομένων. Με βάση το συγκεκριμένο διάγραμμα, συγκαταλέγεται στις ισχυρότερες υποψήφιες κατανομές, χωρίς όμως να επιτυγχάνει πλήρη συμφωνία σε όλα τα τμήματα του εύρους.

Λογαριθμοκανονική κατανομή



Σχήμα 4.4.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Αμφισσα το 2018.

Η λογαριθμοκανονική κατανομή παρουσιάζει επίσης καλή συνολική συμφωνία με τις παρατηρούμενες συχνότητες. Στο χαμηλό και σε σημαντικό μέρος του μεσαίου εύρους, τα σημεία ακολουθούν σχετικά στενά τη διαγώνια γραμμή, γεγονός που δείχνει ότι το μοντέλο αποδίδει αποτελεσματικά τη βασική τάση των εμπειρικών συχνοτήτων.

Στις υψηλότερες συχνότητες η διασπορά αυξάνεται και εμφανίζονται αποκλίσεις και προς τις δύο πλευρές της γραμμής αναφοράς. Σε ορισμένες περιπτώσεις οι παρατηρούμενες συχνότητες είναι μεγαλύτερες από τις αναμενόμενες, ενώ σε άλλες συμβαίνει το αντίθετο. Οι αποκλίσεις αυτές δεν σχηματίζουν τόσο έντονο και γενικευμένο μοτίβο αστοχίας όσο εκείνο που παρατηρείται στην κανονική κατανομή.

Η συνολική διάταξη των σημείων υποστηρίζει την καταλληλότητα της λογαριθμοκανονικής κατανομής για τη δεξιά ασύμμετρη μορφή των συγκεντρώσεων PM_{10} . Η απόδοσή της εμφανίζεται συγκρίσιμη με εκείνη της Weibull και, επομένως, η μεταξύ τους τελική κατάταξη δεν μπορεί να βασιστεί αποκλειστικά στην οπτική εξέταση του συγκεκριμένου διαγράμματος

Συγκριτική αξιολόγηση

Η συγκριτική εξέταση των τεσσάρων διαγνωστικών διαγραμμάτων αναδεικνύει σημαντικές διαφορές ως προς την ποιότητα της προσαρμογής. Η κανονική κατανομή παρουσιάζει τη λιγότερο ικανοποιητική εικόνα, καθώς χαρακτηρίζεται από έντονη μη γραμμικότητα, μεγάλη διασπορά και συστηματικές αποκλίσεις από τη γραμμή $y=x$. Τα ευρήματα αυτά ενισχύουν την εκτίμηση ότι το συμμετρικό αυτό μοντέλο δεν αποδίδει επαρκώς τη δομή των συγκεκριμένων δεδομένων.

Η Gamma εμφανίζει σαφή βελτίωση, ιδιαίτερα στο χαμηλό και σε μέρος του μεσαίου εύρους συχνοτήτων. Εντούτοις, η αυξημένη διασπορά και οι έντονες αποκλίσεις στις μεγαλύτερες αναμενόμενες τιμές περιορίζουν τη συνολική σταθερότητα της προσαρμογής της.

Οι κατανομές Weibull και λογαριθμοκανονική παρουσιάζουν την καλύτερη συνολική εικόνα. Και στις δύο περιπτώσεις η βασική τάση των σημείων ακολουθεί σε σημαντικό βαθμό τη γραμμή αναφοράς, ενώ οι αποκλίσεις είναι περισσότερο τοπικές σε σχέση με εκείνες της κανονικής και της Gamma. Η Weibull παρουσιάζει σχετικά συνεπή γραμμική διάταξη σε μεγάλο μέρος του εύρους, αλλά εμφανίζει ορισμένο συστηματικό μοτίβο αποκλίσεων στο μεσαίο και υψηλό τμήμα. Η λογαριθμοκανονική προσεγγίζει επίσης αποτελεσματικά τη συνολική τάση, με αυξημένη όμως διασπορά σε ορισμένες από τις μεγαλύτερες συχνότητες.

Με βάση αποκλειστικά τα συγκεκριμένα διαγνωστικά διαγράμματα, δεν προκύπτει απόλυτα σαφής υπεροχή της μίας από τις δύο κατανομές. Η Weibull και η

λογαριθμοκανονική αποτελούν τις ισχυρότερες υποψήφιες, η Gamma ακολουθεί, ενώ η κανονική παρουσιάζει τη λιγότερο ικανοποιητική προσαρμογή. Η διατύπωση αυτή είναι ασφαλέστερη από την πρόωρη επιλογή ενός μόνο μοντέλου βάσει της οπτικής εικόνας.

Η τελική κατάταξη των κατανομών προκύπτει από τη συνεκτίμηση των παραπάνω ευρημάτων με τα διαγράμματα υπολοίπων και τη συγκριτική τιμή του στατιστικού χ^2 . Με αυτόν τον τρόπο αξιολογούνται συμπληρωματικά η συνολική σχέση παρατηρούμενων και αναμενόμενων συχνοτήτων, η κατανομή των αποκλίσεων και η συμπεριφορά των μοντέλων στα ποσοστημόρια και στις ουρές της χρονοσειράς.

4.4.3 Διαγράμματα υπολοίπων

Η ποιότητα προσαρμογής των θεωρητικών κατανομών στις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας για το έτος 2018 διερευνήθηκε περαιτέρω μέσω των διαγραμμάτων υπολοίπων. Τα διαγράμματα αυτά αποτυπώνουν όχι μόνο το μέγεθος των αποκλίσεων μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων, αλλά και τον τρόπο με τον οποίο οι αποκλίσεις μεταβάλλονται κατά μήκος της ταξινομημένης σειράς των συγκεντρώσεων.

Για κάθε σημείο της ταξινομημένης σειράς, το απλό υπόλοιπο υπολογίστηκε ως:

$$R_i = O_i - E_i$$

όπου:

- R_i είναι το υπόλοιπο της i -οστής κλάσης,
- O_i η παρατηρούμενη συχνότητα,
- E_i η αναμενόμενη συχνότητα βάσει της αντίστοιχης θεωρητικής κατανομής.

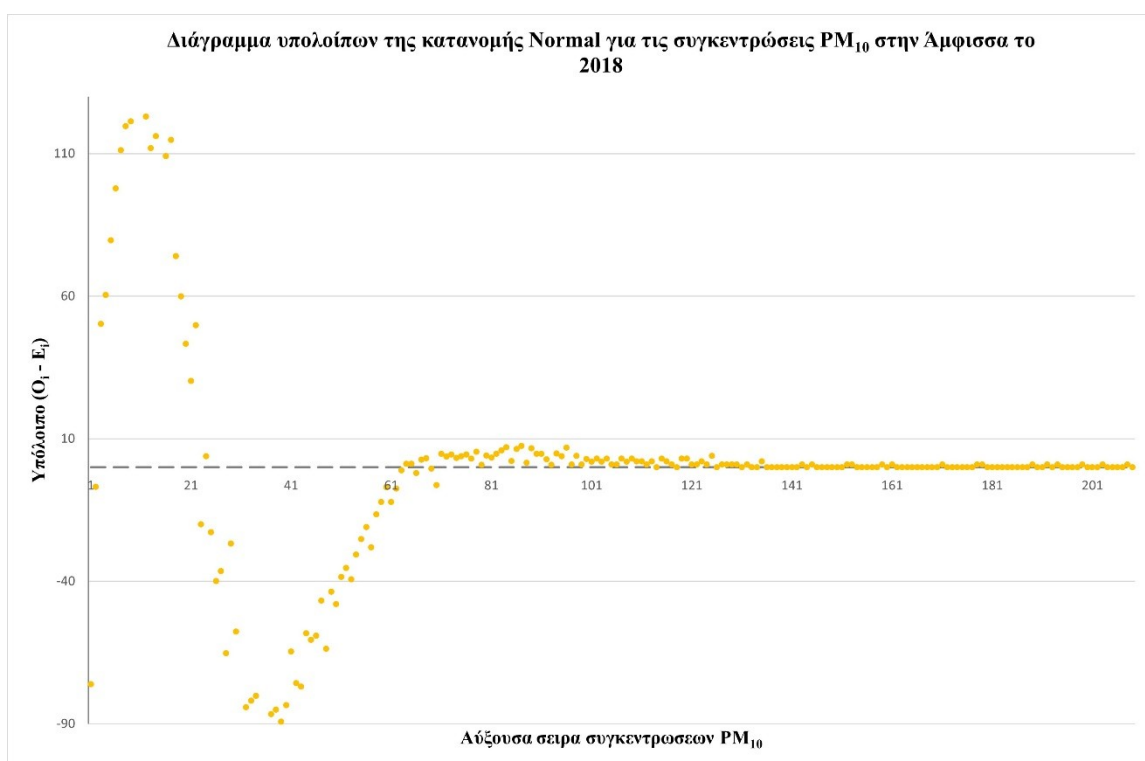
Θετικό υπόλοιπο σημαίνει ότι η παρατηρούμενη συχνότητα είναι μεγαλύτερη από την αναμενόμενη και, επομένως, ότι το θεωρητικό μοντέλο την υποεκτιμά. Αντίθετα, αρνητικό υπόλοιπο σημαίνει ότι η αναμενόμενη συχνότητα είναι μεγαλύτερη από την παρατηρούμενη και ότι το μοντέλο την υπερεκτιμά.

Σε μια ικανοποιητική προσαρμογή, τα υπόλοιπα αναμένεται να βρίσκονται κοντά στη γραμμή $R_i=0$, χωρίς εμφανείς παρατεταμένες ακολουθίες ίδιου προσήμου, έντονες καμπυλότητες ή άλλα συστηματικά πρότυπα. Η ύπαρξη μεγάλων και διαδοχικών θετικών

ή αρνητικών αποκλίσεων αποτελεί ένδειξη ότι το μοντέλο δεν αποδίδει ομοιόμορφα ορισμένα τμήματα της εμπειρικής κατανομής.

Πρέπει, ωστόσο, να επισημανθεί ότι στα παρόντα σχήματα απεικονίζονται απλά και όχι τυποποιημένα υπόλοιπα. Επομένως, το απόλυτο μέγεθός τους επηρεάζεται από το ύψος των αντίστοιχων συχνοτήτων. Για τον λόγο αυτό, η αξιολόγηση βασίζεται κυρίως στη συγκριτική συμπεριφορά των τεσσάρων μοντέλων, στην παρουσία συστηματικών προτύπων και στον βαθμό συγκέντρωσης των υπολοίπων γύρω από τη γραμμή μηδενός.

Κανονική κατανομή



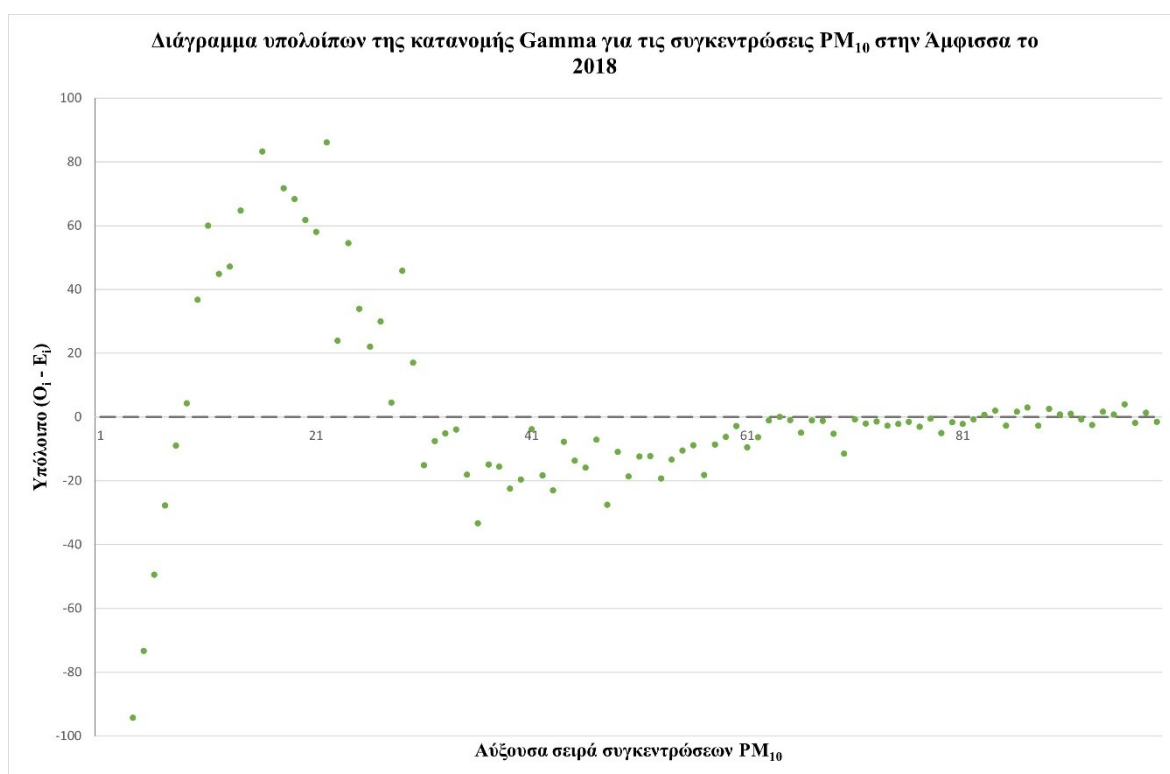
Σχήμα 4.4.3α. Διάγραμμα υπολοίπων (Residual plot) της κανονικής κατανομής (Normal) για τις ωριαίες συγκεντρώσεις PM_{10} της Αμφισσας κατά το έτος 2018.

Η κανονική κατανομή παρουσιάζει το εντονότερο και σαφέστερα συστηματικό πρότυπο υπολοίπων. Στο αρχικό τμήμα της ταξινομημένης σειράς εμφανίζονται πολύ υψηλές θετικές τιμές, οι οποίες υπερβαίνουν το +100. Το εύρημα αυτό δείχνει έντονη υποεκτίμηση των παρατηρούμενων συχνοτήτων από το μοντέλο στο συγκεκριμένο τμήμα των δεδομένων.

Στη συνέχεια, τα υπόλοιπα μεταβάλλονται απότομα προς αρνητικές τιμές και φθάνουν περίπου έως το -90 , γεγονός που αντιστοιχεί σε σημαντική υπερεκτίμηση των παρατηρούμενων συχνοτήτων. Η διαδοχική μετάβαση από μεγάλες θετικές σε μεγάλες αρνητικές αποκλίσεις σχηματίζει έντονη καμπυλότητα και δεν είναι συμβατή με ομοιόμορφη διασπορά γύρω από το μηδέν.

Στο τελευταίο και μεγαλύτερο τμήμα της σειράς, τα υπόλοιπα συγκλίνουν σταδιακά προς τη γραμμή αναφοράς και παραμένουν σχετικά μικρά. Η βελτίωση αυτή, όμως, δεν αναιρεί τις πολύ μεγάλες και δομημένες αποκλίσεις που προηγούνται. Συνεπώς, το διάγραμμα υπολοίπων επιβεβαιώνει ότι η κανονική κατανομή δεν αναπαράγει με συνέπεια τη δομή της συγκεκριμένης χρονοσειράς.

Κατανομή Gamma



Σχήμα 4.4.3β. Διάγραμμα υπολοίπων (Residual plot) της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018.

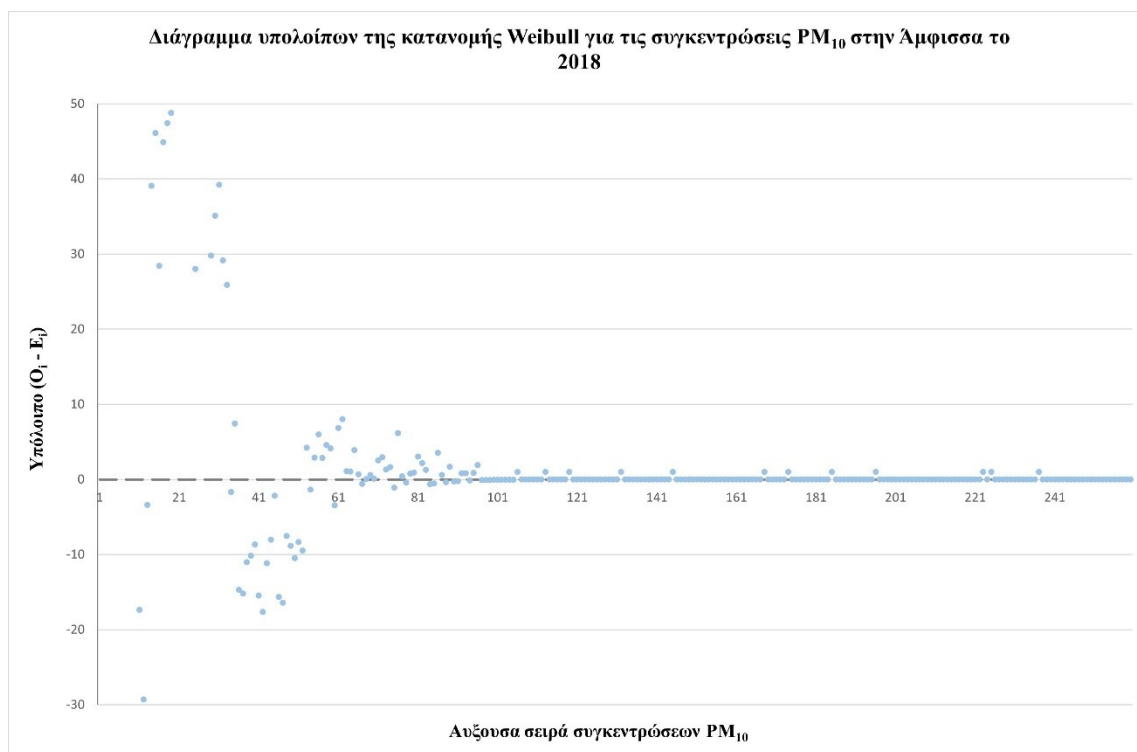
Η κατανομή Gamma παρουσιάζει σαφώς βελτιωμένη συμπεριφορά σε σύγκριση με την κανονική, χωρίς όμως να εξαλείφει πλήρως τα συστηματικά πρότυπα. Στο αρχικό τμήμα

εμφανίζονται τόσο έντονα αρνητικά υπόλοιπα, με τιμές που προσεγγίζουν το -95 , όσο και υψηλές θετικές αποκλίσεις που υπερβαίνουν το $+80$.

Μετά την αρχική αυτή περιοχή, τα υπόλοιπα μειώνονται και ακολουθεί μια εκτεταμένη ακολουθία κυρίως αρνητικών τιμών. Η συγκέντρωση πολλών διαδοχικών σημείων κάτω από τη γραμμή μηδενός υποδηλώνει ότι το μοντέλο υπερεκτιμά συστηματικά τις αντίστοιχες παρατηρούμενες συχνότητες σε σημαντικό τμήμα της ταξινομημένης σειράς.

Προς το ανώτερο τμήμα των συγκεντρώσεων, οι αποκλίσεις περιορίζονται αισθητά και τα σημεία συγκεντρώνονται πιο κοντά στο μηδέν. Παρά τη βελτίωση αυτή, η παρουσία εκτεταμένων ακολουθιών με ίδιο πρόσημο δείχνει ότι η προσαρμογή δεν είναι ομοιόμορφη σε όλο το εύρος. Η Gamma αποτελεί, επομένως, ουσιαστική βελτίωση έναντι της κανονικής κατανομής, αλλά διατηρεί εμφανείς τοπικές αστοχίες.

Κατανομή Weibull



Σχήμα 4.4.3γ. Διάγραμμα υπολοίπων (Residual plot) της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018.

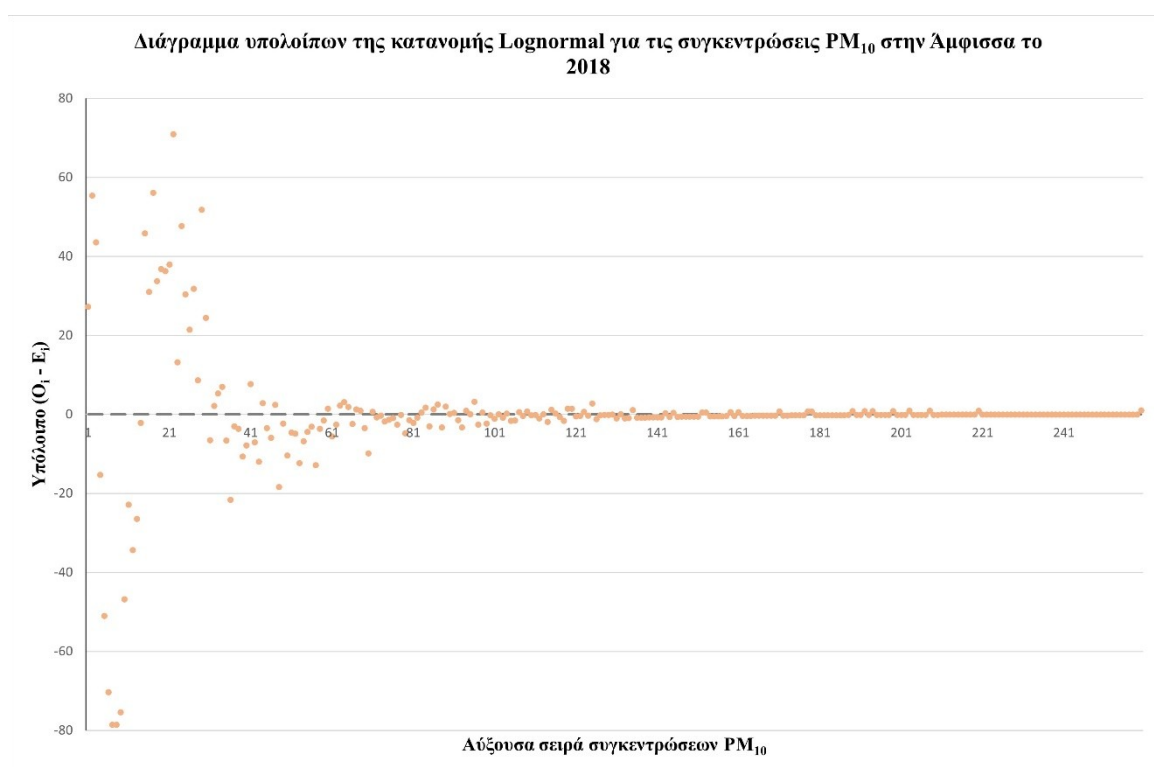
Η κατανομή Weibull εμφανίζει αισθητά μικρότερο εύρος υπολοίπων σε σχέση με την κανονική και την Gamma. Οι μεγαλύτερες αποκλίσεις περιορίζονται κυρίως στο αρχικό

τμήμα της ταξινομημένης σειράς, όπου καταγράφονται θετικά υπόλοιπα έως περίπου +49 και αρνητικά έως περίπου -30.

Στο επόμενο τμήμα παρατηρείται μια ακολουθία κυρίως αρνητικών τιμών, η οποία δείχνει ότι το μοντέλο υπερεκτιμά προσωρινά τις αντίστοιχες παρατηρούμενες συχνότητες. Ωστόσο, η ένταση των αποκλίσεων μειώνεται σχετικά γρήγορα και, μετά το μεσαίο τμήμα της σειράς, τα υπόλοιπα συγκεντρώνονται πολύ κοντά στη γραμμή μηδενός.

Η μακρά συγκέντρωση των σημείων γύρω από το μηδέν υποδηλώνει καλή συμφωνία σε μεγάλο μέρος του εύρους των δεδομένων. Παράλληλα, το αρχικό πρότυπο θετικών και αρνητικών αποκλίσεων δείχνει ότι η προσαρμογή δεν είναι πλήρως ομοιόμορφη. Συνολικά, όμως, η Weibull παρουσιάζει μία από τις πιο συμπαγείς και ευνοϊκές εικόνες υπολοίπων μεταξύ των τεσσάρων κατανομών.

Λογαριθμοκανονική κατανομή



Σχήμα 4.4.3δ. Διάγραμμα υπολοίπων (Residual plot) της λογαριθμοκανονικής κατανομής (Lognormal) για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018.

Η λογαριθμοκανονική κατανομή εμφανίζει επίσης σημαντική βελτίωση σε σχέση με την κανονική και την Gamma. Οι μεγαλύτερες αποκλίσεις εντοπίζονται στο αρχικό τμήμα της ταξινομημένης σειράς, όπου τα υπόλοιπα μεταβάλλονται περίπου από -79 έως $+71$. Η εναλλαγή αυτή δείχνει ότι το μοντέλο δεν αποδίδει με την ίδια ακρίβεια όλα τα σημεία του χαμηλότερου τμήματος.

Στη συνέχεια, τα υπόλοιπα περιορίζονται γρήγορα και κατανέμονται σε στενότερη ζώνη γύρω από το μηδέν. Από το μεσαίο τμήμα και έπειτα, η πλειονότητα των σημείων παρουσιάζει μικρές αποκλίσεις και δεν εμφανίζεται έντονη παρατεταμένη απομάκρυνση από τη γραμμή αναφοράς.

Η εικόνα αυτή δείχνει ότι η λογαριθμοκανονική κατανομή αναπαράγει αποτελεσματικά μεγάλο μέρος της εμπειρικής δομής, αλλά εμφανίζει αισθητές τοπικές αποκλίσεις στο αρχικό τμήμα. Συνεπώς, το διάγραμμα υπολοίπων υποστηρίζει την καταλληλότητά της ως ισχυρού υποψήφιου μοντέλου, χωρίς όμως να τεκμηριώνει από μόνο του σαφή υπεροχή έναντι της Weibull.

Συγκριτική αξιολόγηση

Η σύγκριση των διαγραμμάτων υπολοίπων επιβεβαιώνει ότι η κανονική κατανομή εμφανίζει τη λιγότερο ικανοποιητική προσαρμογή. Οι πολύ μεγάλες θετικές και αρνητικές αποκλίσεις, καθώς και η έντονη καμπυλότητα των υπολοίπων, αποτελούν σαφή ένδειξη συστηματικής ασυμφωνίας με τα εμπειρικά δεδομένα.

Η Gamma περιορίζει εν μέρει τις αποκλίσεις, αλλά διατηρεί μεγάλο εύρος υπολοίπων και εκτεταμένες ακολουθίες σημείων με ίδιο πρόσημο. Το πρότυπο αυτό δείχνει ότι, παρότι προσαρμόζεται καλύτερα από την κανονική, εξακολουθεί να υπερεκτιμά ή να υποεκτιμά συστηματικά ορισμένα τμήματα της κατανομής.

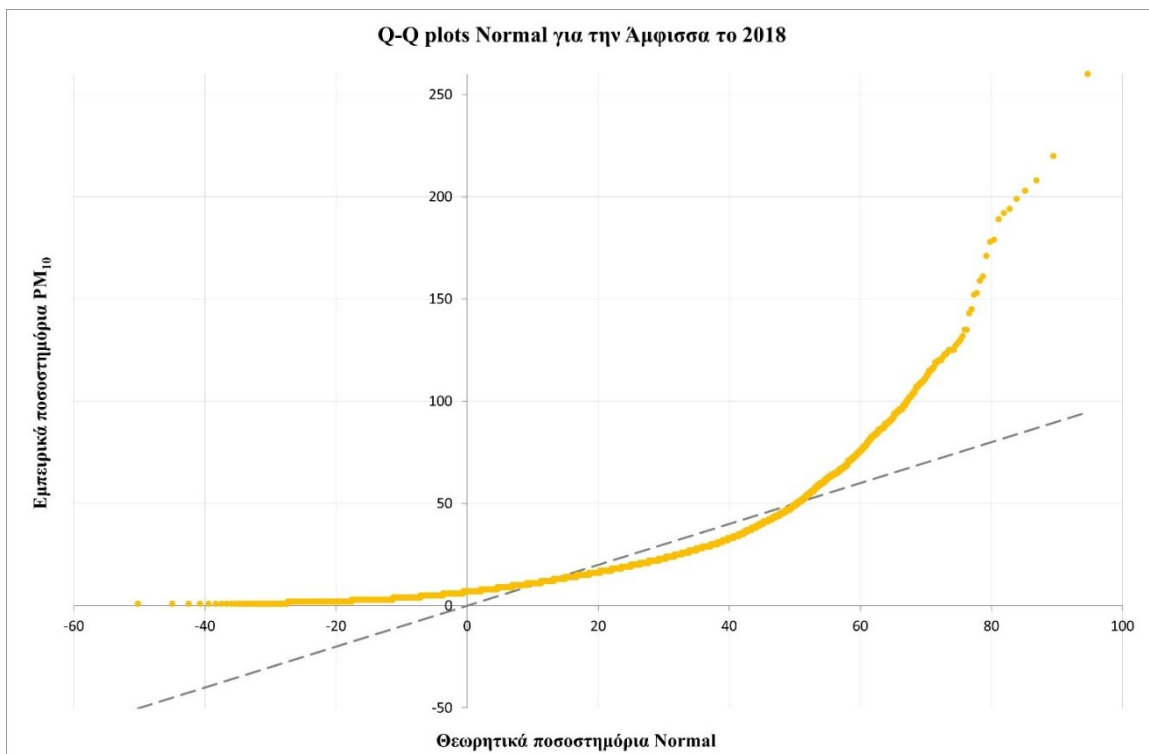
Η Weibull και η λογαριθμοκανονική παρουσιάζουν την καλύτερη συνολική συμπεριφορά. Και στις δύο περιπτώσεις, οι μεγάλες αποκλίσεις περιορίζονται κυρίως στο αρχικό τμήμα, ενώ στο μεγαλύτερο μέρος της ταξινομημένης σειράς τα υπόλοιπα συγκεντρώνονται κοντά στο μηδέν. Η Weibull εμφανίζει μικρότερο συνολικό εύρος ακραίων υπολοίπων και ταχύτερη συγκέντρωση γύρω από τη γραμμή αναφοράς. Η λογαριθμοκανονική παρουσιάζει μεγαλύτερες αρχικές αποκλίσεις, αλλά επίσης γρήγορη σταθεροποίηση και μικρά υπόλοιπα στο μεγαλύτερο τμήμα των δεδομένων.

Με βάση αποκλειστικά τα συγκεκριμένα διαγράμματα, η Weibull παρουσιάζει ελαφρώς ευνοϊκότερη γραφική εικόνα ως προς το εύρος και τη συγκέντρωση των απλών υπολοίπων. Η λογαριθμοκανονική, ωστόσο, παραμένει συγκρίσιμη και εμφανίζει επίσης καλή συμπεριφορά στο μεγαλύτερο μέρος της ταξινομημένης σειράς. Επομένως, δεν είναι επιστημονικά ασφαλές να υποστηριχθεί στο παρόν στάδιο σαφής και οριστική υπεροχή της μίας κατανομής έναντι της άλλης.

Η τελική κατάταξη των δύο κατανομών πρέπει να βασιστεί στη συνεκτίμηση των διαγραμμάτων υπολοίπων με τα υπόλοιπα γραφικά διαγνωστικά και τη συγκριτική τιμή του στατιστικού χ^2 . Με αυτόν τον τρόπο, η επιλογή δεν εξαρτάται αποκλειστικά από το μέγιστο απλό υπόλοιπο, αλλά από τη συνολική κατανομή, τη μορφή και τη σχετική βαρύτητα των αποκλίσεων σε ολόκληρο το εύρος των δεδομένων.

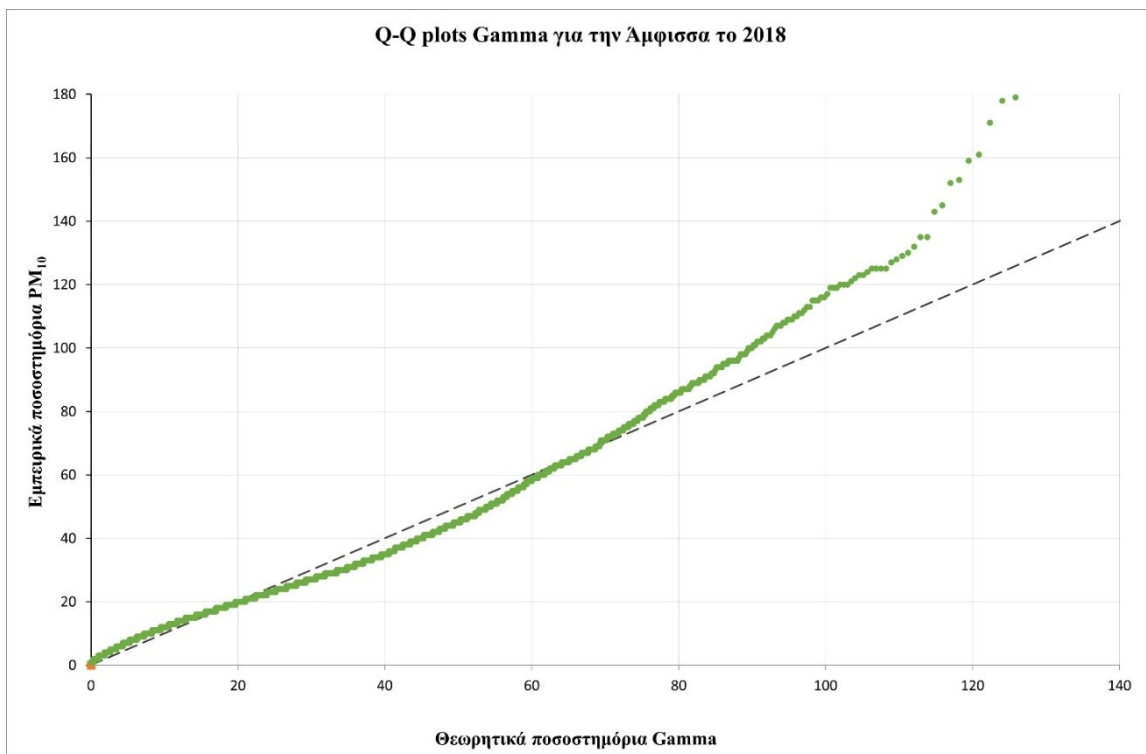
4.4.4 Διαγράμματα Q-Q

Τα διαγράμματα Q-Q χρησιμοποιήθηκαν ως βασικό γραφικό εργαλείο για τη διερεύνηση της συμφωνίας μεταξύ των εμπειρικών ποσοστημορίων των ωριαίων συγκεντρώσεων PM_{10} στην Άμφισσα και των αντίστοιχων θεωρητικών ποσοστημορίων των κατανομών Normal, Gamma, Weibull και Lognormal. Η διαγώνια γραμμή $y=x$ αντιπροσωπεύει την ιδανική περίπτωση πλήρους ταύτισης μεταξύ εμπειρικών και θεωρητικών τιμών. Επομένως, η εγγύτητα των σημείων προς τη γραμμή αυτή αποτελεί ένδειξη καλής προσαρμογής, ενώ η συστηματική καμπυλότητα, η μεταβολή της κλίσης ή η απόκλιση στα άκρα υποδηλώνουν ότι το θεωρητικό μοντέλο δεν αναπαράγει πλήρως τη δομή και τη μεταβλητότητα των δεδομένων. Ιδιαίτερη βαρύτητα δόθηκε στην άνω ουρά της κατανομής, επειδή εκεί αποτυπώνονται τα σπανιότερα αλλά περιβαλλοντικά σημαντικότερα επεισόδια αυξημένων συγκεντρώσεων, τα οποία είναι κρίσιμα για την αξιολόγηση της ποιότητας της προσαρμογής. Η ερμηνεία των διαγραμμάτων δεν βασίστηκε μόνο στην οπτική εγγύτητα των σημείων προς τη διαγώνιο, αλλά και στο αν οι αποκλίσεις ήταν τυχαίες ή παρουσίαζαν συστηματική μορφή σε συγκεκριμένες περιοχές της κατανομής.



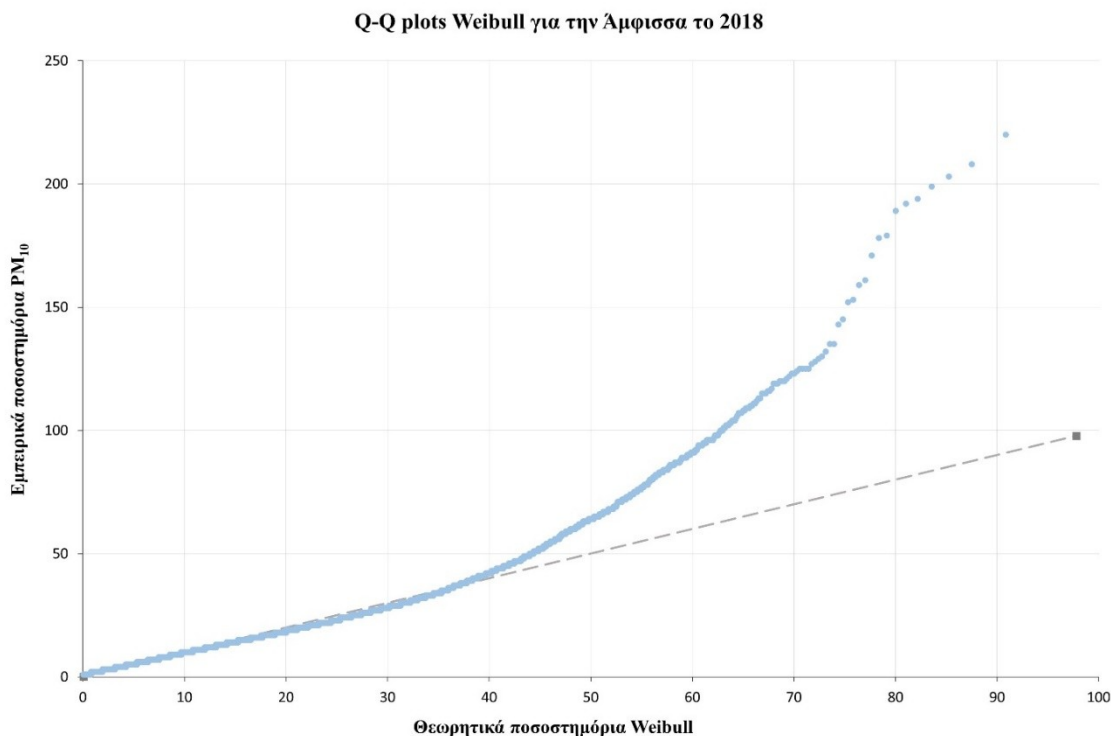
Σχήμα 4.4.4α. Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.

Η Normal εμφανίζει τη σαφέστερη ασυμφωνία με την εμπειρική κατανομή. Η απόκλιση δεν περιορίζεται σε μία μόνο περιοχή του διαγράμματος, αλλά αναπτύσσεται συστηματικά από τα χαμηλά έως τα υψηλά ποσοστημόρια. Στο κατώτερο άκρο, η εμφάνιση αρνητικών θεωρητικών τιμών είναι ασύμβατη με τη μη αρνητική φύση των συγκεντρώσεων PM_{10} , ενώ στο ανώτερο τμήμα τα εμπειρικά ποσοστημόρια αυξάνονται πολύ ταχύτερα από τα θεωρητικά. Η έντονη και συνεχής καμπυλότητα των σημείων αποτυπώνει την αδυναμία της συμμετρικής Normal να περιγράψει τόσο τη θετική ασυμμετρία όσο και τη βαριά δεξιά ουρά της χρονοσειράς. Παράλληλα, η απομάκρυνση από τη διαγώνιο δείχνει ότι η απόκλιση δεν οφείλεται μόνο σε λίγες ακραίες παρατηρήσεις, αλλά σε συνολικότερη ασυμφωνία μεταξύ της θεωρητικής και της εμπειρικής μορφής. Συνεπώς, η κατανομή αυτή δεν αποτελεί κατάλληλη επιλογή για τα δεδομένα της Άμφισσας, καθώς δεν αποδίδει ικανοποιητικά ούτε το κύριο σώμα των παρατηρήσεων ούτε τη συμπεριφορά των υψηλών συγκεντρώσεων.



Σχήμα 4.4.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM₁₀ στην Άμφισσα το 2018.

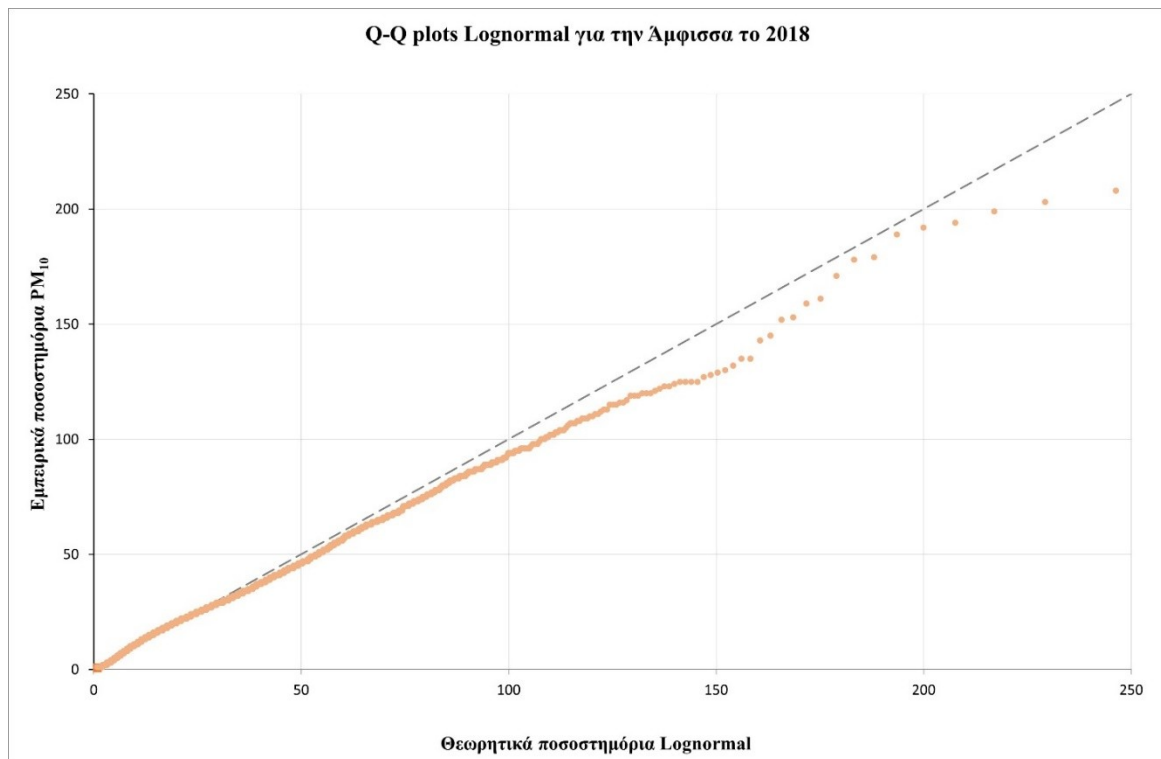
Η Gamma παρουσιάζει αισθητά βελτιωμένη συμφωνία σε σχέση με τη Normal. Στο χαμηλό και στο μεγαλύτερο μέρος του μεσαίου εύρους, τα σημεία ακολουθούν σχετικά ομαλά τη διαγώνιο, γεγονός που δείχνει ότι η κατανομή αποδίδει ικανοποιητικά το κύριο σώμα των παρατηρήσεων. Ωστόσο, όσο αυξάνονται τα ποσοστημόρια, αναπτύσσεται σταδιακή και συστηματική απόκλιση πάνω από τη γραμμή αναφοράς. Η συμπεριφορά αυτή δείχνει ότι η Gamma υποεκτιμά το μέγεθος των υψηλότερων εμπειρικών συγκεντρώσεων και δεν αναπαράγει πλήρως την άνω ουρά. Η απόκλιση δεν περιορίζεται σε ένα μόνο ακραίο σημείο, αλλά εξελίσσεται προοδευτικά, γεγονός που υποδηλώνει ότι η θεωρητική μορφή της κατανομής δεν ακολουθεί πλήρως τον ρυθμό αύξησης των μεγάλων παρατηρήσεων. Παρά τη βελτιωμένη απόδοσή της στο κεντρικό τμήμα, η αδυναμία αυτή μειώνει την καταλληλότητά της για την περιγραφή επεισοδίων υψηλής συγκέντρωσης. Επομένως, η Gamma αποτελεί εύλογη και σαφώς καλύτερη επιλογή από τη Normal, όχι όμως τη βέλτιστη προσέγγιση για τη συγκεκριμένη χρονοσειρά.



Σχήμα 4.4.4γ. Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.

Η Weibull αποδίδει ικανοποιητικά μόνο ένα περιορισμένο τμήμα της εμπειρικής κατανομής. Στα χαμηλά ποσοστημόρια τα σημεία βρίσκονται σχετικά κοντά στη διαγώνιο, όμως η συμφωνία επιδεινώνεται σταδιακά καθώς αυξάνονται οι συγκεντρώσεις. Από το μεσαίο τμήμα και έπειτα, τα εμπειρικά ποσοστημόρια αποκλίνουν έντονα προς τα πάνω, γεγονός που υποδηλώνει συστηματική υποεκτίμηση των υψηλών τιμών από το θεωρητικό μοντέλο. Η απόκλιση αυτή δεν περιορίζεται σε μεμονωμένα ακραία σημεία, αλλά εκφράζει διαφορετική συνολική καμπυλότητα μεταξύ θεωρητικής και εμπειρικής κατανομής. Παράλληλα, η αυξανόμενη απομάκρυνση από τη γραμμή αναφοράς δείχνει ότι η Weibull δεν ακολουθεί επαρκώς τον ρυθμό αύξησης των μεγαλύτερων παρατηρήσεων και δεν αναπαράγει με ακρίβεια τη δομή της άνω ουράς. Η αδυναμία αυτή είναι ιδιαίτερα σημαντική, επειδή τα υψηλά ποσοστημόρια αντιστοιχούν στα σπανιότερα αλλά περιβαλλοντικά κρίσιμότερα επεισόδια αυξημένων συγκεντρώσεων. Για τον λόγο

αυτό, η Weibull δεν περιγράφει επαρκώς τη συνολική δομή των δεδομένων της Άμφισσας και δεν μπορεί να θεωρηθεί κατάλληλη τελική επιλογή για τη συγκεκριμένη χρονοσειρά.



Σχήμα 4.4.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα το 2018.

Η Lognormal εμφανίζει την πιο συνεκτική γραφική συμπεριφορά μεταξύ των τεσσάρων υποψήφιων κατανομών. Τα σημεία ακολουθούν τη διαγώνιο με σχετικά σταθερή εγγύτητα στο χαμηλό και στο μεγαλύτερο μέρος του μεσαίου εύρους, γεγονός που δείχνει ότι η κατανομή αποδίδει αποτελεσματικά το κύριο σώμα της χρονοσειράς και τη θετική ασυμμετρία των συγκεντρώσεων. Στα υψηλότερα ποσοστημόρια παρατηρούνται αποκλίσεις, οι οποίες φανερώνουν ότι ακόμη και η Lognormal δεν αναπαράγει πλήρως τα σπανιότερα επεισόδια πολύ υψηλών συγκεντρώσεων. Ωστόσο, οι αποκλίσεις αυτές εμφανίζονται κυρίως στην άνω ουρά και είναι συγκριτικά ηπιότερες από εκείνες των υπόλοιπων κατανομών. Παράλληλα, η συνολική απόσταση των σημείων από τη γραμμή αναφοράς και η ένταση της καμπυλότητας είναι μικρότερες από εκείνες που εμφανίζονται στις Normal, Gamma και Weibull. Η συμπεριφορά αυτή υποδηλώνει ότι η Lognormal αναπαράγει με μεγαλύτερη συνέπεια τόσο τη μορφή του κεντρικού τμήματος όσο και τη σταδιακή μετάβαση προς τις υψηλότερες συγκεντρώσεις. Συνεπώς, προσεγγίζει

πληρέστερα τη συνολική μορφή της εμπειρικής κατανομής και αποτελεί την καταλληλότερη από τις εξεταζόμενες επιλογές για την Άμφισσα το 2018.

Συγκριτική αποτίμηση

Η συγκριτική ανάγνωση των τεσσάρων Q-Q plots αναδεικνύει σαφή διαφοροποίηση στην ποιότητα της προσαρμογής. Η Normal παρουσιάζει τη μεγαλύτερη και συστηματικότερη ασυμφωνία, καθώς δεν μπορεί να περιγράψει επαρκώς τη μη αρνητικότητα, τη θετική ασυμμετρία και την ένταση της δεξιάς ουράς. Η Weibull εμφανίζει περιορισμένη συμφωνία στις χαμηλότερες τιμές, αλλά αναπτύσσει έντονη και προοδευτική απόκλιση όσο αυξάνονται τα ποσοστημόρια. Η Gamma προσεγγίζει ικανοποιητικότερα το κύριο σώμα των δεδομένων, χωρίς όμως να αποδίδει πλήρως τον ρυθμό αύξησης των υψηλών συγκεντρώσεων. Αντίθετα, η Lognormal παρουσιάζει την πιο ισορροπημένη και σταθερή συνολική συμπεριφορά, συνδυάζοντας καλύτερη ευθυγράμμιση στο κεντρικό τμήμα με συγκριτικά μικρότερες αποκλίσεις στην άνω ουρά. Η υπεροχή της δεν βασίζεται στην απουσία αποκλίσεων, αλλά στο ότι αυτές είναι λιγότερο έντονες και λιγότερο συστηματικές σε σύγκριση με τις υπόλοιπες κατανομές. Η γραφική αυτή κατάταξη συμφωνεί με την αριθμητική αξιολόγηση βάσει του χ^2 και ενισχύει την επιλογή της Lognormal ως της καταλληλότερης από τις εξεταζόμενες κατανομές για τη χρονοσειρά της Άμφισσας το 2018.

4.4.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2

Η γραφική αξιολόγηση των προηγούμενων υποενοτήτων ανέδειξε τα σημεία στα οποία κάθε θεωρητική κατανομή προσεγγίζει ή αποκλίνει από τις εμπειρικές συχνότητες. Για την ποσοτική αποτίμηση της συνολικής απόκλισης μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων χρησιμοποιήθηκε το στατιστικό χ^2 , σύμφωνα με τη διαδικασία που παρουσιάστηκε στο Κεφάλαιο 3.

Το στατιστικό χ^2 υπολογίζεται από τη σχέση:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

όπου:

- O_i είναι η παρατηρούμενη συχνότητα της i -οστής κλάσης,

- E_i είναι η αντίστοιχη αναμενόμενη συχνότητα βάσει της θεωρητικής κατανομής,
- k είναι ο αριθμός των κλάσεων που περιλαμβάνονται στον υπολογισμό.

Μικρότερη τιμή του χ^2 αντιστοιχεί σε μικρότερη συνολική σταθμισμένη απόκλιση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων και, κατά συνέπεια, σε καλύτερη σχετική αριθμητική προσαρμογή. Ο όρος «σχετική προσαρμογή» είναι σημαντικός, καθώς στην παρούσα εργασία το χ^2 χρησιμοποιείται κυρίως για τη συγκριτική κατάταξη των τεσσάρων θεωρητικών μοντέλων και όχι ως αυτοτελές κριτήριο απόλυτης αποδοχής ή απόρριψής τους. Για τη διασφάλιση της συγκρισιμότητας, οι τιμές του χ^2 υπολογίστηκαν με κοινό σύστημα κλάσεων για όλες τις εξεταζόμενες κατανομές. Παράλληλα, η εφαρμογή του κριτηρίου προϋποθέτει ότι οι αναμενόμενες συχνότητες των κλάσεων είναι επαρκείς, ώστε η τελική τιμή να μην επηρεάζεται δυσανάλογα από κλάσεις με πολύ μικρές θεωρητικές συχνότητες. Η διαδικασία σχηματισμού των κλάσεων και οι σχετικές προϋποθέσεις εφαρμογής έχουν παρουσιαστεί στη Μεθοδολογία.

Το χ^2 δεν αποτυπώνει όλες τις επιμέρους πτυχές της προσαρμογής, καθώς συμπυκνώνει τις αποκλίσεις σε μία συνολική αριθμητική τιμή χωρίς να δείχνει άμεσα σε ποιο τμήμα της κατανομής εντοπίζονται. Δύο μοντέλα μπορούν, επομένως, να εμφανίζουν σχετικά κοντινές τιμές χ^2 , αλλά να διαφέρουν ως προς την απόδοσή τους στην εμπειρική κορυφή, στο κεντρικό τμήμα ή στη δεξιά ουρά. Για τον λόγο αυτό, οι αριθμητικές τιμές συνεκτιμώνται με τα διαγράμματα σύγκρισης συχνοτήτων, τα διαγνωστικά διαγράμματα και τα διαγράμματα υπολοίπων.

Κατανομή	χ^2	Σειρά Κατάταξης
Lognormal	96,26	1
Gamma	466,61	2
Weibull	476,42	3
Normal	1634,93	4

Πίνακας 4.4.5.1 Τιμές του στατιστικού χ^2 και κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στην Άμφισσα κατά το έτος 2018.

Όπως προκύπτει από τον Πίνακα 4.4.4.1, η λογαριθμοκανονική κατανομή παρουσιάζει τη μικρότερη τιμή χ^2 , ίση με 96,26. Η τιμή αυτή είναι αισθητά χαμηλότερη από τις αντίστοιχες τιμές των υπόλοιπων μοντέλων και δείχνει ότι, με βάση το συγκεκριμένο σύστημα κλάσεων και τον εφαρμοζόμενο τρόπο υπολογισμού, η λογαριθμοκανονική κατανομή επιτυγχάνει τη μικρότερη συνολική σταθμισμένη απόκλιση από τις παρατηρούμενες συχνότητες.

Η κατανομή Gamma κατατάσσεται δεύτερη, με $\chi^2=466,61$, ενώ η Weibull κατατάσσεται τρίτη, με $\chi^2=476,42$. Η μεταξύ τους διαφορά ανέρχεται σε 9,81 μονάδες, δηλαδή περίπου στο 2% της τιμής της Gamma. Η μικρή αυτή απόσταση δείχνει ότι οι δύο κατανομές παρουσιάζουν παρόμοια συνολική αριθμητική απόδοση. Επομένως, η αριθμητική υπεροχή της Gamma έναντι της Weibull είναι περιορισμένη και δεν υποδηλώνει έντονη ουσιαστική διαφοροποίηση μεταξύ των δύο μοντέλων.

Αντίθετα, η κανονική κατανομή εμφανίζει τιμή $\chi^2=1.634,93$ και κατατάσσεται τελευταία. Η τιμή αυτή είναι πολλαπλάσια των αντίστοιχων τιμών των ασύμμετρων κατανομών και ποσοτικοποιεί τις συστηματικές αποκλίσεις που είχαν ήδη αναδειχθεί μέσω της γραφικής αξιολόγησης. Η κανονική κατανομή δεν αποδίδει ικανοποιητικά ούτε την υψηλή συγκέντρωση παρατηρήσεων στις χαμηλές και μεσαίες τιμές ούτε τη σταδιακή μείωση των συχνοτήτων προς τη δεξιά ουρά.

Η διαφορά μεταξύ της λογαριθμοκανονικής και των υπόλοιπων μοντέλων είναι ιδιαίτερα έντονη. Η τιμή χ^2 της λογαριθμοκανονικής είναι περίπου 79% χαμηλότερη από εκείνες της Gamma και της Weibull και περίπου 94% χαμηλότερη από την τιμή της κανονικής κατανομής. Οι διαφορές αυτές δείχνουν ότι το αριθμητικό πλεονέκτημα της λογαριθμοκανονικής δεν είναι οριακό, αλλά ιδιαίτερα ισχυρό στο πλαίσιο του συγκεκριμένου κριτηρίου και της συγκεκριμένης ομαδοποίησης των δεδομένων.

Η χαμηλότερη τιμή χ^2 , ωστόσο, δεν σημαίνει ότι η λογαριθμοκανονική κατανομή αναπαράγει τέλεια κάθε επιμέρους τμήμα της εμπειρικής κατανομής. Το στατιστικό βασίζεται στις τετραγωνισμένες αποκλίσεις και σταθμίζει καθεμία από αυτές με βάση την αντίστοιχη αναμενόμενη συχνότητα. Κατά συνέπεια, μία σημαντική απόκλιση σε κλάση με σχετικά χαμηλή αναμενόμενη συχνότητα μπορεί να επηρεάσει αισθητά την τελική τιμή. Επιπλέον, επειδή οι αποκλίσεις υψώνονται στο τετράγωνο, οι θετικές και οι αρνητικές διαφορές δεν αλληλοαναιρούνται.

Το χαρακτηριστικό αυτό εξηγεί γιατί η κατάταξη με βάση το χ^2 δεν ταυτίζεται αναγκαστικά με την οπτική εικόνα των απλών υπολοίπων. Στην προηγούμενη υποενότητα, η Weibull παρουσίασε μικρότερο συνολικό εύρος απλών υπολοίπων και σχετικά γρήγορη συγκέντρωση των σημείων γύρω από το μηδέν, ενώ η λογαριθμοκανονική εμφάνισε εντονότερες αποκλίσεις στο αρχικό τμήμα της ταξινομημένης σειράς. Τα απλά υπόλοιπα, όμως, αποτυπώνουν το πρόσημο και το απόλυτο μέγεθος κάθε επιμέρους διαφοράς χωρίς τη στάθμιση που εφαρμόζεται στον υπολογισμό του χ^2 . Τα δύο εργαλεία εξετάζουν, επομένως, διαφορετικές αλλά συμπληρωματικές πτυχές της προσαρμογής.

Η αριθμητική κατάταξη παραμένει συνεπής με το βασικό συμπέρασμα της προηγούμενης γραφικής αξιολόγησης: η κανονική κατανομή παρουσιάζει τη χαμηλότερη καταλληλότητα, ενώ οι ασύμμετρες κατανομές που ορίζονται για θετικές τιμές αποδίδουν αποτελεσματικότερα τη μορφή των συγκεντρώσεων PM_{10} . Η λογαριθμοκανονική αποκτά μέσω του χ^2 σαφές αριθμητικό πλεονέκτημα, ενώ η Gamma και η Weibull σχηματίζουν μία ενδιάμεση ομάδα με αρκετά κοντινή συνολική απόδοση.

Με βάση αποκλειστικά το αριθμητικό κριτήριο, η σχετική κατάταξη των κατανομών είναι:

$$\text{Λογαριθμοκανονική} < \text{Gamma} < \text{Weibull} < \text{Κανονική}$$

όπου η μικρότερη τιμή χ^2 αντιστοιχεί σε καλύτερη σχετική αριθμητική προσαρμογή.

Η παραπάνω σειρά δεν πρέπει να ερμηνεύεται ως απόδειξη ότι οι κατανομές γίνονται στατιστικά αποδεκτές ή απορρίπτονται με την τυπική έννοια ενός ελέγχου υπόθεσης. Για ένα τέτοιο συμπέρασμα θα έπρεπε να προσδιοριστούν και να παρουσιαστούν οι βαθμοί ελευθερίας, η αντίστοιχη κρίσιμη τιμή και το p-value, λαμβάνοντας υπόψη τον αριθμό των παραμέτρων που εκτιμήθηκαν από τα ίδια δεδομένα. Στην παρούσα εργασία, οι τιμές του χ^2 χρησιμοποιούνται ως κοινός δείκτης συγκριτικής αριθμητικής απόδοσης.

Συνολικά, η λογαριθμοκανονική κατανομή συγκεντρώνει την ισχυρότερη αριθμητική ένδειξη καταλληλότητας για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018. Η Gamma και η Weibull παρουσιάζουν σημαντικά μεγαλύτερες και αρκετά κοντινές μεταξύ τους τιμές χ^2 , ενώ η κανονική κατανομή υπολείπεται αισθητά. Το αριθμητικό αποτέλεσμα δεν υποκαθιστά τη γραφική διάγνωση, αλλά τη συμπληρώνει, παρέχοντας ένα κοινό μέτρο της συνολικής σταθμισμένης απόκλισης.

Η τελική επιλογή της καταλληλότερης θεωρητικής κατανομής προκύπτει στην επόμενη υποενότητα από τη σύνθεση όλων των διαθέσιμων γραφικών και αριθμητικών ενδείξεων.

4.4.6 Επιλογή καταλληλότερης κατανομής για την Άμφισσα 2018

Η συνδυαστική αξιολόγηση των αποτελεσμάτων δείχνει ότι η λογαριθμοκανονική κατανομή αποτελεί την καταλληλότερη από τις τέσσερις εξεταζόμενες θεωρητικές κατανομές για την περιγραφή των ωριαίων συγκεντρώσεων PM_{10} στην Άμφισσα κατά το έτος 2018. Η επιλογή αυτή δεν βασίζεται σε ένα μεμονωμένο εύρημα, αλλά στη συνεκτίμηση της σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων, των διαγνωστικών διαγραμμάτων, των διαγραμμάτων υπολοίπων και της αριθμητικής αξιολόγησης μέσω του στατιστικού χ^2 .

Στη σύγκριση των συχνοτήτων ανά κλάση, η λογαριθμοκανονική κατανομή παρουσίασε την πιο ισορροπημένη συμφωνία με την εμπειρική κατανομή. Οι αναμενόμενες συχνότητες της προσέγγισαν ικανοποιητικά τις παρατηρούμενες τιμές στις βασικές κλάσεις 0-10, 10-20 και 20-30 $\mu\text{g}/\text{m}^3$, στις οποίες συγκεντρώνεται το μεγαλύτερο μέρος των παρατηρήσεων. Παράλληλα, απέδωσε σχετικά αποτελεσματικά τη σταδιακή μείωση των συχνοτήτων προς τις υψηλότερες συγκεντρώσεις, χωρίς να εξαλείφει πλήρως τις αποκλίσεις στη δεξιά ουρά.

Τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων έδειξαν ότι η λογαριθμοκανονική και η Weibull αποτέλεσαν τις ισχυρότερες υποψήφιες κατανομές. Και στα δύο μοντέλα η βασική τάση των σημείων ακολούθησε σε σημαντικό βαθμό τη γραμμή $y=x$, παρότι παρατηρήθηκαν τοπικές αποκλίσεις στις μεγαλύτερες συχνότητες. Επομένως, από το συγκεκριμένο γραφικό κριτήριο δεν προέκυψε απόλυτα σαφής υπεροχή της λογαριθμοκανονικής έναντι της Weibull.

Παρόμοια εικόνα προέκυψε και από τα διαγράμματα υπολοίπων. Η Weibull εμφάνισε σχετικά περιορισμένο συνολικό εύρος αποκλίσεων και συγκέντρωση των υπολοίπων κοντά στο μηδέν σε μεγάλο μέρος της ταξινομημένης σειράς. Η λογαριθμοκανονική παρουσίασε μεγαλύτερες αποκλίσεις στο αρχικό τμήμα, αλλά στη συνέχεια τα υπόλοιπα της περιορίστηκαν επίσης σε στενή ζώνη γύρω από τη γραμμή αναφοράς. Τα αποτελέσματα αυτά δείχνουν ότι και οι δύο κατανομές περιγράφουν αποτελεσματικά

σημαντικό μέρος της εμπειρικής δομής, χωρίς τα υπόλοιπα να τεκμηριώνουν από μόνα τους σαφή υπεροχή ενός μοντέλου.

Η αριθμητική αξιολόγηση διαφοροποίησε σαφέστερα τις εναλλακτικές κατανομές. Η λογαριθμοκανονική κατέγραψε τη μικρότερη τιμή χ^2 , ίση με 96,26, αισθητά χαμηλότερη από τις τιμές της Gamma, 466,61, και της Weibull, 476,42. Η διαφορά αυτή δείχνει ότι, υπό το κοινό σύστημα κλάσεων που χρησιμοποιήθηκε, η λογαριθμοκανονική επιτυγχάνει τη μικρότερη συνολική σταθμισμένη απόκλιση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Το αριθμητικό της πλεονέκτημα δεν είναι οριακό, αλλά ιδιαίτερα ισχυρό στο πλαίσιο του συγκεκριμένου κριτηρίου.

Η Gamma και η Weibull συγκροτούν μια ενδιάμεση ομάδα προσαρμογής. Η Gamma παρουσίασε ελαφρώς χαμηλότερη τιμή χ^2 , ενώ η Weibull εμφάνισε ευνοϊκότερη εικόνα σε μέρος των διαγραμμάτων υπολοίπων. Η διαφοροποίηση αυτή επιβεβαιώνει ότι κάθε διαγνωστικό εργαλείο αξιολογεί διαφορετική πτυχή της προσαρμογής. Καμία από τις δύο κατανομές, ωστόσο, δεν συνδύασε τη συνολική γραφική επάρκεια με αριθμητική επίδοση αντίστοιχη εκείνης της λογαριθμοκανονικής.

Η κανονική κατανομή παρουσίασε τη χαμηλότερη συνολική καταλληλότητα. Οι μεγάλες αποκλίσεις στις βασικές κλάσεις συχνοτήτων, η συστηματική απομάκρυνση από τη γραμμή αναφοράς, το έντονο πρότυπο των υπολοίπων και η υψηλή τιμή χ^2 , ίση με 1.634,93, δείχνουν ότι η συμμετρική μορφή της δεν αναπαράγει ικανοποιητικά τη θετική ασυμμετρία και τη δεξιά ουρά της συγκεκριμένης χρονοσειράς.

Με βάση τη συνολική σύνθεση των αποτελεσμάτων, η τελική κατάταξη των εξεταζόμενων μοντέλων είναι:

Λογαριθμοκανονική > Gamma και Weibull > Κανονική

Η Gamma και η Weibull παρουσιάζουν συγκρίσιμη ενδιάμεση απόδοση, με επιμέρους διαφοροποιήσεις ανάλογα με το κριτήριο αξιολόγησης.

Συνεπώς, η λογαριθμοκανονική επιλέγεται ως η καταλληλότερη θεωρητική κατανομή για τις ωριαίες συγκεντρώσεις PM_{10} της Άμφισσας κατά το έτος 2018. Η επιλογή της στηρίζεται στην περισσότερο ισορροπημένη αναπαράσταση των βασικών κλάσεων της εμπειρικής κατανομής, στη συνολικά καλή εικόνα των γραφικών διαγνωστικών και κυρίως στη σαφή αριθμητική υπεροχή που προκύπτει από το στατιστικό χ^2 .

Η επιλογή αυτή αφορά αποκλειστικά τη συγκεκριμένη ετήσια χρονοσειρά και το σύνολο των τεσσάρων εξεταζόμενων μοντέλων. Δεν συνεπάγεται ότι η λογαριθμοκανονική κατανομή αποτελεί καθολικά το βέλτιστο μοντέλο για κάθε έτος ή για κάθε σειρά συγκεντρώσεων PM_{10} στην περιοχή.

4.5 Προσαρμογή θεωρητικών κατανομών στη Λαμία το έτος 2017

Μετά την ολοκλήρωση της ανάλυσης για την Άμφισσα, η ίδια διαδικασία προσαρμογής εφαρμόστηκε στις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας για το έτος 2017. Η εξέταση της συγκεκριμένης χρονοσειράς επιτρέπει να αξιολογηθεί αν το κατανομικό πρότυπο που παρατηρήθηκε στην Άμφισσα εμφανίζεται και σε διαφορετικό αστικό και χρονικό πλαίσιο.

Όπως προέκυψε από την περιγραφική στατιστική και το αντίστοιχο ιστόγραμμα, η χρονοσειρά της Λαμίας χαρακτηρίζεται από μη αρνητικές τιμές, θετική ασυμμετρία και σαφή συγκέντρωση των παρατηρήσεων στις χαμηλές και μεσαίες κλάσεις. Ιδιαίτερα αυξημένο βάρος παρουσιάζουν οι κλάσεις 10-20 και 20-30 $\mu g/m^3$, οι οποίες συγκροτούν το κύριο τμήμα της εμπειρικής κατανομής. Παράλληλα, η σταδιακή μείωση των συχνοτήτων προς τις υψηλότερες συγκεντρώσεις διαμορφώνει δεξιά ουρά.

Για τη μοντελοποίηση της συγκεκριμένης εμπειρικής μορφής εξετάστηκαν η κανονική, η Gamma, η Weibull και η λογαριθμοκανονική κατανομή. Η κανονική χρησιμοποιήθηκε ως συμμετρικό πρότυπο αναφοράς, ενώ οι υπόλοιπες τρεις κατανομές εξετάστηκαν λόγω της δυνατότητάς τους να περιγράψουν μη αρνητικά και δεξιά ασύμμετρα δεδομένα.

Η αξιολόγηση ακολουθεί το γενικό αναλυτικό πλαίσιο της εργασίας και βασίζεται στη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων, στα διαγνωστικά διαγράμματα, στα διαγράμματα υπολοίπων, στα Q-Q διαγράμματα και στο στατιστικό χ^2 . Ιδιαίτερη έμφαση δίνεται στην ικανότητα κάθε μοντέλου να αναπαράγει τις κυρίαρχες κλάσεις 10-30 $\mu g/m^3$, χωρίς να παραβλέπεται η συμπεριφορά του προς τις υψηλότερες συγκεντρώσεις.

Η τελική επιλογή της καταλληλότερης κατανομής προκύπτει από τη σύνθεση των γραφικών και αριθμητικών ενδείξεων και όχι από ένα μεμονωμένο κριτήριο. Με τον

τρόπο αυτό εξετάζεται τόσο η συνολική συμφωνία με την εμπειρική κατανομή όσο και η σταθερότητα της προσαρμογής στα επιμέρους τμήματά της.

4.5.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων

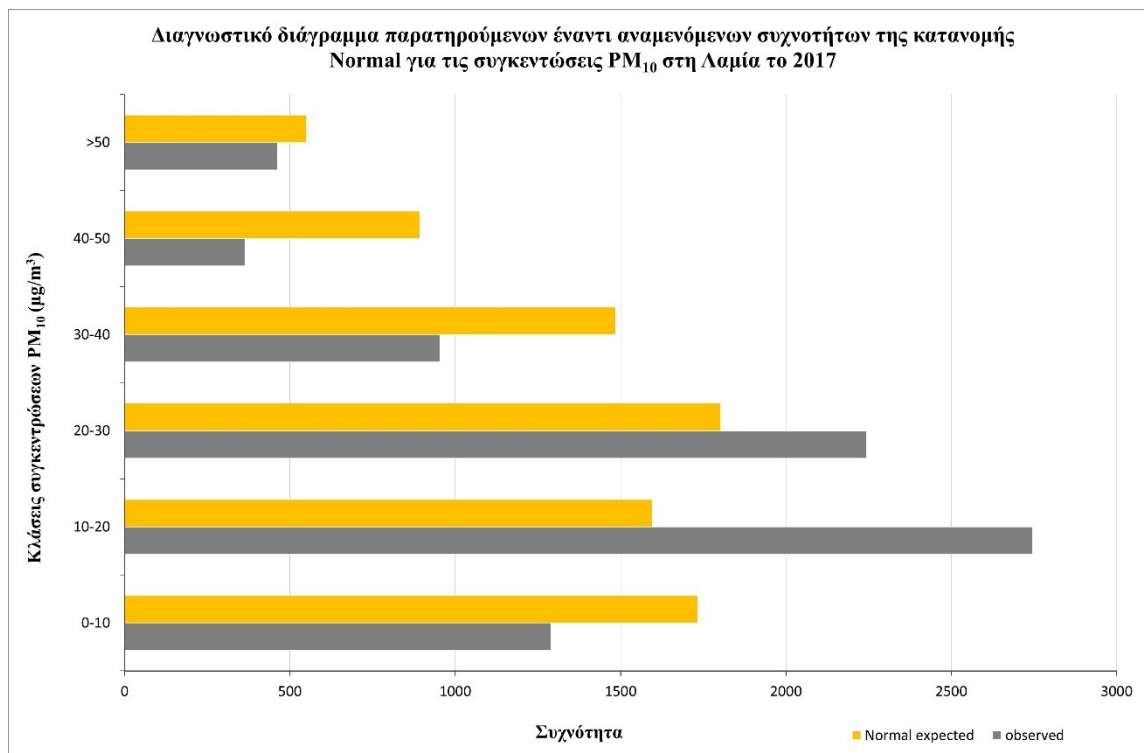
Η σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων ανά κλάση συγκέντρωσης χρησιμοποιείται ως πρώτο στάδιο αξιολόγησης της προσαρμογής των θεωρητικών κατανομών στις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας για το έτος 2017. Καθώς η μεθοδολογική λογική της συγκεκριμένης διαδικασίας παρουσιάστηκε αναλυτικά στην περίπτωση της Άμφισσας, η παρούσα υποενότητα επικεντρώνεται κυρίως στις αποκλίσεις που εμφανίζει κάθε θεωρητικό μοντέλο στις επιμέρους κλάσεις συγκέντρωσης.

Ιδιαίτερη σημασία έχει η απόδοση των κλάσεων 10-20 και 20-30 $\mu g/m^3$, στις οποίες συγκεντρώνεται το μεγαλύτερο μέρος των παρατηρήσεων και διαμορφώνεται η κύρια περιοχή της εμπειρικής κατανομής. Παράλληλα, εξετάζεται η συμπεριφορά των μοντέλων στην κλάση 0-10 $\mu g/m^3$ και στις υψηλότερες κλάσεις, ώστε να αξιολογηθεί κατά πόσο αναπαράγεται η μετάβαση από το κεντρικό τμήμα προς τις μεγαλύτερες συγκεντρώσεις.

Πρέπει να επισημανθεί ότι η σύγκριση βασίζεται σε έξι συγκεντρωτικές κλάσεις και όχι στις επιμέρους ωριαίες τιμές. Ειδικότερα, η ανοικτή κλάση $>50 \mu g/m^3$ περιλαμβάνει όλες τις υψηλότερες συγκεντρώσεις σε μία ενιαία κατηγορία. Κατά συνέπεια, τα διαγράμματα επιτρέπουν την αξιολόγηση της συνολικής κατανομής των συχνοτήτων ανά κλάση, αλλά δεν αποτυπώνουν αναλυτικά τη μορφή και την έκταση της ακραίας δεξιάς ουράς.

Η οπτική αξιολόγηση που ακολουθεί πρέπει, επομένως, να θεωρηθεί προκαταρκτική και να συνεκτιμηθεί με τα υπόλοιπα γραφικά και αριθμητικά διαγνωστικά εργαλεία.

Κανονική κατανομή



Σχήμα 4.5.1α. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

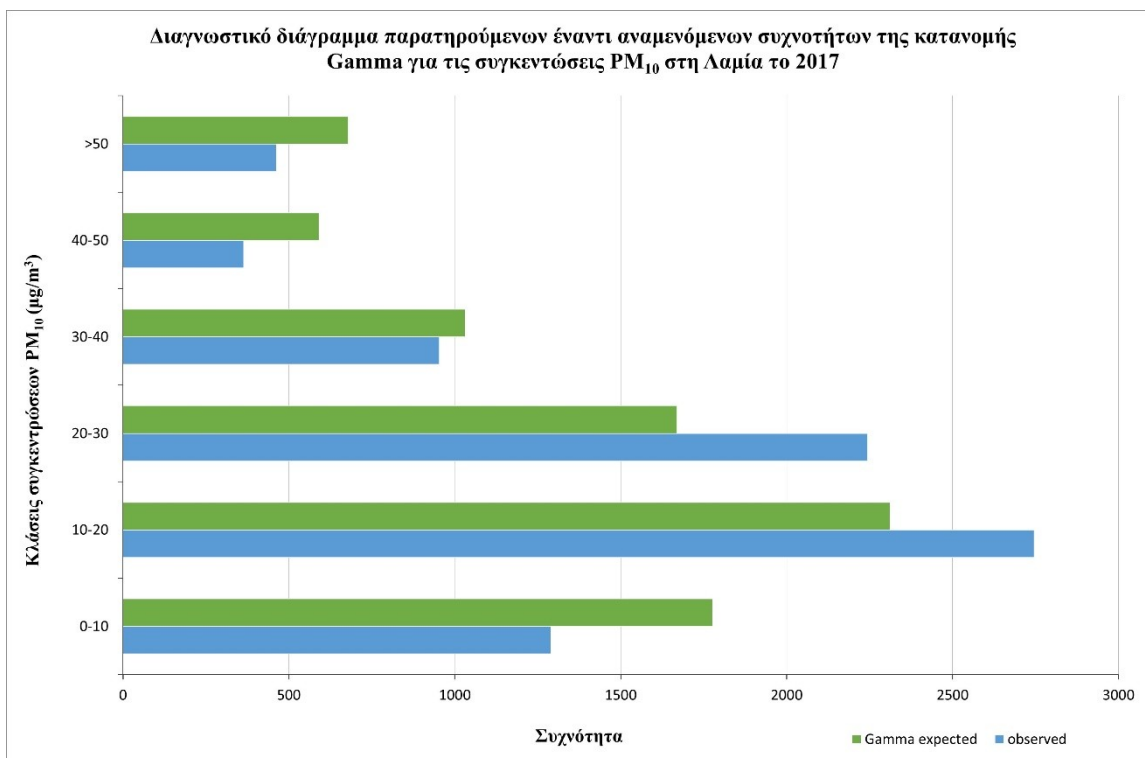
Η κανονική κατανομή εμφανίζει τις εντονότερες συνολικές αποκλίσεις από τις παρατηρούμενες συχνότητες. Η σημαντικότερη αστοχία εντοπίζεται στην κλάση 10-20 $\mu g/m^3$, όπου το μοντέλο υποεκτιμά σε μεγάλο βαθμό την εμπειρική κορυφή. Υποεκτίμηση παρατηρείται και στην κλάση 20-30 $\mu g/m^3$, η οποία αποτελεί επίσης βασικό τμήμα της εμπειρικής κατανομής.

Αντίθετα, η κανονική κατανομή υπερεκτιμά τις συχνότητες στις κλάσεις 0-10, 30-40, 40-50 και >50 $\mu g/m^3$. Ιδιαίτερα εμφανής είναι η υπερεκτίμηση στις κλάσεις 30-40 και 40-50 $\mu g/m^3$. Το μοντέλο κατανέμει, επομένως, μεγαλύτερο μέρος της θεωρητικής πιθανότητας στις γειτονικές και υψηλότερες κλάσεις και μικρότερο στην περιοχή όπου συγκεντρώνεται η πλειονότητα των πραγματικών παρατηρήσεων.

Το πρότυπο αυτό είναι συμβατό με τη συμμετρική μορφή της κανονικής κατανομής, η οποία δεν αναπαράγει αποτελεσματικά τη συγκέντρωση των δεδομένων στις χαμηλές και

μεσαίες τιμές ούτε τη δεξιά ασυμμετρία της χρονοσειράς. Με βάση το συγκεκριμένο διάγραμμα, η κανονική κατανομή παρέχει τη λιγότερο ικανοποιητική οπτική προσαρμογή.

Κατανομή Gamma



Σχήμα 4.5.1β Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

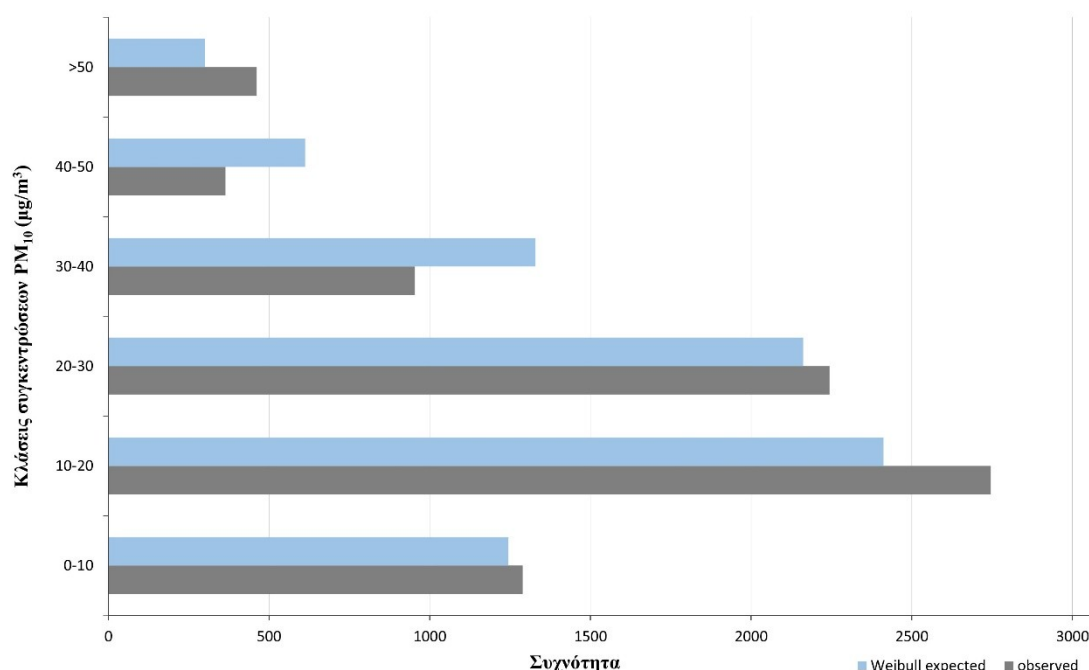
Η κατανομή Gamma παρουσιάζει βελτιωμένη συμπεριφορά σε σχέση με την κανονική, καθώς αποδίδει αποτελεσματικότερα τη γενική ασύμμετρη μορφή της εμπειρικής κατανομής. Ωστόσο, εξακολουθούν να καταγράφονται αισθητές αποκλίσεις στις δύο κλάσεις με τις μεγαλύτερες παρατηρούμενες συχνότητες.

Στις κλάσεις 10-20 και 20-30 $\mu g/m^3$, οι αναμενόμενες συχνότητες είναι χαμηλότερες από τις παρατηρούμενες. Η υποεκτίμηση είναι εντονότερη στην κλάση 20-30 $\mu g/m^3$, γεγονός που δείχνει ότι η Gamma δεν συγκεντρώνει επαρκή θεωρητική πιθανότητα στο κύριο τμήμα της εμπειρικής κατανομής. Στην κλάση 30-40 $\mu g/m^3$ η συμφωνία είναι αισθητά καλύτερη, με σχετικά περιορισμένη υπερεκτίμηση.

Στην κλάση 0-10 $\mu\text{g}/\text{m}^3$ και στις υψηλότερες κλάσεις 40-50 και $>50 \mu\text{g}/\text{m}^3$, η Gamma υπερεκτιμά τις παρατηρούμενες συχνότητες. Επομένως, μέρος της θεωρητικής πιθανότητας μετατοπίζεται προς τα χαμηλότερα και υψηλότερα άκρα εις βάρος των δύο κυρίαρχων κλάσεων. Η Gamma αποτελεί σαφή βελτίωση έναντι της κανονικής κατανομής, αλλά δεν παρέχει πλήρως ισορροπημένη προσαρμογή σε όλο το εξεταζόμενο εύρος.

Κατανομή Weibull

Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις συγκεντρώσεις PM_{10} στη Λαμία το 2017



Σχήμα 4.5.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

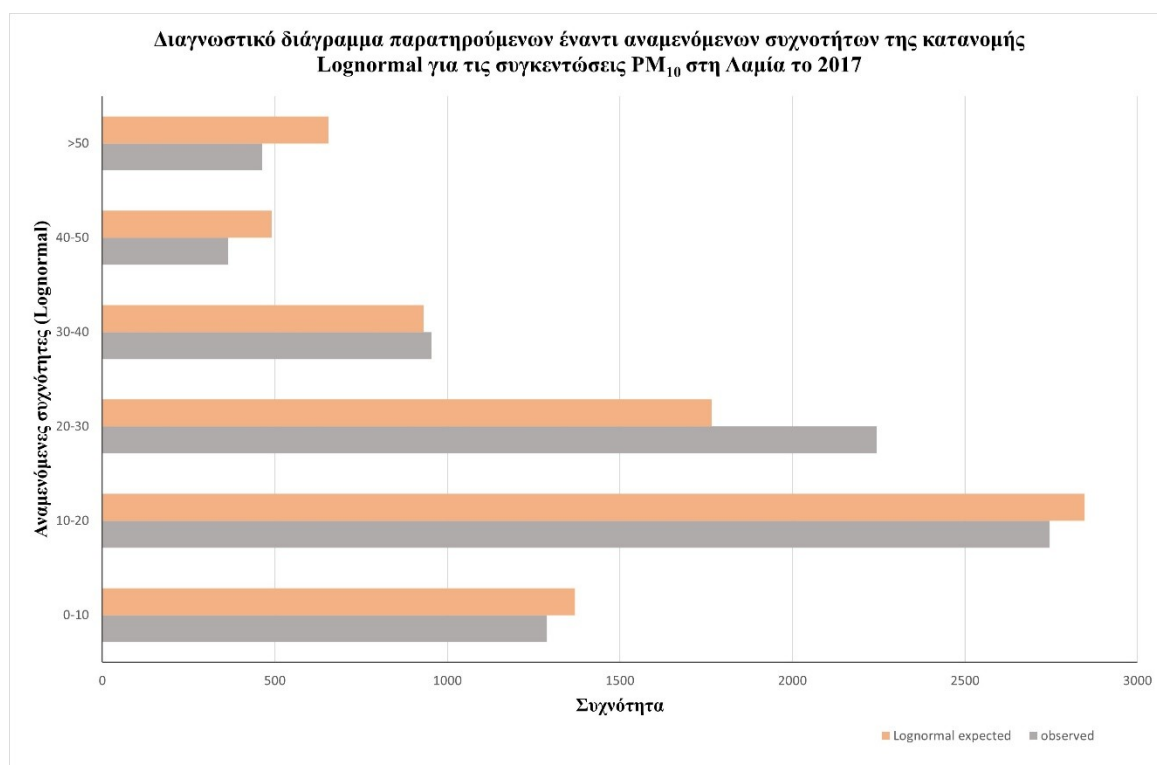
Η κατανομή Weibull εμφανίζει καλή οπτική προσαρμογή στο χαμηλό και κεντρικό τμήμα της εμπειρικής κατανομής. Στην κλάση 0-10 $\mu\text{g}/\text{m}^3$ οι παρατηρούμενες και αναμενόμενες συχνότητες βρίσκονται αρκετά κοντά, ενώ στην κλάση 20-30 $\mu\text{g}/\text{m}^3$ η απόκλιση είναι επίσης περιορισμένη.

Στην κλάση 10-20 $\mu\text{g}/\text{m}^3$ η Weibull υποεκτιμά την εμπειρική κορυφή, αλλά η αστοχία είναι σαφώς μικρότερη από εκείνη της κανονικής κατανομής. Αντίθετα, στις κλάσεις 30-

40 και 40-50 $\mu\text{g}/\text{m}^3$ οι αναμενόμενες συχνότητες υπερβαίνουν τις παρατηρούμενες. Στην ανοικτή κλάση $>50 \mu\text{g}/\text{m}^3$ εμφανίζεται υποεκτίμηση, γεγονός που υποδηλώνει περιορισμένη απόδοση στο ανώτερο τμήμα της κατανομής.

Συνολικά, η Weibull αναπαράγει αρκετά αποτελεσματικά το εύρος 0-30 $\mu\text{g}/\text{m}^3$, στο οποίο εντοπίζεται η πλειονότητα των παρατηρήσεων. Η συμφωνία της περιορίζεται στις αμέσως υψηλότερες κλάσεις και στην ανοικτή ανώτερη κλάση. Παρά τις αποκλίσεις αυτές, αποτελεί μία από τις ισχυρότερες υποψήφιες κατανομές στο συγκεκριμένο στάδιο της αξιολόγησης.

Λογαριθμοκανονική κατανομή



Σχήμα 4.5.1δ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων βάσει της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η λογαριθμοκανονική κατανομή παρουσιάζει την πιο ισορροπημένη συνολική οπτική συμφωνία μεταξύ των τεσσάρων εξεταζόμενων μοντέλων. Στην κλάση 0-10 $\mu\text{g}/\text{m}^3$ εμφανίζει μικρή υπερεκτίμηση, ενώ στην κυρίαρχη κλάση 10-20 $\mu\text{g}/\text{m}^3$ η αναμενόμενη

συχνότητα βρίσκεται σχετικά κοντά στην παρατηρούμενη, παρά την ελαφρά υπερεκτίμησή της.

Στην κλάση $20-30 \mu\text{g}/\text{m}^3$ καταγράφεται αισθητή υποεκτίμηση, η οποία αποτελεί τη σημαντικότερη απόκλιση της λογαριθμοκανονικής στο κεντρικό τμήμα της κατανομής. Αντίθετα, η κλάση $30-40 \mu\text{g}/\text{m}^3$ προσεγγίζεται με ιδιαίτερα μικρή απόκλιση. Στις κλάσεις $40-50$ και $>50 \mu\text{g}/\text{m}^3$, το μοντέλο υπερεκτιμά τις παρατηρούμενες συχνότητες, επομένως η απόδοσή του στο ανώτερο τμήμα δεν είναι πλήρως ακριβής.

Παρά τις επιμέρους αστοχίες, η λογαριθμοκανονική αποφεύγει τις πολύ μεγάλες αποκλίσεις που εμφανίζει η κανονική κατανομή και παρουσιάζει περισσότερο ισορροπημένη κατανομή των σφαλμάτων μεταξύ χαμηλών, μεσαίων και υψηλότερων κλάσεων. Η οπτική εικόνα παρέχει, επομένως, ισχυρή πρώτη ένδειξη καταλληλότητας, χωρίς όμως να αποδεικνύει ότι το μοντέλο αποδίδει εξίσου αποτελεσματικά κάθε τμήμα της εμπειρικής κατανομής.

Συγκριτική αξιολόγηση

Η συγκριτική εξέταση των τεσσάρων διαγραμμάτων αναδεικνύει ουσιαστικές διαφορές στην ικανότητα των μοντέλων να αναπαράγουν τις κυρίαρχες κλάσεις και τη μετάβαση προς τις υψηλότερες συγκεντρώσεις. Η κανονική κατανομή εμφανίζει τη λιγότερο ικανοποιητική συμπεριφορά, καθώς υποεκτιμά έντονα τις κλάσεις $10-20$ και $20-30 \mu\text{g}/\text{m}^3$ και υπερεκτιμά τις περισσότερες από τις υπόλοιπες κλάσεις.

Η Gamma αποδίδει καλύτερα τη γενική ασυμμετρία, αλλά διατηρεί εμφανή υποεκτίμηση στις δύο κυρίαρχες κλάσεις και υπερεκτίμηση στα άκρα. Η Weibull παρουσιάζει καλή συμφωνία κυρίως στο εύρος $0-30 \mu\text{g}/\text{m}^3$ και ιδιαίτερα στην κλάση $20-30 \mu\text{g}/\text{m}^3$, αλλά εμφανίζει αποκλίσεις στις υψηλότερες συγκεντρώσεις.

Η λογαριθμοκανονική κατανομή παρουσιάζει την περισσότερο ισορροπημένη συνολική εικόνα. Προσεγγίζει ικανοποιητικά την κλάση $10-20 \mu\text{g}/\text{m}^3$ και σχεδόν πλήρως την κλάση $30-40 \mu\text{g}/\text{m}^3$, ενώ οι αποκλίσεις της κατανέμονται χωρίς την έντονη ανισορροπία που χαρακτηρίζει την κανονική και τη Gamma. Ωστόσο, η υποεκτίμηση στην κλάση $20-30 \mu\text{g}/\text{m}^3$ και η υπερεκτίμηση στις υψηλότερες κλάσεις δείχνουν ότι η προσαρμογή της δεν είναι τέλεια.

Αξίζει επίσης να σημειωθεί ότι η οπτική εγγύτητα των ράβδων δεν σταθμίζει όλες τις αποκλίσεις με τον ίδιο τρόπο. Μία σχετικά μικρή απόκλιση σε κλάση με πολύ υψηλή παρατηρούμενη συχνότητα μπορεί να είναι περισσότερο σημαντική για τη συνολική

ποιότητα προσαρμογής από μία μεγαλύτερη απόκλιση σε κλάση με περιορισμένο αριθμό παρατηρήσεων. Επομένως, η σύγκριση των ράβδων δεν πρέπει να ερμηνεύεται ανεξάρτητα από την αριθμητική αξιολόγηση και τη συμπεριφορά των υπολοίπων.

Με βάση αποκλειστικά τη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων, η λογαριθμοκανονική και η Weibull εμφανίζονται ως οι ισχυρότερες υποψήφιες κατανομές. Η λογαριθμοκανονική παρουσιάζει ελαφρώς πιο ισορροπημένη συνολική συμπεριφορά, ενώ η Weibull αποδίδει ιδιαίτερα αποτελεσματικά σημαντικό μέρος του κεντρικού εύρους. Η Gamma ακολουθεί, ενώ η κανονική κατανομή εμφανίζει τη χαμηλότερη καταλληλότητα.

Το συμπέρασμα αυτό παραμένει προκαταρκτικό, καθώς βασίζεται στην οπτική σύγκριση περιορισμένου αριθμού συγκεντρωτικών κλάσεων. Η τελική κατάταξη των μοντέλων θα προκύψει από τη συνεκτίμηση των διαγνωστικών διαγραμμάτων παρατηρούμενων έναντι αναμενόμενων συχνοτήτων, των διαγραμμάτων υπολοίπων, των Q-Q διαγραμμάτων και του στατιστικού χ^2 .

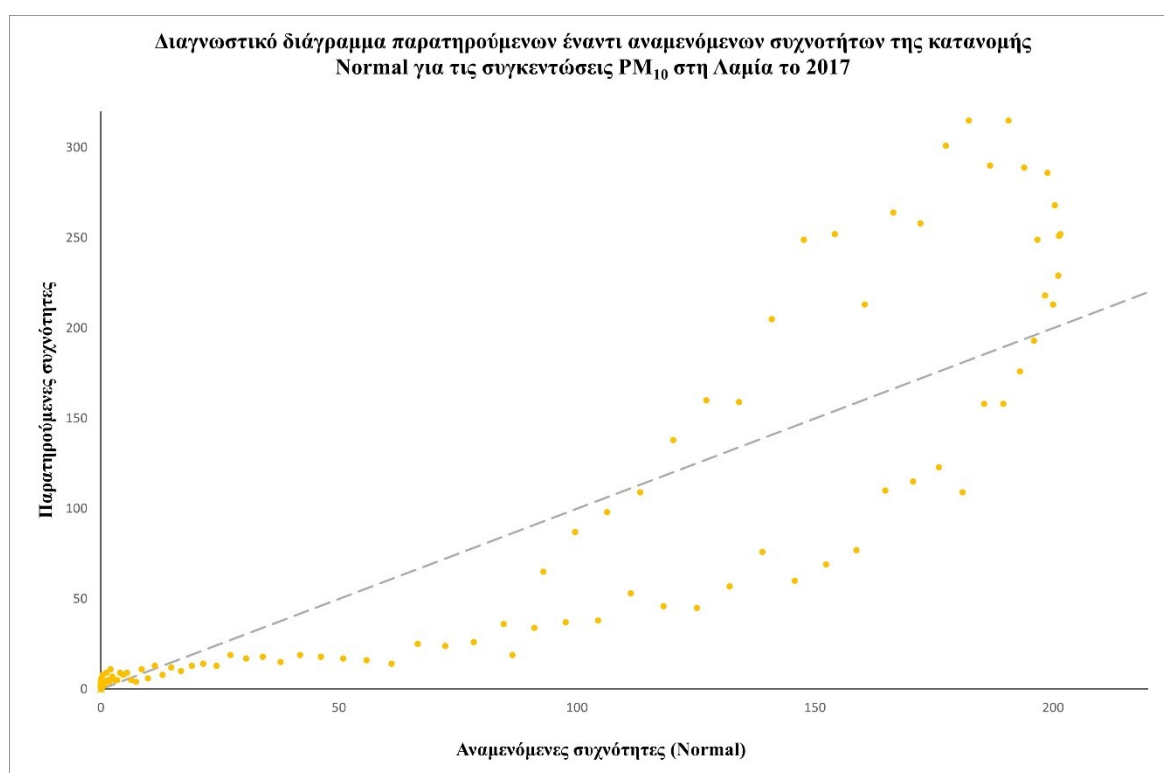
4.5.2 Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων για την πόλη της Λαμίας

Τα διαγνωστικά διαγράμματα των Σχημάτων 4.5.2α-δ παρουσιάζουν τη σχέση μεταξύ των παρατηρούμενων και των αναμενόμενων συχνοτήτων για την κανονική, τη Gamma, τη Weibull και τη λογαριθμοκανονική κατανομή, όπως αυτές προσαρμόστηκαν στις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας για το έτος 2017. Στον οριζόντιο άξονα απεικονίζονται οι αναμενόμενες συχνότητες κάθε θεωρητικού μοντέλου και στον κατακόρυφο άξονα οι αντίστοιχες παρατηρούμενες συχνότητες.

Η διακεκομμένη διαγώνια γραμμή $y=x$ εκφράζει την ιδανική περίπτωση πλήρους συμφωνίας μεταξύ των δύο μεγεθών. Όσο πλησιέστερα βρίσκονται τα σημεία στη γραμμή αναφοράς και όσο μικρότερη είναι η διασπορά τους γύρω από αυτήν, τόσο ισχυρότερη θεωρείται η ένδειξη καλής προσαρμογής. Σημεία πάνω από τη διαγώνιο σημαίνουν ότι η παρατηρούμενη συχνότητα υπερβαίνει την αναμενόμενη και, επομένως, ότι το μοντέλο υποεκτιμά τη συγκεκριμένη τιμή. Αντίθετα, σημεία κάτω από τη γραμμή αντιστοιχούν σε υπερεκτίμηση της παρατηρούμενης συχνότητας από το θεωρητικό μοντέλο.

Η συγκεκριμένη γραφική προσέγγιση συμπληρώνει τη σύγκριση ανά κλάση συγκέντρωσης της προηγούμενης υποενότητας. Τα ραβδογράμματα επιτρέπουν τον άμεσο εντοπισμό των αποκλίσεων σε συγκεκριμένες συγκεντρωτικές κλάσεις, ενώ τα διαγνωστικά διαγράμματα αποτυπώνουν τη συνολικότερη σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Η αξιολόγηση βασίζεται στην εγγύτητα των σημείων προς τη διαγώνιο, στην έκταση της διασποράς και στην παρουσία συστηματικών μη γραμμικών προτύπων.

Κανονική κατανομή



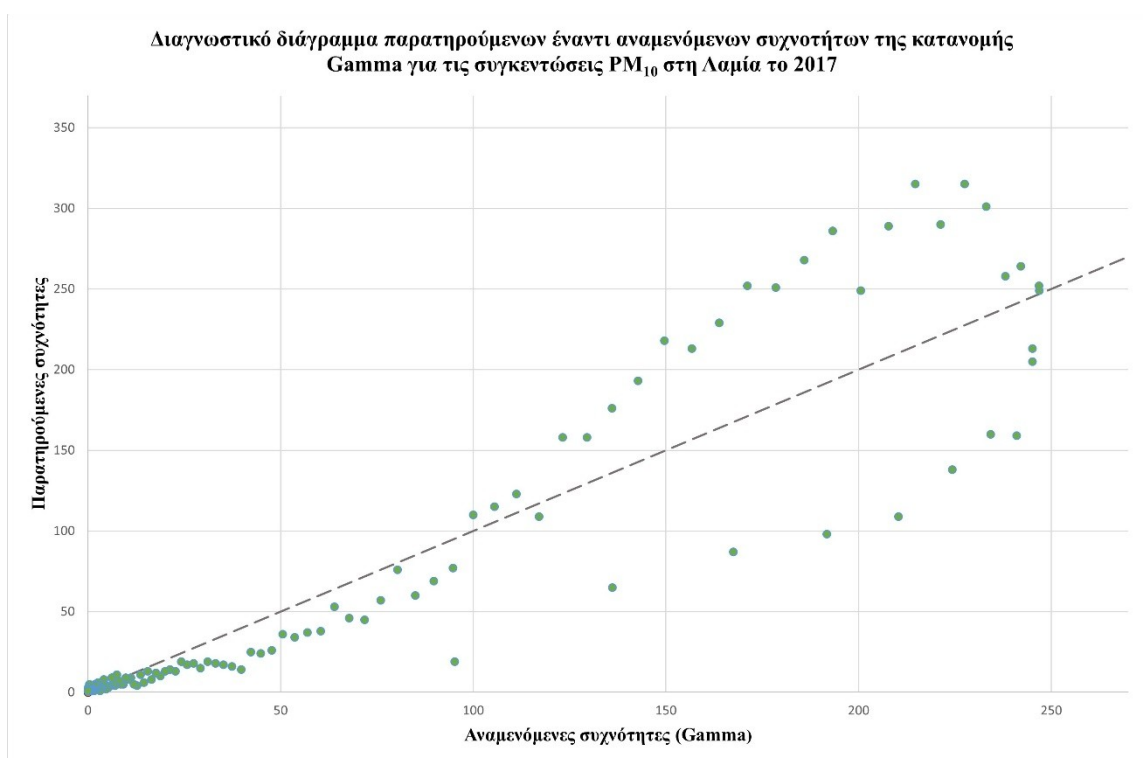
Σχήμα 4.5.2α. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η κανονική κατανομή εμφανίζει τη λιγότερο ικανοποιητική συνολική εικόνα. Στην περιοχή των χαμηλών αναμενόμενων συχνοτήτων αρκετά σημεία βρίσκονται σχετικά κοντά στην αρχή των αξόνων και στη γραμμή αναφοράς. Ωστόσο, με την αύξηση των αναμενόμενων τιμών η διάταξη των σημείων απομακρύνεται έντονα από τη διαγώνιο.

Στο μεσαίο και υψηλότερο τμήμα του διαγράμματος παρατηρείται μεγάλη κατακόρυφη διασπορά. Για παρόμοιες αναμενόμενες συχνότητες εμφανίζεται ευρύ φάσμα παρατηρούμενων τιμών, ενώ τα σημεία βρίσκονται άλλοτε πολύ πάνω και άλλοτε πολύ κάτω από τη γραμμή $y=x$. Το πρότυπο αυτό υποδηλώνει εναλλαγή υποεκτίμησης και υπερεκτίμησης και περιορισμένη ικανότητα του μοντέλου να αναπαράγει με συνέπεια την εμπειρική δομή.

Η εμφανής μη γραμμικότητα είναι συμβατή με τα αποτελέσματα της προηγούμενης υποενότητας, όπου η κανονική κατανομή υποεκτίμησε τις κυρίαρχες κλάσεις και υπερεκτίμησε γειτονικά τμήματα της κατανομής. Συνεπώς, το διάγραμμα επιβεβαιώνει ότι το συμμετρικό αυτό μοντέλο δεν ανταποκρίνεται ικανοποιητικά στη θετική ασυμμετρία της χρονοσειράς.

Κατανομή Gamma



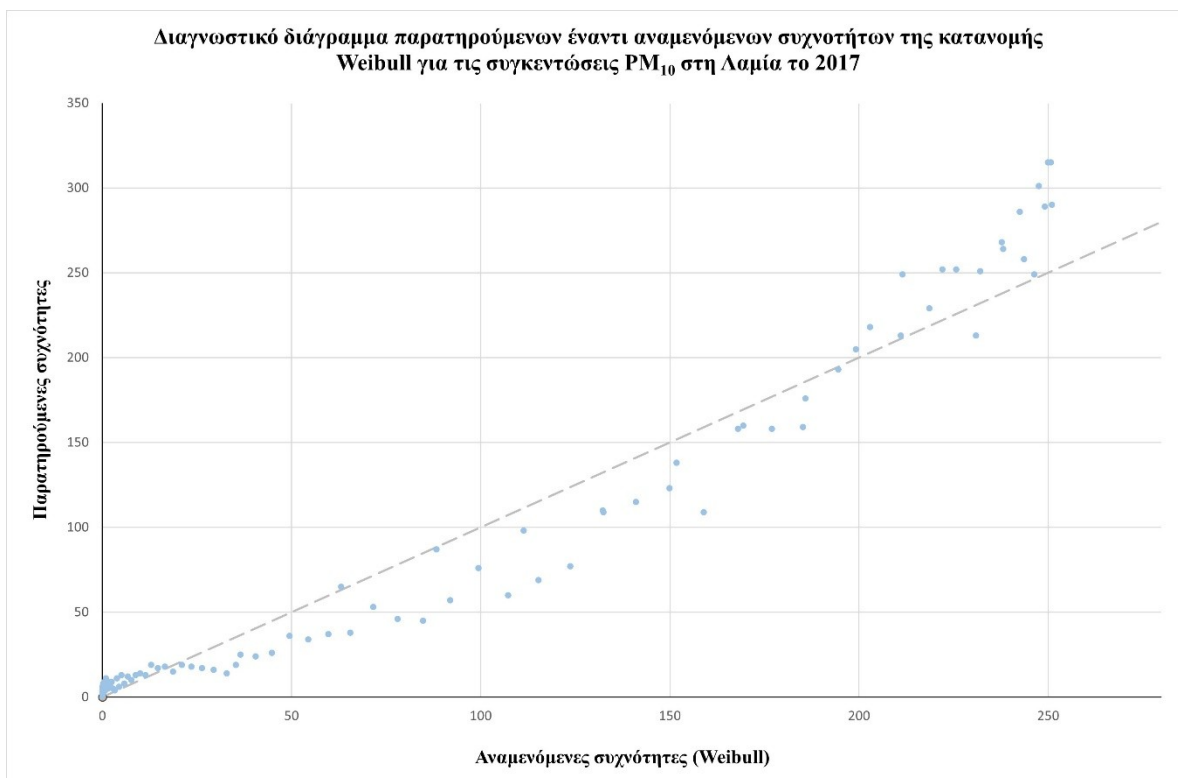
Σχήμα 4.5.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η κατανομή Gamma παρουσιάζει σαφή βελτίωση έναντι της κανονικής. Στο χαμηλό εύρος συχνοτήτων τα σημεία συγκεντρώνονται σχετικά κοντά στη διαγώνιο, ενώ και σε τμήμα του μεσαίου εύρους διατηρείται γενικά θετική σχέση μεταξύ αναμενόμενων και παρατηρούμενων τιμών.

Η συμφωνία, ωστόσο, δεν είναι ομοιόμορφη. Σε αρκετές μεσαίες συχνότητες τα σημεία βρίσκονται κάτω από τη διαγώνιο, γεγονός που υποδηλώνει υπερεκτίμηση από το μοντέλο. Σε άλλο τμήμα, κυρίως καθώς οι αναμενόμενες συχνότητες αυξάνονται, αρκετά σημεία τοποθετούνται αισθητά πάνω από τη γραμμή, υποδηλώνοντας υποεκτίμηση των παρατηρούμενων τιμών.

Στις υψηλότερες αναμενόμενες συχνότητες η διασπορά γίνεται ιδιαίτερα έντονη, με παρατηρούμενες τιμές που εκτείνονται σε μεγάλο κατακόρυφο εύρος για σχετικά παρόμοιες θεωρητικές προβλέψεις. Επομένως, η Gamma αποδίδει καλύτερα τη γενική ασύμμετρη μορφή από ό,τι η κανονική, αλλά δεν διατηρεί σταθερή ακρίβεια σε όλο το εύρος των δεδομένων.

Κατανομή Weibull

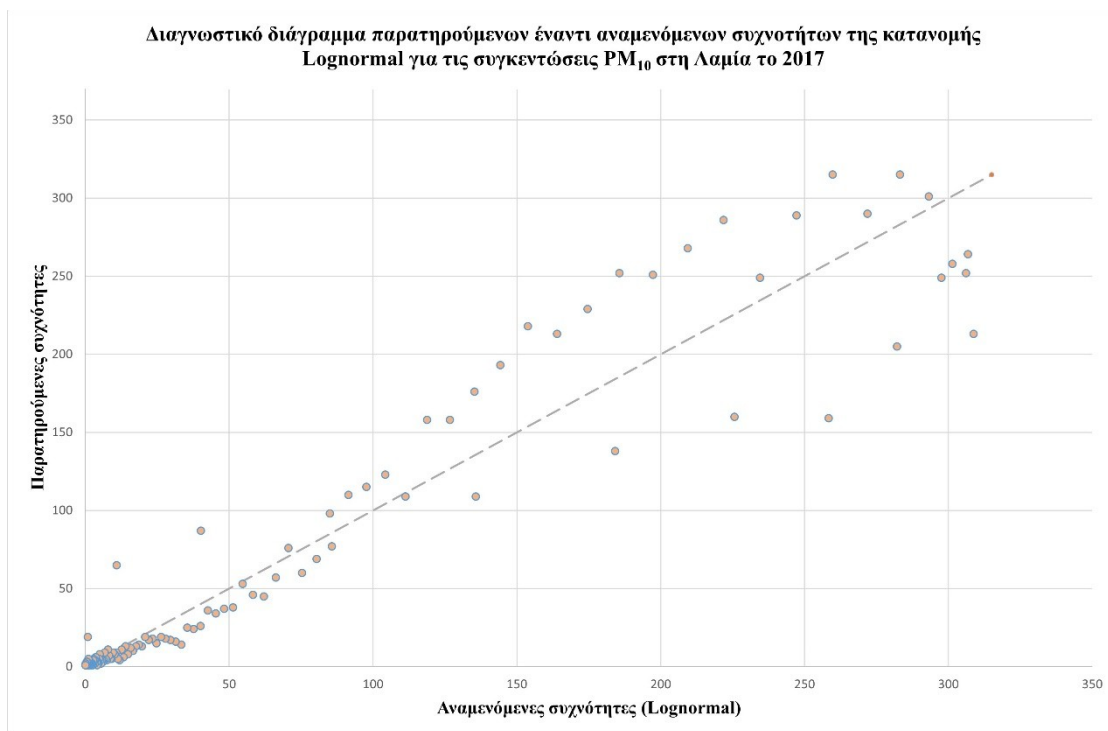


Σχήμα 4.5.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η κατανομή Weibull εμφανίζει την περισσότερο οργανωμένη και σχεδόν γραμμική διάταξη σημείων μεταξύ των τεσσάρων εξεταζόμενων μοντέλων. Στο χαμηλό εύρος οι παρατηρούμενες και αναμενόμενες συχνότητες βρίσκονται σχετικά κοντά στη διαγώνιο, γεγονός που δείχνει καλή συμφωνία για σημαντικό μέρος των μικρότερων τιμών. Στο μεσαίο εύρος αρκετά σημεία βρίσκονται κάτω από τη γραμμή $y=x$, γεγονός που αντιστοιχεί σε τάση υπερεκτίμησης των παρατηρούμενων συχνοτήτων. Καθώς οι αναμενόμενες συχνότητες αυξάνονται, η διάταξη μετατοπίζεται σταδιακά προς τη διαγώνιο και στη συνέχεια αρκετά σημεία τοποθετούνται πάνω από αυτήν, υποδηλώνοντας τάση υποεκτίμησης στο υψηλότερο τμήμα. Παρότι διακρίνεται αυτό το συστηματικό πρότυπο, η συνολική ακολουθία των σημείων παραμένει περισσότερο συνεκτική και λιγότερο διάσπαρτη από εκείνη της Gamma και της κανονικής κατανομής. Το διάγραμμα δείχνει, επομένως, ότι η Weibull αποδίδει αποτελεσματικά τη γενική σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων, χωρίς όμως να εξαλείφει τις τοπικές αποκλίσεις.

Με βάση αποκλειστικά το συγκεκριμένο γραφικό κριτήριο, η Weibull αποτελεί μία από τις ισχυρότερες και πιθανόν την ισχυρότερη οπτική υποψήφια κατανομή για τη χρονοσειρά της Λαμίας

Λογαριθμοκανονική κατανομή



Σχήμα 4.5.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η λογαριθμοκανονική κατανομή παρουσιάζει καλή συμφωνία στο χαμηλό εύρος των συχνοτήτων, όπου μεγάλο μέρος των σημείων βρίσκεται κοντά στη γραμμή αναφοράς. Και στο μεσαίο εύρος η συνολική τάση είναι θετική και ακολουθεί σε γενικές γραμμές τη διαγώνιο.

Ωστόσο, με την αύξηση των αναμενόμενων συχνοτήτων η διασπορά των σημείων διευρύνεται σημαντικά. Σε αρκετές περιπτώσεις οι παρατηρούμενες συχνότητες βρίσκονται αισθητά πάνω από τη διαγώνιο, γεγονός που υποδηλώνει υποεκτίμηση από το μοντέλο. Σε άλλες περιπτώσεις, στο ίδιο περίπου εύρος αναμενόμενων τιμών, τα σημεία βρίσκονται αρκετά κάτω από τη γραμμή, υποδηλώνοντας υπερεκτίμηση.

Η εικόνα αυτή δείχνει ότι η λογαριθμοκανονική αποδίδει ικανοποιητικά τη γενική αύξηση των παρατηρούμενων συχνοτήτων, αλλά δεν διατηρεί ομοιόμορφη ακρίβεια στο ανώτερο τμήμα. Η παρουσία σημαντικής κατακόρυφης διασποράς δεν επιτρέπει να υποστηριχθεί ότι τα περισσότερα σημεία βρίσκονται σταθερά εγγύτερα στη διαγώνιο από ό,τι στην περίπτωση της Weibull.

Συνεπώς, η λογαριθμοκανονική αποτελεί ισχυρή υποψήφια κατανομή, αλλά το συγκεκριμένο διάγραμμα δεν τεκμηριώνει από μόνο του σαφή οπτική υπεροχή έναντι της Weibull.

Συγκριτική αξιολόγηση

Η συγκριτική εξέταση των τεσσάρων διαγνωστικών διαγραμμάτων αναδεικνύει σαφείς διαφοροποιήσεις ως προς τη συνολική γεωμετρική σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Η κανονική κατανομή παρουσιάζει τη χαμηλότερη συνολική καταλληλότητα, καθώς η διάταξη των σημείων αποκλίνει έντονα από τη γραμμή $y=x$, ιδιαίτερα στο μεσαίο και υψηλότερο εύρος συχνοτήτων. Η μεγάλη κατακόρυφη διασπορά και η εναλλαγή σημείων πάνω και κάτω από τη διαγώνιο υποδηλώνουν ότι το μοντέλο δεν αναπαράγει με σταθερότητα τη δομή των εμπειρικών δεδομένων.

Η Gamma αποτελεί ουσιαστική βελτίωση, ιδίως στο χαμηλό εύρος, όπου αρκετά σημεία βρίσκονται σχετικά κοντά στη γραμμή αναφοράς. Ωστόσο, καθώς αυξάνονται οι αναμενόμενες συχνότητες, η διασπορά διευρύνεται και εμφανίζονται σημαντικές αποκλίσεις και προς τις δύο πλευρές της διαγωνίου. Το πρότυπο αυτό δείχνει ότι η Gamma αποδίδει καλύτερα τη γενική ασύμμετρη μορφή της χρονοσειράς από ό,τι η κανονική, χωρίς όμως να διατηρεί ομοιόμορφη ακρίβεια σε όλο το εξεταζόμενο εύρος.

Η λογαριθμοκανονική κατανομή εμφανίζει επίσης καλή συμπεριφορά στο χαμηλό και σε μέρος του μεσαίου εύρους, όπου η συνολική τάση των σημείων ακολουθεί τη διαγώνιο. Στις υψηλότερες αναμενόμενες συχνότητες, όμως, παρατηρείται αισθητή κατακόρυφη διασπορά και τοπική εναλλαγή υπερεκτίμησης και υποεκτίμησης. Επομένως, το διάγραμμα υποστηρίζει την καταλληλότητά της ως ισχυρού υποψήφιου μοντέλου, αλλά δεν τεκμηριώνει από μόνο του σταθερή υπεροχή σε όλο το εύρος των συχνοτήτων.

Η Weibull παρουσιάζει την περισσότερο συνεκτική συνολική διάταξη. Τα σημεία της ακολουθούν εμφανέστερα μια σχεδόν γραμμική πορεία και εμφανίζουν μικρότερη γενική διασπορά σε σχέση με τη Gamma και τη λογαριθμοκανονική. Παράλληλα, όμως, διακρίνεται ένα συστηματικό μοτίβο αποκλίσεων: σε μέρος του μεσαίου εύρους αρκετά σημεία βρίσκονται κάτω από τη διαγώνιο, γεγονός που αντιστοιχεί σε υπερεκτίμηση των παρατηρούμενων συχνοτήτων, ενώ σε ορισμένες υψηλότερες συχνότητες εμφανίζονται σημεία πάνω από τη γραμμή, υποδηλώνοντας υποεκτίμηση. Η Weibull αποδίδει, επομένως, αποτελεσματικά τη γενική τάση, χωρίς να εξαλείφει πλήρως τις τοπικές συστηματικές αποκλίσεις.

Με βάση αποκλειστικά τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων, η Weibull παρέχει την ισχυρότερη οπτική ένδειξη προσαρμογής. Η λογαριθμοκανονική και η Gamma ακολουθούν ως αξιόλογες εναλλακτικές, με τη λογαριθμοκανονική να εμφανίζει καλή γενική τάση αλλά αυξημένη διασπορά στις υψηλότερες συχνότητες και τη Gamma να παρουσιάζει μεγαλύτερη αστάθεια στο ανώτερο τμήμα. Η κανονική κατατάσσεται σαφώς τελευταία.

Η οπτική αυτή κατάταξη δεν πρέπει να θεωρηθεί οριστική, επειδή η εγγύτητα των σημείων προς τη διαγώνιο δεν αποτυπώνει τη σχετική βαρύτητα κάθε απόκλισης. Μια μικρή απόσταση από τη γραμμή σε περιοχή με υψηλή συχνότητα μπορεί να συμβάλλει περισσότερο στη συνολική αστοχία από μια μεγαλύτερη απόκλιση σε σημείο με χαμηλή συχνότητα. Επιπλέον, το διάγραμμα αξιολογεί κυρίως τη συνολική γεωμετρική αντιστοιχία και όχι άμεσα τη συμπεριφορά του μοντέλου στις ουρές ή την κατανομή των σφαλμάτων κατά μήκος της ταξινομημένης σειράς.

Για τον λόγο αυτό, η τελική επιλογή της καταλληλότερης θεωρητικής κατανομής πρέπει να προκύψει από τη συνεκτίμηση των διαγνωστικών διαγραμμάτων με τα διαγράμματα υπολοίπων, τα Q-Q διαγράμματα και την αριθμητική αξιολόγηση μέσω του στατιστικού χ^2 . Η συνδυαστική αυτή προσέγγιση θα δείξει αν η περισσότερο συνεκτική οπτική εικόνα της Weibull επιβεβαιώνεται και από τα υπόλοιπα κριτήρια ή αν κάποιο εναλλακτικό

μοντέλο παρουσιάζει μικρότερη συνολική απόκλιση και καλύτερη συμπεριφορά σε επιμέρους τμήματα της εμπειρικής κατανομής.

4.5.3 Διαγράμματα υπολοίπων

Η αξιολόγηση της ποιότητας προσαρμογής των θεωρητικών κατανομών στις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας για το έτος 2017 συμπληρώθηκε με την εξέταση των διαγραμμάτων υπολοίπων. Τα διαγράμματα αυτά επιτρέπουν την αξιολόγηση όχι μόνο του μεγέθους των αποκλίσεων μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων, αλλά και του τρόπου με τον οποίο οι αποκλίσεις μεταβάλλονται κατά μήκος της ταξινομημένης σειράς των συγκεντρώσεων.

Για κάθε σημείο σύγκρισης, το απλό υπόλοιπο υπολογίστηκε ως:

$$R_i = O_i - E_i$$

όπου:

- R_i είναι το υπόλοιπο,
- O_i η παρατηρούμενη συχνότητα και
- E_i η αναμενόμενη συχνότητα βάσει της αντίστοιχης θεωρητικής κατανομής.

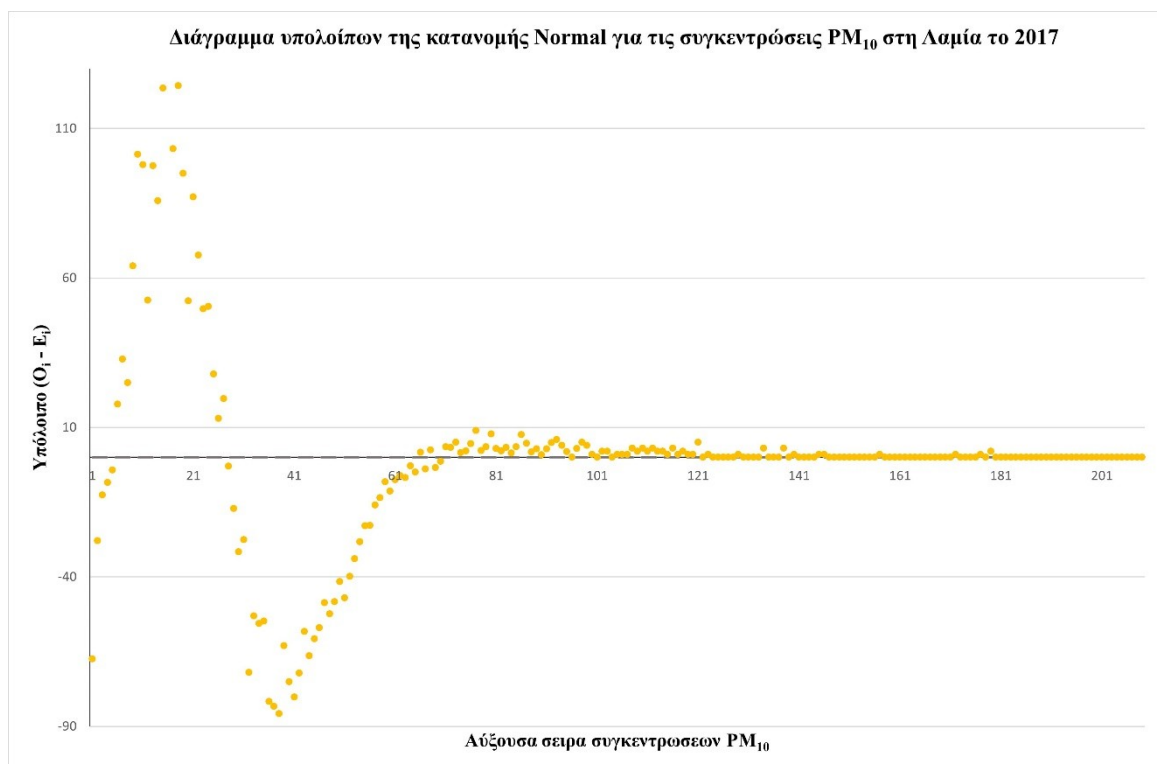
Θετικό υπόλοιπο σημαίνει ότι η παρατηρούμενη συχνότητα είναι μεγαλύτερη από την αναμενόμενη και, επομένως, ότι το μοντέλο την υποεκτιμά. Αντίθετα, αρνητικό υπόλοιπο σημαίνει ότι η θεωρητικά αναμενόμενη συχνότητα υπερβαίνει την παρατηρούμενη και ότι το μοντέλο την υπερεκτιμά.

Σε μια ικανοποιητική προσαρμογή, τα υπόλοιπα αναμένεται να κατανέμονται χωρίς εμφανές συστηματικό πρότυπο γύρω από τη γραμμή $R_i=0$. Παρατεταμένες ακολουθίες σημείων με ίδιο πρόσημο, έντονες καμπυλότητες, κυματοειδείς σχηματισμοί και μεγάλες διαδοχικές μεταβολές αποτελούν ενδείξεις τοπικής ή συστηματικής αστοχίας του θεωρητικού μοντέλου.

Πρέπει να διευκρινιστεί ότι στα παρόντα σχήματα απεικονίζονται απλά και όχι τυποποιημένα υπόλοιπα. Το απόλυτο μέγεθός τους επηρεάζεται, επομένως, από το ύψος των αντίστοιχων συχνοτήτων. Η αξιολόγηση βασίζεται κυρίως στη συγκριτική έκταση

των αποκλίσεων, στην παρουσία οργανωμένων προτύπων και στην ταχύτητα με την οποία τα υπόλοιπα συγκλίνουν προς τη γραμμή μηδενός.

Κανονική κατανομή



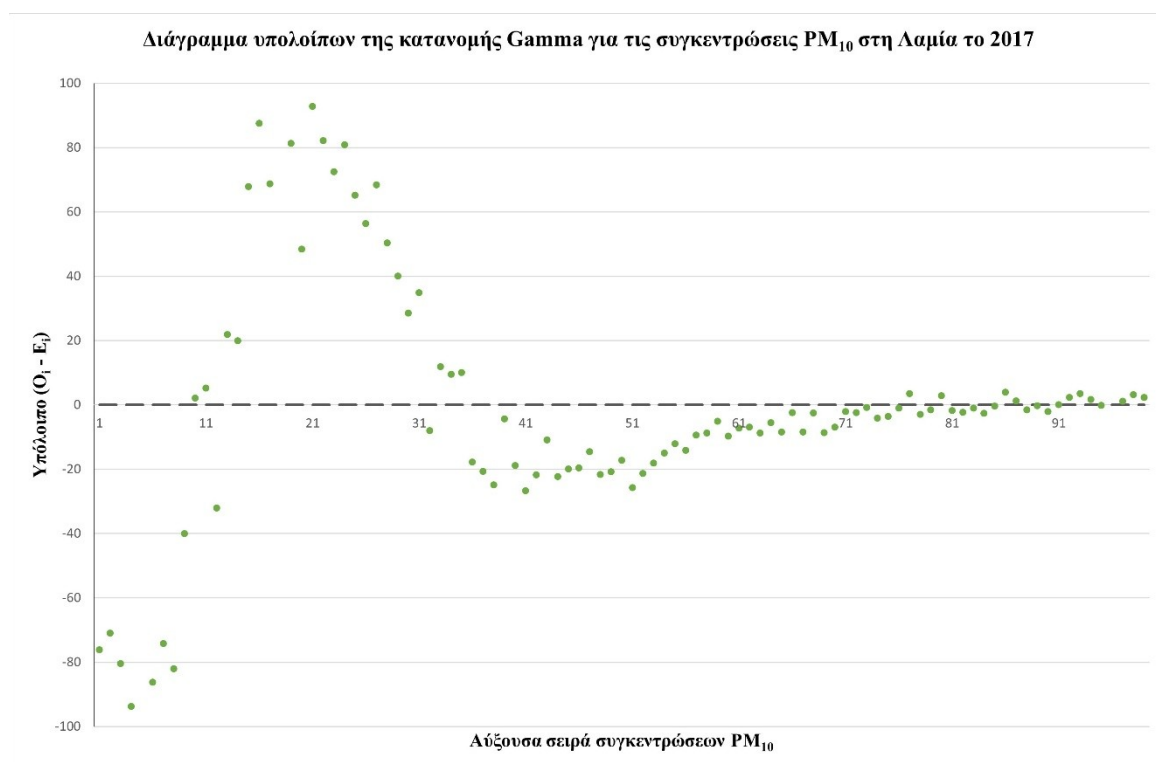
Σχήμα 4.5.3α. Διάγραμμα υπολοίπων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.

Η κανονική κατανομή εμφανίζει το εντονότερο συστηματικό πρότυπο υπολοίπων μεταξύ των τεσσάρων εξεταζόμενων μοντέλων. Στο αρχικό τμήμα της ταξινομημένης σειράς παρατηρούνται πολύ υψηλά θετικά υπόλοιπα, τα οποία σε ορισμένες περιπτώσεις υπερβαίνουν το +120. Οι τιμές αυτές δείχνουν έντονη υποεκτίμηση των αντίστοιχων παρατηρούμενων συχνοτήτων.

Στη συνέχεια, τα υπόλοιπα μεταβάλλονται απότομα προς αρνητικές τιμές και προσεγγίζουν περίπου το -90. Η μεταβολή αυτή αντιστοιχεί σε σημαντική υπερεκτίμηση από το μοντέλο. Η διαδοχική μετάβαση από μεγάλες θετικές σε μεγάλες αρνητικές αποκλίσεις σχηματίζει έντονη καμπυλότητα, η οποία δεν είναι συμβατή με τυχαία διασπορά γύρω από το μηδέν.

Μετά το αρχικό και μεσαίο τμήμα, τα υπόλοιπα συγκλίνουν σταδιακά προς τη γραμμή αναφοράς και παραμένουν σχετικά μικρά στο μεγαλύτερο μέρος του υπόλοιπου διαγράμματος. Η σύγκλιση αυτή, όμως, δεν αναιρεί τις πολύ μεγάλες και οργανωμένες αποκλίσεις που προηγούνται. Το διάγραμμα επιβεβαιώνει, επομένως, ότι η κανονική κατανομή δεν αναπαράγει με συνέπεια τη δεξιά ασύμμετρη δομή της χρονοσειράς.

Κατανομή Gamma



Σχήμα 4.5.3β. Διάγραμμα υπολοίπων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.

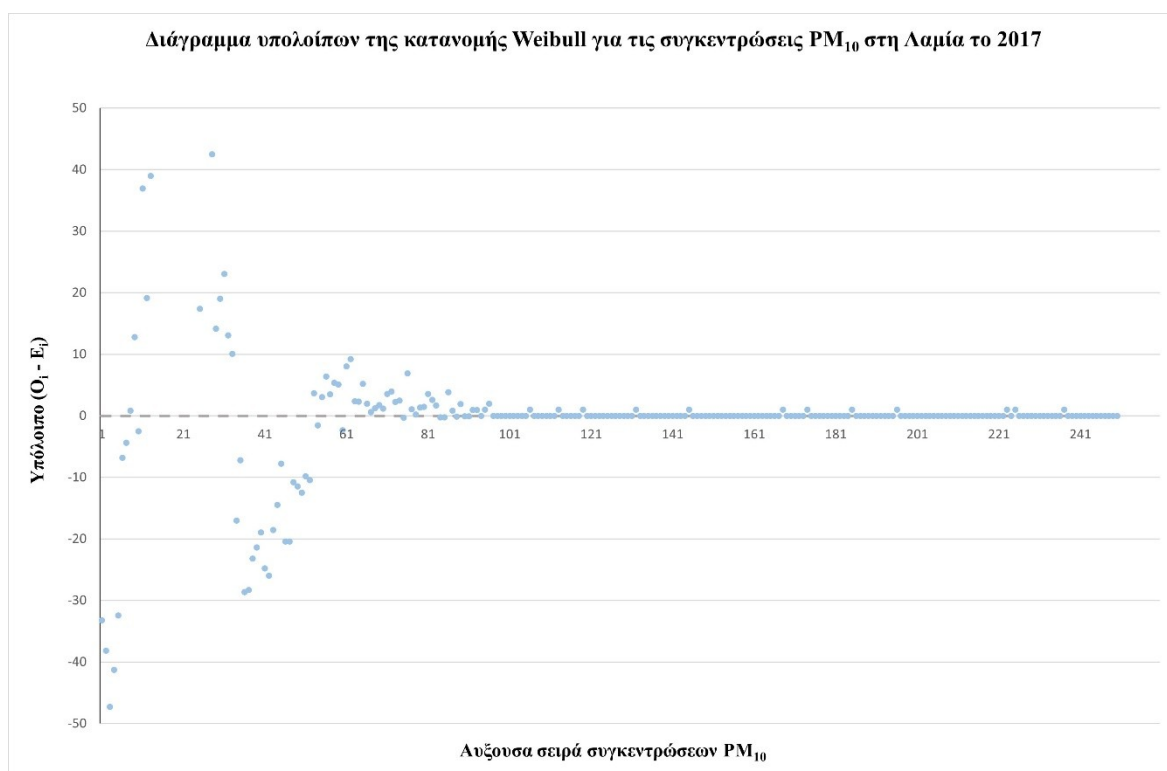
Η κατανομή Gamma παρουσιάζει βελτιωμένη εικόνα σε σχέση με την κανονική, χωρίς όμως να εξαλείφει τα συστηματικά πρότυπα. Στις πρώτες ταξινομημένες παρατηρήσεις καταγράφονται έντονα αρνητικά υπόλοιπα, με τιμές που προσεγγίζουν το -95 , γεγονός που δείχνει σημαντική υπερεκτίμηση από το μοντέλο.

Ακολουθεί περιοχή με μεγάλες θετικές αποκλίσεις, οι οποίες φθάνουν περίπου έως το $+90$. Το μοντέλο μεταβαίνει, επομένως, από έντονη υπερεκτίμηση σε έντονη υποεκτίμηση. Στη συνέχεια, τα υπόλοιπα γίνονται και πάλι αρνητικά και σχηματίζουν

εκτεταμένη ακολουθία κάτω από τη γραμμή μηδενός, με τιμές που κυμαίνονται κυρίως μεταξύ περίπου -5 και -27 .

Το πρότυπο αυτό δεν μπορεί να θεωρηθεί τυχαίο. Δείχνει ότι η Gamma υπερεκτιμά συστηματικά σημαντικό τμήμα της ταξινομημένης σειράς, παρότι στο τελευταίο μέρος τα υπόλοιπα περιορίζονται και κατανέμονται σχετικά κοντά στο μηδέν. Συνεπώς, η Gamma αποτελεί ουσιαστική βελτίωση έναντι της κανονικής, αλλά εξακολουθεί να εμφανίζει εμφανείς τοπικές αστοχίες.

Κατανομή Weibull



Σχήμα 4.5.3γ Διάγραμμα υπολοίπων (Residual plot) της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.

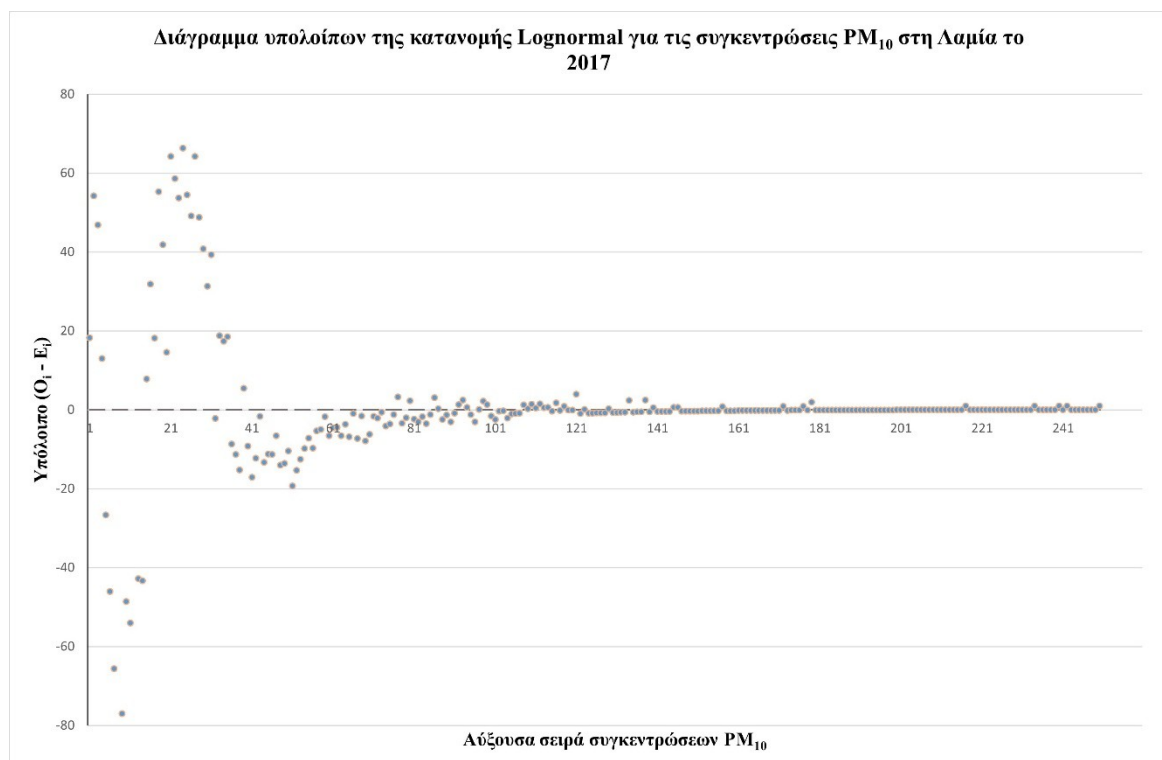
Η κατανομή Weibull παρουσιάζει αισθητά μικρότερο εύρος αποκλίσεων από την κανονική και τη Gamma. Οι μεγαλύτερες αρνητικές τιμές περιορίζονται περίπου έως το -47 , ενώ οι μέγιστες θετικές αποκλίσεις προσεγγίζουν το $+42$. Τα ακραία υπόλοιπα εντοπίζονται κυρίως στο αρχικό τμήμα της ταξινομημένης σειράς.

Στην αρχική περιοχή παρατηρείται εναλλαγή αρνητικών και θετικών τιμών, η οποία δείχνει ότι το μοντέλο δεν αποδίδει με την ίδια ακρίβεια όλες τις χαμηλότερες ταξινομημένες παρατηρήσεις. Ακολουθεί μια περιοχή κυρίως αρνητικών υπολοίπων, περίπου μεταξύ -8 και -29 , που αντιστοιχεί σε συστηματική υπερεκτίμηση από το μοντέλο.

Η ένταση των αποκλίσεων μειώνεται, όμως, σχετικά γρήγορα. Μετά το αρχικό τμήμα, τα υπόλοιπα συγκεντρώνονται σε στενή περιοχή γύρω από το μηδέν και στο μεγαλύτερο μέρος της ταξινομημένης σειράς παραμένουν μηδενικά ή σχεδόν μηδενικά. Η μακρά αυτή σταθεροποίηση υποδηλώνει αποτελεσματική αναπαράσταση σημαντικού μέρους της εμπειρικής δομής.

Η Weibull δεν παρουσιάζει πλήρως τυχαία υπόλοιπα, επειδή στο αρχικό τμήμα διακρίνονται οργανωμένες μεταβολές. Ωστόσο, το σχετικά περιορισμένο εύρος των αποκλίσεων και η γρήγορη σύγκλιση προς το μηδέν συνιστούν ισχυρή ένδειξη καλής προσαρμογής.

Λογαριθμοκανονική κατανομή



Σχήμα 4.5.3δ Διάγραμμα υπολοίπων (Residual plot) της λογαριθμοκανονικής κατανομής (Lognormal) για τις ωριαίες συγκεντρώσεις PM_{10} της Λαμίας κατά το έτος 2017.

Η λογαριθμοκανονική κατανομή παρουσιάζει επίσης σαφή βελτίωση έναντι της κανονικής και της Gamma. Οι μεγαλύτερες αποκλίσεις συγκεντρώνονται κυρίως στο αρχικό τμήμα της ταξινομημένης σειράς, όπου τα υπόλοιπα εκτείνονται περίπου από -77 έως $+66$.

Στην αρχική αυτή περιοχή καταγράφεται εναλλαγή μεγάλων αρνητικών και θετικών τιμών, γεγονός που δείχνει ότι το μοντέλο υπερεκτιμά ορισμένες παρατηρήσεις και υποεκτιμά άλλες. Ακολουθεί τμήμα με κυρίως αρνητικά υπόλοιπα, τα οποία προσεγγίζουν περίπου το -20 , πριν περιοριστούν σταδιακά προς τη γραμμή μηδενός.

Μετά το αρχικό τμήμα, τα υπόλοιπα συγκεντρώνονται σε πολύ στενή ζώνη γύρω από το μηδέν. Στο μεγαλύτερο μέρος της ταξινομημένης σειράς οι αποκλίσεις είναι μικρές και δεν σχηματίζουν εκτεταμένο νέο πρότυπο. Το εύρημα αυτό δείχνει ότι η λογαριθμοκανονική αναπαράγει αποτελεσματικά σημαντικό τμήμα της εμπειρικής δομής. Οι αρχικές αποκλίσεις της είναι, ωστόσο, μεγαλύτερες από εκείνες της Weibull. Επομένως, το συγκεκριμένο διάγραμμα υποστηρίζει τη λογαριθμοκανονική ως ισχυρό υποψήφιο μοντέλο, αλλά δεν τεκμηριώνει ότι παρουσιάζει μικρότερη συνολική διασπορά υπολοίπων από τη Weibull.

Συγκριτική αξιολόγηση

Η συγκριτική εξέταση των τεσσάρων διαγραμμάτων υπολοίπων αναδεικνύει σαφείς διαφοροποιήσεις τόσο ως προς το εύρος των αποκλίσεων όσο και ως προς τον τρόπο με τον οποίο αυτές κατανέμονται κατά μήκος της ταξινομημένης σειράς των συγκεντρώσεων. Η αξιολόγηση δεν βασίζεται μόνο στο μέγιστο θετικό ή αρνητικό υπόλοιπο, αλλά και στην παρουσία παρατεταμένων ακολουθιών ίδιου προσήμου, στην ύπαρξη οργανωμένων καμπυλοτήτων και στην ταχύτητα με την οποία τα σημεία συγκλίνουν προς τη γραμμή $R_i = 0$.

Η κανονική κατανομή εμφανίζει τη λιγότερο ικανοποιητική συμπεριφορά. Τα υπόλοιπα παρουσιάζουν πολύ μεγάλο εύρος, με έντονες θετικές και αρνητικές αποκλίσεις, και σχηματίζουν σαφές κυματοειδές πρότυπο. Η διαδοχική μετάβαση από ισχυρή υποεκτίμηση σε ισχυρή υπερεκτίμηση δείχνει ότι οι αποκλίσεις δεν είναι τυχαίες, αλλά συνδέονται με συστηματική αδυναμία του μοντέλου να αναπαράγει τη μορφή της εμπειρικής κατανομής. Παρότι στο τελευταίο τμήμα της σειράς τα υπόλοιπα συγκλίνουν

προς το μηδέν, η εκτεταμένη προηγούμενη ασυμφωνία καθιστά την κανονική κατανομή σαφώς ανεπαρκή.

Η Gamma περιορίζει μέρος των αποκλίσεων σε σχέση με την κανονική, αλλά εξακολουθεί να εμφανίζει έντονα οργανωμένη συμπεριφορά. Οι μεγάλες αρνητικές τιμές στην αρχή, η μετάβαση σε υψηλά θετικά υπόλοιπα και η επακόλουθη εκτεταμένη ακολουθία αρνητικών τιμών δείχνουν διαδοχικές περιοχές υπερεκτίμησης και υποεκτίμησης. Η σταδιακή σύγκλιση προς το μηδέν στο τελευταίο τμήμα αποτελεί ένδειξη καλύτερης τοπικής προσαρμογής, αλλά δεν αναιρεί τη συστηματική δομή των σφαλμάτων στο αρχικό και μεσαίο εύρος. Επομένως, η Gamma βελτιώνει τη συνολική εικόνα, χωρίς να εξαλείφει πλήρως τις τοπικές αστοχίες.

Η Weibull και η λογαριθμοκανονική παρουσιάζουν αισθητά καλύτερη συμπεριφορά. Και στις δύο περιπτώσεις, οι μεγαλύτερες αποκλίσεις περιορίζονται κυρίως στο αρχικό τμήμα της ταξινομημένης σειράς, ενώ στη συνέχεια τα υπόλοιπα συγκεντρώνονται σε στενή ζώνη γύρω από τη γραμμή μηδενός. Το γεγονός αυτό δείχνει ότι τα δύο μοντέλα αναπαράγουν αποτελεσματικότερα το μεγαλύτερο μέρος της εμπειρικής δομής από ό,τι η κανονική και η Gamma.

Η Weibull εμφανίζει μικρότερο περίπου συνολικό εύρος ακραίων υπολοίπων και ταχύτερη μετάβαση σε περιοχή πολύ μικρών αποκλίσεων. Μετά το αρχικό τμήμα, η πλειονότητα των σημείων παραμένει εξαιρετικά κοντά στο μηδέν, γεγονός που υποδηλώνει σταθερή προσαρμογή σε μεγάλο μέρος της σειράς. Παράλληλα, όμως, στο αρχικό τμήμα παρατηρείται οργανωμένη εναλλαγή αρνητικών και θετικών τιμών, καθώς και μια σύντομη ακολουθία κυρίως αρνητικών υπολοίπων. Συνεπώς, η Weibull δεν εξαλείφει πλήρως τις τοπικές αστοχίες, αλλά περιορίζει αισθητά τη διάρκειά και την έντασή τους.

Η λογαριθμοκανονική εμφανίζει επίσης γρήγορη σταθεροποίηση των υπολοίπων γύρω από το μηδέν και ιδιαίτερα μικρές αποκλίσεις στο μεγαλύτερο μέρος της ταξινομημένης σειράς. Ωστόσο, οι αρχικές θετικές και αρνητικές αποκλίσεις είναι μεγαλύτερες από εκείνες της Weibull. Η παρουσία τιμών που εκτείνονται περίπου από -77 έως $+66$ δείχνει ότι το μοντέλο δυσκολεύεται περισσότερο να αποδώσει με ομοιόμορφη ακρίβεια το αρχικό τμήμα των δεδομένων. Μετά την περιοχή αυτή, πάντως, η συμπεριφορά του είναι ιδιαίτερα σταθερή και τα περισσότερα υπόλοιπα παραμένουν σχεδόν μηδενικά.

Με βάση αποκλειστικά τα απλά υπόλοιπα, η Weibull παρέχει την ισχυρότερη οπτική ένδειξη προσαρμογής. Το πλεονέκτημά της στηρίζεται στο μικρότερο περίπου εύρος ακραίων αποκλίσεων, στη γρηγορότερη σύγκλιση προς τη γραμμή μηδενός και στη μακρά περιοχή μικρών υπολοίπων. Η λογαριθμοκανονική ακολουθεί πολύ κοντά, καθώς εμφανίζει επίσης ιδιαίτερα μικρές αποκλίσεις στο μεγαλύτερο μέρος της σειράς, αλλά παρουσιάζει εντονότερες αρχικές αστοχίες.

Η Gamma κατατάσσεται χαμηλότερα, επειδή διατηρεί περισσότερο εκτεταμένα και συστηματικά πρότυπα υπολοίπων, ενώ η κανονική κατανομή εμφανίζει σαφώς τη χειρότερη εικόνα. Η σχετική κατάταξη που προκύπτει από τα συγκεκριμένα διαγράμματα είναι, επομένως:

$$\text{Weibull} > \text{Λογαριθμοκανονική} > \text{Gamma} > \text{Κανονική}$$

όπου το σύμβολο $>$ δηλώνει ισχυρότερη οπτική ένδειξη προσαρμογής με βάση μόνο τη συμπεριφορά των απλών υπολοίπων.

Η κατάταξη αυτή δεν πρέπει να εκληφθεί ως τελική επιλογή μοντέλου. Τα απλά υπόλοιπα δεν είναι τυποποιημένα και, επομένως, το απόλυτο μέγεθός τους εξαρτάται από το επίπεδο των αντίστοιχων συχνοτήτων. Επιπλέον, η οπτική συγκέντρωση των σημείων γύρω από το μηδέν δεν αποτυπώνει τη σχετική βαρύτητα κάθε απόκλισης ούτε αξιολογεί άμεσα την προσαρμογή στα ποσοστημόρια και στις ουρές.

Για τον λόγο αυτό, η εικόνα των υπολοίπων πρέπει να συνεκτιμηθεί με τα Q-Q διαγράμματα και το στατιστικό χ^2 . Η αριθμητική αξιολόγηση θα δείξει αν η μικρότερη διασπορά των υπολοίπων της Weibull συνοδεύεται και από μικρότερη συνολική σταθμισμένη απόκλιση ή αν η λογαριθμοκανονική υπερέχει σε άλλες πτυχές της προσαρμογής. Με αυτόν τον τρόπο, η τελική επιλογή θα βασιστεί στη σύνθεση συμπληρωματικών ενδείξεων και όχι στην αποκλειστική ερμηνεία ενός μόνο διαγνωστικού εργαλείου.

4.5.4 Διαγράμματα Q-Q

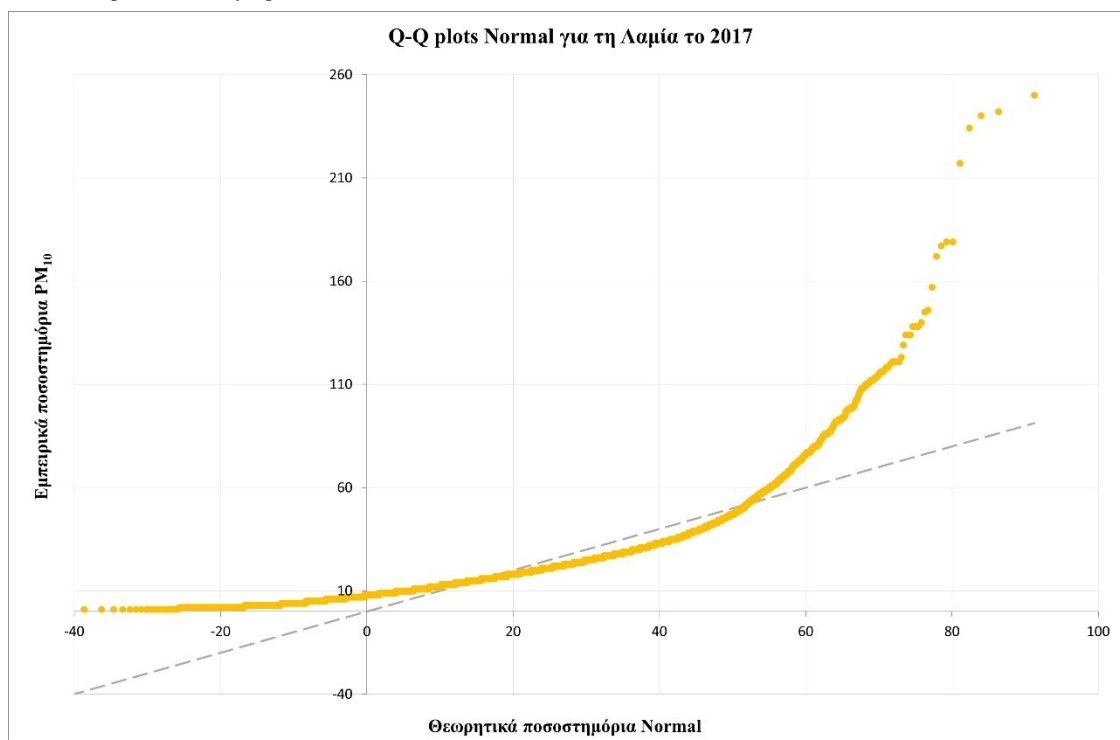
Τα διαγράμματα Q-Q χρησιμοποιήθηκαν για τη σύγκριση των εμπειρικών ποσοστημορίων των ωριαίων συγκεντρώσεων PM_{10} στη Λαμία για το έτος 2017 με τα αντίστοιχα θεωρητικά ποσοστημόρια της κανονικής, της Gamma, της Weibull και της λογαριθμοκανονικής κατανομής. Η συγκεκριμένη γραφική προσέγγιση επιτρέπει να αξιολογηθεί αν κάθε θεωρητικό μοντέλο αναπαράγει με συνέπεια όχι μόνο το κεντρικό τμήμα, αλλά και τα άκρα της εμπειρικής κατανομής.

Σε περίπτωση ικανοποιητικής προσαρμογής, τα σημεία αναμένεται να κατανέμονται κοντά στη διαγώνια γραμμή αναφοράς. Η συστηματική καμπυλότητα, η παρατεταμένη παρουσία σημείων πάνω ή κάτω από τη διαγώνιο και η έντονη απομάκρυνση στα υψηλότερα ποσοστημόρια αποτελούν ενδείξεις ότι το θεωρητικό μοντέλο δεν αποδίδει πλήρως τη μορφή ή το βάρος των ουρών της εμπειρικής κατανομής.

Σημεία που βρίσκονται πάνω από τη γραμμή αναφοράς υποδηλώνουν ότι τα εμπειρικά ποσοστημόρια είναι υψηλότερα από τα αντίστοιχα θεωρητικά. Στην άνω ουρά, η συμπεριφορά αυτή σημαίνει ότι το μοντέλο υποεκτιμά τις υψηλές παρατηρούμενες συγκεντρώσεις. Αντίθετα, σημεία κάτω από τη διαγώνιο δείχνουν ότι τα θεωρητικά ποσοστημόρια υπερβαίνουν τα εμπειρικά.

Ιδιαίτερη έμφαση δίνεται στο ανώτερο τμήμα των διαγραμμάτων, επειδή εκεί αξιολογείται η ικανότητα των μοντέλων να περιγράψουν τη δεξιά ουρά της κατανομής. Ωστόσο, η απόδοση στην ουρά δεν εξετάζεται απομονωμένα, καθώς ένα κατάλληλο μοντέλο πρέπει ταυτόχρονα να αναπαράγει ικανοποιητικά το κύριο σώμα των δεδομένων. Τα συμπεράσματα των Q-Q διαγραμμάτων συνεκτιμώνται με τη σύγκριση των συχνοτήτων, τα διαγνωστικά διαγράμματα, τα υπόλοιπα και το στατιστικό χ^2 .

Κανονική κατανομή

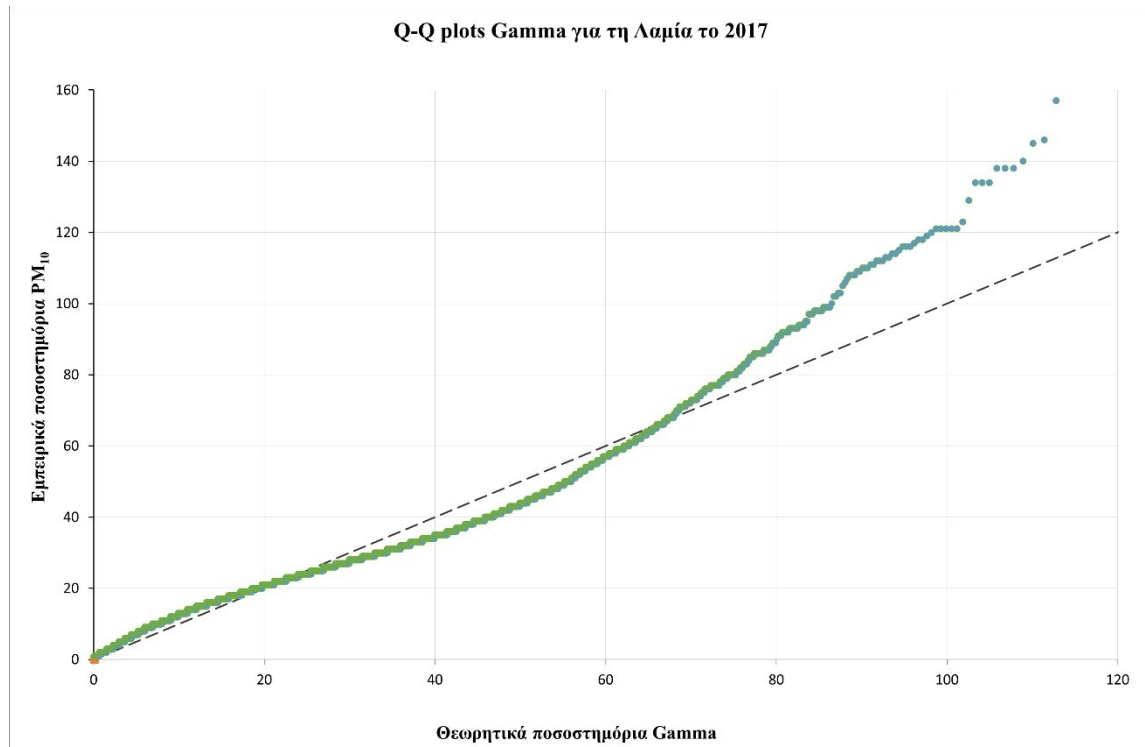


Σχήμα 4.5.4α Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Το διάγραμμα Q-Q της κανονικής κατανομής παρουσιάζει έντονη και συστηματική απόκλιση από τη γραμμή αναφοράς. Τα σημεία δεν ακολουθούν γραμμική διάταξη, αλλά σχηματίζουν σαφή καμπυλότητα σε σχεδόν ολόκληρο το εύρος των ποσοστημορίων. Η εικόνα αυτή αποτελεί ισχυρή ένδειξη ότι η εμπειρική κατανομή δεν είναι συμβατή με τη συμμετρική μορφή της κανονικής. Στο χαμηλό τμήμα, τα εμπειρικά ποσοστημόρια παραμένουν θετικά, ενώ τα αντίστοιχα θεωρητικά ποσοστημόρια της κανονικής εκτείνονται και σε αρνητικές τιμές. Η ασυμφωνία αυτή συνδέεται με το απεριόριστο θεωρητικό πεδίο της κανονικής κατανομής, το οποίο δεν είναι φυσικά συμβατό με τις μη αρνητικές συγκεντρώσεις PM_{10} . Στο μεσαίο τμήμα, τα σημεία εξακολουθούν να αποκλίνουν συστηματικά από τη διαγώνιο και δεν παρουσιάζουν σταθερή ευθυγράμμιση. Η απόκλιση γίνεται εντονότερη στα υψηλότερα ποσοστημόρια, όπου τα εμπειρικά σημεία μετατοπίζονται σημαντικά πάνω από τη γραμμή αναφοράς. Η έντονη ανοδική απομάκρυνση στην άνω ουρά δείχνει ότι η κανονική κατανομή υποεκτιμά σοβαρά τις υψηλές και ακραίες συγκεντρώσεις. Συνεπώς, το συγκεκριμένο διάγραμμα επιβεβαιώνει

ότι η κανονική κατανομή δεν μπορεί να αναπαράγει ούτε την ασυμμετρία ούτε το βάρος της δεξιάς ουράς της χρονοσειράς της Λαμίας.

Κατανομή Gamma



Σχήμα 4.5.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

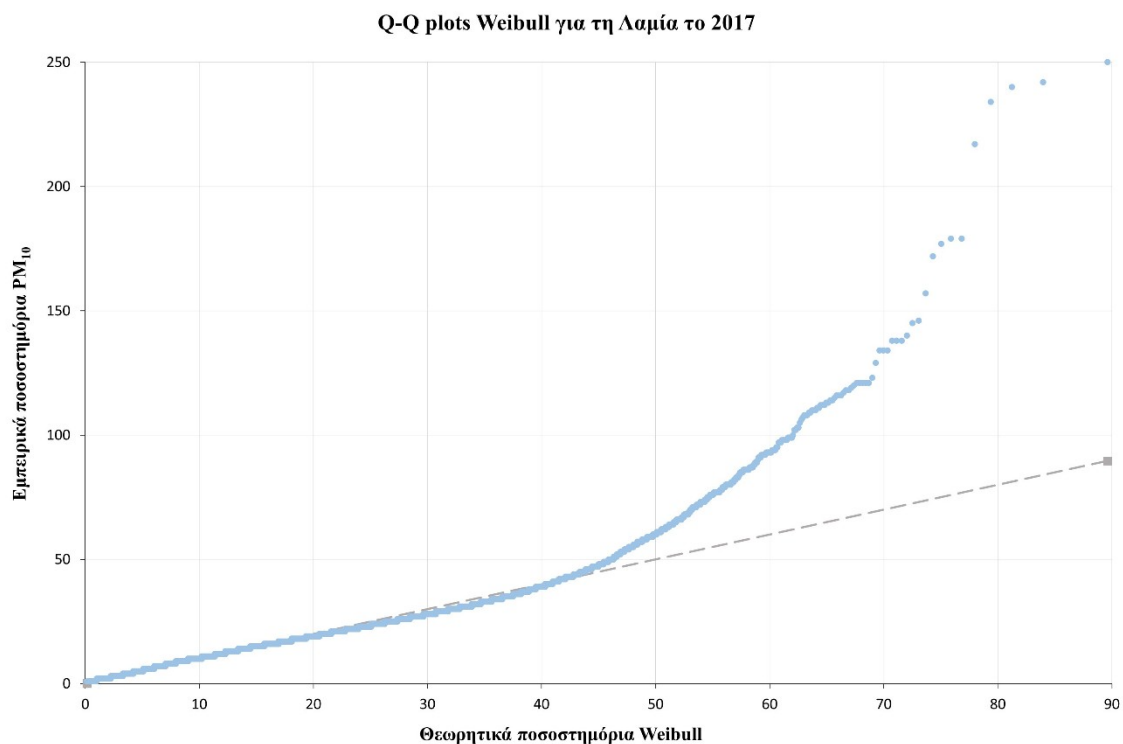
Η κατανομή Gamma παρουσιάζει σαφώς καλύτερη συμφωνία από την κανονική, αλλά η διάταξη των σημείων εξακολουθεί να εμφανίζει συστηματική καμπυλότητα. Στα χαμηλά ποσοστημόρια τα σημεία βρίσκονται αρχικά λίγο πάνω από τη γραμμή αναφοράς, ενώ σε τμήμα του χαμηλού και μεσαίου εύρους μετατοπίζονται σταδιακά κάτω από αυτήν.

Η παρουσία σημείων κάτω από τη διαγώνιο στο μεσαίο τμήμα δείχνει ότι η Gamma τείνει να παράγει υψηλότερα θεωρητικά ποσοστημόρια από εκείνα που παρατηρούνται εμπειρικά. Η εικόνα αυτή αντιστοιχεί σε τοπική υπερεκτίμηση του αντίστοιχου τμήματος της κατανομής. Περίπου στο ανώτερο μεσαίο τμήμα, τα σημεία διασταυρώνουν τη γραμμή αναφοράς και στη συνέχεια μετατοπίζονται προοδευτικά πάνω από αυτήν. Η ανοδική αυτή απόκλιση γίνεται περισσότερο εμφανής όσο αυξάνονται τα ποσοστημόρια,

γεγονός που δείχνει ότι η Gamma υποεκτιμά τις υψηλότερες συγκεντρώσεις και δεν διαθέτει επαρκώς βαριά άνω ουρά για να αποδώσει τις μεγαλύτερες παρατηρήσεις.

Παρά τις αποκλίσεις, η καμπυλότητα της Gamma είναι σαφώς ηπιότερη από εκείνη της κανονικής κατανομής και σημαντικό μέρος των σημείων βρίσκεται σχετικά κοντά στη διαγώνιο. Επομένως, η Gamma αναπαράγει ικανοποιητικότερα το κύριο σώμα της εμπειρικής κατανομής, αλλά εμφανίζει σαφή αδυναμία στην άνω ουρά.

Κατανομή Weibull

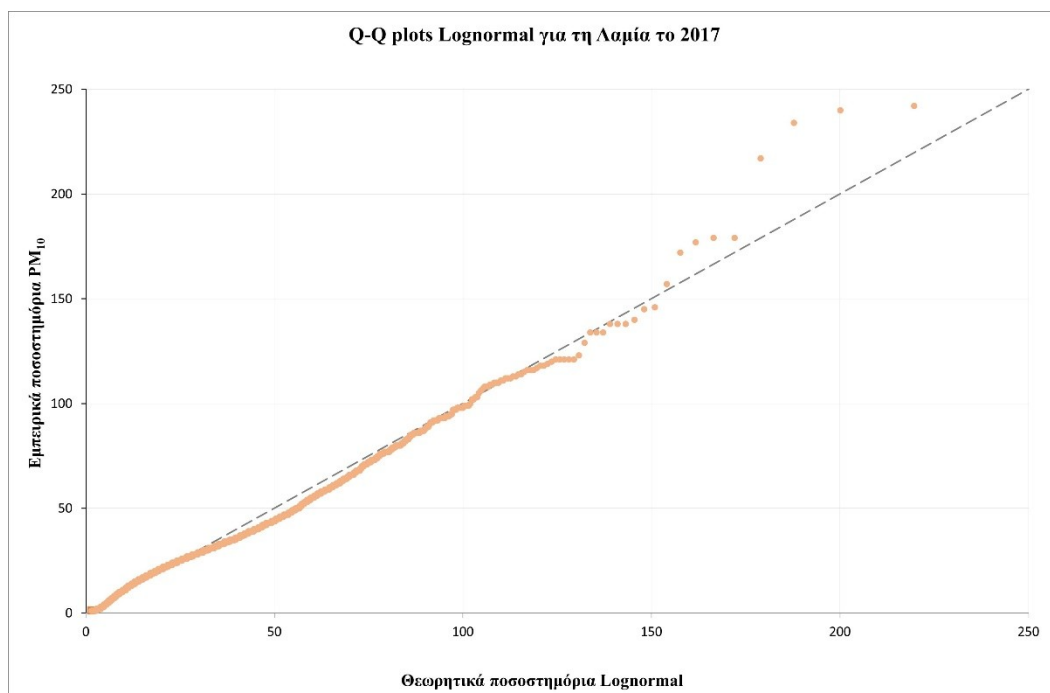


Σχήμα 4.5.4γ Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η κατανομή Weibull εμφανίζει σχετικά καλή συμφωνία μόνο στο χαμηλό και σε μέρος του μεσαίου εύρους των ποσοστημορίων. Στην αρχική περιοχή τα σημεία βρίσκονται αρκετά κοντά στη γραμμή αναφοράς, στοιχείο που δείχνει ότι το μοντέλο αποδίδει ικανοποιητικά τις μικρότερες και ορισμένες από τις συνηθέστερες συγκεντρώσεις. Στη συνέχεια, όμως, αναπτύσσεται σαφής συστηματική καμπυλότητα. Σε ένα ενδιαμέσο τμήμα τα σημεία βρίσκονται κάτω από τη διαγώνιο, υποδηλώνοντας ότι η Weibull υπερεκτιμά τα αντίστοιχα εμπειρικά ποσοστημόρια. Η τάση αυτή αντιστρέφεται περίπου στο κεντρικό τμήμα και τα σημεία αρχίζουν να μετατοπίζονται πάνω από τη γραμμή

αναφοράς. Από τα ανώτερα μεσαία ποσοστημόρια και μετά, η ανοδική απόκλιση αυξάνεται με ταχύ ρυθμό. Τα υψηλότερα εμπειρικά ποσοστημόρια υπερβαίνουν κατά πολύ τα θεωρητικά, ενώ οι τελευταίες παρατηρήσεις απομακρύνονται έντονα προς τα επάνω. Η συμπεριφορά αυτή δείχνει σημαντική υποεκτίμηση της δεξιάς ουράς. Παρότι η Weibull είχε παρουσιάσει ευνοϊκή εικόνα στα διαγνωστικά διαγράμματα και στα απλά υπόλοιπα, το Q-Q διάγραμμα αποκαλύπτει σοβαρή αδυναμία στην αναπαράσταση των υψηλών και ακραίων συγκεντρώσεων. Το εύρημα αυτό καταδεικνύει τη σημασία της συνδυαστικής αξιολόγησης: ένα μοντέλο μπορεί να αποδίδει ικανοποιητικά τις συχνότητες ή μεγάλο μέρος της ταξινομημένης σειράς, αλλά να μην αναπαράγει σωστά τα ανώτερα ποσοστημόρια.

Λογαριθμοκανονική κατανομή



Σχήμα 4.5.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η λογαριθμοκανονική κατανομή παρουσιάζει την καλύτερη συνολική ευθυγράμμιση με τη γραμμή αναφοράς μεταξύ των τεσσάρων εξεταζόμενων μοντέλων. Στο χαμηλό και στο μεγαλύτερο μέρος του μεσαίου εύρους, τα σημεία ακολουθούν στενά τη διαγώνιο και εμφανίζουν σχετικά ομαλή, σχεδόν γραμμική διάταξη.

Παρατηρούνται ορισμένες ήπιες συστηματικές αποκλίσεις στο χαμηλό και μεσαίο τμήμα, καθώς τα σημεία μετατοπίζονται ελαφρώς πάνω ή κάτω από τη γραμμή σε επιμέρους περιοχές. Ωστόσο, η ένταση και η διάρκεια των αποκλίσεων είναι σαφώς μικρότερες από εκείνες που εμφανίζονται στα διαγράμματα της κανονικής, της Gamma και της Weibull.

Η καλή ευθυγράμμιση διατηρείται μέχρι σχετικά υψηλά ποσοστημόρια. Στην περιοχή περίπου των 130-150 $\mu\text{g}/\text{m}^3$ εμφανίζονται μικρές αποκλίσεις, αλλά αρκετά σημεία παραμένουν κοντά στη γραμμή αναφοράς. Αυτό δείχνει ότι η λογαριθμοκανονική αναπαράγει αποτελεσματικότερα όχι μόνο το κύριο σώμα, αλλά και σημαντικό μέρος του ανώτερου τμήματος της κατανομής.

Στα πλέον ακραία ποσοστημόρια, τα τελευταία σημεία μετατοπίζονται αισθητά πάνω από τη διαγώνιο. Το μοτίβο αυτό δείχνει ότι και η λογαριθμοκανονική υποεκτιμά τις μεγαλύτερες παρατηρούμενες συγκεντρώσεις και δεν αναπαράγει πλήρως το ακραίο άκρο της δεξιάς ουράς.

Η απόκλιση, όμως, εμφανίζεται αργότερα και αφορά μικρότερο αριθμό σημείων σε σχέση με την κανονική και τη Weibull. Συγκριτικά με τη Gamma, η λογαριθμοκανονική διατηρεί στενότερη ευθυγράμμιση σε μεγαλύτερο εύρος ποσοστημορίων και εμφανίζει λιγότερο έντονη συστηματική καμπυλότητα. Επομένως, το διάγραμμα παρέχει την ισχυρότερη γραφική ένδειξη συνολικής κατανομικής καταλληλότητας.

Συγκριτική αξιολόγηση

Η συγκριτική εξέταση των τεσσάρων Q-Q διαγραμμάτων αναδεικνύει ουσιαστικές διαφορές τόσο στην απόδοση του κύριου σώματος όσο και στην αναπαράσταση της δεξιάς ουράς. Κοινό χαρακτηριστικό όλων των μοντέλων είναι ότι οι υψηλότερες εμπειρικές παρατηρήσεις τείνουν να υπερβαίνουν τα αντίστοιχα θεωρητικά ποσοστημόρια. Επομένως, καμία από τις εξεταζόμενες κατανομές δεν αναπαράγει πλήρως τις πλέον ακραίες συγκεντρώσεις της χρονοσειράς.

Η κανονική κατανομή παρουσιάζει τη δυσμενέστερη εικόνα. Η ισχυρή καμπυλότητα σε όλο σχεδόν το εύρος, η θεωρητική παραγωγή αρνητικών τιμών και η πολύ μεγάλη ανοδική απόκλιση στην άνω ουρά καταδεικνύουν συνολική ασυμβατότητα με τη μη αρνητική και δεξιά ασύμμετρη φύση των δεδομένων.

Η Gamma εμφανίζει σημαντική βελτίωση και προσεγγίζει αρκετά ικανοποιητικά μεγάλο τμήμα των χαμηλών και μεσαίων ποσοστημορίων. Παράλληλα, όμως, παρουσιάζει σαφές

συστηματικό πρότυπο: αρχική θετική απόκλιση, ενδιάμεση μετατόπιση κάτω από τη γραμμή και τελική ανοδική απόκλιση. Το πρότυπο αυτό δείχνει ότι η προσαρμογή της δεν είναι ομοιόμορφη σε όλη την κατανομή.

Η Weibull αποδίδει σχετικά ικανοποιητικά το χαμηλότερο τμήμα, αλλά η απόκλιση στην άνω ουρά ξεκινά νωρίτερα και αυξάνεται πολύ εντονότερα. Η μεγάλη απομάκρυνση των υψηλότερων εμπειρικών ποσοστημορίων από τη διαγώνιο αποκαλύπτει ότι το μοντέλο υποεκτιμά σημαντικά το μέγεθος των υψηλών συγκεντρώσεων. Συνεπώς, η καλή εικόνα της Weibull στα υπόλοιπα δεν συνοδεύεται από αντίστοιχα καλή αναπαράσταση των ακραίων ποσοστημορίων.

Η λογαριθμοκανονική παρουσιάζει την περισσότερο ισορροπημένη συνολική συμπεριφορά. Τα σημεία παραμένουν κοντά στη γραμμή αναφοράς σε μεγαλύτερο τμήμα της κατανομής, η συστηματική καμπυλότητα είναι περιορισμένη και η σημαντική ανοδική απόκλιση εμφανίζεται κυρίως στις ελάχιστες υψηλότερες παρατηρήσεις. Δεν αποδίδει τέλεια την ακραία ουρά, αλλά αναπαράγει αποτελεσματικότερα τον συνδυασμό κύριου σώματος και υψηλών ποσοστημορίων.

Με βάση αποκλειστικά τα Q-Q διαγράμματα, η σχετική κατάταξη είναι:

Λογαριθμοκανονική > Gamma > Weibull > Κανονική

όπου το σύμβολο > δηλώνει ισχυρότερη γραφική ένδειξη προσαρμογής ως προς τη συνολική ευθυγράμμιση των εμπειρικών και θεωρητικών ποσοστημορίων.

Η κατάταξη αυτή διαφοροποιείται από εκείνη που προέκυψε από τα απλά υπόλοιπα, στα οποία η Weibull είχε εμφανίσει μικρότερο εύρος αποκλίσεων. Η διαφοροποίηση δεν αποτελεί αντίφαση, αλλά αντανakλά το γεγονός ότι τα δύο διαγνωστικά εργαλεία αξιολογούν διαφορετικές πτυχές. Τα υπόλοιπα εξετάζουν τις αποκλίσεις των συχνοτήτων κατά μήκος της ταξινομημένης σειράς, ενώ τα Q-Q διαγράμματα αξιολογούν άμεσα την αναπαράσταση των ποσοστημορίων και είναι ιδιαίτερα αποκαλυπτικά ως προς τη συμπεριφορά των ουρών.

Επομένως, τα Q-Q διαγράμματα παρέχουν ισχυρή ένδειξη υπέρ της λογαριθμοκανονικής κατανομής, αλλά η τελική επιλογή δεν πρέπει να βασιστεί αποκλειστικά σε αυτά. Η αριθμητική αξιολόγηση μέσω του στατιστικού χ^2 θα δείξει αν η καλύτερη συνολική ευθυγράμμιση της λογαριθμοκανονικής συνοδεύεται και από μικρότερη σταθμισμένη απόκλιση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων.

4.5.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2

Μετά την ολοκλήρωση της γραφικής διερεύνησης, η σύγκριση των τεσσάρων θεωρητικών κατανομών συμπληρώθηκε με την αριθμητική αποτίμηση των αποκλίσεων τους από τις παρατηρούμενες συχνότητες. Για τον σκοπό αυτό χρησιμοποιήθηκε το στατιστικό χ^2 , σύμφωνα με τη διαδικασία που παρουσιάστηκε αναλυτικά στην αντίστοιχη ενότητα της Άμφισσας. Η επανάληψη του μαθηματικού ορισμού δεν κρίθηκε αναγκαία, καθώς στην παρούσα υποενότητα το ενδιαφέρον επικεντρώνεται στην ερμηνεία της αριθμητικής κατάταξης για τη συγκεκριμένη χρονοσειρά.

Στην περίπτωση της Λαμίας, η αριθμητική αξιολόγηση αποκτά ιδιαίτερη σημασία, επειδή τα επιμέρους γραφικά εργαλεία δεν ανέδειξαν απόλυτα ταυτόσημη σειρά κατάταξης. Τα διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων και τα διαγράμματα υπολοίπων ανέδειξαν ευνοϊκή συμπεριφορά της Weibull σε σημαντικό μέρος των δεδομένων, ενώ τα Q-Q διαγράμματα έδειξαν πληρέστερη συνολική αντιστοίχιση των εμπειρικών και θεωρητικών ποσοστημορίων από τη λογαριθμοκανονική κατανομή. Το χ^2 χρησιμοποιείται, επομένως, για να διερευνηθεί ποιο από τα δύο μοντέλα εμφανίζει μικρότερη συνολική σταθμισμένη απόκλιση, όταν οι επιμέρους αποκλίσεις συνεκτιμώνται στο ίδιο σύστημα κλάσεων.

Οι τιμές του χ^2 λειτουργούν ως σχετικό κριτήριο σύγκρισης μεταξύ των τεσσάρων μοντέλων. Μικρότερη τιμή αντιστοιχεί σε μικρότερη συνολική σταθμισμένη απόκλιση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων και, συνεπώς, σε ισχυρότερη αριθμητική ένδειξη προσαρμογής. Η κατάταξη δεν ερμηνεύεται ως αυτοτελής απόφαση αποδοχής ή απόρριψης μιας θεωρητικής κατανομής, αλλά ως ποσοτική συνιστώσα της συνολικής διαγνωστικής αξιολόγησης.

Η χρήση κοινού συστήματος κλάσεων και του ίδιου συνόλου παρατηρήσεων διασφαλίζει ότι οι τιμές των τεσσάρων κατανομών είναι άμεσα συγκρίσιμες. Παράλληλα, πρέπει να λαμβάνεται υπόψη ότι το χ^2 δεν αξιολογεί όλες τις περιοχές της εμπειρικής κατανομής με τον ίδιο τρόπο, αλλά σταθμίζει τις τετραγωνισμένες αποκλίσεις ως προς τις αντίστοιχες αναμενόμενες συχνότητες. Για τον λόγο αυτό, η αριθμητική κατάταξη μπορεί να διαφοροποιείται από την οπτική εντύπωση των απλών υπολοίπων ή των Q-Q διαγραμμάτων, χωρίς αυτό να συνιστά αντίφαση.

Η συνδυαστική ερμηνεία των αποτελεσμάτων επιτρέπει, επομένως, να διακριθεί η τοπική καλή συμπεριφορά ενός μοντέλου σε συγκεκριμένο τμήμα των δεδομένων από τη

συνολική του απόδοση σε ολόκληρη την κατανομή. Με αυτό το σκεπτικό, οι τιμές του χ^2 που ακολουθούν χρησιμοποιούνται για να ποσοτικοποιήσουν και να συμπληρώσουν τις γραφικές ενδείξεις, όχι για να τις αντικαταστήσουν.

Κατανομή	χ^2	Θέση κατάταξης
Lognormal	227,85	1
Weibull	346,07	2
Gamma	574,16	3
Normal	1568,23	4

Πίνακας 4.5.5 Τιμές του στατιστικού χ^2 και σχετική κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία το 2017.

Η λογαριθμοκανονική κατανομή καταγράφει τη μικρότερη τιμή χ^2 , ίση με 227,85, και καταλαμβάνει την πρώτη θέση. Το αποτέλεσμα δείχνει ότι, στο πλαίσιο των κοινών κλάσεων που χρησιμοποιήθηκαν, το συγκεκριμένο μοντέλο επιτυγχάνει τη μικρότερη συνολική σταθμισμένη απόκλιση από τις εμπειρικές συχνότητες.

Η Weibull ακολουθεί με $\chi^2 = 346,07$. Η απόσταση των 118,22 μονάδων από τη λογαριθμοκανονική δείχνει ότι η διαφορά μεταξύ των δύο μοντέλων δεν είναι αμελητέα. Παρότι η Weibull εμφάνισε συνεκτική οπτική διάταξη στα διαγνωστικά διαγράμματα και μικρότερο περίπου εύρος απλών υπολοίπων, η υψηλότερη τιμή χ^2 υποδηλώνει ότι οι αποκλίσεις της είναι συνολικά δυσμενέστερα κατανεμημένες στις κλάσεις που συμμετέχουν στον υπολογισμό.

Η Gamma κατατάσσεται τρίτη, με $\chi^2 = 574,16$. Η τιμή της είναι αισθητά υψηλότερη από εκείνη της Weibull και δείχνει σαφώς ασθενέστερη αριθμητική προσαρμογή. Το εύρημα συμφωνεί με τα οργανωμένα πρότυπα που παρατηρήθηκαν στα υπόλοιπά της και με τη μεταβαλλόμενη θέση των σημείων γύρω από τη γραμμή αναφοράς στα διαγνωστικά και Q-Q διαγράμματα.

Η κανονική κατανομή εμφανίζει τη μεγαλύτερη τιμή, $\chi^2 = 1.568,23$, και κατατάσσεται τελευταία με μεγάλη διαφορά. Η επίδοσή της επιβεβαιώνει αριθμητικά την αδυναμία ενός συμμετρικού μοντέλου να αποδώσει τη συγκέντρωση των παρατηρήσεων στις χαμηλές και μεσαίες τιμές, καθώς και τη θετική ασυμμετρία και τη βαριά δεξιά ουρά της χρονοσειράς.

Η λογαριθμοκανονική παρουσιάζει τιμή χ^2 περίπου κατά 34% χαμηλότερη από τη Weibull, κατά 60% χαμηλότερη από τη Gamma και κατά 85% χαμηλότερη από την κανονική κατανομή. Οι διαφορές αυτές δείχνουν ότι η πρώτη θέση της δεν προκύπτει από οριακή μόνο υπεροχή, αλλά από σαφώς ευνοϊκότερη συνολική αριθμητική συμπεριφορά.

Η περίπτωση της Λαμίας παρουσιάζει ιδιαίτερο ερμηνευτικό ενδιαφέρον, επειδή τα επιμέρους διαγνωστικά δεν κατέταξαν όλες τις κατανομές με ακριβώς την ίδια σειρά. Η Weibull εμφάνισε ισχυρή εικόνα στα απλά υπόλοιπα και σχετικά οργανωμένη πορεία στα διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων. Αντίθετα, η λογαριθμοκανονική παρουσίασε καλύτερη συνολική ευθυγράμμιση των εμπειρικών και θεωρητικών ποσοστημορίων και σαφώς χαμηλότερη τιμή χ^2 .

Η διαφοροποίηση αυτή δεν συνιστά ασυνέπεια. Τα διαγράμματα υπολοίπων αξιολογούν κυρίως το πρόσημο, το εύρος και τη διαδοχική δομή των αποκλίσεων, ενώ τα Q-Q διαγράμματα εξετάζουν άμεσα την αντιστοίχιση των ποσοστημορίων και τη συμπεριφορά των ουρών. Το χ^2 , από την άλλη πλευρά, συμπυκνώνει σε μία τιμή τις σταθμισμένες αποκλίσεις των συχνοτήτων ανά κλάση. Επομένως, ένα μοντέλο μπορεί να παρουσιάζει μικρότερο μέγιστο απλό υπόλοιπο, αλλά ταυτόχρονα μεγαλύτερη συνολική τιμή χ^2 .

Στη συγκεκριμένη χρονοσειρά, η Weibull φαίνεται να αποδίδει αποτελεσματικά εκτεταμένο τμήμα του κεντρικού εύρους, όμως το Q-Q διάγραμμα αποκάλυψε έντονη υποεκτίμηση των υψηλότερων ποσοστημορίων. Η λογαριθμοκανονική, παρότι δεν εξαλείφει όλες τις τοπικές αποκλίσεις, διατηρεί περισσότερο ισορροπημένη συμπεριφορά μεταξύ κύριου σώματος και άνω ουράς. Η χαμηλότερη τιμή χ^2 ενισχύει αυτή την εικόνα, δείχνοντας ότι η συνολική κατανομή των αποκλίσεων της είναι ευνοϊκότερη.

Με βάση το αριθμητικό κριτήριο, η σχετική κατάταξη είναι:

Λογαριθμοκανονική < Weibull < Gamma < Κανονική

όπου η μικρότερη τιμή χ^2 αντιστοιχεί σε καλύτερη σχετική προσαρμογή.

Η πρώτη θέση της λογαριθμοκανονικής δεν σημαίνει ότι το μοντέλο αναπαράγει τέλεια κάθε τμήμα της εμπειρικής κατανομής. Στη σύγκριση ανά κλάση εμφανίστηκε αισθητή υποεκτίμηση στην περιοχή $20-30 \mu\text{g}/\text{m}^3$ και υπερεκτίμηση σε ορισμένες υψηλότερες κλάσεις, ενώ στο Q-Q διάγραμμα οι ακραίες εμπειρικές τιμές εξακολουθούν να υπερβαίνουν τα αντίστοιχα θεωρητικά ποσοστημόρια. Η αριθμητική της υπεροχή αφορά, επομένως, τη συνολική απόδοση μεταξύ των τεσσάρων εξεταζόμενων μοντέλων και όχι απόλυτη ταύτιση με τα δεδομένα.

Αντίστοιχα, η δεύτερη θέση της Weibull δεν αναιρεί την καλή συμπεριφορά που εμφάνισε στα υπόλοιπα. Δείχνει, όμως, ότι η αποτελεσματική προσαρμογή της σε μεγάλο μέρος της ταξινομημένης σειράς δεν επαρκεί για να αντισταθμίσει τις αστοχίες της στην άνω ουρά και τις σταθμισμένες αποκλίσεις στις επιμέρους κλάσεις.

Συνολικά, η αριθμητική αξιολόγηση μετατοπίζει καθαρά το βάρος της τελικής επιλογής προς τη λογαριθμοκανονική κατανομή. Η Weibull παραμένει ισχυρή εναλλακτική, κυρίως ως προς τη συμπεριφορά των υπολοίπων και την απόδοση στο κεντρικό εύρος. Η Gamma παρουσιάζει εμφανώς ασθενέστερη συνολική προσαρμογή, ενώ η κανονική κατανομή υπολείπεται σημαντικά.

Η τελική επιλογή για τη Λαμία δεν βασίζεται, συνεπώς, σε μηχανική αποδοχή της μικρότερης τιμής χ^2 , αλλά στη σύνθεση τριών διαφορετικών διαστάσεων της προσαρμογής: της συμφωνίας των συχνοτήτων, της δομής των υπολοίπων και της αντιστοίχισης των ποσοστημορίων. Η ολοκληρωμένη αυτή αποτίμηση παρουσιάζεται στην επόμενη υποενότητα.

4.5.6 Επιλογή καταλληλότερης κατανομής για τη Λαμία 2017

Η τελική επιλογή της θεωρητικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λαμία κατά το έτος 2017 βασίστηκε στη συνδυαστική αξιολόγηση όλων των γραφικών και αριθμητικών κριτηρίων. Η διαδικασία αυτή ήταν ιδιαίτερα σημαντική, καθώς τα επιμέρους διαγνωστικά εργαλεία δεν κατέταξαν τις εξεταζόμενες κατανομές με απολύτως ταυτόσημο τρόπο, αλλά ανέδειξαν διαφορετικές πτυχές της ποιότητας προσαρμογής.

Η κανονική κατανομή εμφάνισε με συνέπεια τη χαμηλότερη καταλληλότητα. Στα διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων παρουσίασε μεγάλες και συστηματικές αποκλίσεις, ενώ στα διαγράμματα υπολοίπων σχημάτισε έντονο μη

τυχαίο πρότυπο. Παράλληλα, το Q-Q διάγραμμα ανέδειξε σαφή καμπυλότητα και σημαντική υποεκτίμηση των υψηλών εμπειρικών ποσοστημορίων. Η πολύ υψηλή τιμή του στατιστικού χ^2 , ίση με 1.568,23, επιβεβαίωσε αριθμητικά την ασθενή συνολική απόδοσή της.

Η Gamma αποτέλεσε ουσιαστική βελτίωση σε σχέση με την κανονική κατανομή, καθώς ανταποκρίθηκε καλύτερα στη μη αρνητική και θετικά ασύμμετρη μορφή των δεδομένων. Ωστόσο, η προσαρμογή της δεν ήταν ομοιόμορφη. Στα διαγνωστικά διαγράμματα παρατηρήθηκε αυξημένη διασπορά στις μεγαλύτερες αναμενόμενες συχνότητες, ενώ στα υπόλοιπα εμφανίστηκαν οργανωμένες περιοχές υπερεκτίμησης και υποεκτίμησης. Η τιμή $\chi^2 = 574,16$ την κατέταξε στην τρίτη θέση.

Η Weibull παρουσίασε ιδιαίτερα ανταγωνιστική συμπεριφορά σε αρκετά στάδια της ανάλυσης. Στα διαγνωστικά διαγράμματα εμφάνισε συνεκτική διάταξη σημείων, ενώ στα διαγράμματα υπολοίπων κατέγραψε μικρότερο περίπου εύρος ακραίων αποκλίσεων και ταχύτερη σταθεροποίηση γύρω από τη γραμμή μηδενός. Τα ευρήματα αυτά δείχνουν ότι αποδίδει αποτελεσματικά σημαντικό μέρος του κεντρικού εύρους των συγκεντρώσεων.

Η εικόνα της Weibull διαφοροποιήθηκε, όμως, στο Q-Q διάγραμμα. Παρότι η συμφωνία ήταν ικανοποιητική στα χαμηλά και σε μέρος των μεσαίων ποσοστημορίων, η απόκλιση από τη γραμμή αναφοράς άρχισε σχετικά νωρίς στο ανώτερο τμήμα και εξελίχθηκε σε έντονη υποεκτίμηση των υψηλών παρατηρήσεων. Η δεύτερη θέση της στο αριθμητικό κριτήριο, με $\chi^2 = 346,07$, επιβεβαιώνει ότι αποτελεί ισχυρή εναλλακτική, αλλά ότι η συνολική απόκλισή της παραμένει μεγαλύτερη από εκείνη της λογαριθμοκανονικής.

Η λογαριθμοκανονική κατανομή δεν υπερείχε σε κάθε επιμέρους γραφικό κριτήριο. Στα διαγράμματα υπολοίπων, για παράδειγμα, εμφάνισε εντονότερες αρχικές αποκλίσεις από τη Weibull, ενώ και στη σύγκριση ανά κλάση δεν απέδωσε με την ίδια ακρίβεια όλες τις περιοχές συγκέντρωσης. Παρά τις τοπικές αυτές αστοχίες, παρουσίασε την πιο ισορροπημένη συνολική συμπεριφορά.

Στο Q-Q διάγραμμα διατήρησε την πλησιέστερη ευθυγράμμιση με τη γραμμή αναφοράς σε μεγαλύτερο εύρος ποσοστημορίων και απέδωσε αποτελεσματικότερα τόσο το κύριο σώμα όσο και σημαντικό μέρος του ανώτερου τμήματος της εμπειρικής κατανομής. Οι αποκλίσεις έγιναν εντονότερες κυρίως στις ελάχιστες υψηλότερες παρατηρήσεις, γεγονός που δείχνει ότι, αν και δεν αναπαράγει πλήρως το ακραίο άκρο της δεξιάς ουράς, εμφανίζει καλύτερη συνολική κατανομική αντιστοίχιση από τα υπόλοιπα μοντέλα.

Η υπεροχή αυτή ενισχύθηκε από την αριθμητική αξιολόγηση.

Η λογαριθμοκανονική κατέγραψε τη μικρότερη τιμή του στατιστικού χ^2 :

$$\chi^2=227,85$$

έναντι 346,07 για τη Weibull, 574,16 για τη Gamma και 1.568,23 για την κανονική κατανομή. Η τιμή της ήταν περίπου κατά 34% χαμηλότερη από εκείνη της Weibull, γεγονός που δείχνει ότι το αριθμητικό της πλεονέκτημα δεν είναι οριακό.

Η διαφοροποίηση μεταξύ των γραφικών και αριθμητικών αποτελεσμάτων δεν συνιστά αντίφαση. Η Weibull εμφάνισε καλύτερη τοπική συμπεριφορά σε ορισμένα τμήματα της ταξινομημένης σειράς, ενώ η λογαριθμοκανονική υπερείχε ως προς τη συνολική αντιστοίχιση των ποσοστημορίων και τη σταθμισμένη απόκλιση των συχνοτήτων. Επομένως, τα δύο μοντέλα παρουσιάζουν διαφορετικά πλεονεκτήματα, αλλά η λογαριθμοκανονική επιτυγχάνει πιο ισορροπημένη συνολική απόδοση.

Η συνεκτίμηση των αποτελεσμάτων μετατοπίζει τελικά το βάρος της επιλογής προς τη λογαριθμοκανονική κατανομή. Παρότι η Weibull παρουσίασε καλύτερη τοπική συμπεριφορά σε ορισμένα γραφικά κριτήρια, η λογαριθμοκανονική συνδύασε πληρέστερη αντιστοίχιση των ποσοστημορίων με τη μικρότερη συνολική σταθμισμένη απόκλιση. Παράλληλα, διατήρησε ικανοποιητική προσαρμογή τόσο στην περιοχή όπου συγκεντρώνεται η κύρια μάζα των παρατηρήσεων όσο και σε σημαντικό μέρος των υψηλότερων συγκεντρώσεων. Για τον λόγο αυτό επιλέγεται ως η καταλληλότερη, σε συγκριτικό επίπεδο, κατανομή για τη Λαμία το 2017.

Η επιλογή της λογαριθμοκανονικής είναι συμβατή και με τα βασικά μορφολογικά χαρακτηριστικά της εμπειρικής κατανομής, η οποία χαρακτηρίζεται από μη αρνητικές τιμές, θετική ασυμμετρία, συκέντρωση μεγάλου αριθμού παρατηρήσεων στις χαμηλές και μεσαίες συγκεντρώσεις και παρουσία περιορισμένου αριθμού πολύ υψηλών τιμών. Η θεωρητική μορφή της λογαριθμοκανονικής επιτρέπει την αποτελεσματικότερη αναπαράσταση αυτών των χαρακτηριστικών από τη συμμετρική κανονική κατανομή και, στη συγκεκριμένη χρονοσειρά, συνολικά καλύτερα από τις Gamma και Weibull.

Πρέπει, πάντως, να επισημανθεί ότι η επιλογή αφορά αποκλειστικά τη χρονοσειρά της Λαμίας για το έτος 2017 και το συγκεκριμένο σύνολο εξεταζόμενων κατανομών. Δεν συνεπάγεται ότι η λογαριθμοκανονική αποτελεί καθολικά κατάλληλο μοντέλο για κάθε χρονική περίοδο ή γεωγραφική περιοχή. Επιπλέον, η πρώτη θέση της είναι συγκριτική και δεν σημαίνει πλήρη αναπαράσταση όλων των εμπειρικών χαρακτηριστικών.

Το αποτέλεσμα για τη Λαμία συμφωνεί με την επιλογή που προέκυψε προηγουμένως για την Άμφισσα. Η επανεμφάνιση της λογαριθμοκανονικής ως καταλληλότερης κατανομής σε δύο διαφορετικές ετήσιες χρονοσειρές αποτελεί ένδειξη ότι μπορεί να ανταποκρίνεται αποτελεσματικά σε δεδομένα ωριαίων συγκεντρώσεων PM_{10} με σαφή δεξιά ασυμμετρία. Ωστόσο, οι δύο περιπτώσεις δεν επαρκούν για γενίκευση και το συμπέρασμα πρέπει να αξιολογηθεί περαιτέρω στις υπόλοιπες πόλεις.

4.6 Προσαρμογή θεωρητικών κατανομών στη Χαλκίδα το 2017

Η ανάλυση συνεχίζεται με τη χρονοσειρά των ωριαίων συγκεντρώσεων PM_{10} στη Χαλκίδα για το έτος 2017. Η συγκεκριμένη περίπτωση παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς τα περιγραφικά στατιστικά και η εμπειρική κατανομή ανέδειξαν υψηλότερο επίπεδο συγκεντρώσεων, αυξημένη μεταβλητότητα και εντονότερη θετική ασυμμετρία σε σύγκριση με τις προηγούμενες πόλεις.

Η μορφή αυτή υποδηλώνει ότι η εμπειρική κατανομή χαρακτηρίζεται όχι μόνο από συγκέντρωση μεγάλου αριθμού παρατηρήσεων στις χαμηλές και μεσαίες τιμές, αλλά και από εκτεταμένη δεξιά ουρά, στην οποία περιλαμβάνονται λιγότερες αλλά ιδιαίτερα υψηλές συγκεντρώσεις. Κατά συνέπεια, η αξιολόγηση των θεωρητικών μοντέλων αποκτά αυξημένες απαιτήσεις, καθώς μια ικανοποιητική προσαρμογή πρέπει να αποδίδει ταυτόχρονα το κύριο σώμα της κατανομής, τη μετάβαση προς τις υψηλότερες τιμές και, στον βαθμό που είναι δυνατό, τις ακραίες παρατηρήσεις.

Στην παρούσα ενότητα εφαρμόζεται η ίδια διαδικασία προσαρμογής για την κανονική, τη Gamma, τη Weibull και τη λογαριθμοκανονική κατανομή. Η συγκριτική αξιολόγηση βασίζεται στις παρατηρούμενες και αναμενόμενες συχνότητες ανά κλάση συγκέντρωσης, στα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων, στα διαγράμματα υπολοίπων, στα Q-Q διαγράμματα και στην αριθμητική κατάταξη μέσω του στατιστικού χ^2 .

Η εφαρμογή πολλαπλών και συμπληρωματικών κριτηρίων είναι ιδιαίτερα σημαντική για τη Χαλκίδα. Ένα μοντέλο μπορεί να περιγράψει ικανοποιητικά την περιοχή όπου συγκεντρώνεται η πλειονότητα των παρατηρήσεων, αλλά να υποεκτιμά σημαντικά την άνω ουρά. Αντίστροφα, μια κατανομή με μεγαλύτερη ευελιξία στις υψηλές τιμές μπορεί

να εμφανίζει αστοχίες στις κυριότερες κλάσεις συγκέντρωσης. Επομένως, η τελική επιλογή δεν θα βασιστεί σε ένα μεμονωμένο διάγραμμα ή αριθμητικό μέτρο, αλλά στη συνολική ισορροπία της προσαρμογής.

Η Χαλκίδα αποτελεί έτσι μια περισσότερο απαιτητική δοκιμασία για τις τέσσερις θεωρητικές κατανομές. Η σύγκρισή τους θα δείξει σε ποιον βαθμό μπορούν να αναπαραστήσουν μια χρονοσειρά με υψηλή διασπορά, έντονη δεξιά ασυμμετρία και παρουσία πολύ υψηλών ωριαίων συγκεντρώσεων PM_{10} , χωρίς η καλή απόδοση σε ένα τμήμα της κατανομής να αποκρύπτει σημαντικές αποκλίσεις σε κάποιο άλλο.

4.6.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων

Η σύγκριση των παρατηρούμενων με τις θεωρητικά αναμενόμενες συχνότητες αποτελεί το πρώτο στάδιο αξιολόγησης της προσαρμογής των τεσσάρων θεωρητικών κατανομών στις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας για το έτος 2017. Στα Σχήματα 4.6.1α-δ παρουσιάζονται, για κοινές κλάσεις συγκέντρωσης, οι εμπειρικές συχνότητες και οι αντίστοιχες τιμές που προκύπτουν από την κανονική, τη Gamma, τη Weibull και τη λογαριθμοκανονική κατανομή.

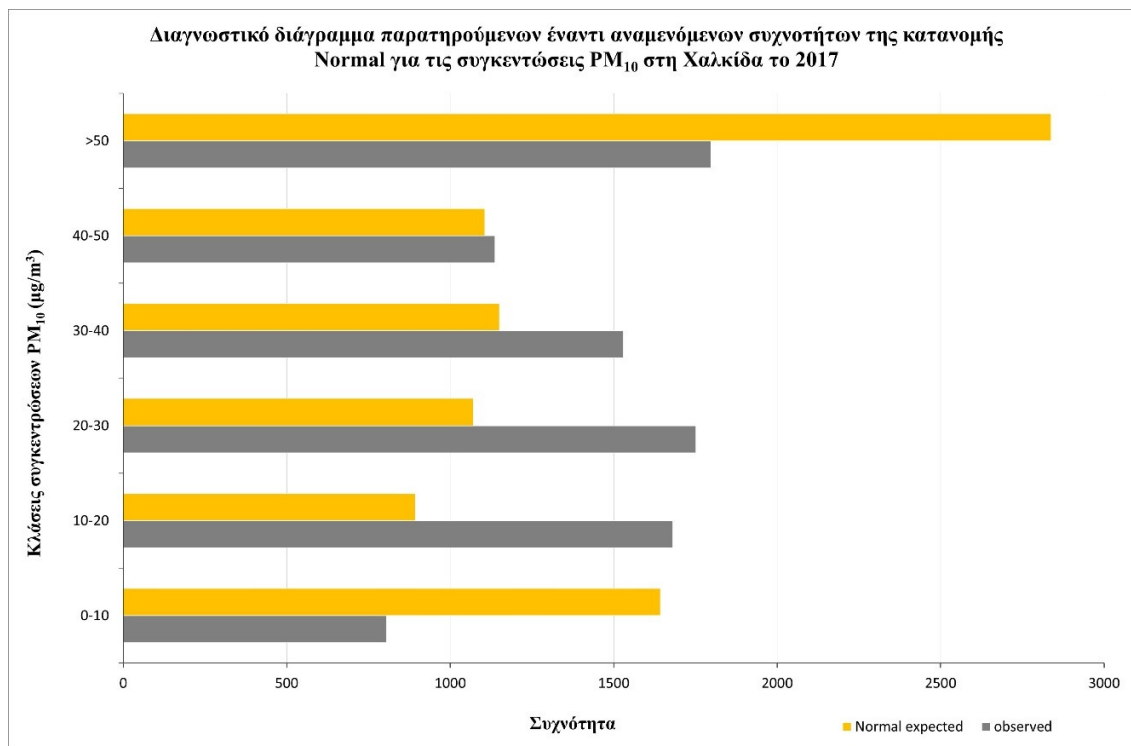
Η συγκεκριμένη απεικόνιση επιτρέπει τον άμεσο εντοπισμό των κλάσεων στις οποίες κάθε μοντέλο προσεγγίζει ικανοποιητικά τα δεδομένα, καθώς και εκείνων στις οποίες εμφανίζει υποεκτίμηση ή υπερεκτίμηση. Όταν η αναμενόμενη συχνότητα είναι μικρότερη από την παρατηρούμενη, το μοντέλο υποεκτιμά τη συγκεκριμένη κλάση. Αντίθετα, μεγαλύτερη αναμενόμενη συχνότητα αντιστοιχεί σε υπερεκτίμηση.

Στην περίπτωση της Χαλκίδας, η αξιολόγηση παρουσιάζει αυξημένο ενδιαφέρον, επειδή σημαντικό μέρος των παρατηρήσεων κατανέμεται στις μεσαίες και υψηλότερες κλάσεις, ενώ η κατηγορία άνω των $50 \mu g/m^3$ συγκεντρώνει επίσης μεγάλο αριθμό ωριαίων τιμών. Ένα κατάλληλο θεωρητικό μοντέλο πρέπει, επομένως, να αποδίδει όχι μόνο τις χαμηλές και μεσαίες συγκεντρώσεις, αλλά και τη σημαντική εμπειρική μάζα στο δεξιό τμήμα της κατανομής.

Η παρούσα σύγκριση παρέχει μια πρώτη συνοπτική εικόνα και δεν επαρκεί από μόνη της για την τελική επιλογή μοντέλου. Οι έξι κλάσεις συμπυκνώνουν μεγάλο αριθμό διαφορετικών τιμών, ενώ η τελευταία κλάση είναι ανοικτή και περιλαμβάνει όλες τις

συγκεντρώσεις άνω των $50 \mu\text{g}/\text{m}^3$. Συνεπώς, η συμπεριφορά των μοντέλων μέσα στην άνω ουρά δεν μπορεί να αξιολογηθεί πλήρως μόνο από τα συγκεκριμένα ραβδογράμματα.

Κανονική κατανομή



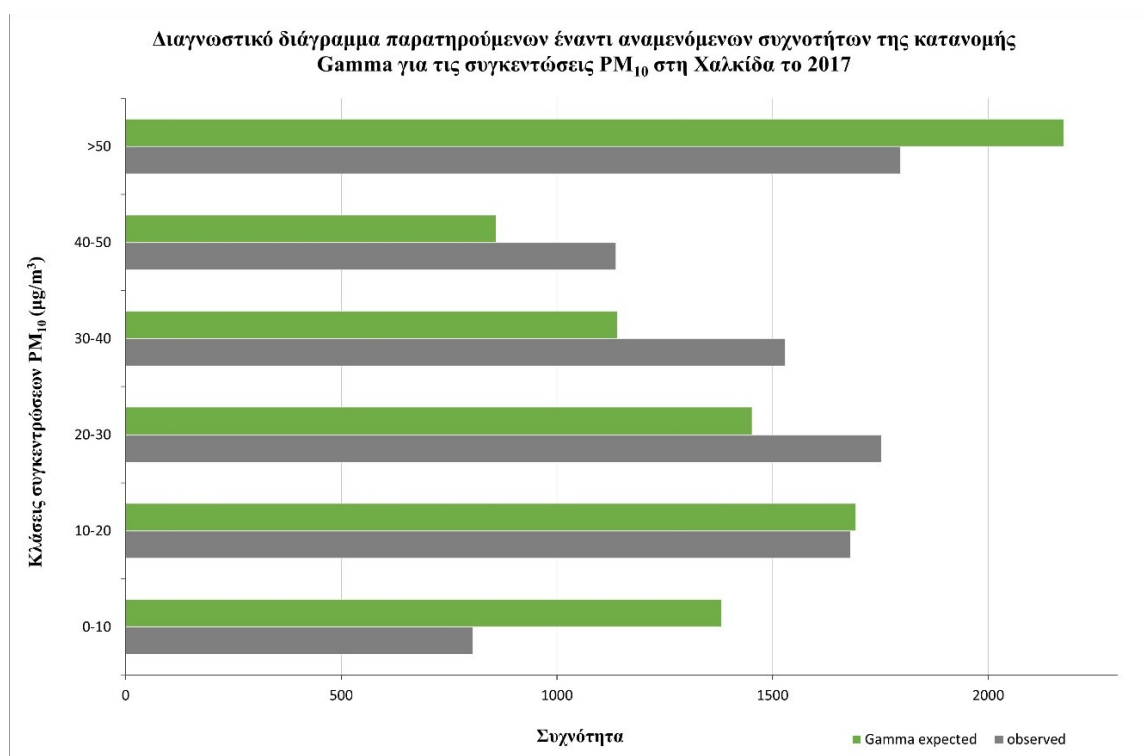
Σχήμα 4.6.1α Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η κανονική κατανομή εμφανίζει τη λιγότερο ικανοποιητική οπτική προσαρμογή μεταξύ των τεσσάρων εξεταζόμενων μοντέλων. Στην κλάση $0-10 \mu\text{g}/\text{m}^3$ υπερεκτιμά σημαντικά την παρατηρούμενη συχνότητα. Αντίθετα, στις κλάσεις $10-20$, $20-30$ και $30-40 \mu\text{g}/\text{m}^3$ οι αναμενόμενες συχνότητες είναι αισθητά χαμηλότερες από τις εμπειρικές, γεγονός που δείχνει συστηματική υποεκτίμηση του κύριου σώματος της κατανομής.

Στην κλάση $40-50 \mu\text{g}/\text{m}^3$ η απόκλιση είναι σχετικά περιορισμένη, με την κανονική κατανομή να υποεκτιμά ελαφρώς την παρατηρούμενη συχνότητα. Ωστόσο, στην ανοικτή κλάση άνω των $50 \mu\text{g}/\text{m}^3$ η θεωρητική συχνότητα υπερβαίνει κατά πολύ την παρατηρούμενη. Επομένως, το μοντέλο μετατοπίζει υπερβολικά μεγάλη θεωρητική μάζα τόσο προς την κατώτερη όσο και προς την ανώτερη κλάση, ενώ ταυτόχρονα υποεκτιμά τις ενδιάμεσες περιοχές.

Η διαδοχή αυτή δεν αποτελεί αποτέλεσμα μεμονωμένης αστοχίας, αλλά συστηματικού προτύπου. Η κανονική κατανομή δεν αναπαράγει σωστά ούτε τη θέση της κύριας εμπειρικής μάζας ούτε την κατανομή των παρατηρήσεων προς τις υψηλότερες συγκεντρώσεις. Η συμμετρική μορφή της δεν ανταποκρίνεται στην έντονα δεξιά ασύμμετρη δομή των δεδομένων της Χαλκίδας και, με βάση το παρόν κριτήριο, αποτελεί τη λιγότερο κατάλληλη υποψήφια κατανομή.

Κατανομή Gamma



Σχήμα 4.6.1β. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

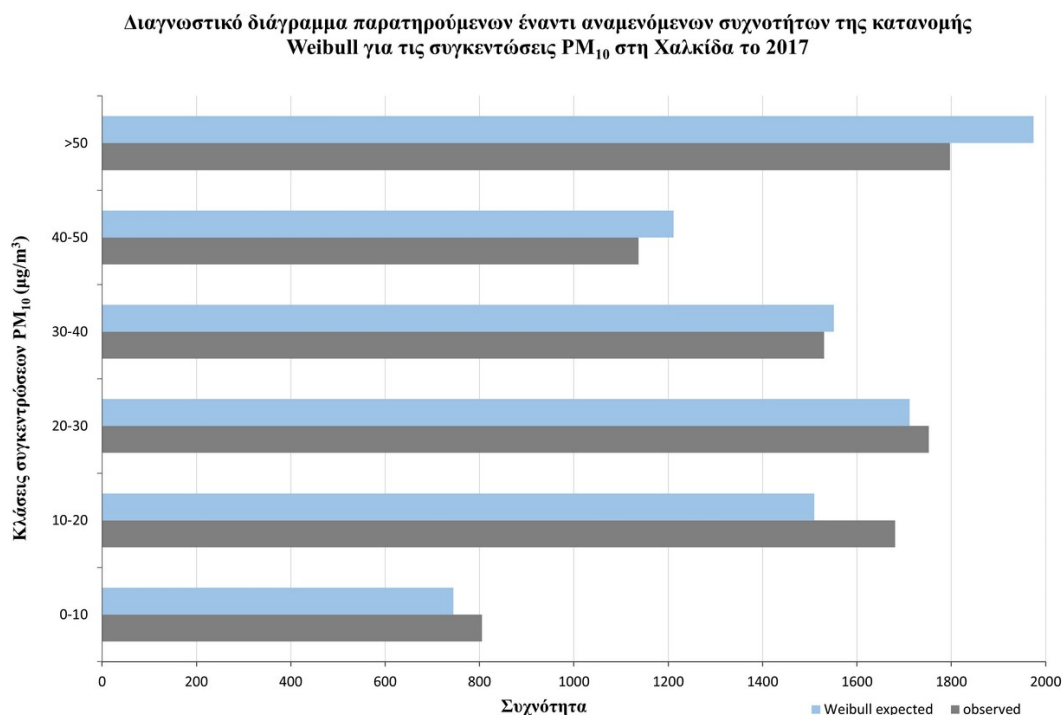
Η κατανομή Gamma παρουσιάζει ουσιαστική βελτίωση έναντι της κανονικής, κυρίως επειδή μπορεί να περιγράψει μη αρνητικά και θετικά ασύμμετρα δεδομένα. Η καλύτερη συμφωνία εμφανίζεται στην κλάση $10-20 \mu g/m^3$, όπου η αναμενόμενη συχνότητα προσεγγίζει πολύ στενά την παρατηρούμενη. Το εύρημα αυτό δείχνει ότι η Gamma αποδίδει αποτελεσματικά μία από τις βασικές περιοχές συγκέντρωσης.

Η εικόνα δεν είναι εξίσου καλή στις υπόλοιπες κλάσεις. Στην περιοχή 0-10 $\mu\text{g}/\text{m}^3$ παρατηρείται έντονη υπερεκτίμηση, καθώς η θεωρητική συχνότητα είναι πολύ μεγαλύτερη από την εμπειρική. Αντίθετα, στις κλάσεις 20-30, 30-40 και 40-50 $\mu\text{g}/\text{m}^3$ η Gamma υποεκτιμά συστηματικά τις παρατηρούμενες συχνότητες. Οι αποκλίσεις είναι ιδιαίτερα εμφανείς στις κλάσεις 20-30 και 30-40 $\mu\text{g}/\text{m}^3$, οι οποίες περιλαμβάνουν σημαντικό μέρος της εμπειρικής κατανομής.

Στην κλάση άνω των 50 $\mu\text{g}/\text{m}^3$ η Gamma υπερεκτιμά επίσης την παρατηρούμενη συχνότητα. Το συνολικό πρότυπο δείχνει ότι το μοντέλο αποδίδει υπερβολικά μεγάλη πιθανότητα στις δύο ακραίες συγκεντρωτικές κατηγορίες, ενώ αφαιρεί θεωρητική μάζα από αρκετές ενδιάμεσες κλάσεις.

Συνεπώς, η Gamma περιγράφει αποτελεσματικότερα τη γενική ασύμμετρη μορφή των δεδομένων από την κανονική κατανομή, αλλά η προσαρμογή της δεν είναι ισορροπημένη σε όλο το εύρος. Η πολύ καλή απόδοση στην κλάση 10-20 $\mu\text{g}/\text{m}^3$ δεν επαρκεί για να αντισταθμίσει τις μεγαλύτερες αποκλίσεις στις χαμηλότερες, μεσαίες και υψηλότερες κατηγορίες.

Κατανομή Weibull



Σχήμα 4.6.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η Weibull εμφανίζει την περισσότερο ισορροπημένη οπτική προσαρμογή στο συγκεκριμένο διαγνωστικό στάδιο. Οι αναμενόμενες συχνότητες ακολουθούν σε σημαντικό βαθμό τη γενική κατανομή των εμπειρικών τιμών, ενώ οι αποκλίσεις είναι μικρότερες και περισσότερο ομοιόμορφα κατανεμημένες από εκείνες της κανονικής και της Gamma.

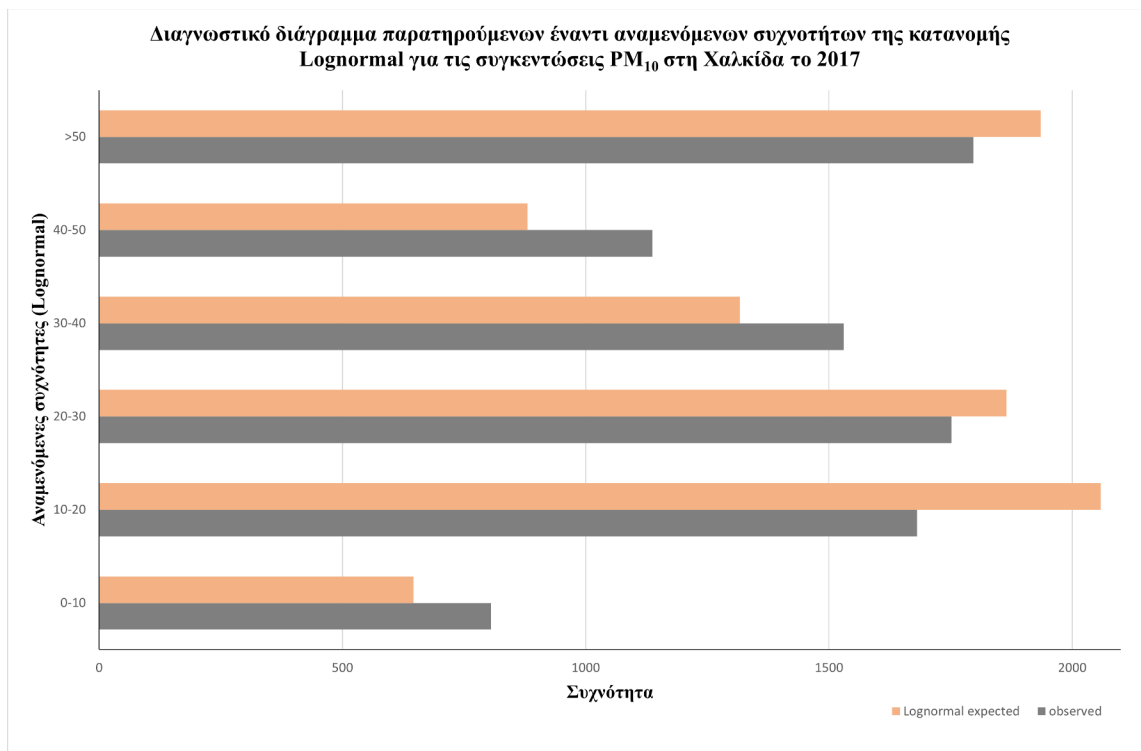
Στις κλάσεις 0-10 και 10-20 $\mu\text{g}/\text{m}^3$ η Weibull υποεκτιμά τις παρατηρούμενες συχνότητες. Η απόκλιση στην πρώτη κλάση είναι σχετικά περιορισμένη, ενώ στην κλάση 10-20 $\mu\text{g}/\text{m}^3$ γίνεται περισσότερο εμφανής. Επομένως, το μοντέλο δεν αναπαράγει με απόλυτη ακρίβεια την εμπειρική μάζα στο χαμηλό τμήμα της κατανομής.

Στην κλάση 20-30 $\mu\text{g}/\text{m}^3$ η συμφωνία είναι ιδιαίτερα καλή, με μικρή μόνο υποεκτίμηση. Αντίστοιχα, στην κλάση 30-40 $\mu\text{g}/\text{m}^3$ οι θεωρητικές και οι παρατηρούμενες συχνότητες βρίσκονται πολύ κοντά, με τη Weibull να παρουσιάζει ελαφρά υπερεκτίμηση. Η καλή απόδοση στις δύο αυτές κλάσεις είναι σημαντική, καθώς σε αυτές συγκεντρώνεται μεγάλο μέρος των εμπειρικών παρατηρήσεων.

Στην κλάση 40-50 $\mu\text{g}/\text{m}^3$ παρατηρείται επίσης υπερεκτίμηση, η οποία είναι περισσότερο αισθητή αλλά παραμένει μικρότερη από τις αντίστοιχες αποκλίσεις άλλων μοντέλων. Στην ανοικτή κλάση άνω των 50 $\mu\text{g}/\text{m}^3$ η Weibull αποδίδει μεγαλύτερη συχνότητα από την παρατηρούμενη, χωρίς όμως η διαφορά να λαμβάνει την έκταση που εμφανίζεται στην κανονική ή στη Gamma.

Η Weibull δεν επιτυγχάνει πλήρη ταύτιση σε καμία περιοχή, παρουσιάζει όμως καλή ισορροπία μεταξύ των επιμέρους αποκλίσεων. Δεν εμφανίζει μία μόνο κλάση με εξαιρετική συμφωνία και ταυτόχρονα πολύ μεγάλες αστοχίες στις υπόλοιπες, αλλά διατηρεί σχετικά σταθερή απόδοση στο μεγαλύτερο μέρος του εξεταζόμενου εύρους. Για τον λόγο αυτό αναδεικνύεται, με βάση αποκλειστικά τη σύγκριση ανά κλάση, ως η ισχυρότερη υποψήφια κατανομή.

Λογαριθμοκανονική κατανομή



Σχήμα 4.6.1δ Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η λογαριθμοκανονική κατανομή αποδίδει σε γενικές γραμμές τη θετική ασυμμετρία των δεδομένων, αλλά εμφανίζει αισθητές μετατοπίσεις της θεωρητικής μάζας μεταξύ γειτονικών κλάσεων. Στην περιοχή $0-10 \mu\text{g}/\text{m}^3$ υποεκτιμά την παρατηρούμενη συχνότητα, δείχνοντας ότι προβλέπει μικρότερο αριθμό πολύ χαμηλών συγκεντρώσεων από εκείνον που καταγράφηκε.

Στην κλάση $10-20 \mu\text{g}/\text{m}^3$ εμφανίζεται η εντονότερη αστοχία του μοντέλου, καθώς η αναμενόμενη συχνότητα υπερβαίνει σημαντικά την παρατηρούμενη. Υπερεκτίμηση παρατηρείται και στην κλάση $20-30 \mu\text{g}/\text{m}^3$, αλλά σε σαφώς μικρότερο βαθμό. Η εικόνα αυτή δείχνει ότι η λογαριθμοκανονική συγκεντρώνει υπερβολικά μεγάλο μέρος της θεωρητικής πιθανότητας στις χαμηλομεσαίες συγκεντρώσεις.

Στις κλάσεις $30-40$ και $40-50 \mu\text{g}/\text{m}^3$ οι αναμενόμενες συχνότητες είναι χαμηλότερες από τις παρατηρούμενες. Η υποεκτίμηση είναι εντονότερη στην κλάση $40-50 \mu\text{g}/\text{m}^3$, γεγονός

που δείχνει ότι το μοντέλο δεν αποδίδει πλήρως τη σημαντική εμπειρική παρουσία στο μεσαίο προς υψηλό τμήμα.

Στην κλάση άνω των $50 \mu\text{g}/\text{m}^3$ η λογαριθμοκανονική παρουσιάζει σχετικά καλή συμφωνία, αν και εξακολουθεί να υπερεκτιμά την παρατηρούμενη συχνότητα. Η απόκλιση είναι σαφώς μικρότερη από εκείνη της κανονικής και της Gamma και παρόμοιας γενικής τάξης με την αντίστοιχη απόκλιση της Weibull.

Η σχετικά καλή συμφωνία στην ανοικτή τελευταία κλάση δεν σημαίνει κατ' ανάγκη ότι η λογαριθμοκανονική αναπαράγει με ακρίβεια ολόκληρη την άνω ουρά. Η κατηγορία περιλαμβάνει όλες τις τιμές άνω των $50 \mu\text{g}/\text{m}^3$ και, συνεπώς, μπορεί να αποκρύπτει διαφοροποιήσεις μεταξύ μέτρια υψηλών και εξαιρετικά ακραίων συγκεντρώσεων. Η συμπεριφορά αυτή θα εξεταστεί πληρέστερα μέσα από τα Q-Q διαγράμματα.

Συνολικά, η λογαριθμοκανονική αποτελεί αξιόλογη υποψήφια κατανομή, αλλά στο συγκεκριμένο στάδιο εμφανίζει λιγότερο ισορροπημένη προσαρμογή από τη Weibull. Οι σημαντικές αποκλίσεις στις κλάσεις 10-20 και $40\text{-}50 \mu\text{g}/\text{m}^3$ περιορίζουν την οπτική της υπεροχή, παρά τη σχετικά καλή απόδοση στην ανώτερη συγκεντρωτική κατηγορία.

Συγκριτική αξιολόγηση

Η σύγκριση των τεσσάρων μοντέλων αναδεικνύει διαφορετικά πρότυπα αστοχίας. Η κανονική κατανομή παρουσιάζει τις μεγαλύτερες και περισσότερο συστηματικές αποκλίσεις. Υπερεκτιμά έντονα τις δύο ακραίες κατηγορίες και υποεκτιμά το μεγαλύτερο μέρος των ενδιάμεσων κλάσεων. Η εικόνα αυτή δείχνει ότι δεν μπορεί να αναπαραστήσει αποτελεσματικά τη θέση και τη μορφή της εμπειρικής κατανομής.

Η Gamma αποτελεί σαφή βελτίωση, ιδίως λόγω της σχεδόν πλήρους συμφωνίας στην κλάση 10-20 $\mu\text{g}/\text{m}^3$. Εντούτοις, η πολύ μεγάλη υπερεκτίμηση της κλάσης 0-10 $\mu\text{g}/\text{m}^3$ και η υποεκτίμηση των κλάσεων 20-50 $\mu\text{g}/\text{m}^3$ δείχνουν ότι η θεωρητική μάζα δεν κατανέμεται σωστά σε όλο το εύρος. Η υπερεκτίμηση της κατηγορίας άνω των $50 \mu\text{g}/\text{m}^3$ επιβαρύνει επιπλέον τη συνολική εικόνα.

Η λογαριθμοκανονική αποδίδει συγκριτικά καλύτερα το ανώτερο συγκεντρωτικό τμήμα, αλλά εμφανίζει σημαντική υπερεκτίμηση στην κλάση 10-20 $\mu\text{g}/\text{m}^3$ και υποεκτίμηση στις κλάσεις 30-50 $\mu\text{g}/\text{m}^3$. Επομένως, παρά τη συμβατότητά της με τη θετική ασυμμετρία των δεδομένων, δεν επιτυγχάνει ομοιόμορφη συμφωνία στις κύριες κλάσεις της συγκεκριμένης χρονοσειράς.

Η Weibull παρουσιάζει την πιο σταθερή ισορροπία μεταξύ χαμηλών, μεσαίων και υψηλότερων συγκεντρώσεων. Οι αποκλίσεις της στις κλάσεις 0-20 $\mu\text{g}/\text{m}^3$ είναι κυρίως αρνητικές, ενώ στις κλάσεις άνω των 30 $\mu\text{g}/\text{m}^3$ τείνει να υπερεκτιμά τις συχνότητες. Ωστόσο, οι διαφορές παραμένουν συγκριτικά περιορισμένες και δεν εμφανίζεται μία ιδιαίτερα μεγάλη τοπική αστοχία, όπως συμβαίνει με τις άλλες κατανομές. Ιδιαίτερα στις κλάσεις 20-30 και 30-40 $\mu\text{g}/\text{m}^3$, η συμφωνία είναι πολύ καλή.

Με βάση αποκλειστικά τη σύγκριση των παρατηρούμενων και αναμενόμενων συχνοτήτων ανά κλάση, η Weibull παρέχει την ισχυρότερη οπτική ένδειξη προσαρμογής. Η λογαριθμοκανονική και η Gamma ακολουθούν, αλλά με διαφορετικά πλεονεκτήματα και αδυναμίες: η λογαριθμοκανονική παρουσιάζει συγκριτικά καλύτερη συμφωνία στην τελευταία κλάση, ενώ η Gamma αποδίδει σχεδόν ακριβώς την κλάση 10-20 $\mu\text{g}/\text{m}^3$. Η κανονική κατατάσσεται σαφώς τελευταία.

Η σχετική αυτή κατάταξη είναι προσωρινή. Η χρήση μόνο έξι ευρύτερων κλάσεων δεν επιτρέπει να εξεταστεί η λεπτομερής συμπεριφορά των μοντέλων στο εσωτερικό κάθε διαστήματος, ενώ η ανοικτή κλάση άνω των 50 $\mu\text{g}/\text{m}^3$ συμπυκνώνει ένα ιδιαίτερα μεγάλο εύρος συγκεντρώσεων. Ένα μοντέλο μπορεί, συνεπώς, να προσεγγίζει σωστά τη συνολική συχνότητα της τελευταίας κλάσης, αλλά να μην αποδίδει τις επιμέρους υψηλές και ακραίες τιμές.

Για τον λόγο αυτό, η οπτική υπεροχή της Weibull στο παρόν στάδιο πρέπει να ελεγχθεί με τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων, τα διαγράμματα υπολοίπων, τα Q-Q διαγράμματα και το στατιστικό χ^2 . Η συνδυαστική αξιολόγηση θα δείξει αν η ισορροπημένη συμπεριφορά της ανά κλάση διατηρείται και σε άλλες πτυχές της προσαρμογής ή αν κάποιο εναλλακτικό μοντέλο αποδίδει αποτελεσματικότερα τη συνολική κατανομική μορφή και την εκτεταμένη δεξιά ουρά της Χαλκίδας.

4.6.2 Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων

Τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων χρησιμοποιήθηκαν για τη συνολική οπτική αξιολόγηση της προσαρμογής της κανονικής, της Gamma, της Weibull και της λογαριθμοκανονικής κατανομής στις ωριαίες

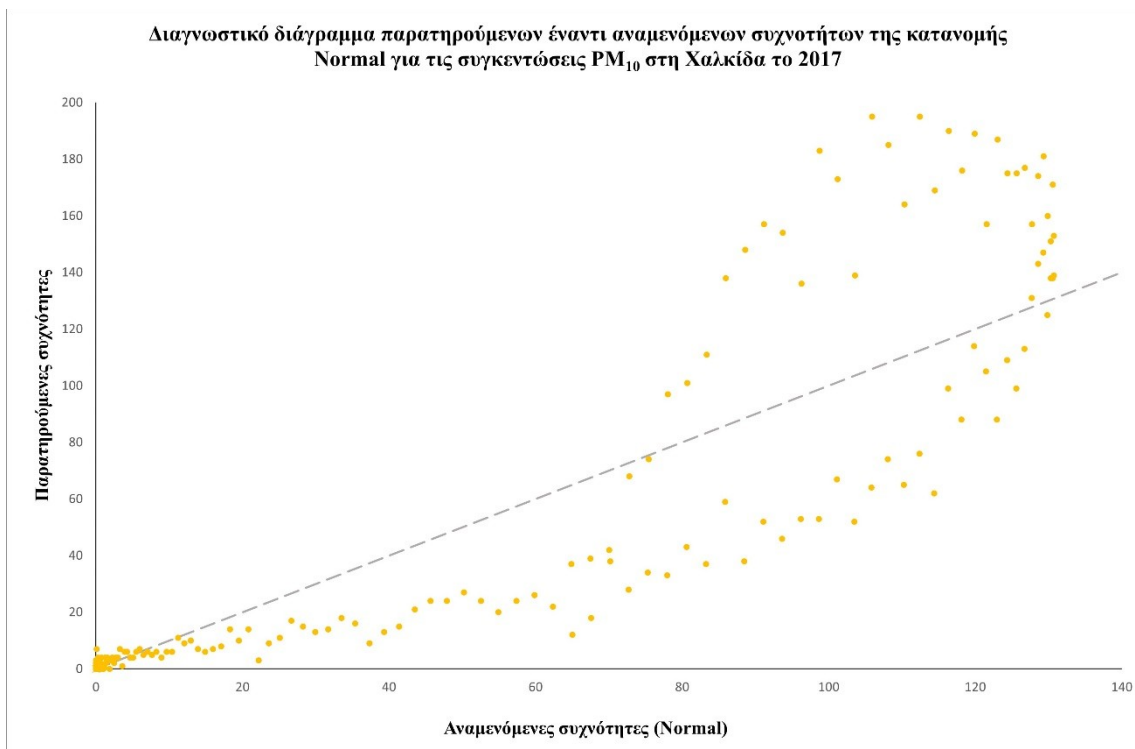
συγκεντρώσεις PM_{10} της Χαλκίδας κατά το έτος 2017. Σε κάθε διάγραμμα, στον οριζόντιο άξονα απεικονίζονται οι θεωρητικά αναμενόμενες συχνότητες και στον κατακόρυφο άξονα οι αντίστοιχες παρατηρούμενες συχνότητες.

Η διακεκομμένη διαγώνια γραμμή $y=x$ αντιπροσωπεύει την ιδανική περίπτωση πλήρους συμφωνίας. Όσο πλησιέστερα βρίσκονται τα σημεία προς τη γραμμή αναφοράς και όσο μικρότερη είναι η διασπορά τους γύρω από αυτήν, τόσο ισχυρότερη είναι η οπτική ένδειξη καλής προσαρμογής. Σημεία πάνω από τη διαγώνιο σημαίνουν ότι η παρατηρούμενη συχνότητα είναι μεγαλύτερη από την αναμενόμενη και, επομένως, ότι το θεωρητικό μοντέλο υποεκτιμά τη συγκεκριμένη συχνότητα. Αντίθετα, σημεία κάτω από τη γραμμή αντιστοιχούν σε υπερεκτίμηση από το μοντέλο.

Η αξιολόγηση δεν περιορίζεται στη μέση απόσταση των σημείων από τη διαγώνιο. Εξετάζεται επίσης αν οι αποκλίσεις παρουσιάζουν συστηματικό πρότυπο, αν η διασπορά αυξάνεται σε συγκεκριμένο εύρος αναμενόμενων συχνοτήτων και αν εμφανίζονται παρατεταμένες περιοχές στις οποίες τα σημεία βρίσκονται κυρίως στην ίδια πλευρά της γραμμής αναφοράς.

Πρέπει να διευκρινιστεί ότι οι μεγαλύτερες τιμές του οριζόντιου άξονα αντιστοιχούν σε μεγαλύτερες αναμενόμενες συχνότητες και όχι κατ' ανάγκη στις υψηλότερες συγκεντρώσεις PM_{10} . Τα συγκεκριμένα διαγράμματα αξιολογούν τη σχέση μεταξύ συχνοτήτων, ενώ η συμπεριφορά των μοντέλων στα υψηλά ποσοστημόρια και στη δεξιά ουρά εξετάζεται αναλυτικότερα μέσω των Q-Q διαγραμμάτων.

Κανονική κατανομή



Σχήμα 4.6.2α. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η κανονική κατανομή παρουσιάζει τη λιγότερο ικανοποιητική εικόνα μεταξύ των τεσσάρων εξεταζόμενων μοντέλων. Στο εύρος των πολύ χαμηλών αναμενόμενων συχνοτήτων ορισμένα σημεία βρίσκονται σχετικά κοντά στη γραμμή αναφοράς, χωρίς όμως να διαμορφώνεται σταθερή γραμμική σχέση. Ήδη από τις χαμηλές και μεσαίες συχνότητες εμφανίζεται οργανωμένη απομάκρυνση των σημείων από τη διαγώνιο.

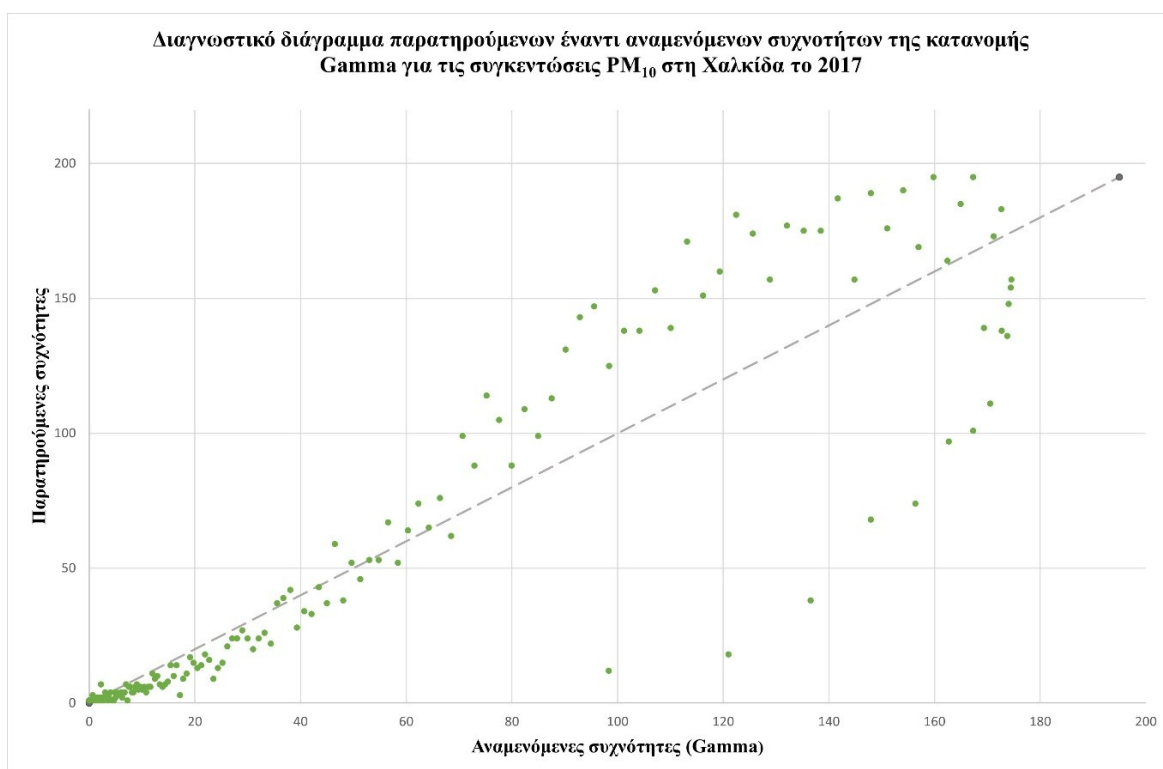
Σε σημαντικό τμήμα του μεσαίου εύρους τα σημεία βρίσκονται κυρίως κάτω από τη γραμμή $y=x$, στοιχείο που υποδηλώνει υπερεκτίμηση των αντίστοιχων παρατηρούμενων συχνοτήτων. Στη συνέχεια, αρκετά σημεία μετατοπίζονται έντονα πάνω από τη διαγώνιο, υποδηλώνοντας μετάβαση σε σημαντική υποεκτίμηση. Στις μεγαλύτερες αναμενόμενες συχνότητες παρατηρείται πολύ μεγάλη κατακόρυφη διασπορά, καθώς για παραπλήσιες θεωρητικές τιμές εμφανίζεται ευρύ φάσμα παρατηρούμενων συχνοτήτων.

Η διάταξη αυτή δεν συνιστά τυχαία διασπορά γύρω από τη γραμμή ιδανικής συμφωνίας. Αντίθετα, σχηματίζει σαφές μη γραμμικό και συστηματικό πρότυπο, το οποίο δείχνει ότι η

κανονική κατανομή εναλλάσσεται μεταξύ υπερεκτίμησης και υποεκτίμησης σε διαφορετικές περιοχές των δεδομένων.

Συνολικά, η κανονική κατανομή δεν αναπαράγει με συνέπεια τη σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Το αποτέλεσμα συμφωνεί με τη σύγκριση ανά κλάση συγκέντρωσης, όπου είχαν ήδη διαπιστωθεί μεγάλες αποκλίσεις στο μεγαλύτερο μέρος της εμπειρικής κατανομής.

Κατανομή Gamma



Σχήμα 4.6.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η κατανομή Gamma παρουσιάζει σαφή βελτίωση έναντι της κανονικής. Στις χαμηλές αναμενόμενες συχνοτήτες τα σημεία εμφανίζουν γενικά θετική και σχετικά συνεκτική σχέση, ενώ αρκετά από αυτά βρίσκονται κοντά στη διαγώνιο. Το στοιχείο αυτό δείχνει ότι η Gamma αποδίδει ικανοποιητικότερα σημαντικό μέρος των μικρότερων συχνοτήτων.

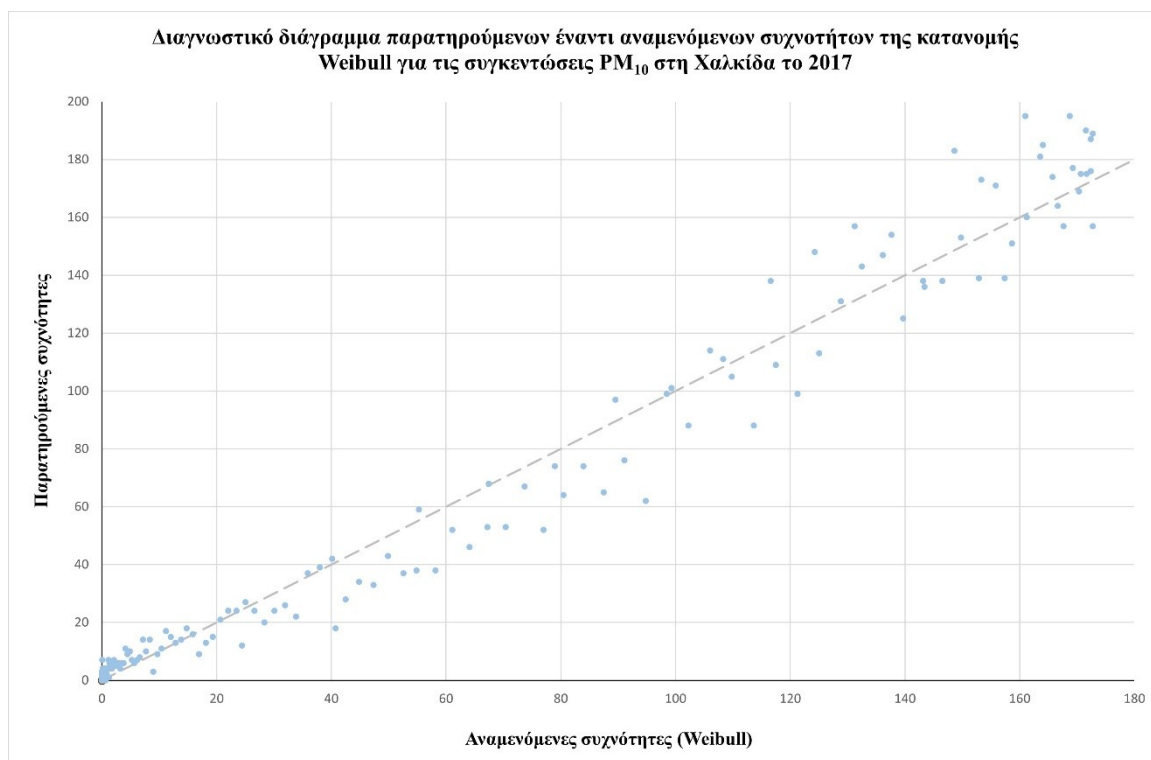
Καθώς αυξάνονται οι αναμενόμενες τιμές, όμως, η συμφωνία γίνεται λιγότερο σταθερή. Στο μεσαίο εύρος αρκετά σημεία βρίσκονται πάνω από τη γραμμή αναφοράς,

υποδηλώνοντας υποεκτίμηση των αντίστοιχων παρατηρούμενων συχνοτήτων. Παράλληλα, εμφανίζονται και μεμονωμένα σημεία πολύ κάτω από τη διαγώνιο, τα οποία αντιστοιχούν σε σημαντική τοπική υπερεκτίμηση από το μοντέλο.

Στις μεγαλύτερες αναμενόμενες συχνότητες η κατακόρυφη διασπορά αυξάνεται αισθητά. Για παρόμοιες θεωρητικές τιμές καταγράφονται παρατηρούμενες συχνότητες που διαφέρουν σημαντικά μεταξύ τους. Η εικόνα αυτή δείχνει ότι η Gamma δεν διατηρεί ομοιόμορφη ακρίβεια σε ολόκληρο το εξεταζόμενο εύρος.

Παρότι η διασπορά και οι τοπικές αποκλίσεις είναι αισθητές, η συνολική εικόνα παραμένει σαφώς καλύτερη από εκείνη της κανονικής κατανομής. Η Gamma αναπαράγει αποτελεσματικότερα τη γενική τάση, αλλά η παρουσία αρκετών ακραίων αποκλίσεων περιορίζει τη συνοχή της προσαρμογής της.

Κατανομή Weibull



Σχήμα 4.6.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η Weibull εμφανίζει την περισσότερο συνεκτική συνολική διάταξη σημείων μεταξύ των τεσσάρων κατανομών. Από τις χαμηλές έως και τις μεγαλύτερες αναμενόμενες συχνότητες, τα σημεία ακολουθούν σε σημαντικό βαθμό τη γενική κατεύθυνση της διαγώνιου και σχηματίζουν σαφέστερη σχεδόν γραμμική σχέση.

Στο χαμηλό εύρος οι περισσότερες παρατηρήσεις βρίσκονται σχετικά κοντά στη γραμμή αναφοράς, αν και εμφανίζονται μικρές αποκλίσεις και προς τις δύο πλευρές της.

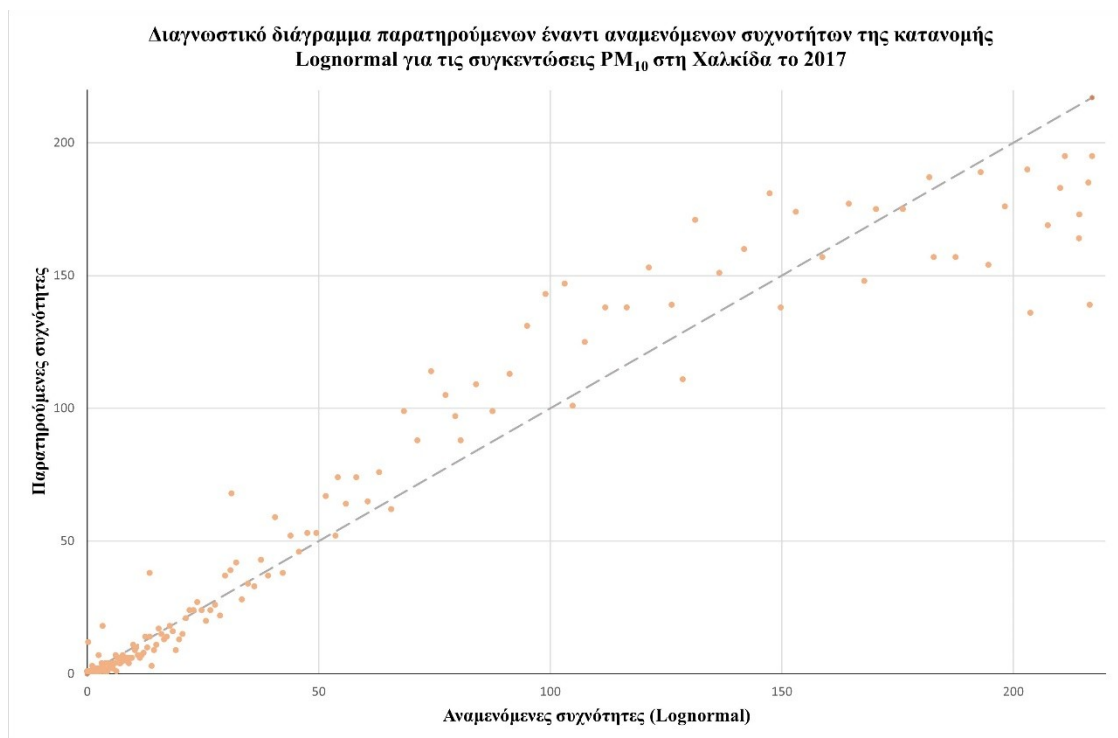
Στο χαμηλό προς μεσαίο εύρος παρατηρείται περιοχή στην οποία αρκετά σημεία βρίσκονται κάτω από τη διαγώνιο, υποδηλώνοντας τάση υπερεκτίμησης από το μοντέλο. Η απόκλιση αυτή είναι οργανωμένη, αλλά η έντασή της παραμένει συγκριτικά περιορισμένη.

Στο μεσαίο και υψηλότερο εύρος τα σημεία κατανέμονται και στις δύο πλευρές της γραμμής, χωρίς να εμφανίζεται η εκτεταμένη κατακόρυφη διασπορά που παρατηρείται στην Gamma και κυρίως στην κανονική κατανομή. Ορισμένες τοπικές αποκλίσεις παραμένουν, όμως η γενική διάταξη διατηρείται συνεκτική μέχρι το ανώτερο τμήμα του διαγράμματος.

Η Weibull δεν επιτυγχάνει πλήρη ταύτιση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Ωστόσο, περιορίζει τόσο τη γενική διασπορά όσο και την έκταση των ακραίων τοπικών αστοχιών. Με βάση το συγκεκριμένο διαγνωστικό κριτήριο, παρέχει την ισχυρότερη οπτική ένδειξη προσαρμογής.

Το αποτέλεσμα αυτό είναι συνεπές με τη σύγκριση ανά κλάση συγκέντρωσης, στην οποία η Weibull είχε επίσης εμφανίσει την πιο ισορροπημένη απόδοση. Η συμφωνία των δύο πρώτων διαγνωστικών σταδίων ενισχύει τη θέση της ως βασικής υποψήφιας κατανομής για τη Χαλκίδα.

Λογαριθμοκανονική κατανομή



Σχήμα 4.6.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η λογαριθμοκανονική κατανομή παρουσιάζει επίσης ικανοποιητική συνολική σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Στις χαμηλές τιμές τα σημεία συγκεντρώνονται σχετικά κοντά στη διαγώνιο, γεγονός που δείχνει ότι το μοντέλο αποδίδει αποτελεσματικά σημαντικό μέρος των μικρότερων συχνοτήτων.

Στο μεσαίο εύρος αρκετά σημεία μετατοπίζονται πάνω από τη γραμμή αναφοράς. Η συμπεριφορά αυτή δείχνει ότι η λογαριθμοκανονική υποεκτιμά τις αντίστοιχες παρατηρούμενες συχνότητες. Παρότι ορισμένα σημεία παραμένουν κοντά στη διαγώνιο, η θετική απόκλιση διατηρείται σε αρκετά εκτεταμένο τμήμα του διαγράμματος και αποτελεί οργανωμένο πρότυπο ασυμφωνίας.

Στις μεγαλύτερες αναμενόμενες συχνότητες η τάση μεταβάλλεται. Πολλά σημεία βρίσκονται κάτω από τη διαγώνιο, γεγονός που αντιστοιχεί σε υπερεκτίμηση από το

θεωρητικό μοντέλο. Η μετάβαση από περιοχή κυρίως θετικών αποκλίσεων σε περιοχή κυρίως αρνητικών αποκλίσεων δείχνει ότι η λογαριθμοκανονική δεν διατηρεί σταθερή ακρίβεια σε όλο το εύρος.

Η κατακόρυφη διασπορά της λογαριθμοκανονικής είναι μικρότερη από εκείνη της κανονικής και σε αρκετές περιοχές μικρότερη από της Gamma. Ωστόσο, η συστηματική θέση των σημείων πάνω από τη γραμμή στο μεσαίο εύρος και κάτω από αυτήν στις υψηλότερες αναμενόμενες συχνότητες δημιουργεί περισσότερο εμφανές καμπύλο πρότυπο από ό,τι στην περίπτωση της Weibull.

Συνολικά, η λογαριθμοκανονική αποτελεί ισχυρή υποψήφια κατανομή και παρουσιάζει σαφώς καλύτερη συμπεριφορά από την κανονική. Ωστόσο, με βάση αποκλειστικά το συγκεκριμένο διάγραμμα, η προσαρμογή της είναι λιγότερο ισορροπημένη από εκείνη της Weibull.

Συγκριτική αξιολόγηση

Η συγκριτική εξέταση των τεσσάρων διαγνωστικών διαγραμμάτων επιβεβαιώνει ότι η κανονική κατανομή παρουσιάζει τη χαμηλότερη συνολική καταλληλότητα. Η μεγάλη διασπορά, η έντονη μη γραμμικότητα και οι διαδοχικές περιοχές υπερεκτίμησης και υποεκτίμησης δείχνουν ότι το μοντέλο δεν αναπαράγει με συνέπεια τη σχέση μεταξύ εμπειρικών και θεωρητικών συχνοτήτων.

Η Gamma βελτιώνει αισθητά τη γενική εικόνα και ακολουθεί σχετικά καλά τη διαγώνιο στο χαμηλό εύρος. Ωστόσο, στις μεσαίες και μεγαλύτερες αναμενόμενες συχνότητες εμφανίζει αυξημένη διασπορά, καθώς και ορισμένες πολύ μεγάλες τοπικές αποκλίσεις. Η προσαρμογή της είναι επομένως καλύτερη από της κανονικής, αλλά όχι ομοιόμορφη.

Η λογαριθμοκανονική παρουσιάζει περισσότερο συνεκτική σχέση από την Gamma σε σημαντικό μέρος του διαγράμματος, αλλά εμφανίζει σαφές συστηματικό πρότυπο. Στο μεσαίο εύρος αρκετά σημεία βρίσκονται πάνω από τη διαγώνιο, ενώ στις μεγαλύτερες αναμενόμενες συχνότητες πολλά μετατοπίζονται κάτω από αυτήν. Η εναλλαγή αυτή υποδηλώνει ότι το μοντέλο μεταβαίνει από υποεκτίμηση σε υπερεκτίμηση, ανάλογα με το επίπεδο της θεωρητικής συχνότητας.

Η Weibull εμφανίζει την πιο σταθερή και συνεκτική συνολική διάταξη. Παρότι δεν απουσιάζουν οι τοπικές αποκλίσεις, τα σημεία ακολουθούν περισσότερο οργανωμένη σχεδόν γραμμική πορεία, ενώ η γενική τους διασπορά παραμένει μικρότερη. Σε αντίθεση

με τη Gamma και τη λογαριθμοκανονική, δεν παρατηρείται εξίσου έντονη συγκέντρωση μεγάλου αριθμού σημείων σε μία μόνο πλευρά της γραμμής αναφοράς για εκτεταμένο μέρος του διαγράμματος.

Με βάση αποκλειστικά τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων, η Weibull παρέχει την ισχυρότερη οπτική ένδειξη προσαρμογής. Η λογαριθμοκανονική ακολουθεί ως σημαντική εναλλακτική, η Gamma παρουσιάζει μεγαλύτερη αστάθεια και η κανονική κατατάσσεται σαφώς τελευταία.

Η εικόνα αυτή ενισχύει το αποτέλεσμα της προηγούμενης υποενοήτας, στην οποία η Weibull είχε επίσης εμφανίσει την πιο ισορροπημένη συμφωνία ανά κλάση. Η σύγκλιση των δύο πρώτων γραφικών κριτηρίων αποτελεί σημαντική ένδειξη υπέρ της, χωρίς όμως να θεμελιώνει ακόμη οριστική επιλογή.

Η αξιολόγηση πρέπει να ερμηνευθεί με προσοχή, επειδή το διάγραμμα εξετάζει τη γεωμετρική σχέση μεταξύ των συχνοτήτων και όχι άμεσα την αντιστοίχιση των ποσοστημορίων ή την απόδοση στην άνω ουρά. Επιπλέον, δεν σταθμίζει τις αποκλίσεις ανάλογα με το μέγεθος της αναμενόμενης συχνότητας και δεν παρέχει ενιαίο αριθμητικό μέτρο της συνολικής αστοχίας.

Για τον λόγο αυτό, η προσωρινή υπεροχή της Weibull θα ελεγχθεί στη συνέχεια μέσω των διαγραμμάτων υπολοίπων, των Q-Q διαγραμμάτων και του στατιστικού χ^2 . Η συνδυαστική εξέταση θα δείξει αν η περισσότερο συνεκτική οπτική της εικόνα διατηρείται και ως προς τη δομή των σφαλμάτων, την αντιστοίχιση των ποσοστημορίων και τη συνολική σταθμισμένη απόκλιση.

4.6.3 Διαγράμματα υπολοίπων

Στη Χαλκίδα, τα διαγράμματα υπολοίπων έχουν ιδιαίτερη διαγνωστική αξία, επειδή επιτρέπουν να εξεταστεί αν η καλή ή κακή προσαρμογή των θεωρητικών κατανομών περιορίζεται σε μεμονωμένες περιοχές ή διατηρείται σε ολόκληρη την ταξινομημένη σειρά των συγκεντρώσεων PM_{10} . Η συγκεκριμένη χρονοσειρά χαρακτηρίζεται από μεγαλύτερη διασπορά και εντονότερη παρουσία υψηλών τιμών σε σχέση με τις προηγούμενες πόλεις, γεγονός που καθιστά πιθανότερη την εμφάνιση τοπικών αστοχιών ακόμη και για μοντέλα με γενικά καλή προσαρμογή.

Για κάθε σημείο, το υπόλοιπο υπολογίστηκε ως:

$$R_i = O_i - E_i$$

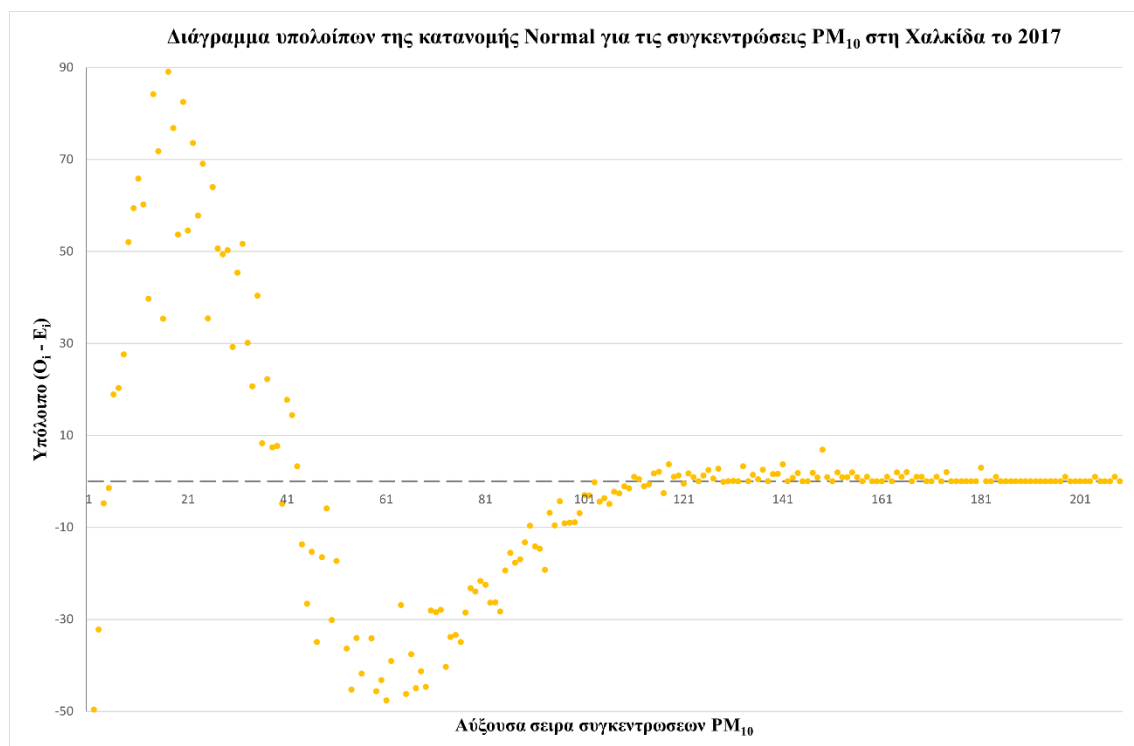
όπου:

- O_i είναι η παρατηρούμενη
- E_i η θεωρητικά αναμενόμενη συχνότητα.

Θετικό υπόλοιπο δηλώνει ότι το μοντέλο υποεκτιμά την παρατηρούμενη συχνότητα, ενώ αρνητικό υπόλοιπο ότι την υπερεκτιμά.

Η αξιολόγηση δεν βασίζεται μόνο στο μέγιστο απόλυτο υπόλοιπο. Εξετάζεται επίσης αν οι αποκλίσεις αλλάζουν συστηματικά πρόσημο, αν παραμένουν για μεγάλο διάστημα στην ίδια πλευρά της γραμμής μηδενός και πόσο γρήγορα σταθεροποιούνται. Επειδή πρόκειται για απλά και όχι τυποποιημένα υπόλοιπα, η έμφαση δίνεται κυρίως στη μορφή και στη διάρκεια των αποκλίσεων και όχι αποκλειστικά στο αριθμητικό τους μέγεθος.

Κανονική κατανομή



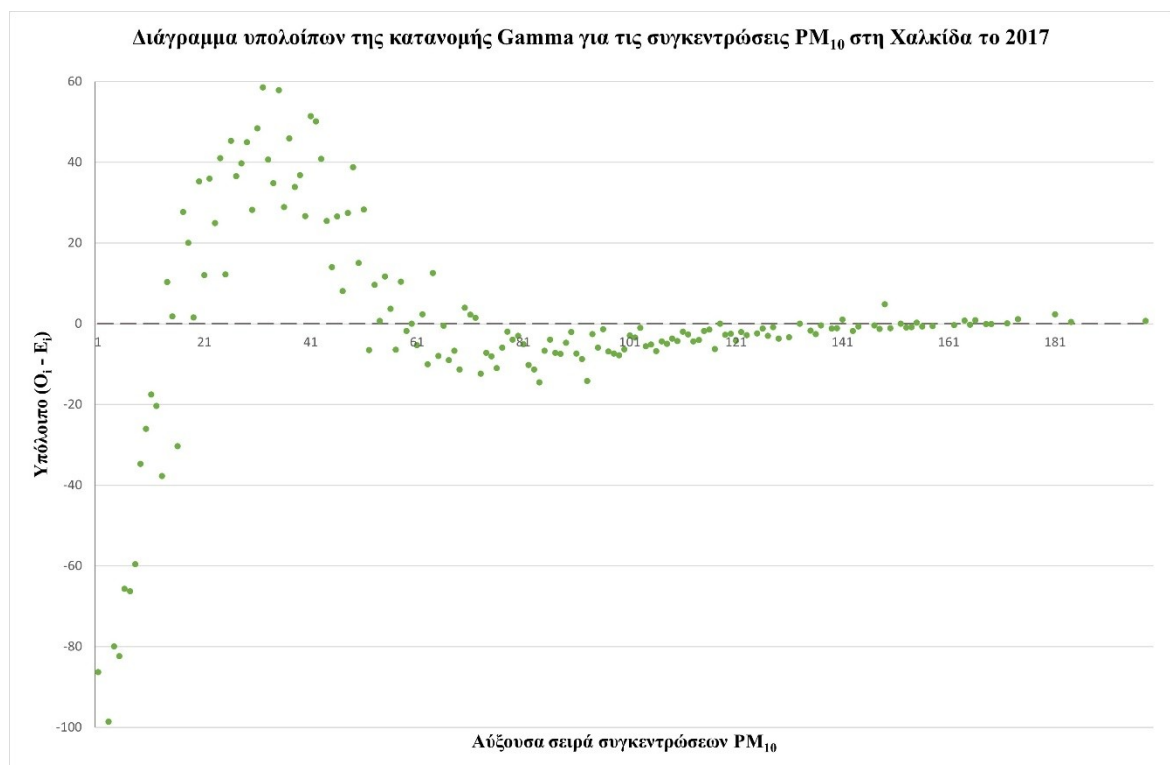
Σχήμα 4.6.3α. Διάγραμμα υπολοίπων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας το 2017.

Η κανονική κατανομή εμφανίζει έντονη εναλλαγή μεταξύ υποεκτίμησης και υπερεκτίμησης. Μετά από ορισμένες αρχικές αρνητικές τιμές, τα υπόλοιπα αυξάνονται απότομα και φθάνουν περίπου έως το +90. Η περιοχή αυτή δείχνει ισχυρή υποεκτίμηση των αντίστοιχων παρατηρούμενων συχνοτήτων.

Ακολουθεί σταδιακή αλλά εκτεταμένη μετατόπιση προς αρνητικές τιμές, οι οποίες προσεγγίζουν το -50. Το μοντέλο μεταβαίνει, επομένως, από έντονη υποεκτίμηση σε σημαντική υπερεκτίμηση. Η διαδοχή αυτή σχηματίζει σαφές κυματοειδές πρότυπο και δείχνει ότι οι αποκλίσεις δεν είναι τυχαίες.

Μόνο μετά από μεγάλο μέρος της ταξινομημένης σειράς τα υπόλοιπα περιορίζονται γύρω από το μηδέν. Η καθυστερημένη αυτή σταθεροποίηση δεν αναιρεί τη συστηματική αστοχία που έχει προηγηθεί. Η κανονική κατανομή αδυνατεί, επομένως, να αποδώσει με σταθερό τρόπο τη δομή της χρονοσειράς της Χαλκίδας.

Κατανομή Gamma

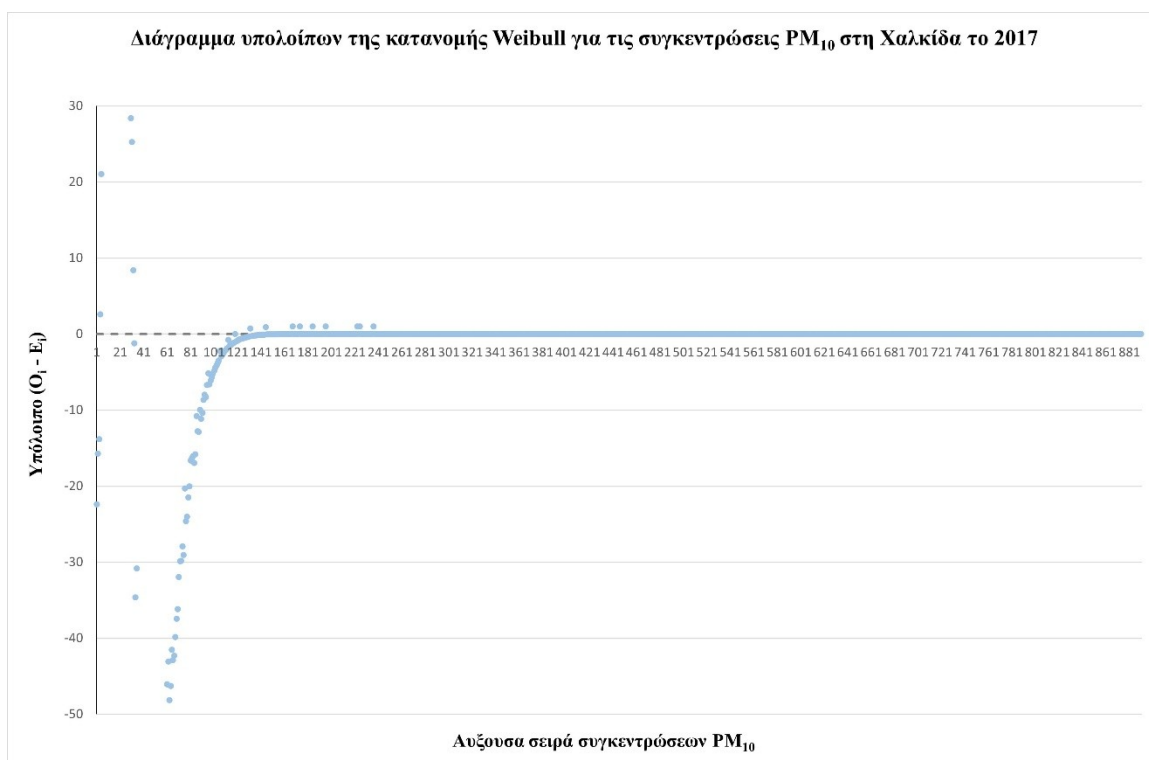


Σχήμα 4.6.3β Διάγραμμα υπολοίπων (Residual plot) της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας κατά το έτος 2017.

Η Gamma διαφοροποιείται από την κανονική, αλλά εξακολουθεί να παρουσιάζει έντονα οργανωμένο αρχικό πρότυπο. Οι πρώτες αποκλίσεις είναι πολύ αρνητικές και προσεγγίζουν το -100 , υποδηλώνοντας μεγάλη υπερεκτίμηση από το μοντέλο.

Στη συνέχεια, τα υπόλοιπα μετακινούνται απότομα σε θετικές τιμές, οι οποίες φθάνουν περίπου έως το $+60$. Η απότομη αυτή αντιστροφή δείχνει ότι η Gamma ανακατανέμει λανθασμένα τη θεωρητική μάζα μεταξύ διαδοχικών τμημάτων της εμπειρικής κατανομής. Αργότερα, οι αποκλίσεις μειώνονται σε ένταση και παραμένουν κυρίως ελαφρώς αρνητικές, πριν συγκλίνουν κοντά στο μηδέν. Η συμπεριφορά αυτή δείχνει ότι η Gamma βελτιώνει σημαντικά την προσαρμογή στο μεγαλύτερο μέρος της σειράς, αλλά διατηρεί σοβαρή τοπική αστοχία στο αρχικό της τμήμα.

Κατανομή Weibull



Σχήμα 4.6.3γ. Διάγραμμα υπολοίπων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας το 2017.

Η Weibull παρουσιάζει διαφορετική εικόνα από τις δύο προηγούμενες κατανομές. Οι αποκλίσεις της είναι συγκεντρωμένες σχεδόν αποκλειστικά στο αρχικό τμήμα, ενώ το

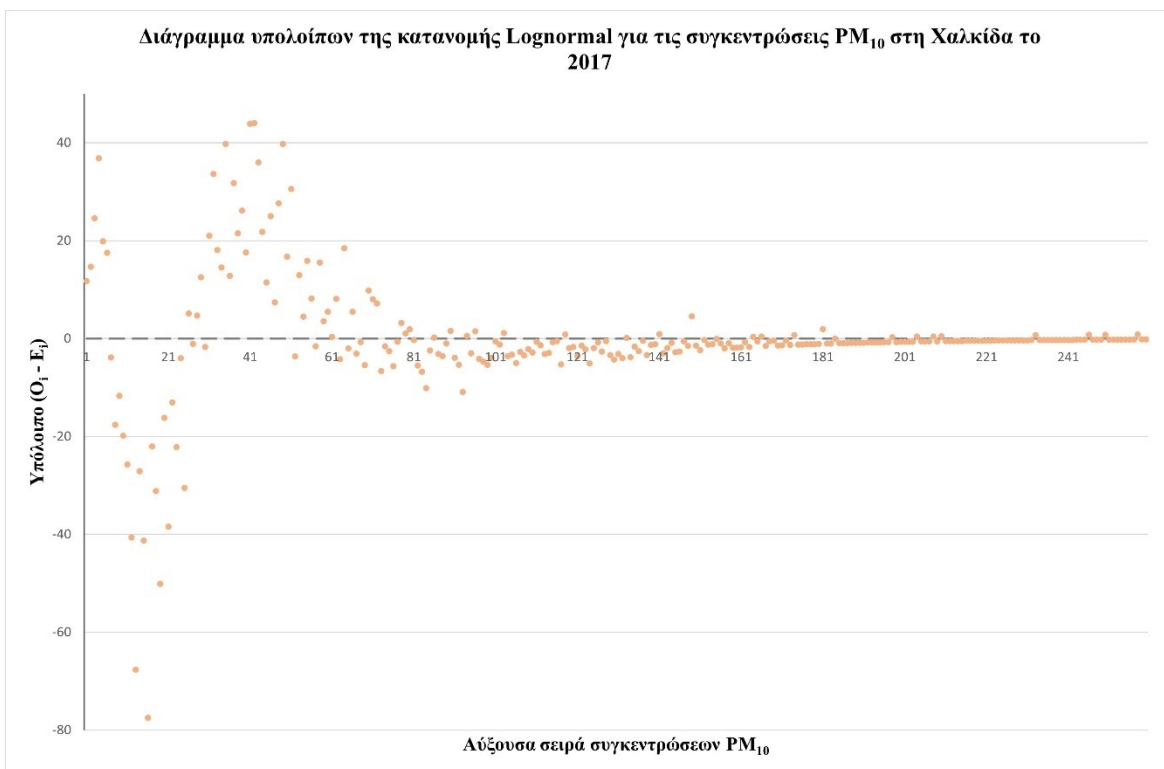
μεγαλύτερο μέρος της ταξινομημένης σειράς χαρακτηρίζεται από σχεδόν μηδενικά υπόλοιπα.

Στις πρώτες παρατηρήσεις εμφανίζονται ορισμένες θετικές αποκλίσεις, με τιμές έως περίπου +28, και στη συνέχεια μία εντονότερη περιοχή αρνητικών υπολοίπων, με ελάχιστο κοντά στο -48. Το πρότυπο αυτό δείχνει ότι η Weibull δεν αποδίδει με την ίδια ακρίβεια το αρχικό τμήμα των δεδομένων και τείνει εκεί να υπερεκτιμά ορισμένες συχνότητες.

Η διαφοροποίηση εμφανίζεται αμέσως μετά. Τα υπόλοιπα μειώνονται γρήγορα και από σχετικά πρώιμο σημείο της σειράς παραμένουν σχεδόν πάνω στη γραμμή μηδενός. Η μακρά αυτή περιοχή σταθερότητας αποτελεί ιδιαίτερα ισχυρή ένδειξη ότι το μοντέλο αποδίδει με συνέπεια μεγάλο μέρος της εμπειρικής δομής.

Το πλεονέκτημα της Weibull δεν είναι ότι εξαλείφει πλήρως τις τοπικές αστοχίες, αλλά ότι τις περιορίζει σε μικρότερο τμήμα της σειράς και δεν τις μεταφέρει στο υπόλοιπο εύρος των δεδομένων. Η εικόνα αυτή είναι συμβατή με την καλή συμπεριφορά της στα προηγούμενα διαγνωστικά στάδια.

Λογαριθμοκανονική κατανομή



Σχήμα 4.6.3δ. Διάγραμμα υπολοίπων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} της Χαλκίδας το 2017.

Η λογαριθμοκανονική παρουσιάζει επίσης σαφή συγκέντρωση των μεγαλύτερων αποκλίσεων στο αρχικό τμήμα. Σε αντίθεση, όμως, με τη Weibull, οι μεταβολές του προσήμου είναι περισσότερες και το συνολικό εύρος των υπολοίπων μεγαλύτερο.

Οι θετικές τιμές φθάνουν περίπου έως το +44, ενώ οι αρνητικές προσεγγίζουν το -78. Η έντονη εναλλαγή δείχνει ότι το μοντέλο υποεκτιμά ορισμένες συχνότητες και υπερεκτιμά άλλες σε διαδοχικά τμήματα της αρχικής περιοχής.

Στη συνέχεια, οι αποκλίσεις μειώνονται σταδιακά και στο μεσαίο τμήμα παραμένουν κυρίως μικρές και αρνητικές. Προς το τέλος της σειράς συγκεντρώνονται πολύ κοντά στο μηδέν. Η λογαριθμοκανονική εμφανίζει, επομένως, ιδιαίτερα καλή συμπεριφορά σε μεγάλο μέρος των δεδομένων, αλλά χρειάζεται μεγαλύτερο διάστημα για να σταθεροποιηθεί σε σχέση με τη Weibull.

Η βασική της αδυναμία δεν βρίσκεται στο τελικό τμήμα, αλλά στην εντονότερη αρχική αστάθεια. Για τον λόγο αυτό, το διάγραμμα την αναδεικνύει ως ισχυρή εναλλακτική, όχι όμως ως καθαρά ανώτερη επιλογή.

Συγκριτική αποτίμηση

Η Χαλκίδα παρουσιάζει μια πιο καθαρή διαφοροποίηση μεταξύ των μοντέλων από εκείνη που είχε παρατηρηθεί στις προηγούμενες πόλεις. Η κανονική και η Gamma εμφανίζουν εκτεταμένες και οργανωμένες αποκλίσεις, ενώ η Weibull και η λογαριθμοκανονική περιορίζουν τις σημαντικές αστοχίες κυρίως στο αρχικό τμήμα της ταξινομημένης σειράς.

Η κανονική κατανομή εμφανίζει παρατεταμένο κυματοειδές πρότυπο και καθυστερεί σημαντικά να συγκλίνει προς το μηδέν. Η Gamma παρουσιάζει ακόμη μεγαλύτερη αρχική αρνητική απόκλιση, ακολουθούμενη από έντονη θετική περιοχή. Παρότι σταθεροποιείται νωρίτερα από την κανονική, η αρχική της συμπεριφορά παραμένει ιδιαίτερα ασυνεπής.

Η ουσιαστική σύγκριση αφορά τη Weibull και τη λογαριθμοκανονική. Και οι δύο κατανομές παρουσιάζουν μικρά υπόλοιπα στο μεγαλύτερο μέρος της σειράς, αλλά διαφέρουν ως προς την έκταση της αρχικής αστοχίας. Η Weibull εμφανίζει μικρότερο περίπου εύρος αποκλίσεων και πολύ ταχύτερη μετάβαση σε περιοχή σχεδόν μηδενικών υπολοίπων. Η λογαριθμοκανονική χρειάζεται μεγαλύτερο τμήμα της σειράς για να σταθεροποιηθεί και παρουσιάζει εντονότερες εναλλαγές θετικών και αρνητικών τιμών.

Το αποτέλεσμα αυτό είναι ιδιαίτερα σημαντικό, επειδή ενισχύει την ίδια ένδειξη που είχε ήδη προκύψει από τη σύγκριση ανά κλάση και από τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων. Στη Χαλκίδα, επομένως, η Weibull

δεν εμφανίζεται ευνοϊκή μόνο σε ένα μεμονωμένο διάγραμμα, αλλά παρουσιάζει επαναλαμβανόμενα συνεπή συμπεριφορά στα τρία πρώτα στάδια της αξιολόγησης. Η συνέπεια αυτή αυξάνει τη διαγνωστική βαρύτητα της υποψηφιότητάς της, καθώς υποδηλώνει ότι η καλή προσαρμογή της δεν εξαρτάται αποκλειστικά από μία συγκεκριμένη μορφή γραφικής απεικόνισης.

Παράλληλα, η λογαριθμοκανονική δεν παύει να αποτελεί ισχυρή εναλλακτική. Μετά το αρχικό τμήμα της σειράς, τα υπόλοιπά της περιορίζονται αισθητά και παραμένουν κοντά στο μηδέν σε μεγάλο μέρος των παρατηρήσεων. Η βασική διαφορά της από τη Weibull εντοπίζεται στην εντονότερη αρχική αστάθεια και στο μεγαλύτερο εύρος των ακραίων αποκλίσεων. Επομένως, τα διαγράμματα υπολοίπων δεν αποκλείουν τη λογαριθμοκανονική, αλλά προσδίδουν σαφέστερο πλεονέκτημα στη Weibull ως προς τη σταθερότητα και την ταχύτητα σύγκλισης των σφαλμάτων.

Με βάση αποκλειστικά τη δομή των απλών υπολοίπων, η σχετική κατάταξη είναι:

$$\text{Weibull} > \text{Λογαριθμοκανονική} > \text{Gamma} > \text{Κανονική}$$

Η κατάταξη αυτή πρέπει να ερμηνευθεί ως συγκριτική ένδειξη και όχι ως οριστική απόφαση. Τα απλά υπόλοιπα δεν είναι τυποποιημένα και, επομένως, το μέγεθός τους επηρεάζεται από το επίπεδο των αντίστοιχων συχνοτήτων. Επιπλέον, η μεγάλη συγκέντρωση σημείων πολύ κοντά στη γραμμή μηδενός μπορεί να δημιουργεί ιδιαίτερα ευνοϊκή οπτική εικόνα χωρίς να αποκαλύπτει πλήρως τη συμπεριφορά του μοντέλου στα υψηλότερα ποσοστημόρια ή στις ακραίες παρατηρήσεις.

Ιδιαίτερη προσοχή απαιτεί και η ερμηνεία του τελικού τμήματος των διαγραμμάτων. Η σχεδόν μηδενική τιμή των υπολοίπων σε μεγάλο μέρος της ταξινομημένης σειράς δείχνει καλή συμφωνία των αντίστοιχων συχνοτήτων, αλλά δεν τεκμηριώνει αυτομάτως ακριβή αναπαράσταση της δεξιάς ουράς. Ένα μοντέλο μπορεί να εμφανίζει μικρές αποκλίσεις στις συχνότητες και ταυτόχρονα να υποεκτιμά το μέγεθος των υψηλότερων εμπειρικών ποσοστημορίων. Η διάκριση αυτή θα εξεταστεί πληρέστερα στα Q-Q διαγράμματα.

Συνεπώς, τα διαγράμματα υπολοίπων μετατοπίζουν ακόμη περισσότερο το βάρος της προσωρινής αξιολόγησης προς τη Weibull. Η επαναλαμβανόμενη καλή της εικόνα στις συγκρίσεις συχνοτήτων, στα observed-expected διαγράμματα και στα υπόλοιπα δημιουργεί ισχυρή και συνεκτική ένδειξη υπέρ της. Η οριστική επιλογή, ωστόσο, θα εξαρτηθεί από το αν το πλεονέκτημα αυτό διατηρείται και στην αντιστοίχιση των

θεωρητικών με τα εμπειρικά ποσοστημόρια, καθώς και από το αν επιβεβαιώνεται αριθμητικά μέσω του στατιστικού χ^2 . Με αυτόν τον τρόπο θα διαπιστωθεί αν η Weibull υπερέχει συνολικά ή αν η λογαριθμοκανονική εμφανίζει καλύτερη συμπεριφορά σε πτυχές της προσαρμογής που δεν αποτυπώνονται πλήρως από τα απλά υπόλοιπα.

4.6.4 Διαγράμματα Q-Q

Τα διαγράμματα Q-Q χρησιμοποιούνται για την άμεση σύγκριση των εμπειρικών ποσοστημορίων των ωριαίων συγκεντρώσεων PM_{10} στη Χαλκίδα το 2017 με τα αντίστοιχα ποσοστημόρια των τεσσάρων προσαρμοσμένων θεωρητικών κατανομών. Σε αντίθεση με τις προηγούμενες γραφικές προσεγγίσεις, οι οποίες επικεντρώθηκαν στις συχνότητες και στα υπόλοιπά τους, τα Q-Q διαγράμματα εξετάζουν κατά πόσο κάθε μοντέλο αναπαράγει το μέγεθος των παρατηρήσεων σε διαδοχικές θέσεις της εμπειρικής κατανομής.

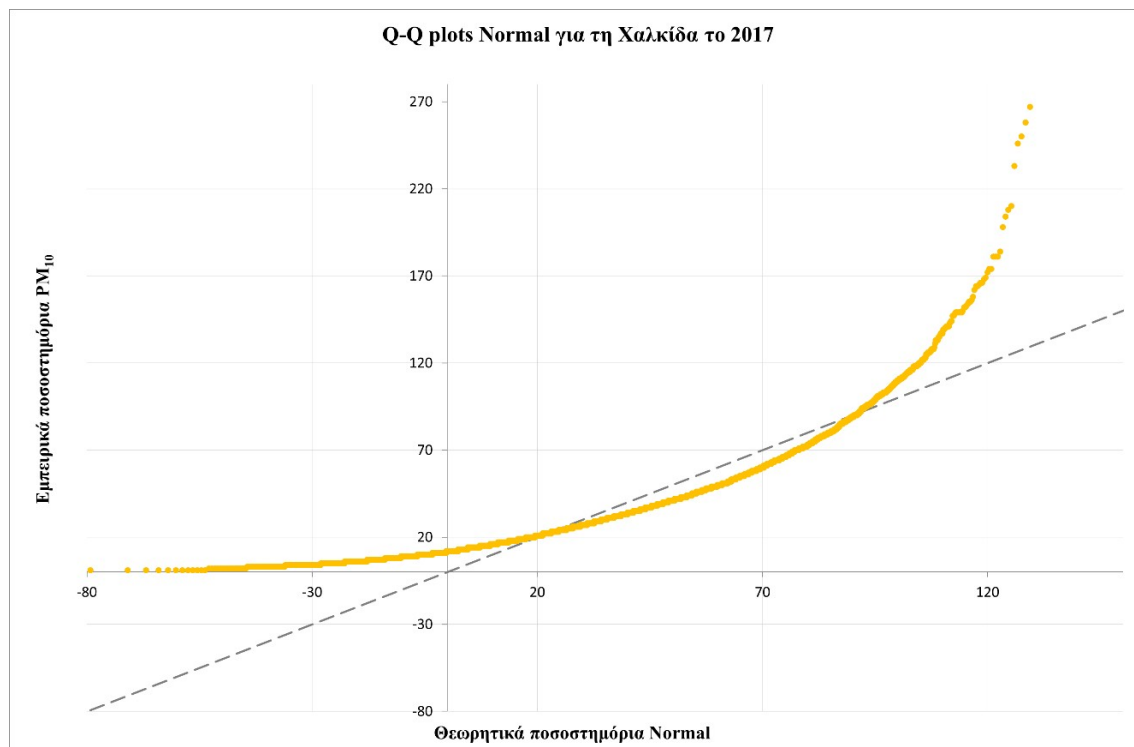
Η διαγώνια γραμμή αναφοράς εκφράζει την ιδανική περίπτωση ισότητας μεταξύ θεωρητικών και εμπειρικών ποσοστημορίων. Όταν τα σημεία ακολουθούν στενά τη γραμμή, το θεωρητικό μοντέλο θεωρείται ότι αποδίδει ικανοποιητικά το αντίστοιχο τμήμα της κατανομής. Αντίθετα, η παρατεταμένη παρουσία σημείων πάνω ή κάτω από τη διαγώνιο και η εμφάνιση συστηματικής καμπυλότητας υποδηλώνουν ασυμφωνία ως προς τη θέση, τη διασπορά ή το βάρος των ουρών.

Σημεία πάνω από τη γραμμή αναφοράς σημαίνουν ότι τα εμπειρικά ποσοστημόρια είναι μεγαλύτερα από τα θεωρητικά. Στην άνω ουρά, η συμπεριφορά αυτή αντιστοιχεί σε υποεκτίμηση των υψηλών συγκεντρώσεων από το μοντέλο. Αντίθετα, σημεία κάτω από τη διαγώνιο δείχνουν ότι τα θεωρητικά ποσοστημόρια είναι υψηλότερα και, επομένως, ότι το μοντέλο υπερεκτιμά τις αντίστοιχες εμπειρικές τιμές.

Η εφαρμογή των Q-Q διαγραμμάτων είναι ιδιαίτερα χρήσιμη στη Χαλκίδα, καθώς η συγκεκριμένη χρονοσειρά παρουσιάζει υψηλή μεταβλητότητα, έντονη θετική ασυμμετρία και εκτεταμένη δεξιά ουρά. Συνεπώς, η αξιολόγηση δεν περιορίζεται στην προσαρμογή του κεντρικού τμήματος, αλλά επεκτείνεται και στην ικανότητα κάθε μοντέλου να περιγράψει τη μετάβαση προς τις υψηλές συγκεντρώσεις.

Πρέπει, ωστόσο, να επισημανθεί ότι τα απεικονιζόμενα Q-Q διαγράμματα φθάνουν σε εμπειρικές τιμές περίπου 250-270 $\mu\text{g}/\text{m}^3$, ενώ από την περιγραφική στατιστική είχε καταγραφεί υψηλότερη μέγιστη συγκέντρωση για τη Χαλκίδα. Επομένως, οι παρακάτω διαπιστώσεις για την άνω ουρά αφορούν κυρίως το εύρος που εμφανίζεται στα σχήματα και όχι κατ' ανάγκη το σύνολο των πλέον ακραίων παρατηρήσεων.

Κανονική κατανομή



Σχήμα 4.6.4α. Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η κανονική κατανομή παρουσιάζει έντονη και συστηματική απόκλιση από τη γραμμή αναφοράς σε ολόκληρο σχεδόν το εύρος των ποσοστημορίων. Στο κατώτερο τμήμα, τα θεωρητικά ποσοστημόρια λαμβάνουν αρνητικές τιμές, ενώ οι συγκεντρώσεις PM_{10} είναι εκ φύσεως μη αρνητικές. Η ασυμφωνία αυτή αναδεικνύει έναν βασικό δομικό περιορισμό της κανονικής κατανομής για τη συγκεκριμένη μεταβλητή.

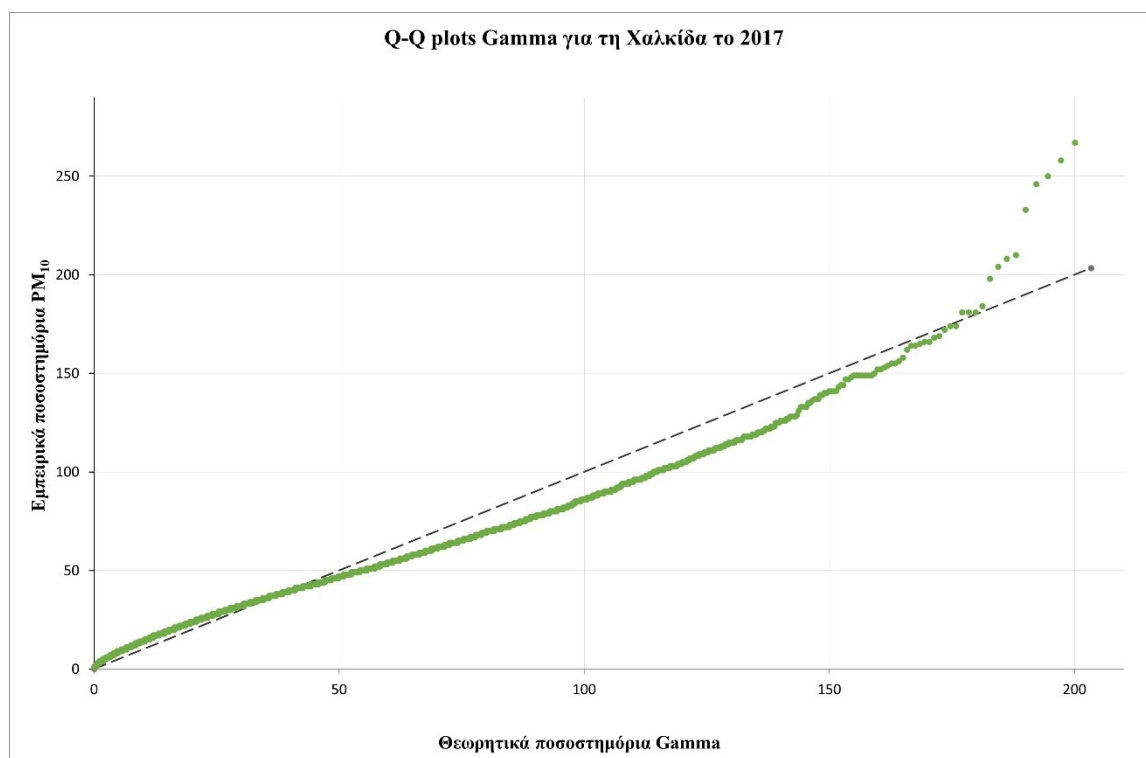
Στο χαμηλό και πρώτο μεσαίο τμήμα, τα σημεία βρίσκονται αρχικά πάνω από τη γραμμή και στη συνέχεια την προσεγγίζουν μόνο για περιορισμένο εύρος. Η προσωρινή αυτή

εγγύτητα δεν διατηρείται. Καθώς αυξάνονται τα θεωρητικά ποσοστημόρια, εμφανίζεται έντονη καμπυλότητα, με τα εμπειρικά σημεία να μετατοπίζονται πρώτα κάτω και κατόπιν πολύ πάνω από τη διαγώνιο.

Στην άνω περιοχή η ανοδική απόκλιση αυξάνεται ταχύτατα. Τα εμπειρικά ποσοστημόρια υπερβαίνουν σημαντικά τα θεωρητικά, γεγονός που δείχνει ότι η κανονική κατανομή υποεκτιμά το μέγεθος των υψηλών συγκεντρώσεων και δεν διαθέτει αρκετά βαριά δεξιά ουρά για να αναπαραγάγει τα δεδομένα.

Η συνολική μορφή του διαγράμματος επιβεβαιώνει ότι η κανονική κατανομή δεν μπορεί να αποδώσει ούτε τη μη αρνητικότητα ούτε τη θετική ασυμμετρία της χρονοσειράς. Η περιορισμένη τοπική προσέγγιση της γραμμής σε ένα μικρό τμήμα δεν αντισταθμίζει τη γενικευμένη ασυμφωνία που εμφανίζεται στα υπόλοιπα ποσοστημόρια.

Κατανομή Gamma



Σχήμα 4.6.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

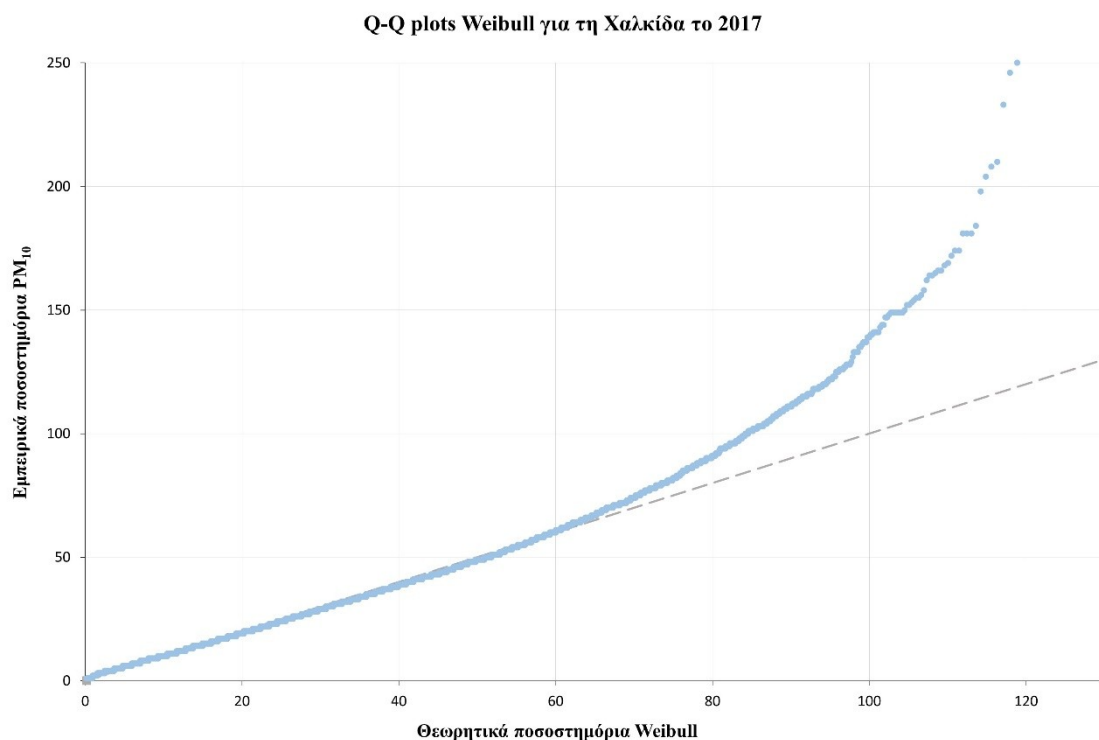
Η κατανομή Gamma εμφανίζει σαφώς βελτιωμένη και περισσότερο συνεκτική εικόνα. Στα χαμηλά ποσοστημόρια τα σημεία βρίσκονται κοντά στη διαγώνιο, αν και αρχικά

εμφανίζονται ελαφρώς πάνω από αυτήν. Η συμπεριφορά αυτή δείχνει ότι το μοντέλο αποδίδει ικανοποιητικά το κατώτερο τμήμα της εμπειρικής κατανομής. Από τα χαμηλά προς τα μεσαία ποσοστημόρια, τα σημεία διασταυρώνουν τη γραμμή αναφοράς και στη συνέχεια τοποθετούνται συστηματικά κάτω από αυτήν. Η απόκλιση αυξάνεται σταδιακά, χωρίς όμως απότομες μεταβολές. Το πρότυπο αυτό σημαίνει ότι η Gamma τείνει να παράγει μεγαλύτερα θεωρητικά ποσοστημόρια από τα αντίστοιχα εμπειρικά και, επομένως, να υπερεκτιμά μέρος του κεντρικού και του μεσαίου προς υψηλό εύρους.

Προς τα ανώτερα ποσοστημόρια, τα σημεία προσεγγίζουν εκ νέου τη διαγώνιο. Στις τελευταίες παρατηρήσεις, όμως, η τάση αντιστρέφεται και τα εμπειρικά ποσοστημόρια μετατοπίζονται αισθητά πάνω από τη γραμμή. Η ανοδική αυτή απόκλιση δείχνει ότι η Gamma υποεκτιμά τις υψηλότερες τιμές του απεικονιζόμενου εύρους.

Παρά τις συστηματικές αποκλίσεις, η Gamma διατηρεί σχετικά ομαλή πορεία και προσεγγίζει τη γραμμή αναφοράς σε μεγαλύτερο μέρος της κατανομής από ό,τι η κανονική και η Weibull. Δεν επιτυγχάνει τέλεια προσαρμογή, αλλά το διάγραμμα την αναδεικνύει ως μία από τις ισχυρότερες υποψήφιες κατανομές, ιδιαίτερα ως προς την ισορροπία μεταξύ κεντρικού τμήματος και άνω ουράς.

Κατανομή Weibull



Σχήμα 4.6.4γ. Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η Weibull παρουσιάζει πολύ καλή ευθυγράμμιση στα χαμηλά και σε σημαντικό μέρος των μεσαίων ποσοστημορίων. Μέχρι περίπου το κεντρικό τμήμα του διαγράμματος, τα σημεία ακολουθούν στενά τη γραμμή αναφοράς, γεγονός που δείχνει αποτελεσματική αναπαράσταση των χαμηλότερων και συνηθέστερων συγκεντρώσεων.

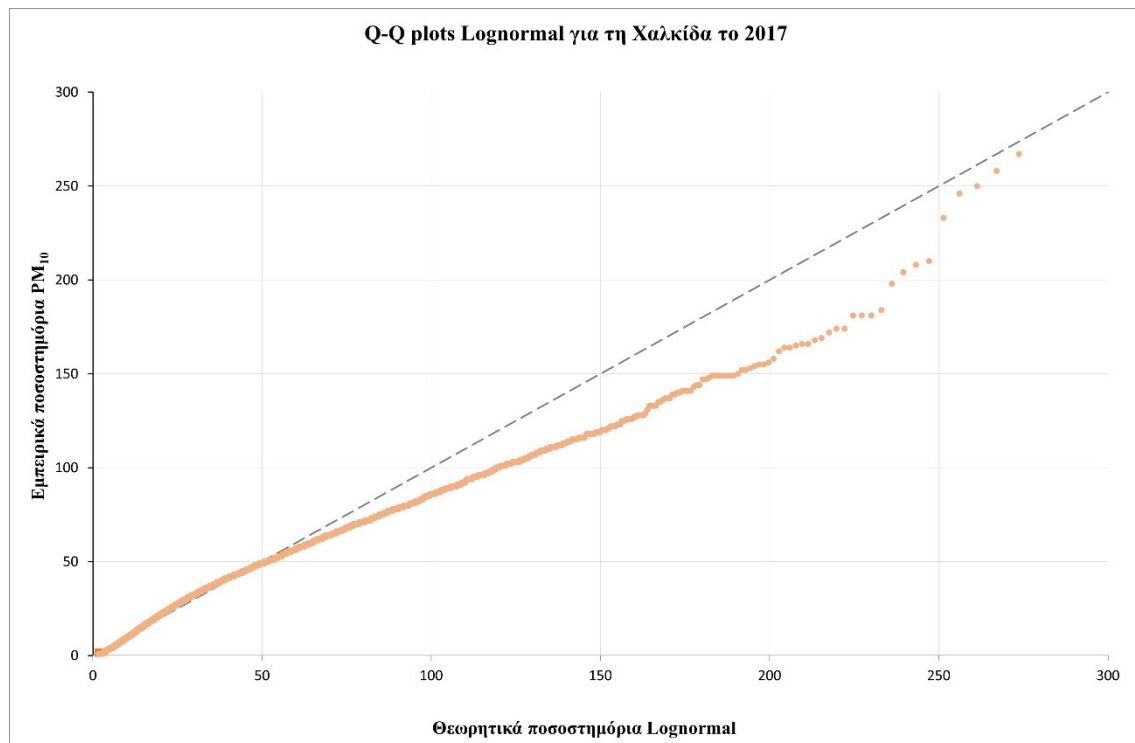
Η εικόνα διαφοροποιείται μετά το μέσο περίπου της κατανομής. Τα σημεία αρχίζουν να μετατοπίζονται σταδιακά πάνω από τη διαγώνιο και η απόκλιση αυξάνεται συνεχώς με την άνοδο των θεωρητικών ποσοστημορίων. Η συστηματική αυτή καμπυλότητα δείχνει ότι η Weibull υποεκτιμά προοδευτικά τις εμπειρικές τιμές του ανώτερου τμήματος.

Στα υψηλότερα ποσοστημόρια η απόσταση από τη γραμμή γίνεται ιδιαίτερα μεγάλη. Τα εμπειρικά σημεία αυξάνονται πολύ ταχύτερα από τα θεωρητικά και οι τελευταίες παρατηρήσεις απομακρύνονται απότομα προς τα επάνω. Επομένως, η Weibull δεν διαθέτει επαρκώς βαριά δεξιά ουρά για την αναπαράσταση των υψηλών συγκεντρώσεων της Χαλκίδας.

Το εύρημα αυτό διαφοροποιεί ουσιαστικά την εικόνα που είχε προκύψει από τα προηγούμενα διαγνωστικά εργαλεία. Η Weibull είχε εμφανίσει πολύ καλή συμφωνία ανά κλάση, συνεκτική σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων και ιδιαίτερα περιορισμένα υπόλοιπα στο μεγαλύτερο μέρος της σειράς. Το Q-Q διάγραμμα αποκαλύπτει, όμως, ότι η καλή αυτή απόδοση δεν επεκτείνεται με την ίδια αποτελεσματικότητα στην άνω ουρά.

Συνεπώς, η Weibull αναπαράγει πολύ καλά το κύριο σώμα της εμπειρικής κατανομής, αλλά η έντονη και πρόωμη ανοδική απόκλιση περιορίζει τη συνολική κατανομική της καταλληλότητα.

Λογαριθμοκανονική κατανομή



Σχήμα 4.6.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η λογαριθμοκανονική κατανομή παρουσιάζει ικανοποιητική ευθυγράμμιση στα χαμηλά ποσοστημόρια. Τα σημεία ακολουθούν αρχικά τη γενική κατεύθυνση της διαγωνίου και αποδίδουν σχετικά αποτελεσματικά το κατώτερο τμήμα της εμπειρικής κατανομής.

Μετά το χαμηλό εύρος, τα σημεία μετατοπίζονται σταδιακά κάτω από τη γραμμή αναφοράς. Η απόκλιση διατηρείται σε μεγάλο μέρος του μεσαίου και του μεσαίου προς υψηλό τμήματος, γεγονός που δείχνει ότι τα θεωρητικά ποσοστημόρια της λογαριθμοκανονικής είναι συστηματικά υψηλότερα από τα εμπειρικά. Το μοντέλο τείνει, επομένως, να υπερεκτιμά τις αντίστοιχες συγκεντρώσεις.

Η αρνητική απόκλιση δεν είναι τοπική, αλλά εκτείνεται σε σημαντικό εύρος των ποσοστημορίων. Παρότι η πορεία των σημείων είναι ομαλή και δεν εμφανίζει την έντονη ανοδική καμπυλότητα της Weibull, η παρατεταμένη απόσταση κάτω από τη διαγώνιο υποδηλώνει συστηματική ασυμφωνία ως προς την κλίμακα του κεντρικού και ανώτερου τμήματος.

Στα υψηλότερα απεικονιζόμενα ποσοστημόρια τα σημεία προσεγγίζουν σταδιακά τη γραμμή, αλλά τα περισσότερα παραμένουν κάτω από αυτήν. Συνεπώς, με βάση το συγκεκριμένο διάγραμμα, η λογαριθμοκανονική δεν υποεκτιμά γενικά την άνω ουρά όπως η Weibull· αντίθετα, σε μεγάλο μέρος του ανώτερου εύρους τείνει να προβλέπει υψηλότερα θεωρητικά ποσοστημόρια από εκείνα που παρατηρούνται εμπειρικά.

Η λογαριθμοκανονική εμφανίζει ομαλή και σταθερή καμπύλη, αλλά η ομαλότητα αυτή δεν πρέπει να συγχέεται με εγγύτητα προς τη διαγώνιο. Η συστηματική θέση των σημείων κάτω από τη γραμμή δείχνει ότι η προσαρμογή της δεν είναι πλήρως ισορροπημένη. Παρ' όλα αυτά, η απουσία απότομης ανοδικής απόκλισης στα υψηλότερα ποσοστημόρια την καθιστά σημαντική εναλλακτική ως προς την περιγραφή της δεξιάς ουράς.

Συγκριτική αξιολόγηση

Η συγκριτική εξέταση των τεσσάρων Q-Q διαγραμμάτων μεταβάλλει σε σημαντικό βαθμό την εικόνα που είχε προκύψει από τα προηγούμενα διαγνωστικά στάδια. Η κανονική κατανομή παραμένει σαφώς η λιγότερο κατάλληλη, ενώ μεταξύ των τριών ασύμμετρων μοντέλων εμφανίζονται διαφορετικά και διακριτά πρότυπα απόκλισης.

Η κανονική παρουσιάζει συστηματική ασυμφωνία στο σύνολο σχεδόν της κατανομής. Η εμφάνιση αρνητικών θεωρητικών ποσοστημορίων, η έντονη καμπυλότητα και η μεγάλη ανοδική απόκλιση στο ανώτερο τμήμα επιβεβαιώνουν ότι δεν αποτελεί κατάλληλο μοντέλο για τις μη αρνητικές και έντονα δεξιά ασύμμετρες συγκεντρώσεις της Χαλκίδας.

Η Weibull εμφανίζει την καλύτερη ευθυγράμμιση στο χαμηλό και κεντρικό εύρος, αλλά η προσαρμογή της επιδεινώνεται αισθητά μετά τα μεσαία ποσοστημόρια. Η πρώτη και προοδευτικά αυξανόμενη ανοδική απόκλιση δείχνει έντονη υποεκτίμηση του ανώτερου τμήματος. Το Q-Q διάγραμμα περιορίζει, επομένως, την ισχυρή υποψηφιότητα που είχε αποκτήσει από τα διαγράμματα συχνοτήτων και υπολοίπων.

Η λογαριθμοκανονική παρουσιάζει διαφορετική αστοχία. Δεν εμφανίζει αντίστοιχη απότομη ανοδική απόκλιση, αλλά διατηρεί μεγάλο μέρος των σημείων κάτω από τη γραμμή αναφοράς. Συνεπώς, τείνει να υπερεκτιμά το μέγεθος των μεσαίων και αρκετών υψηλότερων ποσοστημορίων. Η πορεία της είναι ομαλότερη από εκείνη της Weibull, αλλά η παρατεταμένη απόσταση από τη διαγώνιο δεν επιτρέπει να χαρακτηριστεί ως η κατανομή με την καλύτερη συνολική ευθυγράμμιση.

Η Gamma εμφανίζει επίσης συστηματική καμπυλότητα, αλλά διατηρεί σχετικά μικρότερη απόσταση από τη γραμμή σε σημαντικό μέρος του εύρους. Αρχικά υπερεκτιμά ελαφρώς τις χαμηλές τιμές, στη συνέχεια υπερεκτιμά περισσότερο μέρος του μεσαίου και μεσαίου προς υψηλού εύρους και τελικά υποεκτιμά τις υψηλότερες παρατηρήσεις. Παρά τη μεταβολή αυτή, η συνολική απόκλιση της αναπτύσσεται περισσότερο ομαλά και η ευθυγράμμιση της είναι συγκριτικά καλύτερη από εκείνη της κανονικής και της Weibull. Με βάση αποκλειστικά τα Q-Q διαγράμματα, η Gamma φαίνεται να παρέχει την περισσότερο ισορροπημένη συνολική αντιστοίχιση μεταξύ θεωρητικών και εμπειρικών ποσοστημορίων. Η λογαριθμοκανονική ακολουθεί ως ισχυρή εναλλακτική, λόγω της ομαλής συμπεριφοράς και της απουσίας απότομης υποεκτίμησης της άνω ουράς, αλλά παρουσιάζει παρατεταμένη υπερεκτίμηση σε μεγάλο μέρος του εύρους. Η Weibull αποδίδει πολύ καλά το κεντρικό τμήμα, όμως η έντονη αστοχία της στα υψηλότερα ποσοστημόρια επιβαρύνει σημαντικά τη συνολική αξιολόγησή της.

Η προσωρινή κατάταξη με βάση το συγκεκριμένο γραφικό κριτήριο μπορεί να αποδοθεί ως:

Gamma > Λογαριθμοκανονική > Weibull > Κανονική

Η κατάταξη αυτή διαφοροποιείται αισθητά από εκείνη των προηγούμενων σταδίων, στα οποία η Weibull είχε εμφανίσει σταθερό πλεονέκτημα. Η διαφοροποίηση δεν αποτελεί αντίφαση, αλλά δείχνει ότι η Weibull αποδίδει αποτελεσματικά τις συχνότητες και το κύριο σώμα των δεδομένων, χωρίς να αναπαράγει αντίστοιχα καλά το μέγεθος των υψηλότερων παρατηρήσεων. Αντίστοιχα, η Gamma και η λογαριθμοκανονική εμφανίζουν μεγαλύτερες αποκλίσεις σε ορισμένα διαγράμματα συχνότητας, αλλά παρουσιάζουν συγκριτικά πιο ελεγχόμενη συμπεριφορά στα ποσοστημόρια του ανώτερου εύρους. Επομένως, η Χαλκίδα αποτελεί χαρακτηριστική περίπτωση στην οποία διαφορετικά κριτήρια προσαρμογής αναδεικνύουν διαφορετικά πλεονεκτήματα των θεωρητικών μοντέλων.

Η τελική επιλογή δεν μπορεί ακόμη να θεωρηθεί δεδομένη. Η αριθμητική αξιολόγηση μέσω του στατιστικού χ^2 θα δείξει αν η καλή συνολική εικόνα της Gamma στα Q-Q διαγράμματα συνοδεύεται και από μικρότερη σταθμισμένη απόκλιση συχνοτήτων ή αν η υπεροχή της Weibull στα προηγούμενα διαγνωστικά στάδια αποδεικνύεται ισχυρότερη στη συνολική αριθμητική αποτίμηση.

4.6.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2

Η αριθμητική αποτίμηση της προσαρμογής στη Χαλκίδα παρουσιάζει ιδιαίτερο ενδιαφέρον, επειδή τα προηγούμενα διαγνωστικά εργαλεία δεν ανέδειξαν απολύτως ταυτόσημη εικόνα. Η Weibull εμφάνισε σαφές πλεονέκτημα στη σύγκριση των συχνοτήτων και στα διαγράμματα υπολοίπων, ενώ τα Q-Q διαγράμματα ανέδειξαν εντονότερη αδυναμία της στο ανώτερο τμήμα των ποσοστημορίων. Για τον λόγο αυτό, το στατιστικό χ^2 χρησιμοποιείται εδώ για να διαπιστωθεί ποιο μοντέλο παρουσιάζει τη μικρότερη συνολική σταθμισμένη απόκλιση στις κοινές κλάσεις συγκέντρωσης.

Η μεθοδολογία υπολογισμού του χ^2 έχει ήδη παρουσιαστεί στις προηγούμενες ενότητες και δεν επαναλαμβάνεται αναλυτικά. Στην παρούσα υποενότητα ο δείκτης χρησιμοποιείται ως συγκριτικό μέτρο κατάταξης και όχι ως αυτοτελές κριτήριο αποδοχής ή απόρριψης μιας θεωρητικής κατανομής. Μικρότερη τιμή αντιστοιχεί σε μικρότερη συνολική απόκλιση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων, υπό την προϋπόθεση ότι για όλα τα μοντέλα χρησιμοποιούνται το ίδιο σύνολο δεδομένων και το ίδιο σύστημα κλάσεων.

Παρά τη χρησιμότητά του, το χ^2 δεν αποτυπώνει άμεσα σε ποια περιοχή της κατανομής εντοπίζονται οι αστοχίες. Επομένως, ένα μοντέλο μπορεί να εμφανίζει πολύ χαμηλή συνολική τιμή και ταυτόχρονα να δυσκολεύεται στην άνω ουρά ή σε συγκεκριμένες κλάσεις. Η αριθμητική αξιολόγηση πρέπει, συνεπώς, να ερμηνεύεται σε συνδυασμό με τα προηγούμενα γραφικά ευρήματα.

Κατανομή	χ^2	Θέση κατάταξης
Weibull	46,09	1
Lognormal	234,53	2
Gamma	591,19	3
Normal	2062,81	4

Πίνακας 4.6.5. Τιμές του στατιστικού χ^2 και σχετική κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα το 2017.

Η Weibull παρουσιάζει τη μικρότερη τιμή χ^2 , ίση με 46,09, και καταλαμβάνει με μεγάλη διαφορά την πρώτη θέση. Το αποτέλεσμα δείχνει ότι, στο κοινό σύστημα κλάσεων που χρησιμοποιήθηκε, οι αναμενόμενες συχνότητές της βρίσκονται συνολικά πολύ πιο κοντά στις παρατηρούμενες από ό,τι οι αντίστοιχες τιμές των άλλων μοντέλων.

Η αριθμητική υπεροχή της Weibull δεν είναι οριακή. Η τιμή της λογαριθμοκανονικής είναι περίπου πενταπλάσια, της Gamma σχεδόν δεκατριπλάσια και της κανονικής περίπου σαρανταπενταπλάσια. Η μεγάλη αυτή απόσταση ενισχύει σημαντικά την ένδειξη ότι η Weibull αποδίδει με ιδιαίτερη αποτελεσματικότητα τη συνολική κατανομή των συχνοτήτων στις εξεταζόμενες κλάσεις.

Η λογαριθμοκανονική ακολουθεί στη δεύτερη θέση με $\chi^2 = 234,53$. Η τιμή της είναι αισθητά μεγαλύτερη από εκείνη της Weibull, αλλά σαφώς χαμηλότερη από τις τιμές της Gamma και της κανονικής κατανομής. Το αποτέλεσμα επιβεβαιώνει ότι αποτελεί ισχυρή εναλλακτική, χωρίς όμως να επιτυγχάνει την ίδια συνολική αριθμητική συμφωνία στις κλάσεις συγκέντρωσης.

Η Gamma κατατάσσεται τρίτη, με $\chi^2 = 591,19$. Η τιμή αυτή είναι συμβατή με τις οργανωμένες αποκλίσεις που παρατηρήθηκαν τόσο στα ραβδογράμματα όσο και στα διαγράμματα υπολοίπων. Παρότι το Q-Q διάγραμμα της Gamma εμφάνισε σχετικά ισορροπημένη συμπεριφορά σε σημαντικό μέρος των ποσοστημορίων, η συνολική απόκλιση των συχνοτήτων ανά κλάση παραμένει αισθητά μεγαλύτερη από εκείνη των Weibull και λογαριθμοκανονικής.

Η κανονική κατανομή εμφανίζει τη μεγαλύτερη τιμή, $\chi^2 = 2.062,81$, και κατατάσσεται τελευταία με πολύ μεγάλη διαφορά. Η επίδοσή της ποσοτικοποιεί τις συστηματικές αστοχίες που είχαν ήδη αναδειχθεί γραφικά. Η συμμετρική της μορφή δεν μπορεί να αναπαραστήσει αποτελεσματικά την έντονη θετική ασυμμετρία, τη μεγάλη διασπορά και την εκτεταμένη δεξιά ουρά της χρονοσειράς της Χαλκίδας.

Η αριθμητική κατάταξη συμφωνεί σε μεγάλο βαθμό με τα τρία πρώτα διαγνωστικά στάδια. Η Weibull είχε ήδη εμφανίσει την πιο ισορροπημένη συμφωνία στις επιμέρους κλάσεις, την περισσότερο συνεκτική διάταξη στα διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων και το μικρότερο περίπου εύρος υπολοίπων. Η πολύ χαμηλή τιμή χ^2 επιβεβαιώνει ότι τα επιμέρους αυτά πλεονεκτήματα συνδυάζονται και σε επίπεδο συνολικής αριθμητικής απόδοσης.

Η μόνη ουσιαστική επιφύλαξη προκύπτει από το Q-Q διάγραμμα, όπου η Weibull παρουσίασε έντονη ανοδική απόκλιση στα υψηλότερα ποσοστημόρια και, επομένως,

υποεκτίμηση των υψηλών εμπειρικών τιμών. Το εύρημα αυτό δεν αναιρεί τη χαμηλή τιμή χ^2 , αλλά δείχνει ότι τα δύο κριτήρια αξιολογούν διαφορετικές πτυχές της προσαρμογής. Το χ^2 εξετάζει τις σταθμισμένες αποκλίσεις των συχνοτήτων στις επιλεγμένες κλάσεις, ενώ το Q-Q διάγραμμα αξιολογεί άμεσα την αντιστοίχιση του μεγέθους των ποσοστημορίων και αναδεικνύει εντονότερα τη συμπεριφορά της ουράς.

Η λογαριθμοκανονική παρουσιάζει την αντίστροφη περίπου εικόνα. Δεν προσεγγίζει τις συχνότητες ανά κλάση με την ίδια ακρίβεια με τη Weibull, αλλά εμφανίζει ομαλότερη συμπεριφορά σε τμήμα του ανώτερου εύρους των ποσοστημορίων. Αυτό εξηγεί γιατί παραμένει σοβαρή υποψήφια, παρά τη σημαντικά υψηλότερη τιμή χ^2 .

Η Gamma, από την άλλη πλευρά, εμφάνισε σχετικά ευνοϊκή εικόνα στο Q-Q διάγραμμα, αλλά η τρίτη θέση της στο χ^2 δείχνει ότι η συμφωνία των ποσοστημορίων δεν συνοδεύεται από εξίσου καλή αναπαράσταση των συνολικών συχνοτήτων ανά κλάση. Η διαφοροποίηση αυτή υπογραμμίζει ότι η ομαλή γεωμετρική πορεία ενός Q-Q διαγράμματος δεν αρκεί από μόνη της για να θεμελιώσει συνολική υπεροχή.

Με βάση το στατιστικό χ^2 , η σχετική κατάταξη είναι:

$$\text{Weibull} < \text{Λογαριθμοκανονική} < \text{Gamma} < \text{Κανονική}$$

όπου η μικρότερη τιμή αντιστοιχεί σε καλύτερη σχετική αριθμητική προσαρμογή.

Η πολύ μεγάλη διαφορά μεταξύ της Weibull και των υπόλοιπων μοντέλων προσδίδει ιδιαίτερη βαρύτητα στο αποτέλεσμα. Δεν πρόκειται απλώς για μικρή βελτίωση, αλλά για σαφώς μικρότερη συνολική σταθμισμένη απόκλιση. Παράλληλα, η Weibull έχει ήδη παρουσιάσει επαναλαμβανόμενα καλή εικόνα σε περισσότερα από ένα γραφικά κριτήρια. Επομένως, η αριθμητική αξιολόγηση δεν αποτελεί απομονωμένο εύρημα, αλλά ενισχύει μια τάση που είχε διαμορφωθεί από τα προηγούμενα στάδια.

Ωστόσο, η πρώτη θέση της Weibull δεν πρέπει να ερμηνευθεί ως τέλεια αναπαράσταση ολόκληρης της εμπειρικής κατανομής. Η έντονη απόκλιση στο ανώτερο τμήμα του Q-Q διαγράμματος δείχνει ότι το μοντέλο εξακολουθεί να υποεκτιμά το μέγεθος ορισμένων υψηλών συγκεντρώσεων. Η υπεροχή της αφορά, συνεπώς, την καλύτερη συνολική απόδοση μεταξύ των τεσσάρων εξεταζόμενων κατανομών και όχι την πλήρη εξάλειψη όλων των τοπικών αστοχιών.

Συνολικά, η αριθμητική αξιολόγηση μετατοπίζει καθαρά το βάρος της τελικής επιλογής προς τη Weibull. Η λογαριθμοκανονική παραμένει η πλησιέστερη εναλλακτική, κυρίως

λόγω της ομαλότερης συμπεριφοράς της σε ορισμένα τμήματα των ποσοστημορίων. Η Gamma παρουσιάζει ενδιάμεση αλλά αισθητά ασθενέστερη συνολική απόδοση, ενώ η κανονική κατανομή υπολείπεται σημαντικά σε όλα σχεδόν τα κριτήρια.

Η οριστική επιλογή στην επόμενη υποενότητα πρέπει, επομένως, να βασιστεί στη σύνθεση δύο βασικών ευρημάτων: αφενός της πολύ ισχυρής αριθμητικής και γραφικής υπεροχής της Weibull στις συχνότητες και στα υπόλοιπα και αφετέρου της διαπιστωμένης αδυναμίας της στην αναπαράσταση των υψηλότερων ποσοστημορίων. Αυτή η σύνθεση θα καθορίσει αν η συνολική της υπεροχή υπερκαλύπτει τη συγκεκριμένη αστοχία στην άνω ουρά.

4.6.6 Επιλογή καταλληλότερης κατανομής για τη Χαλκίδα 2017

Η τελική επιλογή της θεωρητικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Χαλκίδα κατά το έτος 2017 βασίστηκε στη συνδυαστική ερμηνεία όλων των γραφικών και αριθμητικών διαγνωστικών κριτηρίων. Η συγκεκριμένη περίπτωση ήταν περισσότερο σύνθετη από την Άμφισσα και τη Λαμία, καθώς τα επιμέρους εργαλεία δεν ανέδειξαν απολύτως ταυτόσημη εικόνα. Η Weibull εμφάνισε σαφές πλεονέκτημα στη σύγκριση των συχνοτήτων, στα διαγνωστικά διαγράμματα και στα υπόλοιπα, ενώ τα Q-Q διαγράμματα αποκάλυψαν περιορισμούς στην αναπαράσταση των υψηλότερων ποσοστημορίων.

Η κανονική κατανομή παρουσίασε με συνέπεια τη χαμηλότερη καταλληλότητα. Στη σύγκριση ανά κλάση εμφάνισε μεγάλες αποκλίσεις, ενώ στα διαγνωστικά διαγράμματα τα σημεία χαρακτηρίζονταν από έντονη διασπορά και συστηματική απομάκρυνση από τη γραμμή ιδανικής συμφωνίας. Στα υπόλοιπα διαμορφώθηκε σαφές κυματοειδές πρότυπο και το Q-Q διάγραμμα ανέδειξε τη δομική ασυμβατότητα της συμμετρικής κανονικής κατανομής με τα μη αρνητικά και έντονα δεξιά ασύμμετρα δεδομένα της Χαλκίδας. Η πολύ υψηλή τιμή $\chi^2 = 2.062,81$ επιβεβαίωσε αριθμητικά την ασθενή συνολική απόδοσή της.

Η Gamma αποτέλεσε ουσιαστική βελτίωση έναντι της κανονικής, αλλά η προσαρμογή της δεν ήταν ομοιόμορφη. Στα ραβδογράμματα υπερεκτίμησε τις ακραίες συγκεντρωτικές κατηγορίες και υποεκτίμησε αρκετές ενδιάμεσες κλάσεις. Στα διαγράμματα υπολοίπων εμφανίστηκαν μεγάλες αρχικές αποκλίσεις και διαδοχικές περιοχές υπερεκτίμησης και

υποεκτίμησης. Το Q-Q διάγραμμα παρουσίασε σχετικά ισορροπημένη πορεία σε σημαντικό μέρος των ποσοστημορίων, χωρίς όμως να εξαλείφει τις αποκλίσεις στο μεσαίο και ανώτερο εύρος. Η τιμή $\chi^2 = 591,19$ την κατέταξε στην τρίτη θέση.

Η λογαριθμοκανονική κατανομή εμφάνισε σαφώς καλύτερη συμπεριφορά από τη Gamma και την κανονική. Στα διαγράμματα συχνοτήτων απέδωσε ικανοποιητικά τη γενική μορφή της εμπειρικής κατανομής, αλλά παρουσίασε αισθητές αποκλίσεις σε ορισμένες χαμηλομεσαίες και μεσαίες κλάσεις. Στα διαγράμματα υπολοίπων οι μεγαλύτερες αστοχίες περιορίστηκαν κυρίως στο αρχικό τμήμα, ενώ στη συνέχεια τα υπόλοιπα συγκεντρώθηκαν κοντά στη γραμμή μηδενός.

Στο Q-Q διάγραμμα η λογαριθμοκανονική παρουσίασε ομαλή, αλλά όχι πλήρως ευθυγραμμισμένη πορεία. Σε σημαντικό μέρος του μεσαίου και ανώτερου εύρους τα σημεία βρίσκονταν κάτω από τη γραμμή αναφοράς, γεγονός που υποδηλώνει συστηματική υπερεκτίμηση των αντίστοιχων εμπειρικών ποσοστημορίων. Παρότι η απουσία απότομης ανοδικής απόκλισης αποτελεί θετικό στοιχείο, η παρατεταμένη απόσταση από τη διαγώνιο δεν επιτρέπει να θεωρηθεί ότι η λογαριθμοκανονική παρουσιάζει την καλύτερη συνολική προσαρμογή. Η τιμή $\chi^2 = 234,53$ την καθιστά τη σημαντικότερη εναλλακτική κατανομή.

Η Weibull εμφάνισε την πιο συνεπή υπεροχή στα περισσότερα στάδια της αξιολόγησης. Στη σύγκριση ανά κλάση παρουσίασε τις μικρότερες και πιο ισορροπημένες αποκλίσεις, ιδίως στις κλάσεις 20-30 και 30-40 $\mu\text{g}/\text{m}^3$. Στα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων τα σημεία της ακολούθησαν την περισσότερο οργανωμένη και σχεδόν γραμμική πορεία γύρω από τη γραμμή αναφοράς.

Αντίστοιχα, στα διαγράμματα υπολοίπων η Weibull εμφάνισε μικρότερο περίπου εύρος ακραίων αποκλίσεων και ταχύτερη σταθεροποίηση γύρω από το μηδέν. Οι σημαντικότερες αστοχίες της περιορίστηκαν στο αρχικό τμήμα της ταξινομημένης σειράς, ενώ στο μεγαλύτερο μέρος του διαγράμματος τα υπόλοιπα παρέμειναν εξαιρετικά μικρά.

Η βασική επιφύλαξη προέκυψε από το Q-Q διάγραμμα. Παρότι η συμφωνία ήταν πολύ καλή στα χαμηλά και μεσαία ποσοστημόρια, στα υψηλότερα σημεία εμφανίστηκε έντονη και προοδευτικά αυξανόμενη ανοδική απόκλιση. Το εύρημα αυτό δείχνει ότι η Weibull υποεκτιμά το μέγεθος των υψηλών συγκεντρώσεων και δεν αποδίδει πλήρως την εκτεταμένη δεξιά ουρά της Χαλκίδας.

Η αδυναμία αυτή είναι ουσιαστική, αλλά δεν αναιρεί τη συνολική υπεροχή της στα υπόλοιπα κριτήρια. Η αριθμητική αξιολόγηση παρείχε ιδιαίτερα ισχυρή επιβεβαίωση, καθώς η Weibull κατέγραψε: $\chi^2=46,09$

έναντι 234,53 για τη λογαριθμοκανονική, 591,19 για τη Gamma και 2.062,81 για την κανονική κατανομή. Η τιμή της λογαριθμοκανονικής είναι περίπου πενταπλάσια, γεγονός που δείχνει ότι η αριθμητική υπεροχή της Weibull δεν είναι οριακή.

Η διαφοροποίηση μεταξύ των Q-Q διαγραμμάτων και των υπόλοιπων κριτηρίων δεν αποτελεί αντίφαση. Τα ραβδογράμματα, τα διαγνωστικά διαγράμματα και το χ^2 αξιολογούν κυρίως τη συμφωνία των συχνοτήτων, ενώ τα Q-Q διαγράμματα εξετάζουν την αντιστοίχιση των ποσοστημορίων και είναι ιδιαίτερα ευαίσθητα στην απόδοση της άνω ουράς. Συνεπώς, η Weibull μπορεί να αναπαράγει εξαιρετικά τη συνολική κατανομή των συχνοτήτων και ταυτόχρονα να υποεκτιμά ορισμένες από τις μεγαλύτερες παρατηρήσεις.

Με βάση τη σύνθεση όλων των ευρημάτων, η Weibull επιλέγεται ως η καταλληλότερη μεταξύ των τεσσάρων εξεταζόμενων κατανομών για τη Χαλκίδα το 2017. Η επιλογή της στηρίζεται στη συνεπή υπεροχή της στα τρία πρώτα γραφικά στάδια και, κυρίως, στη σαφώς χαμηλότερη τιμή του στατιστικού χ^2 . Η λογαριθμοκανονική αποτελεί την ισχυρότερη εναλλακτική, λόγω της σχετικά ομαλής συμπεριφοράς της σε μεγάλο μέρος της κατανομής.

Η τελική επιλογή δεν προκύπτει, επομένως, από μια απλή αριθμητική σειρά, αλλά από τη διαφορετική βαρύτητα των επιμέρους διαγνωστικών ευρημάτων. Η Weibull αναδεικνύεται ως το μοντέλο με την πιο ολοκληρωμένη συνολική απόδοση, η λογαριθμοκανονική διατηρεί ισχυρή αλλά δευτερεύουσα υποψηφιότητα, η Gamma παρουσιάζει περιορισμένη συνέπεια και η κανονική κατανομή υπολείπεται σαφώς.

Η επιλογή της Weibull πρέπει να ερμηνευθεί συγκριτικά και όχι ως ένδειξη τέλει προσαρμογής. Η απόκλιση στα υψηλότερα ποσοστημόρια δείχνει ότι το μοντέλο δεν αναπαράγει πλήρως τις πλέον υψηλές συγκεντρώσεις. Ωστόσο, στο σύνολο των εξεταζόμενων κριτηρίων παρέχει την καλύτερη ισορροπία μεταξύ συμφωνίας των συχνοτήτων, περιορισμού των υπολοίπων και αριθμητικής προσαρμογής.

Το αποτέλεσμα διαφοροποιείται από εκείνα της Άμφισσας και της Λαμίας, όπου είχε επιλεγεί η λογαριθμοκανονική κατανομή. Η διαφοροποίηση αυτή δείχνει ότι κοινά χαρακτηριστικά, όπως η θετική ασυμμετρία και η παρουσία δεξιάς ουράς, δεν

συνεπάγονται αναγκαστικά την ίδια βέλτιστη θεωρητική κατανομή. Κατά συνέπεια, η επιλογή του μοντέλου πρέπει να πραγματοποιείται χωριστά για κάθε χρονοσειρά και να στηρίζεται στη σύνθεση συμπληρωματικών διαγνωστικών εργαλείων.

4.7 Προσαρμογή θεωρητικών κατανομών στη Λειβαδιά για το έτος 2019

Η ανάλυση συνεχίζεται με τη χρονοσειρά των ωριαίων συγκεντρώσεων PM_{10} στη Λειβαδιά για το έτος 2019. Η εμπειρική κατανομή παρουσιάζει σαφή θετική ασυμμετρία, υψηλή κύρτωση και παρουσία περιορισμένου αριθμού πολύ μεγάλων τιμών. Τα χαρακτηριστικά αυτά υποδηλώνουν συγκέντρωση των περισσότερων παρατηρήσεων στις χαμηλές και μεσαίες τιμές, σε συνδυασμό με εκτεταμένη δεξιά ουρά.

Η συγκεκριμένη χρονοσειρά είναι ιδιαίτερα απαιτητική ως προς την προσαρμογή θεωρητικών μοντέλων. Μια κατανομή μπορεί να αποδίδει ικανοποιητικά το κύριο σώμα των δεδομένων, αλλά να υποεκτιμά τις υψηλότερες παρατηρήσεις. Αντίστροφα, ένα μοντέλο με μεγαλύτερη ευελιξία στην άνω ουρά μπορεί να εμφανίζει αποκλίσεις στις κλάσεις όπου συγκεντρώνεται η πλειονότητα των τιμών. Επομένως, η αξιολόγηση πρέπει να λαμβάνει υπόψη τόσο το κεντρικό τμήμα όσο και το ανώτερο άκρο της κατανομής.

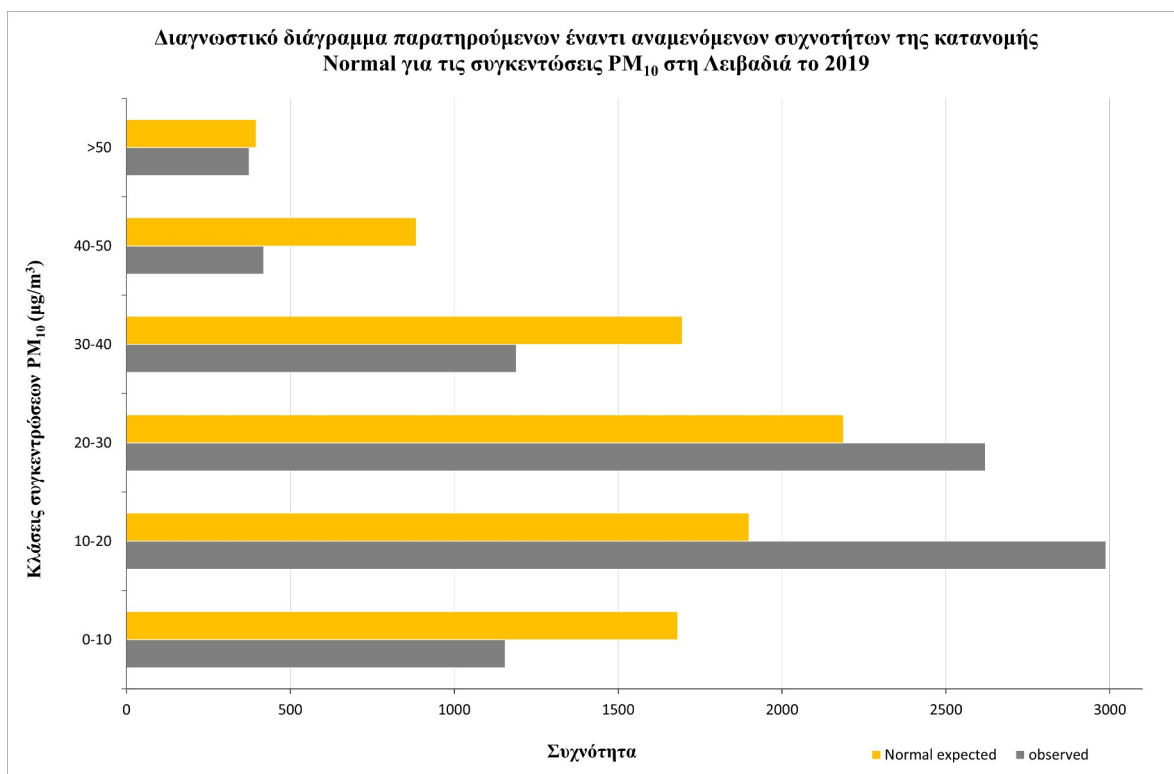
Στην παρούσα ενότητα εξετάζονται η κανονική, η Gamma, η Weibull και η λογαριθμοκανονική κατανομή, με την ίδια μεθοδολογική διαδικασία που εφαρμόστηκε στις προηγούμενες πόλεις. Η κοινή προσέγγιση εξασφαλίζει τη συγκρισιμότητα των αποτελεσμάτων, χωρίς να προϋποθέτει ότι η ίδια κατανομή θα είναι καταλληλότερη σε κάθε χρονοσειρά.

Η σύγκριση βασίζεται στις παρατηρούμενες και αναμενόμενες συχνότητες ανά κλάση, στα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων, στα διαγράμματα υπολοίπων, στα Q-Q διαγράμματα και στην αριθμητική αξιολόγηση μέσω του στατιστικού χ^2 . Η τελική επιλογή θα προκύψει από τη συνολική ισορροπία μεταξύ της προσαρμογής στις συνηθέστερες συγκεντρώσεις και της ικανότητας κάθε μοντέλου να αναπαραστήσει την έντονη δεξιά ουρά της Λειβαδιάς.

4.7.1 Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων

Η πρώτη ένδειξη για την ποιότητα προσαρμογής των τεσσάρων θεωρητικών κατανομών προκύπτει από την άμεση αντιπαραβολή των παρατηρούμενων και αναμενόμενων συχνοτήτων στις έξι κλάσεις συγκεντρώσεως. Στη Λειβαδιά, η αξιολόγηση επικεντρώνεται κυρίως στις κλάσεις 10-20 και 20-30 $\mu\text{g}/\text{m}^3$, στις οποίες συγκεντρώνεται το μεγαλύτερο μέρος των ωριαίων παρατηρήσεων. Παράλληλα, εξετάζεται αν κάθε μοντέλο αποδίδει ικανοποιητικά τη σταδιακή μείωση των συχνοτήτων προς τις υψηλότερες συγκεντρώσεις. Η συγκεκριμένη σύγκριση παρέχει μια συνοπτική εικόνα της κατανομής της θεωρητικής πιθανότητας μεταξύ των κλάσεων. Δεν αποκαλύπτει, ωστόσο, τη μορφή των αποκλίσεων στο εσωτερικό κάθε διαστήματος. Ιδίως η ανοικτή κλάση άνω των 50 $\mu\text{g}/\text{m}^3$ συγκεντρώνει όλες τις υψηλότερες τιμές και, επομένως, επιτρέπει την αξιολόγηση της συνολικής συχνότητάς τους, όχι όμως της ακριβούς μορφής της δεξιάς ουράς.

Κανονική κατανομή

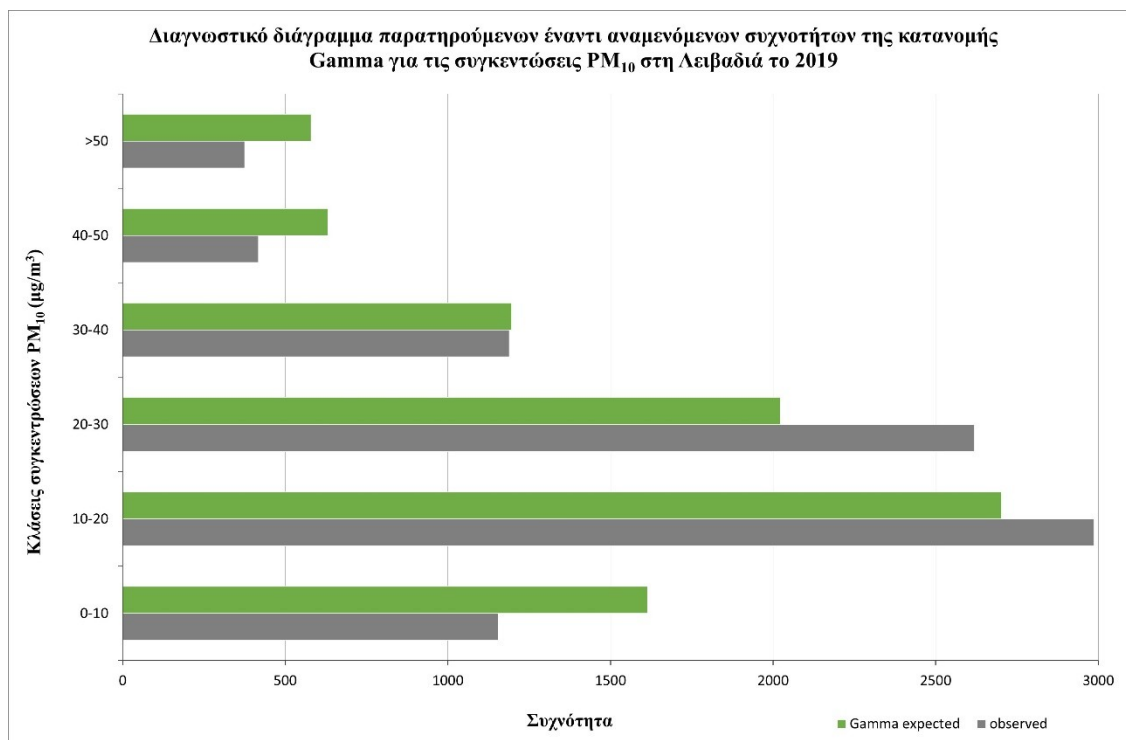


Σχήμα 4.7.1α. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η κανονική κατανομή παρουσιάζει εκτεταμένη ανακατανομή των θεωρητικών συχνοτήτων μακριά από τις δύο κυρίαρχες κλάσεις. Στην κλάση 10-20 $\mu\text{g}/\text{m}^3$ η αναμενόμενη συχνότητα είναι αισθητά χαμηλότερη από την παρατηρούμενη, ενώ υποεκτίμηση εμφανίζεται και στην κλάση 20-30 $\mu\text{g}/\text{m}^3$. Επομένως, το μοντέλο δεν αποδίδει με επάρκεια το τμήμα στο οποίο συγκεντρώνεται η πλειονότητα των μετρήσεων. Αντίθετα, στις κλάσεις 0-10, 30-40 και 40-50 $\mu\text{g}/\text{m}^3$ οι αναμενόμενες συχνότητες είναι υψηλότερες από τις παρατηρούμενες. Ιδιαίτερα εμφανής είναι η υπερεκτίμηση στις κλάσεις 30-40 και 40-50 $\mu\text{g}/\text{m}^3$, γεγονός που δείχνει ότι η κανονική κατανομή μεταφέρει υπερβολική θεωρητική μάζα από το κύριο σώμα προς τις γειτονικές υψηλότερες κλάσεις. Στην ανοικτή κλάση άνω των 50 $\mu\text{g}/\text{m}^3$ η διαφορά είναι μικρότερη, αλλά το μοντέλο εξακολουθεί να εμφανίζει ελαφρά υπερεκτίμηση.

Η εικόνα αυτή είναι χαρακτηριστική ενός συμμετρικού μοντέλου που δυσκολεύεται να αναπαραστήσει την απότομη συγκέντρωση των δεδομένων στις χαμηλομεσαίες τιμές και τη σταδιακή ανάπτυξη μιας δεξιάς ουράς. Συνεπώς, η κανονική κατανομή παρουσιάζει τη λιγότερο ισορροπημένη προσαρμογή στο συγκεκριμένο στάδιο.

Κατανομή Gamma



Σχήμα 4.7.1β. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

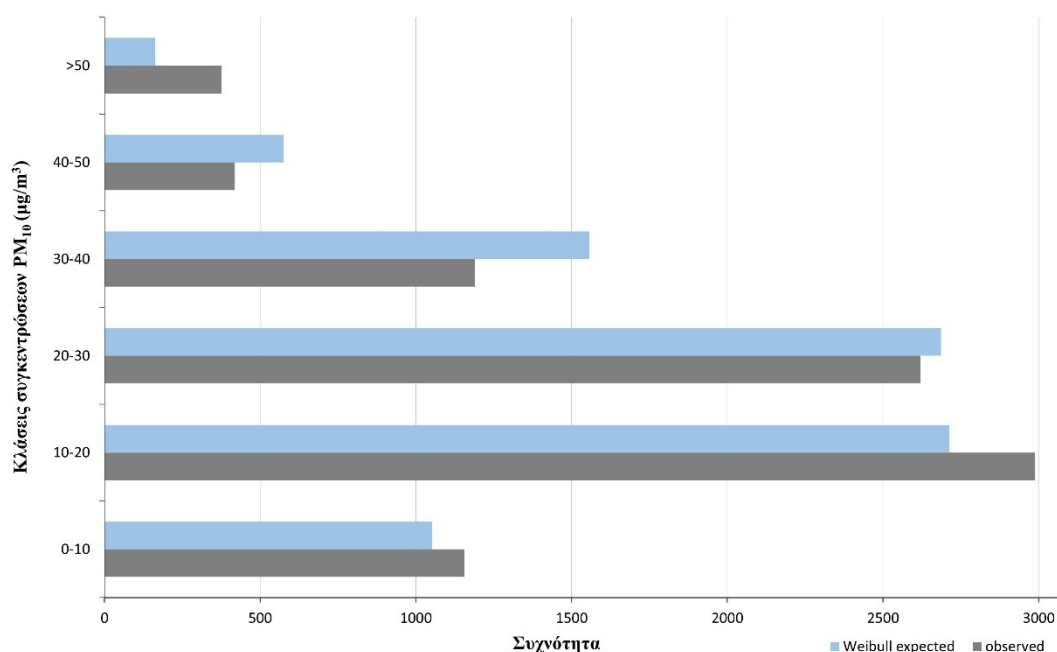
Η Gamma ανταποκρίνεται καλύτερα στη θετική ασυμμετρία των δεδομένων, αλλά εμφανίζει συγκεκριμένο πρότυπο αποκλίσεων. Η καλύτερη αντιστοίχιση παρατηρείται στην κλάση 30-40 $\mu\text{g}/\text{m}^3$, όπου οι δύο συχνότητες είναι σχεδόν ίσες. Στην κλάση 10-20 $\mu\text{g}/\text{m}^3$ η υποεκτίμηση είναι σχετικά περιορισμένη, αλλά στην κλάση 20-30 $\mu\text{g}/\text{m}^3$ γίνεται εντονότερη.

Παράλληλα, η Gamma υπερεκτιμά την κλάση 0-10 $\mu\text{g}/\text{m}^3$, καθώς και τις δύο ανώτερες κλάσεις 40-50 και άνω των 50 $\mu\text{g}/\text{m}^3$. Η συμπεριφορά αυτή δείχνει ότι το μοντέλο κατανέμει μεγαλύτερη θεωρητική πιθανότητα στα δύο άκρα του εξεταζόμενου εύρους και μικρότερη στο βασικό τμήμα 20-30 $\mu\text{g}/\text{m}^3$.

Η Gamma αποτελεί σαφή βελτίωση έναντι της κανονικής κατανομής, καθώς αποδίδει αποτελεσματικότερα τη γενική ασύμμετρη δομή. Ωστόσο, η ταυτόχρονη υποεκτίμηση του κύριου σώματος και υπερεκτίμηση των υψηλότερων κλάσεων δεν επιτρέπει να χαρακτηριστεί η προσαρμογή της πλήρως ισορροπημένη.

Κατανομή Weibull

Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019



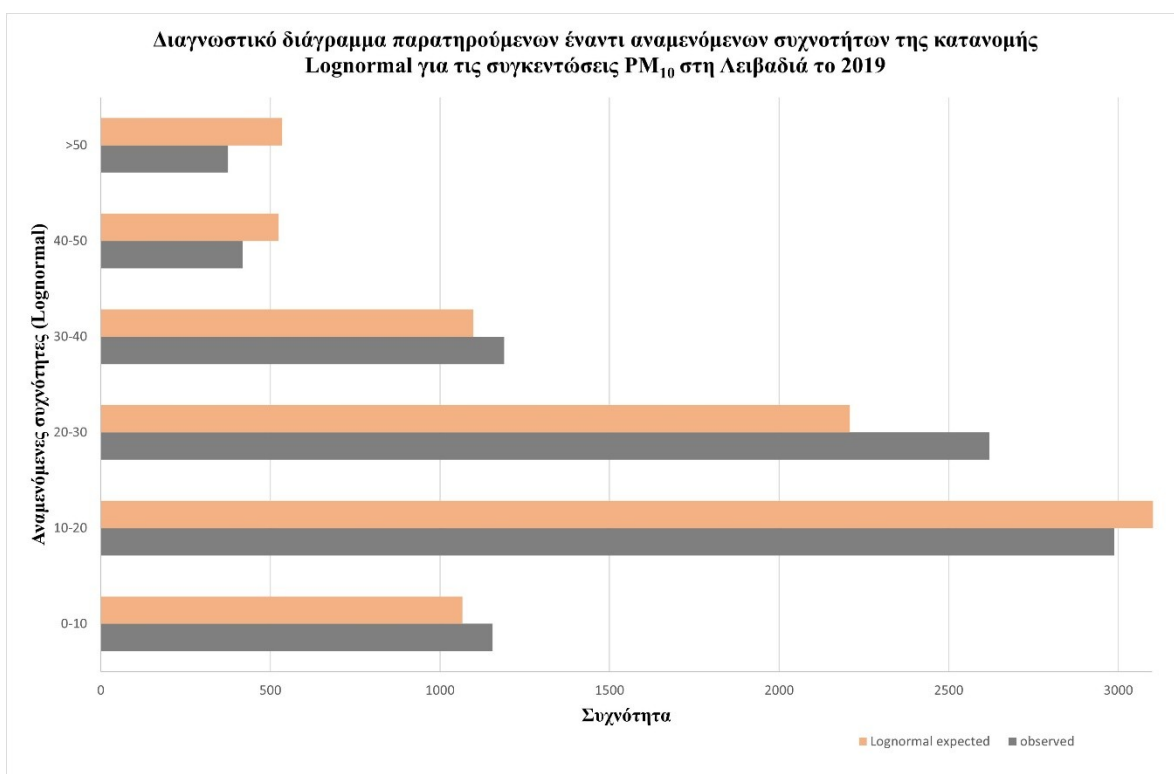
Σχήμα 4.7.1γ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η Weibull εμφανίζει καλή συμφωνία σε σημαντικό μέρος του κύριου σώματος της κατανομής. Στην κλάση 20-30 $\mu\text{g}/\text{m}^3$ η αναμενόμενη συχνότητα βρίσκεται πολύ κοντά στην παρατηρούμενη, ενώ σχετικά περιορισμένη είναι και η απόκλιση στην κλάση 0-10 $\mu\text{g}/\text{m}^3$.

Στην πολυπληθέστερη κλάση 10-20 $\mu\text{g}/\text{m}^3$, όμως, το μοντέλο υποεκτιμά την παρατηρούμενη συχνότητα. Αντίθετα, στις κλάσεις 30-40 και 40-50 $\mu\text{g}/\text{m}^3$ εμφανίζει υπερεκτίμηση, περισσότερο έντονη στην πρώτη από αυτές. Το σημαντικότερο χαρακτηριστικό της προσαρμογής της εντοπίζεται στην ανοικτή κλάση άνω των 50 $\mu\text{g}/\text{m}^3$, όπου η θεωρητική συχνότητα είναι αισθητά χαμηλότερη από την παρατηρούμενη.

Το τελευταίο εύρημα υποδηλώνει ότι η Weibull δεν αποδίδει επαρκώς το συνολικό πλήθος των υψηλών συγκεντρώσεων, παρά την καλή της συμπεριφορά σε ορισμένες κεντρικές κλάσεις. Επομένως, προσφέρει ικανοποιητική περιγραφή του κύριου σώματος, αλλά παρουσιάζει σαφή αδυναμία προς το ανώτερο τμήμα της εμπειρικής κατανομής.

Λογαριθμοκανονική κατανομή



Σχήμα 4.7.1δ. Σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η λογαριθμοκανονική κατανομή παρουσιάζει την περισσότερη ισορροπημένη εικόνα στο σύνολο των κλάσεων. Στην κλάση $10-20 \mu\text{g}/\text{m}^3$, όπου καταγράφεται η μεγαλύτερη συχνότητα, η θεωρητική τιμή βρίσκεται πολύ κοντά στην παρατηρούμενη, με μικρή μόνο υπερεκτίμηση. Η απόκλιση είναι επίσης περιορισμένη στις κλάσεις $0-10$ και $30-40 \mu\text{g}/\text{m}^3$, όπου το μοντέλο υποεκτιμά ελαφρά τις πραγματικές συχνότητες.

Η μεγαλύτερη αστοχία εμφανίζεται στην κλάση $20-30 \mu\text{g}/\text{m}^3$, όπου η λογαριθμοκανονική υποεκτιμά αισθητά τις παρατηρήσεις. Στις κλάσεις $40-50$ και άνω των $50 \mu\text{g}/\text{m}^3$ η κατεύθυνση της απόκλισης αντιστρέφεται και το μοντέλο προβλέπει περισσότερες παρατηρήσεις από όσες καταγράφηκαν.

Παρά τις αποκλίσεις αυτές, η λογαριθμοκανονική αποφεύγει τις πολύ μεγάλες αστοχίες της κανονικής κατανομής και δεν εμφανίζει την έντονη υποεκτίμηση της ανώτερης κλάσης που χαρακτηρίζει τη Weibull. Η συμφωνία της στην πολυπληθέστερη κλάση, σε συνδυασμό με τις σχετικά ελεγχόμενες αποκλίσεις στις γειτονικές περιοχές, της προσδίδει την πληρέστερη οπτική εικόνα στο παρόν στάδιο.

Συγκριτική αξιολόγηση

Τα τέσσερα διαγράμματα αναδεικνύουν διαφορετικούς τρόπους κατανομής της θεωρητικής πιθανότητας μεταξύ των επιμέρους κλάσεων. Η κανονική κατανομή υποεκτιμά τις δύο βασικές περιοχές συγκέντρωσης και υπερεκτιμά αρκετές γειτονικές κλάσεις, με αποτέλεσμα η προσαρμογή της να χαρακτηρίζεται από εκτεταμένη και συστηματική ανακατανομή των συχνοτήτων. Η εικόνα αυτή επιβεβαιώνει ότι η συμμετρική μορφή της δεν ανταποκρίνεται επαρκώς στη θετικά ασύμμετρη δομή των δεδομένων της Λειβαδιάς.

Η Gamma ακολουθεί αποτελεσματικότερα τη γενική ασυμμετρία, αλλά κατανέμει υπερβολική θεωρητική μάζα τόσο στην κατώτερη όσο και στις ανώτερες κλάσεις. Παράλληλα, υποεκτιμά το βασικό τμήμα $20-30 \mu\text{g}/\text{m}^3$, όπου συγκεντρώνεται σημαντικό μέρος των παρατηρήσεων. Επομένως, η βελτίωσή της έναντι της κανονικής είναι σαφής, χωρίς όμως να συνοδεύεται από ομοιόμορφη συμφωνία σε όλο το εύρος.

Η Weibull εμφανίζει ιδιαίτερα καλή αντιστοίχιση στην κλάση $20-30 \mu\text{g}/\text{m}^3$ και σχετικά περιορισμένη απόκλιση στις χαμηλότερες συγκεντρώσεις. Ωστόσο, υποεκτιμά την πολυπληθέστερη κλάση $10-20 \mu\text{g}/\text{m}^3$ και, κυρίως, το συνολικό πλήθος των παρατηρήσεων άνω των $50 \mu\text{g}/\text{m}^3$. Η τελευταία απόκλιση είναι κρίσιμη, επειδή δείχνει ότι η καλή

προσαρμογή στο κύριο σώμα δεν επεκτείνεται με την ίδια αποτελεσματικότητα προς το ανώτερο τμήμα της εμπειρικής κατανομής.

Η λογαριθμοκανονική δεν επιτυγχάνει την καλύτερη συμφωνία σε κάθε επιμέρους κλάση, αλλά παρουσιάζει την πιο ομοιόμορφη συνολική συμπεριφορά. Προσεγγίζει ιδιαίτερα καλά την κλάση 10-20 $\mu\text{g}/\text{m}^3$, στην οποία καταγράφεται η μεγαλύτερη συχνότητα, και διατηρεί σχετικά ελεγχόμενες αποκλίσεις στις περισσότερες γειτονικές κλάσεις. Η βασική της αδυναμία εντοπίζεται στην υποεκτίμηση της κλάσης 20-30 $\mu\text{g}/\text{m}^3$ και στην υπερεκτίμηση των δύο ανώτερων κλάσεων. Παρ' όλα αυτά, οι αποκλίσεις της δεν συγκεντρώνονται αποκλειστικά σε ένα μόνο κρίσιμο τμήμα, αλλά κατανέμονται περισσότερο ισορροπημένα στο συνολικό εύρος.

Με βάση αποκλειστικά τη σύγκριση των ραβδογραμμάτων, η λογαριθμοκανονική παρέχει την πληρέστερη οπτική προσαρμογή. Το πλεονέκτημά της δεν οφείλεται σε τέλεια αντιστοίχιση όλων των κλάσεων, αλλά στο ότι συνδυάζει καλή απόδοση στην πολυπληθέστερη περιοχή με απουσία εξαιρετικά μεγάλης αστοχίας σε μία μόνο κεντρική κλάση. Η Weibull αποτελεί ισχυρή εναλλακτική ως προς το κύριο σώμα, αλλά επιβαρύνεται από την έντονη υποεκτίμηση της ανοικτής ανώτερης κλάσης. Η Gamma εμφανίζει ενδιάμεση απόδοση, ενώ η κανονική κατανομή υπολείπεται σαφώς.

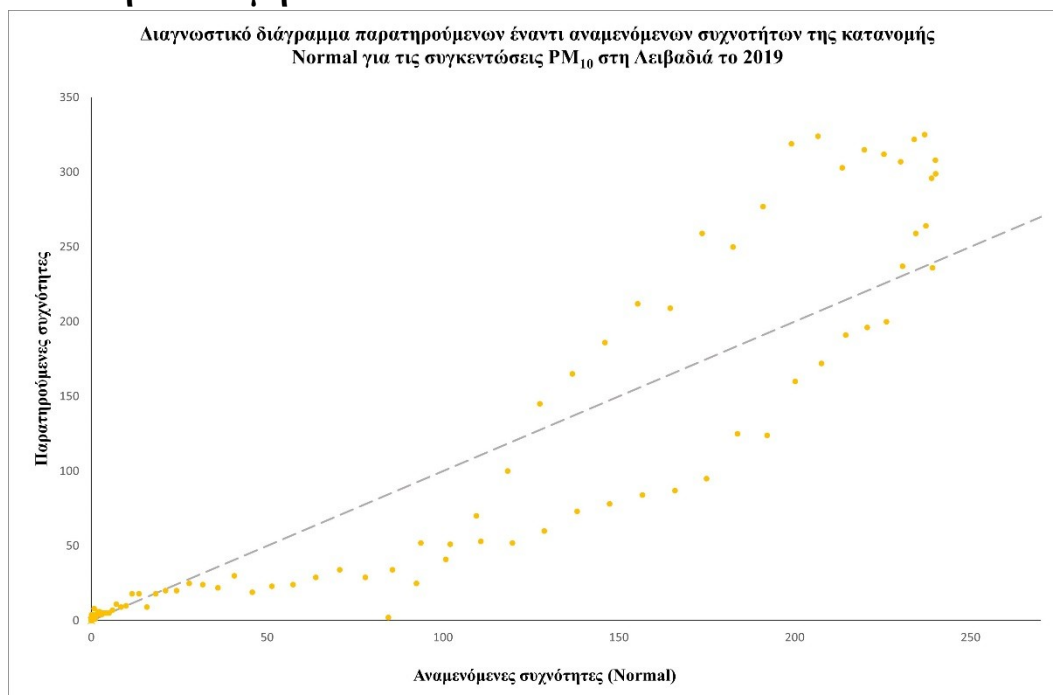
Η συγκεκριμένη εικόνα υποδηλώνει ότι, στη Λειβαδιά, η αξιολόγηση δεν εξαρτάται μόνο από την ακριβή προσαρμογή στην κλάση με τη μεγαλύτερη συχνότητα, αλλά και από το κατά πόσο το μοντέλο αποδίδει ομαλά τη μετάβαση από το κύριο σώμα προς τις υψηλότερες συγκεντρώσεις. Υπό αυτή την έννοια, η λογαριθμοκανονική εμφανίζει στο παρόν στάδιο την καλύτερη ισορροπία μεταξύ κεντρικής προσαρμογής και περιγραφής της δεξιάς πλευράς της κατανομής.

Το συμπέρασμα παραμένει προσωρινό. Η ανοικτή κλάση άνω των 50 $\mu\text{g}/\text{m}^3$ συγκεντρώνει τιμές με πολύ διαφορετικό μέγεθος και δεν επιτρέπει να διαπιστωθεί αν ένα μοντέλο αποδίδει ικανοποιητικά τις μέτρια υψηλές συγκεντρώσεις αλλά αποτυγχάνει στις πλέον ακραίες ή το αντίστροφο. Η διάκριση αυτή θα εξεταστεί πληρέστερα στα διαγράμματα υπολοίπων και, κυρίως, στα Q-Q διαγράμματα. Η τελική αξιολόγηση θα προκύψει από τη σύνθεση των γραφικών ευρημάτων με την αριθμητική αποτίμηση μέσω του στατιστικού χ^2 .

4.7.2 Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων

Τα διαγνωστικά διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων χρησιμοποιούνται για τη διερεύνηση της συνολικής συμφωνίας μεταξύ των εμπειρικών δεδομένων της Λειβαδιάς και των τιμών που προκύπτουν από τις τέσσερις θεωρητικές κατανομές. Η διακεκομμένη διαγώνιος παριστάνει την ιδανική σχέση $O_i = E_i$, όπου η παρατηρούμενη συχνότητα ισούται με την αντίστοιχη αναμενόμενη. Όσο πλησιέστερα βρίσκονται τα σημεία στη γραμμή αυτή, τόσο ισχυρότερη είναι η οπτική ένδειξη καλής προσαρμογής. Η αξιολόγηση δεν περιορίζεται στην απόσταση μεμονωμένων σημείων από τη διαγώνιο. Εξετάζονται επίσης η συνοχή της συνολικής διάταξης, η εμφάνιση συστηματικής καμπυλότητας και η μεταβολή της διασποράς καθώς αυξάνονται οι αναμενόμενες συχνότητες. Σημεία πάνω από τη γραμμή υποδηλώνουν ότι η θεωρητική κατανομή υποεκτιμά την αντίστοιχη παρατηρούμενη συχνότητα, ενώ σημεία κάτω από αυτήν υποδηλώνουν υπερεκτίμηση. Πρέπει να διευκρινιστεί ότι ο οριζόντιος άξονας παριστάνει αναμενόμενες συχνότητες και όχι συγκεντρώσεις PM_{10} . Επομένως, τα σημεία που βρίσκονται στο δεξιό τμήμα των διαγραμμάτων δεν αντιστοιχούν κατ' ανάγκη στις υψηλότερες συγκεντρώσεις, αλλά στις περιοχές με τις μεγαλύτερες θεωρητικά αναμενόμενες συχνότητες. Η συμπεριφορά της άνω ουράς θα αξιολογηθεί αμεσότερα μέσω των Q-Q διαγραμμάτων.

Κανονική κατανομή



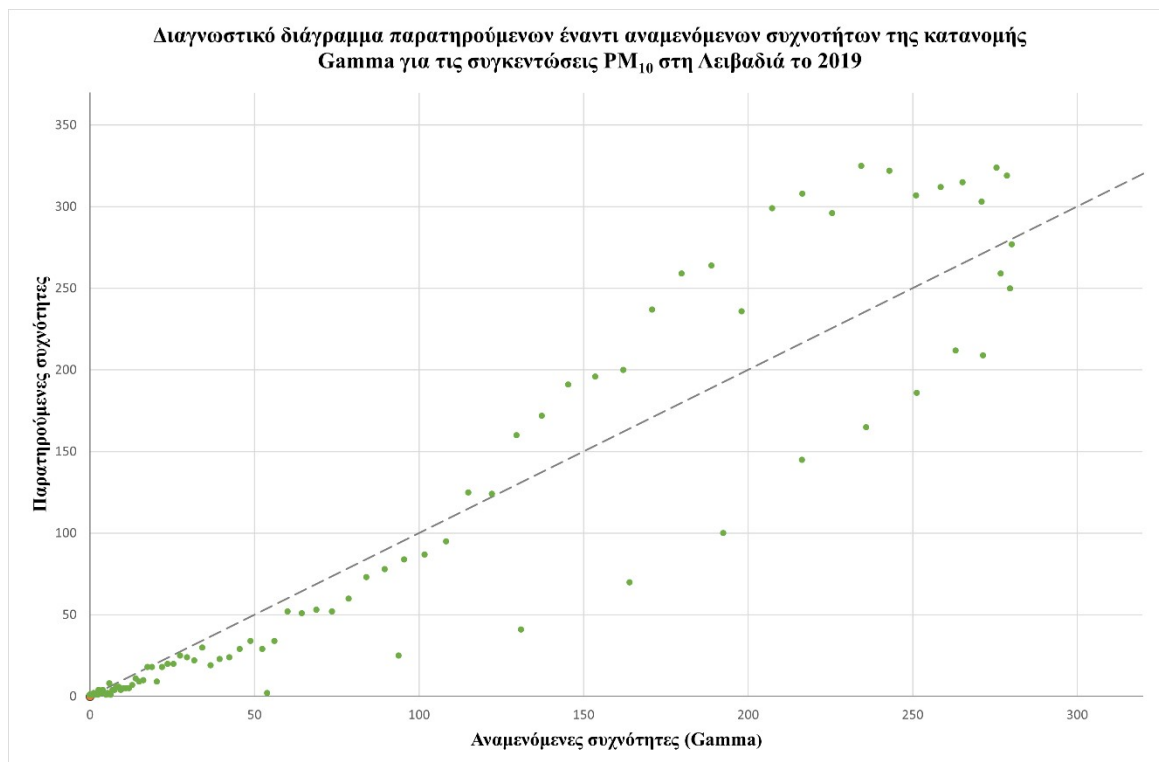
Σχήμα 4.7.2α. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η κανονική κατανομή παρουσιάζει τη λιγότερο συνεκτική διάταξη σημείων μεταξύ των τεσσάρων μοντέλων. Στις πολύ χαμηλές αναμενόμενες συχνότητες παρατηρείται σχετική συγκέντρωση κοντά στη διαγώνιο, χωρίς όμως η συμφωνία αυτή να διατηρείται καθώς αυξάνεται η κλίμακα των συχνοτήτων.

Στο μεσαίο τμήμα εμφανίζονται διαδοχικές ομάδες σημείων τόσο πάνω όσο και κάτω από τη γραμμή αναφοράς. Οι αποκλίσεις αυξάνονται έντονα και η διάταξη δεν ακολουθεί σταθερή γραμμική πορεία. Για παρόμοια επίπεδα αναμενόμενων συχνοτήτων εμφανίζονται πολύ διαφορετικές παρατηρούμενες τιμές, γεγονός που δείχνει ότι το μοντέλο δεν αναπαράγει με συνέπεια τη συχνοτική δομή των δεδομένων.

Η μεγάλη κατακόρυφη διασπορά στις υψηλότερες αναμενόμενες συχνότητες επιβεβαιώνει ότι η κανονική κατανομή μεταβαίνει επανειλημμένα από υπερεκτίμηση σε υποεκτίμηση. Η συμπεριφορά αυτή είναι συμβατή με την αδυναμία ενός συμμετρικού μοντέλου να αποδώσει την έντονα θετικά ασύμμετρη εμπειρική κατανομή της Λειβαδιάς. Συνεπώς, η κανονική κατανομή παρέχει τη χαμηλότερη οπτική ποιότητα προσαρμογής στο συγκεκριμένο διαγνωστικό στάδιο.

Κατανομή Gamma



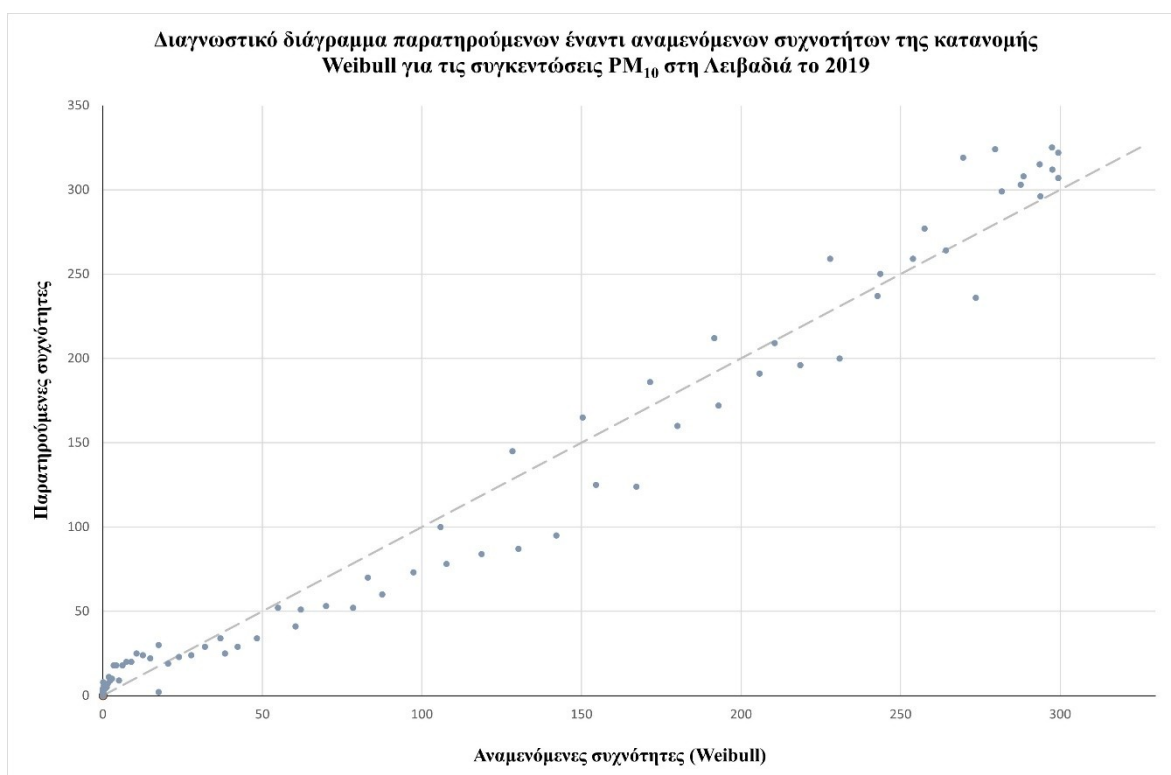
Σχήμα 4.7.2β. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η Gamma εμφανίζει σαφή βελτίωση σε σχέση με την κανονική κατανομή. Στις χαμηλές αναμενόμενες συχνότητες τα σημεία σχηματίζουν σχετικά συνεκτική ανοδική διάταξη και βρίσκονται κοντά στη γραμμή ιδανικής συμφωνίας. Η εικόνα αυτή υποδηλώνει ικανοποιητική αντιστοίχιση σε σημαντικό μέρος των λιγότερο συχνών παρατηρήσεων.

Στις μεσαίες αναμενόμενες συχνότητες η γραμμική σχέση διατηρείται εν μέρει, αλλά εμφανίζονται πλέον σαφέστερες αποκλίσεις και εναλλαγές γύρω από τη διαγώνιο. Ορισμένα σημεία βρίσκονται αισθητά πάνω από τη γραμμή, υποδηλώνοντας υποεκτίμηση των παρατηρούμενων συχνοτήτων, ενώ άλλα μετατοπίζονται αρκετά κάτω από αυτήν.

Η διασπορά γίνεται ιδιαίτερα έντονη στις μεγαλύτερες αναμενόμενες συχνότητες. Για παρόμοιες θεωρητικές τιμές εμφανίζονται παρατηρούμενες συχνότητες με μεγάλη μεταξύ τους απόσταση, στοιχείο που δείχνει ότι η προσαρμογή δεν είναι σταθερή σε όλο το εύρος. Επομένως, η Gamma αποδίδει αποτελεσματικότερα από την κανονική τη γενική συχνοτική δομή, αλλά εξακολουθεί να εμφανίζει σημαντικές τοπικές αστοχίες.

Κατανομή Weibull

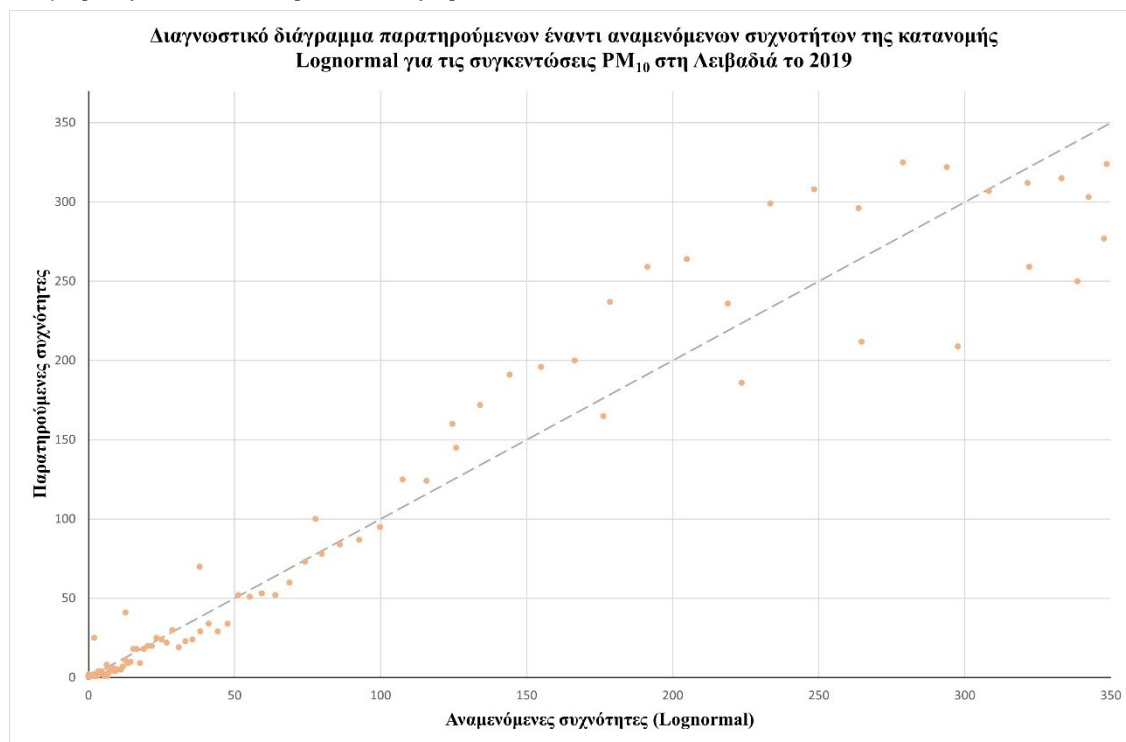


Σχήμα 4.7.2γ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η Weibull παρουσιάζει την περισσότερη συνεκτική και σχεδόν γραμμική διάταξη των σημείων. Στο χαμηλό εύρος οι παρατηρούμενες και αναμενόμενες συχνότητες βρίσκονται γενικά κοντά στη διαγώνιο, με ορισμένες τοπικές αποκλίσεις μικρού ή μέτριου μεγέθους. Καθώς αυξάνονται οι αναμενόμενες συχνότητες, τα σημεία εξακολουθούν να ακολουθούν τη γενική κατεύθυνση της γραμμής αναφοράς. Στο μεσαίο εύρος παρατηρούνται ομάδες σημείων κάτω από τη διαγώνιο, οι οποίες υποδηλώνουν τοπική υπερεκτίμηση από το μοντέλο, καθώς και ορισμένα σημεία πάνω από αυτήν. Ωστόσο, οι αποκλίσεις αυτές εντάσσονται σε μια σαφώς οργανωμένη συνολική πορεία και δεν διαμορφώνουν τη μεγάλη κατακόρυφη διασπορά που χαρακτηρίζει τη Normal και τη Gamma. Στις μεγαλύτερες αναμενόμενες συχνότητες τα περισσότερα σημεία παραμένουν σχετικά κοντά στη διαγώνιο. Παρατηρείται μια ήπια τάση τοποθέτησης αρκετών σημείων πάνω από τη γραμμή, η οποία αντιστοιχεί σε υποεκτίμηση των παρατηρούμενων συχνοτήτων, αλλά η συνολική απόσταση παραμένει περισσότερο ελεγχόμενη.

Με βάση το συγκεκριμένο διάγραμμα, η Weibull εμφανίζει την ισχυρότερη οπτική προσαρμογή. Το πλεονέκτημά της δεν συνίσταται στην απουσία αποκλίσεων, αλλά στη σταθερότερη γραμμική σχέση και στη μικρότερη συνολική διασπορά γύρω από τη γραμμή ιδανικής συμφωνίας.

Λογαριθμοκανονική κατανομή



Σχήμα 4.7.2δ. Διαγνωστικό διάγραμμα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η λογαριθμοκανονική παρουσιάζει καλή συμφωνία στις χαμηλές αναμενόμενες συχνότητες, όπου τα περισσότερα σημεία συγκεντρώνονται κοντά στη διαγώνιο. Η αρχική αυτή εικόνα είναι συμβατή με την ικανότητα της κατανομής να περιγράψει μη αρνητικά και θετικά ασύμμετρα δεδομένα.

Στο χαμηλό προς μεσαίο εύρος εμφανίζονται σχετικά περιορισμένες αποκλίσεις, ενώ η γενική ανοδική σχέση διατηρείται. Καθώς, όμως, αυξάνονται οι αναμενόμενες συχνότητες, η διασπορά διευρύνεται. Αρκετά σημεία τοποθετούνται πάνω από τη γραμμή αναφοράς, υποδηλώνοντας υποεκτίμηση από το μοντέλο, ενώ άλλα βρίσκονται σημαντικά κάτω από αυτήν.

Στο ανώτερο εύρος των αναμενόμενων συχνοτήτων η λογαριθμοκανονική εμφανίζει αισθητή κατακόρυφη διασπορά. Για παρόμοιες θεωρητικές τιμές καταγράφονται παρατηρούμενες συχνότητες με σημαντικές διαφορές, με αποτέλεσμα η διάταξη να είναι λιγότερο συμπαγής από εκείνη της Weibull.

Η λογαριθμοκανονική παρέχει, επομένως, καλή αλλά όχι την καλύτερη οπτική προσαρμογή. Ακολουθεί σε σημαντικό βαθμό τη γενική σχέση μεταξύ αναμενόμενων και παρατηρούμενων συχνοτήτων, αλλά οι αποκλίσεις της αυξάνονται στο υψηλότερο τμήμα του οριζόντιου άξονα και περιορίζουν τη συνολική της συνέπεια.

Συγκριτική αξιολόγηση

Η σύγκριση των τεσσάρων διαγραμμάτων διαφοροποιεί εν μέρει την αρχική εικόνα των ραβδογραμμάτων. Στη σύγκριση ανά κλάση, η λογαριθμοκανονική είχε παρουσιάσει την πιο ισορροπημένη συνολική αντιστοίχιση, κυρίως λόγω της καλής της απόδοσης στην πολυπληθέστερη κλάση. Στα διαγνωστικά διαγράμματα, όμως, η Weibull εμφανίζει πιο συμπαγή και περισσότερο συνεκτική γραμμική σχέση σε ευρύτερο εύρος συχνοτήτων.

Η κανονική κατανομή παρουσιάζει τη μεγαλύτερη διασπορά και τις εντονότερες εναλλαγές μεταξύ υποεκτίμησης και υπερεκτίμησης. Η Gamma περιορίζει μέρος των αποκλίσεων, αλλά στις μεγαλύτερες αναμενόμενες συχνότητες εξακολουθεί να εμφανίζει μεγάλη κατακόρυφη μεταβλητότητα. Η λογαριθμοκανονική ακολουθεί ικανοποιητικά τη διαγώνιο στο χαμηλό και σε τμήμα του μεσαίου εύρους, αλλά η εικόνα της γίνεται λιγότερο σταθερή όσο αυξάνονται οι αναμενόμενες συχνότητες.

Η Weibull διατηρεί την πιο συνεκτική ανοδική διάταξη και παρουσιάζει μικρότερη συνολική διασπορά γύρω από τη διαγώνιο. Παρά τις τοπικές αποκλίσεις, η σχέση

παρατηρούμενων και αναμενόμενων συχνοτήτων παραμένει περισσότερο οργανωμένη από ό,τι στις υπόλοιπες κατανομές. Συνεπώς, με βάση αποκλειστικά το συγκεκριμένο διαγνωστικό εργαλείο, η Weibull εμφανίζει την ισχυρότερη οπτική ένδειξη προσαρμογής. Η λογαριθμοκανονική ακολουθεί ως η πλησιέστερη εναλλακτική, ενώ η σχετική θέση της Gamma είναι ασθενέστερη εξαιτίας της εντονότερης διασποράς. Η κανονική κατανομή κατατάσσεται τελευταία. Η προσωρινή αυτή αξιολόγηση δεν αναιρεί το συμπέρασμα των ραβδογραμμάτων, αλλά δείχνει ότι η λογαριθμοκανονική αποδίδει καλύτερα ορισμένες συγκεντρωτικές κλάσεις, ενώ η Weibull παρουσιάζει περισσότερο συνεπή σχέση στο σύνολο των επιμέρους συχνοτήτων.

Η διαφοροποίηση αυτή επιβεβαιώνει ότι τα δύο γραφικά εργαλεία εξετάζουν διαφορετικές πτυχές της προσαρμογής. Τα ραβδογράμματα συνοψίζουν τις αποκλίσεις σε έξι ευρείες κλάσεις, ενώ τα διαγνωστικά διαγράμματα αποτυπώνουν περισσότερα επιμέρους ζεύγη παρατηρούμενων και αναμενόμενων συχνοτήτων. Για τον λόγο αυτό, μια κατανομή μπορεί να εμφανίζει καλή συνολική αντιστοίχιση ανά κλάση, αλλά μεγαλύτερη διασπορά όταν εξετάζονται αναλυτικότερα οι επιμέρους συχνότητες.

Το συμπέρασμα παραμένει προσωρινό. Η μεγάλη τιμή του οριζόντιου άξονα δεν ταυτίζεται με υψηλή συγκέντρωση PM_{10} και, επομένως, τα διαγράμματα αυτά δεν αξιολογούν άμεσα τη δεξιά ουρά. Η συμπεριφορά των μοντέλων στα υψηλότερα ποσοστημόρια θα εξεταστεί στα Q-Q διαγράμματα, ενώ η δομή και η κατεύθυνση των αποκλίσεων θα διερευνηθούν προηγουμένως μέσω των διαγραμμάτων υπολοίπων.

4.7.3 Διαγράμματα υπολοίπων

Τα διαγράμματα υπολοίπων χρησιμοποιούνται για να εξεταστεί σε ποια σημεία της ταξινομημένης σειράς οι θεωρητικές κατανομές χάνουν την επαφή τους με τα εμπειρικά δεδομένα.

Για κάθε σημείο ισχύει:

$$R_i = O_i - E_i$$

όπου θετική τιμή δηλώνει υποεκτίμηση από το μοντέλο και αρνητική τιμή υπερεκτίμηση. Θετική τιμή του υπολοίπου δηλώνει ότι το μοντέλο υποεκτιμά την παρατηρούμενη συχνότητα, ενώ αρνητική τιμή υποδηλώνει υπερεκτίμηση. Η γραμμή μηδενός λειτουργεί,

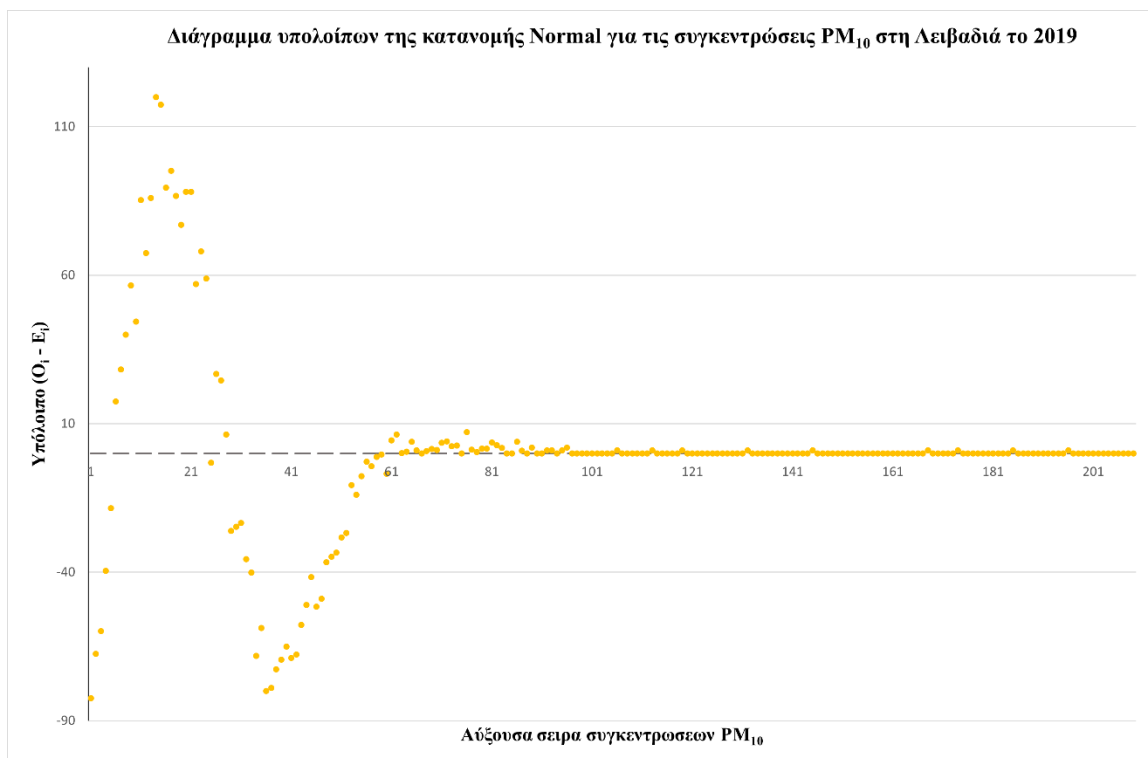
επομένως, ως σημείο αναφοράς για την αξιολόγηση της κατεύθυνσης και της έντασης των αποκλίσεων.

Η ερμηνεία των διαγραμμάτων δεν περιορίζεται μόνο στο μέγιστο απόλυτο μέγεθος των υπολοίπων. Ιδιαίτερη σημασία έχει η συνολική μορφή που σχηματίζουν τα σημεία: αν οι αποκλίσεις είναι τοπικές ή παρατεταμένες, αν εμφανίζονται διαδοχικές ζώνες ίδιου προσήμου, αν μεταβάλλονται συστηματικά από θετικές σε αρνητικές τιμές και πόσο γρήγορα συγκλίνουν προς τη γραμμή μηδενός. Μια σχετικά τυχαία και περιορισμένη διασπορά γύρω από το μηδέν αποτελεί ένδειξη ικανοποιητικής προσαρμογής. Αντίθετα, η εμφάνιση κυματοειδών σχηματισμών ή εκτεταμένων περιοχών υπερεκτίμησης και υποεκτίμησης δείχνει ότι το θεωρητικό μοντέλο δεν αποδίδει ομοιόμορφα τη δομή των εμπειρικών συχνοτήτων.

Στη Λειβαδιά, τα τέσσερα διαγράμματα διαφοροποιούνται όχι μόνο ως προς το εύρος των σφαλμάτων, αλλά και ως προς τη διάρκεια και τη θέση τους μέσα στην ταξινομημένη σειρά. Ορισμένα μοντέλα εμφανίζουν μεγάλες και εναλλασσόμενες αποκλίσεις σε εκτεταμένο τμήμα του διαγράμματος, ενώ άλλα περιορίζουν τις σημαντικότερες αστοχίες κυρίως στις πρώτες παρατηρήσεις και σταθεροποιούνται νωρίτερα γύρω από το μηδέν. Το στοιχείο αυτό είναι ιδιαίτερα χρήσιμο, επειδή επιτρέπει να διακριθεί αν μια κατανομή παρουσιάζει γενικευμένη αδυναμία προσαρμογής ή αν η ασυμφωνία της είναι περισσότερο τοπική.

Ο οριζόντιος άξονας εκφράζει την αύξουσα σειρά των συγκεντρώσεων PM_{10} και όχι τις ίδιες τις κλάσεις συχνοτήτων. Επομένως, τα δεξιότερα σημεία αντιστοιχούν σε υψηλότερες ταξινομημένες συγκεντρώσεις, χωρίς όμως το διάγραμμα να αποτυπώνει άμεσα το ακριβές μέγεθός τους. Για τον λόγο αυτό, τα residual plots αξιολογούν κυρίως τη δομή των σφαλμάτων και λειτουργούν συμπληρωματικά προς τα Q-Q διαγράμματα, τα οποία θα εξετάσουν πιο άμεσα τη συμφωνία στα υψηλότερα ποσοστημόρια και στη δεξιά ουρά της εμπειρικής κατανομής.

Κανονική κατανομή

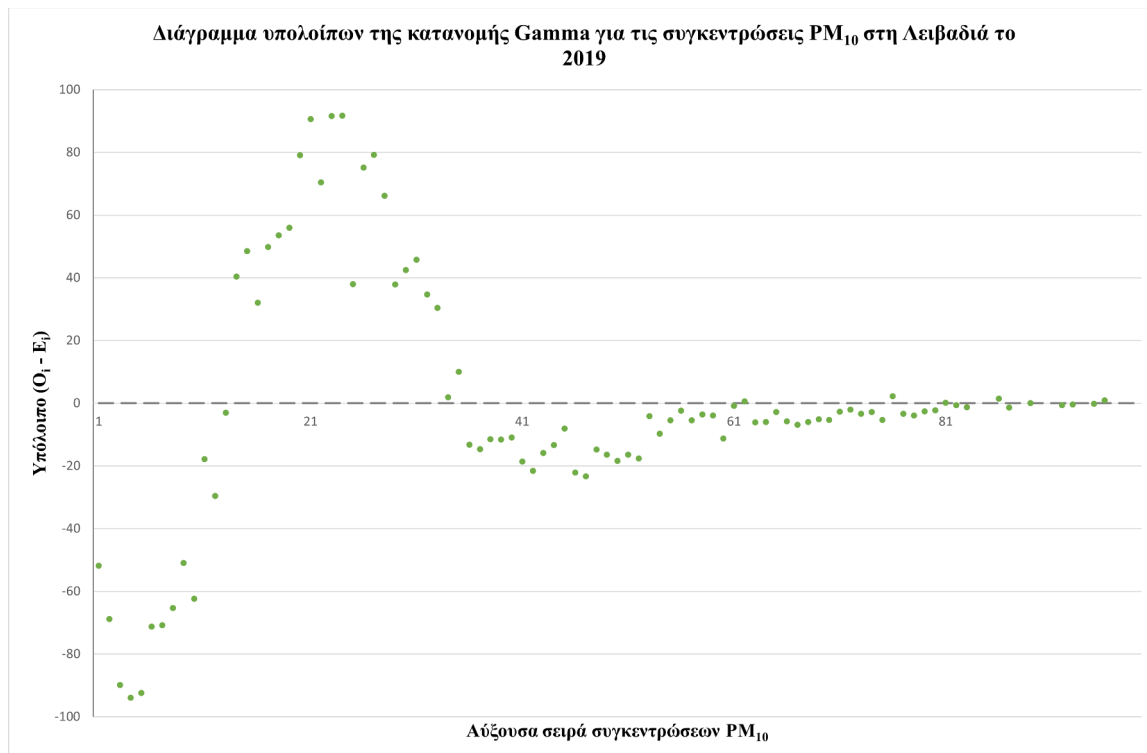


Σχήμα 4.7.3α. Διάγραμμα υπολοίπων της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η κανονική κατανομή παρουσιάζει την πιο έντονη εναλλαγή μεταξύ υπερεκτίμησης και υποεκτίμησης. Στις πρώτες ταξινομημένες τιμές καταγράφονται μεγάλα αρνητικά υπόλοιπα, τα οποία ακολουθούνται από πολύ υψηλές θετικές αποκλίσεις, άνω του +110. Στη συνέχεια, το πρόσημο αντιστρέφεται εκ νέου και σχηματίζεται δεύτερη εκτεταμένη περιοχή αρνητικών τιμών.

Η διαδοχή αυτή δείχνει ότι το μοντέλο δεν αστοχεί μόνο σε μεμονωμένες παρατηρήσεις, αλλά μετακινεί συστηματικά τη θεωρητική συχνότητα από τη μία περιοχή της κατανομής στην άλλη. Η μεταγενέστερη σύγκλιση προς το μηδέν εμφανίζεται μόνο αφού έχει προηγηθεί μεγάλο τμήμα με έντονες αποκλίσεις. Επομένως, η τελική σταθεροποίηση δεν μεταβάλλει τη συνολικά δυσμενή εικόνα.

Κατανομή Gamma



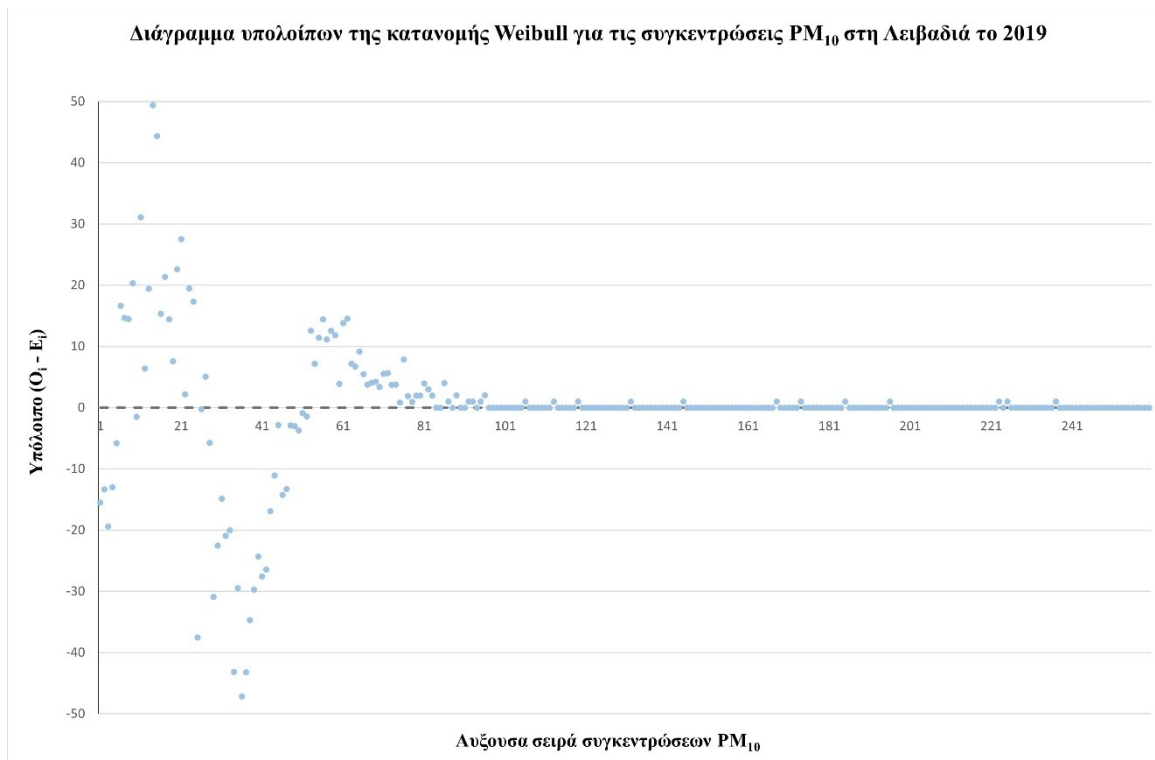
Σχήμα 4.7.3β. Διάγραμμα υπολοίπων της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η Gamma εμφανίζει διαφορετικό, αλλά επίσης έντονα οργανωμένο, πρότυπο. Οι πρώτες αποκλίσεις είναι πολύ αρνητικές και πλησιάζουν το -95 , ενώ στη συνέχεια ακολουθεί απότομη μετάβαση σε θετικές τιμές που φθάνουν περίπου στο $+90$.

Το μοντέλο μεταβαίνει, επομένως, από έντονη υπερεκτίμηση σε έντονη υποεκτίμηση. Αργότερα, τα υπόλοιπα επιστρέφουν σε αρνητικές τιμές μικρότερου μεγέθους και σχηματίζουν παρατεταμένη ζώνη κάτω από τη γραμμή μηδενός. Η τελική σύγκλιση είναι σαφής, αλλά προηγείται ένα μεγάλο τμήμα με συστηματική δομή.

Η Gamma αποδίδει καλύτερα από την κανονική το τελευταίο μέρος της σειράς, όμως οι αρχικές και μεσαίες αποκλίσεις παραμένουν πολύ μεγάλες ώστε να θεωρηθεί η προσαρμογή ομοιόμορφη.

Κατανομή Weibull



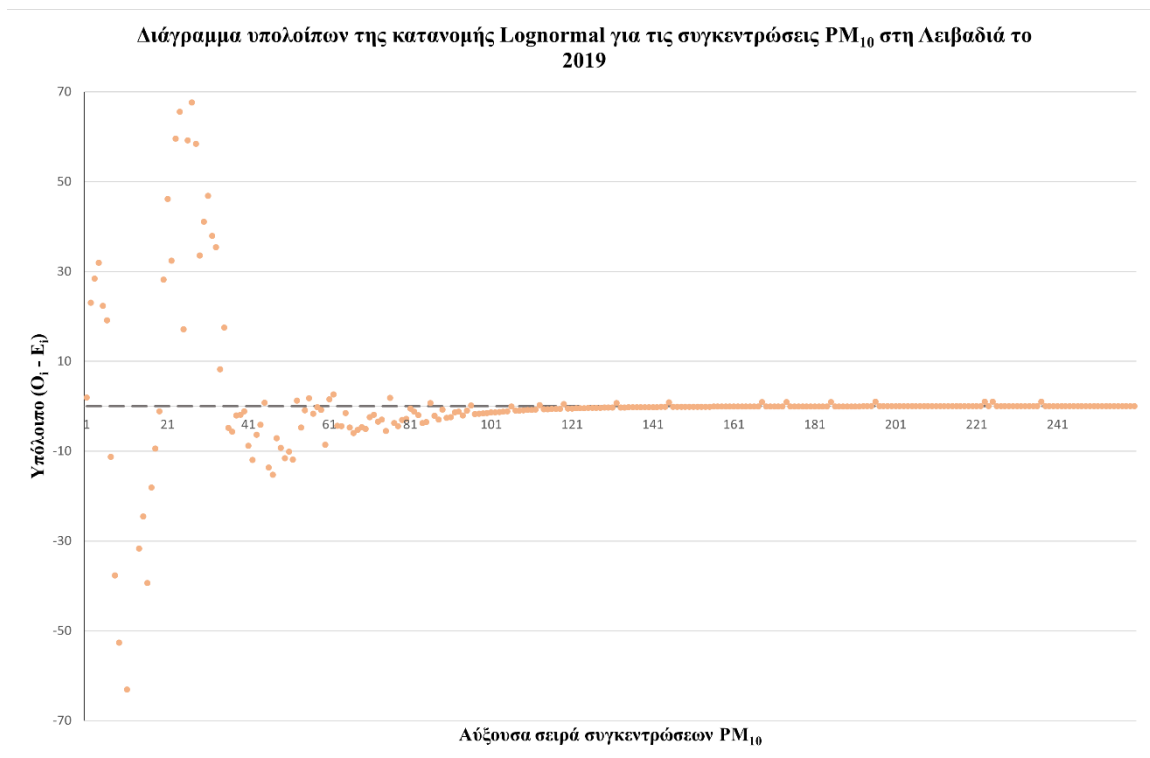
Σχήμα 4.7.3γ. Διάγραμμα υπολοίπων της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Στη Weibull οι σημαντικές αποκλίσεις είναι περισσότερο συγκεντρωμένες χρονικά. Στο αρχικό τμήμα εμφανίζονται τόσο θετικές όσο και αρνητικές τιμές, με εύρος περίπου από -48 έως $+49$. Η εναλλαγή αυτή δείχνει ότι το μοντέλο δεν αποδίδει με την ίδια ακρίβεια τις πρώτες ταξινομημένες παρατηρήσεις.

Η διαφορά από τις Normal και Gamma είναι ότι η ένταση των υπολοίπων μειώνεται πολύ νωρίτερα. Μετά το αρχικό τμήμα, εμφανίζεται μόνο μία μικρότερη περιοχή θετικών αποκλίσεων και στη συνέχεια τα σημεία συγκεντρώνονται σχεδόν πλήρως πάνω στη γραμμή μηδενός.

Η Weibull δεν εξαλείφει τις τοπικές αστοχίες, αλλά τις περιορίζει σε μικρότερο μέρος της σειράς. Αυτό αποτελεί σημαντικό πλεονέκτημα, επειδή δείχνει ότι η ασυμφωνία δεν μεταφέρεται στο μεγαλύτερο μέρος της εμπειρικής κατανομής.

Λογαριθμοκανονική κατανομή



Σχήμα 4.7.3δ. Διάγραμμα υπολοίπων της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η λογαριθμοκανονική παρουσιάζει επίσης συγκέντρωση των μεγαλύτερων αποκλίσεων στο αρχικό τμήμα. Οι τιμές κυμαίνονται περίπου από -64 έως $+68$, άρα το αρχικό εύρος είναι μεγαλύτερο από εκείνο της Weibull.

Μετά τη θετική κορύφωση, τα υπόλοιπα μειώνονται γρήγορα και ακολουθεί περιοχή μικρών κυρίως αρνητικών τιμών. Η συμπεριφορά αυτή υποδηλώνει ήπια, αλλά σχετικά παρατεταμένη, υπερεκτίμηση σε τμήμα της σειράς. Στη συνέχεια, τα σημεία συγκεντρώνονται πολύ κοντά στο μηδέν και παραμένουν σταθερά έως το τέλος.

Η λογαριθμοκανονική παρουσιάζει, επομένως, πολύ καλή συμπεριφορά στο μεγαλύτερο μέρος των δεδομένων. Η βασική της αδυναμία εντοπίζεται στην εντονότερη αρχική αστάθεια και στη μεγαλύτερη διάρκεια των μικρών αρνητικών αποκλίσεων σε σχέση με τη Weibull.

Συγκριτική αξιολόγηση

Η πιο ουσιαστική διαφοροποίηση μεταξύ των τεσσάρων μοντέλων δεν βρίσκεται στο γεγονός ότι όλα τελικά συγκλίνουν προς το μηδέν, αλλά στο πόσο γρήγορα και με ποιο κόστος το επιτυγχάνουν.

Η κανονική και η Gamma χρειάζονται μεγάλο μέρος της σειράς για να σταθεροποιηθούν και προηγουμένως εμφανίζουν πολύ μεγάλες και διαδοχικές εναλλαγές προσήμου. Οι αποκλίσεις τους είναι εκτεταμένες και οργανωμένες, στοιχείο που δείχνει ότι η αστοχία αφορά τη συνολική μορφή της προσαρμογής και όχι μόνο μεμονωμένα σημεία.

Αντίθετα, στη Weibull και στη λογαριθμοκανονική οι μεγαλύτερες αποκλίσεις συγκεντρώνονται κυρίως στην αρχή και εξασθενούν πολύ νωρίτερα. Η σύγκριση μεταξύ των δύο είναι πιο λεπτή. Η Weibull εμφανίζει μικρότερο αρχικό εύρος και ταχύτερη μετάβαση σε σχεδόν μηδενικά υπόλοιπα. Η λογαριθμοκανονική παρουσιάζει μεγαλύτερες αρχικές διακυμάνσεις και μια μικρή παρατεταμένη αρνητική ζώνη, αλλά στη συνέχεια διατηρεί επίσης εξαιρετικά σταθερή εικόνα.

Με βάση αποκλειστικά τη συμπεριφορά των υπολοίπων, η Weibull φαίνεται να διαθέτει ελαφρύ οπτικό πλεονέκτημα ως προς τον περιορισμό και τη διάρκεια των αποκλίσεων. Η λογαριθμοκανονική παραμένει πολύ κοντά και παρουσιάζει ισχυρή συνολική συμπεριφορά, αλλά δεν προκύπτει από τα συγκεκριμένα διαγράμματα σαφής υπεροχή της. Το αποτέλεσμα αυτό διαφοροποιεί ελαφρά την εικόνα των ραβδογραμμάτων, όπου η λογαριθμοκανονική είχε εμφανιστεί περισσότερο ισορροπημένη. Τα residual plots δείχνουν ότι η Weibull διαχειρίζεται αποτελεσματικότερα τη μετάβαση από τις αρχικές αστοχίες προς μια σταθερή περιοχή σχεδόν μηδενικών σφαλμάτων.

Η αξιολόγηση αυτή παραμένει προσωρινή. Τα μικρά υπόλοιπα στο δεξιό τμήμα της σειράς δεν αποδεικνύουν από μόνα τους ακριβή περιγραφή της άνω ουράς. Η συμφωνία στα υψηλότερα ποσοστημόρια θα εξεταστεί στα Q-Q διαγράμματα, όπου θα φανεί αν η καλή σταθερότητα των Weibull και λογαριθμοκανονικής συνοδεύεται και από ικανοποιητική αναπαράσταση των μεγαλύτερων συγκεντρώσεων.

4.7.4 Διαγράμματα Q-Q

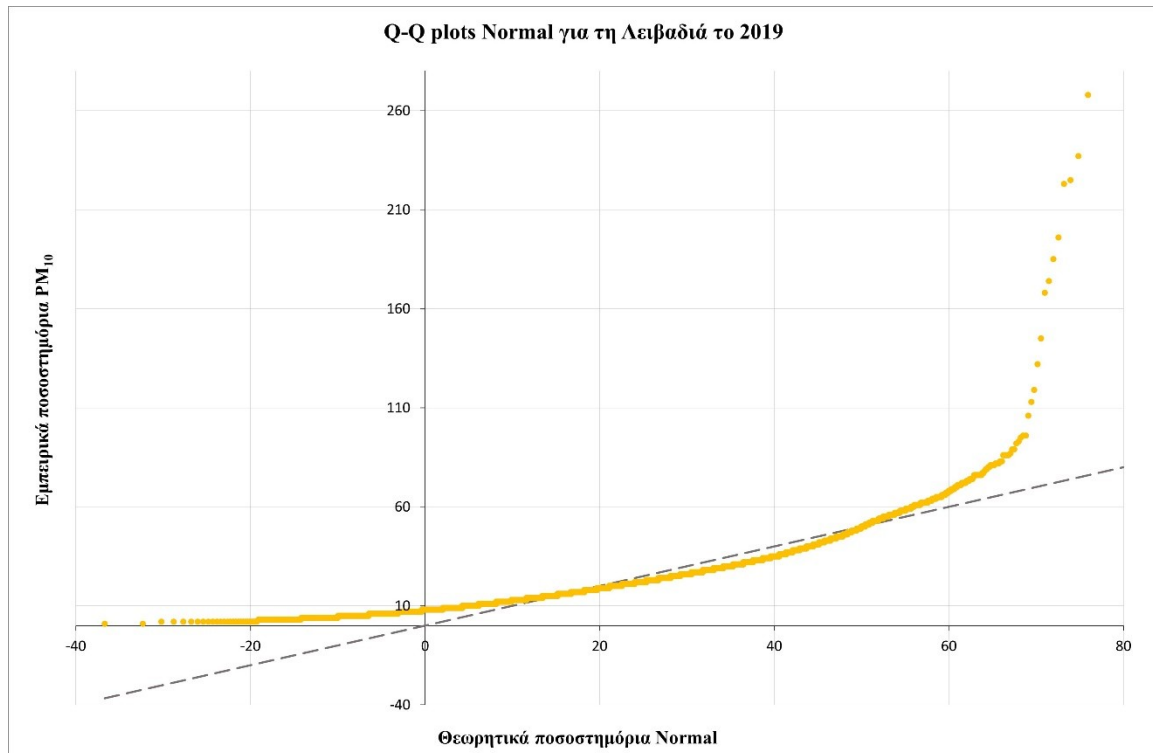
Τα διαγράμματα Q-Q εξετάζουν κατά πόσο τα εμπειρικά ποσοστημόρια των ωριαίων συγκεντρώσεων PM_{10} συμφωνούν με τα αντίστοιχα ποσοστημόρια των προσαρμοσμένων θεωρητικών κατανομών. Η διαγώνια γραμμή αναφοράς εκφράζει την ιδανική ισότητα μεταξύ θεωρητικών και εμπειρικών τιμών. Όσο περισσότερα σημεία ακολουθούν τη γραμμή αυτή, τόσο ισχυρότερη είναι η ένδειξη κατανομικής συμφωνίας.

Η θέση των σημείων ως προς τη διαγώνιο παρέχει και πληροφορία για την κατεύθυνση της απόκλισης. Σημεία πάνω από τη γραμμή σημαίνουν ότι τα εμπειρικά ποσοστημόρια είναι μεγαλύτερα από τα θεωρητικά και, επομένως, ότι το μοντέλο υποεκτιμά τις αντίστοιχες συγκεντρώσεις. Αντίθετα, σημεία κάτω από τη γραμμή υποδηλώνουν ότι το μοντέλο προβλέπει υψηλότερες τιμές από εκείνες που παρατηρούνται εμπειρικά.

Η συγκεκριμένη εξέταση είναι ιδιαίτερα σημαντική για τη Λειβαδιά, καθώς η εμπειρική κατανομή χαρακτηρίζεται από έντονη δεξιά ασυμμετρία και περιορισμένο αριθμό πολύ υψηλών συγκεντρώσεων. Ένα μοντέλο μπορεί να περιγράψει ικανοποιητικά το κύριο σώμα των δεδομένων, αλλά να αποτυγχάνει ακριβώς στο ανώτερο τμήμα, όπου εμφανίζονται οι μεγαλύτερες αποκλίσεις.

Πριν από την οριστική σύγκριση απαιτείται, ωστόσο, ένας τεχνικός έλεγχος των διαγραμμάτων. Τα εμπειρικά ποσοστημόρια πρέπει να προκύπτουν από το ίδιο δείγμα και τις ίδιες πιθανότητες σε όλες τις κατανομές και, συνεπώς, ο κατακόρυφος άξονας θα πρέπει να περιλαμβάνει τις ίδιες εμπειρικές τιμές. Στα παρουσιαζόμενα σχήματα το μέγιστο εμπειρικό ποσοστημόριο διαφέρει αισθητά μεταξύ των μοντέλων. Αυτό ενδέχεται να οφείλεται σε διαφορετική επιλογή ποσοστημορίων, περικοπή σημείων ή διαφορετική κλίμακα υπολογισμού. Για τον λόγο αυτό, η παρακάτω αξιολόγηση αφορά την οπτική εικόνα των συγκεκριμένων σχημάτων και πρέπει να θεωρηθεί προσωρινή έως ότου επιβεβαιωθεί ότι όλα κατασκευάστηκαν με κοινή διαδικασία.

Κανονική κατανομή



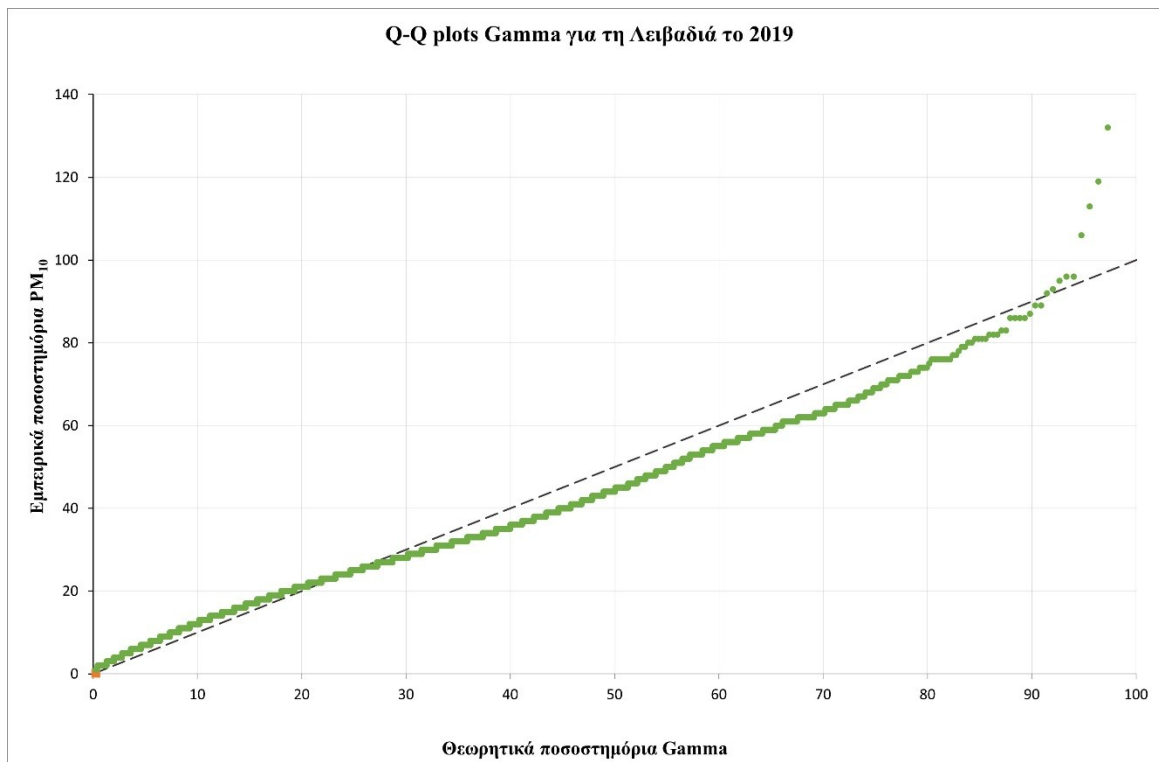
Σχήμα 4.7.4α. Διάγραμμα Q-Q της κανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η κανονική κατανομή παρουσιάζει εκτεταμένη και συστηματική καμπυλότητα. Στο κατώτερο τμήμα εμφανίζονται αρνητικά θεωρητικά ποσοστημόρια, ενώ οι συγκεντρώσεις PM_{10} δεν μπορούν να λάβουν αρνητικές τιμές. Η εικόνα αυτή αποκαλύπτει εξ αρχής έναν δομικό περιορισμό της κανονικής κατανομής για τη συγκεκριμένη μεταβλητή.

Στο χαμηλό και μεσαίο εύρος τα σημεία προσεγγίζουν τη διαγώνιο μόνο τοπικά. Στη συνέχεια, όμως, απομακρύνονται προοδευτικά προς τα επάνω. Η απόκλιση γίνεται ιδιαίτερα έντονη στο ανώτερο τμήμα, όπου οι εμπειρικές τιμές αυξάνονται πολύ ταχύτερα από τις θεωρητικές.

Η έντονη αυτή ανοδική καμπυλότητα δείχνει ότι η κανονική κατανομή υποεκτιμά σημαντικά τις υψηλές συγκεντρώσεις και δεν διαθέτει επαρκώς εκτεταμένη δεξιά ουρά. Το διάγραμμα επιβεβαιώνει, επομένως, ότι η συμμετρική μορφή της δεν ανταποκρίνεται στη θετικά ασύμμετρη κατανομική δομή της Λειβαδιάς.

Κατανομή Gamma

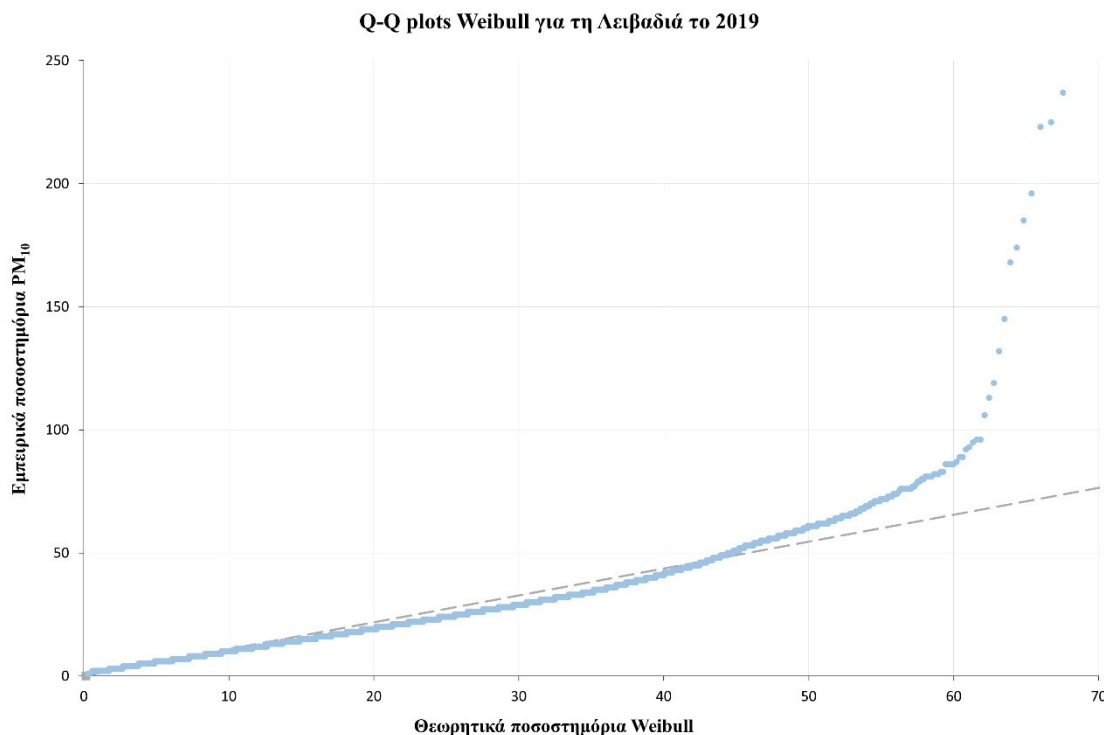


Σχήμα 4.7.4β. Διάγραμμα Q-Q της κατανομής Gamma για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η Gamma παρουσιάζει πολύ καλή ευθυγράμμιση στο χαμηλό και σε σημαντικό μέρος του μεσαίου εύρους. Τα πρώτα σημεία βρίσκονται ελαφρώς πάνω από τη γραμμή αναφοράς, όμως η απόκλιση είναι περιορισμένη και μειώνεται καθώς αυξάνονται τα ποσοστημόρια. Στη συνέχεια, μεγάλο μέρος των σημείων τοποθετείται λίγο κάτω από τη διαγώνιο. Το πρότυπο αυτό υποδηλώνει ήπια υπερεκτίμηση των αντίστοιχων εμπειρικών ποσοστημορίων από το μοντέλο. Παρά τη συστηματική κατεύθυνση της απόκλισης, η απόσταση από τη γραμμή παραμένει σχετικά μικρή και εξελίσσεται ομαλά.

Μόνο στα τελευταία σημεία παρατηρείται σαφής ανοδική μετατόπιση. Οι μεγαλύτερες εμπειρικές τιμές υπερβαίνουν τις θεωρητικές προβλέψεις, γεγονός που υποδηλώνει υποεκτίμηση του ακραίου δεξιού άκρου. Με βάση το παρουσιαζόμενο διάγραμμα, η Gamma αποδίδει ιδιαίτερα αποτελεσματικά το κύριο σώμα και τα περισσότερα ποσοστημόρια, ενώ η βασική της αστοχία περιορίζεται στις πλέον υψηλές παρατηρήσεις.

Κατανομή Weibull



Σχήμα 4.7.4γ. Διάγραμμα Q-Q της κατανομής Weibull για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

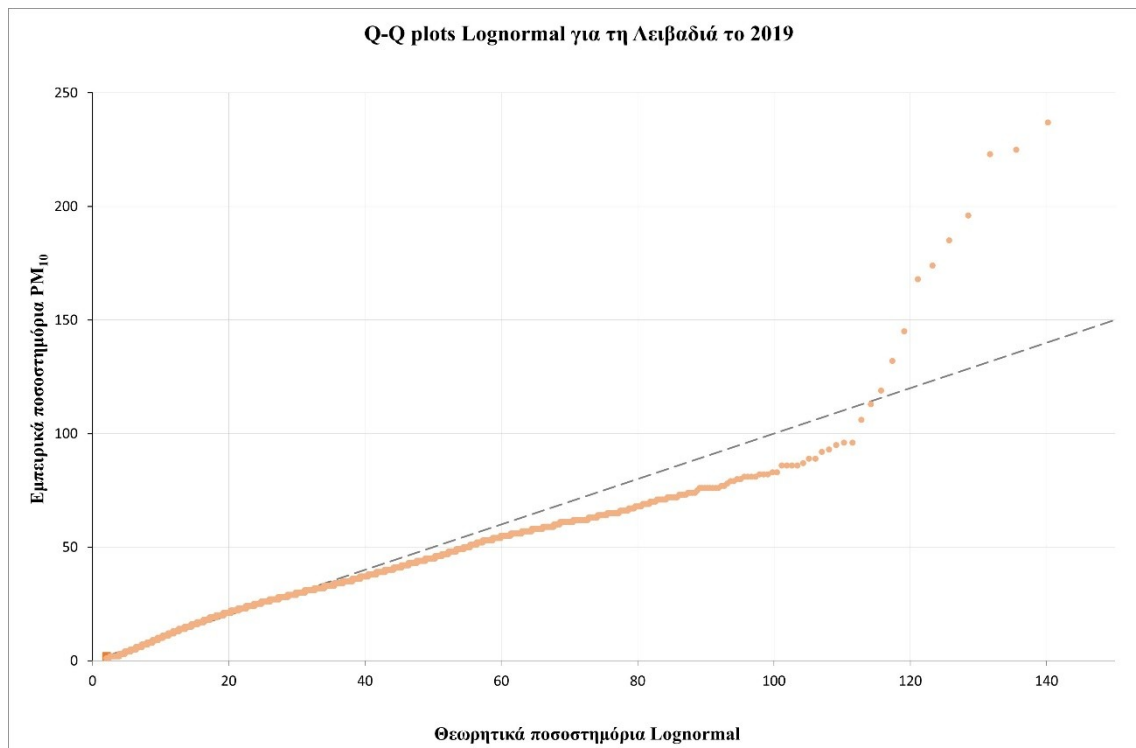
Η Weibull ακολουθεί σχετικά καλά τη γραμμή αναφοράς στις χαμηλότερες τιμές. Στο μεσαίο τμήμα τα σημεία μετατοπίζονται αρχικά κάτω από τη διαγώνιο, γεγονός που υποδηλώνει ότι το μοντέλο τείνει να υπερεκτιμά μέρος των αντίστοιχων εμπειρικών ποσοστημορίων.

Η τάση αυτή αντιστρέφεται καθώς αυξάνονται τα θεωρητικά ποσοστημόρια. Τα σημεία διασταυρώνουν τη γραμμή και στη συνέχεια απομακρύνονται σταδιακά προς τα επάνω. Στο ανώτερο τμήμα η απόκλιση γίνεται απότομη και πολύ μεγάλη, με τις εμπειρικές συγκεντρώσεις να υπερβαίνουν σημαντικά τις θεωρητικές.

Η μορφή αυτή δείχνει ότι η Weibull προσαρμόζεται ικανοποιητικά σε μεγάλο μέρος των χαμηλών και μεσαίων τιμών, αλλά υποεκτιμά έντονα την άνω ουρά. Το εύρημα είναι συμβατό με τα ραβδογράμματα, όπου είχε καταγραφεί σημαντική υποεκτίμηση της συχνότητας στην ανοικτή κλάση άνω των $50 \mu g/m^3$.

Επομένως, η καλή συμπεριφορά της Weibull στα διαγράμματα παρατηρούμενων και αναμενόμενων συχνοτήτων και στα υπόλοιπα δεν επεκτείνεται με την ίδια αποτελεσματικότητα στην αναπαράσταση του μεγέθους των υψηλότερων συγκεντρώσεων.

Λογαριθμοκανονική κατανομή



Σχήμα 4.7.4δ. Διάγραμμα Q-Q της λογαριθμοκανονικής κατανομής για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η λογαριθμοκανονική κατανομή παρουσιάζει πολύ καλή συμφωνία στις χαμηλές τιμές, όπου τα σημεία ακολουθούν στενά τη διαγώνιο. Η ευθυγράμμιση παραμένει ικανοποιητική έως το πρώτο μεσαίο τμήμα, γεγονός που υποδηλώνει αποτελεσματική περιγραφή των συνηθέστερων χαμηλότερων συγκεντρώσεων.

Στο ευρύτερο μεσαίο εύρος τα σημεία μετατοπίζονται σταδιακά κάτω από τη γραμμή αναφοράς. Η λογαριθμοκανονική τείνει, επομένως, να υπερεκτιμά τα αντίστοιχα

εμπειρικά ποσοστημόρια. Η απόκλιση είναι συστηματική, αλλά εξελίσσεται ομαλά και χωρίς τις απότομες μεταβολές που εμφανίζει η Weibull.

Προς το ανώτερο τμήμα τα σημεία προσεγγίζουν εκ νέου τη διαγώνιο και στη συνέχεια την υπερβαίνουν. Οι τελευταίες εμπειρικές τιμές βρίσκονται σημαντικά πάνω από τη γραμμή, υποδηλώνοντας ότι και η λογαριθμοκανονική υποεκτιμά τις πιο ακραίες συγκεντρώσεις. Η ανοδική απόκλιση είναι έντονη, αλλά εμφανίζεται αργότερα από ό,τι στη Weibull.

Η λογαριθμοκανονική παρέχει, συνεπώς, ισχυρή προσαρμογή στο χαμηλό και σε μεγάλο μέρος του κεντρικού εύρους, αλλά παρουσιάζει δύο διακριτές αδυναμίες: υπερεκτίμηση σε μέρος των μεσαίων ποσοστημορίων και υποεκτίμηση του ακραίου δεξιού άκρου.

Συγκριτική αξιολόγηση

Τα Q-Q διαγράμματα αποκαλύπτουν ότι οι τέσσερις κατανομές δεν αστοχούν με τον ίδιο τρόπο. Η κανονική παρουσιάζει γενικευμένη ασυμφωνία, η Weibull χάνει γρήγορα την επαφή με την άνω ουρά, ενώ οι Gamma και λογαριθμοκανονική διατηρούν περισσότερο ελεγχόμενη συμπεριφορά στο μεγαλύτερο μέρος των ποσοστημορίων.

Η κανονική κατανομή δεν αποδίδει ούτε το κατώτερο ούτε το ανώτερο τμήμα με επάρκεια. Η Weibull προσεγγίζει ικανοποιητικά τις χαμηλές και μεσαίες τιμές, αλλά η έντονη ανοδική καμπυλότητα αρχίζει σχετικά νωρίς και οδηγεί σε μεγάλη υποεκτίμηση των υψηλότερων συγκεντρώσεων.

Η λογαριθμοκανονική παρουσιάζει πιο σύνθετη συμπεριφορά. Αποδίδει πολύ καλά το αρχικό τμήμα, υπερεκτιμά ένα σημαντικό μέρος των μεσαίων ποσοστημορίων και τελικά υποεκτιμά τις πλέον υψηλές παρατηρήσεις. Παρά τις αποκλίσεις αυτές, η μετάβασή της είναι περισσότερο ομαλή και η απομάκρυνση από τη διαγώνιο εμφανίζεται αργότερα από ό,τι στη Weibull.

Η Gamma εμφανίζει, με βάση το συγκεκριμένο σχήμα, τη στενότερη συνολική σχέση με τη διαγώνιο. Η απόκλισή της στο μεσαίο εύρος είναι ήπια και σχετικά ομοιόμορφη, ενώ η έντονη ανοδική μετατόπιση περιορίζεται στα τελευταία ποσοστημόρια. Υπό αυτή την έννοια, παρέχει την ισχυρότερη οπτική εικόνα στο συγκεκριμένο διαγνωστικό στάδιο.

Το αποτέλεσμα αυτό διαφοροποιείται από τα ραβδογράμματα, όπου η λογαριθμοκανονική είχε εμφανίσει πιο ισορροπημένη συνολική αντιστοίχιση, και από τα διαγράμματα observed-expected και υπολοίπων, όπου η Weibull είχε παρουσιάσει ιδιαίτερα συνεκτική

συμπεριφορά. Η διαφοροποίηση δεν συνιστά αντίφαση. Τα προηγούμενα εργαλεία αξιολογούν κυρίως τη συμφωνία των συχνοτήτων, ενώ τα Q-Q διαγράμματα εξετάζουν άμεσα το μέγεθος των εμπειρικών ποσοστημορίων και αναδεικνύουν με μεγαλύτερη σαφήνεια τις αστοχίες στην άνω ουρά.

Κοινό χαρακτηριστικό των τριών ασύμμετρων κατανομών είναι ότι τα τελευταία σημεία βρίσκονται πάνω από τη γραμμή αναφοράς. Καμία από αυτές δεν αναπαράγει πλήρως τις πλέον υψηλές συγκεντρώσεις της Λειβαδιάς. Η διαφορά αφορά κυρίως το πόσο νωρίς αρχίζει η απόκλιση και πόσο γρήγορα αυξάνεται. Στη Weibull η αστοχία είναι εντονότερη και εμφανίζεται νωρίτερα, στη λογαριθμοκανονική εκδηλώνεται αργότερα, ενώ στη Gamma περιορίζεται περισσότερο στα τελευταία σημεία του απεικονιζόμενου εύρους.

Με βάση τα παρουσιαζόμενα Q-Q διαγράμματα, η Gamma παρέχει την περισσότερο ισορροπημένη αντιστοίχιση ποσοστημορίων, με τη λογαριθμοκανονική να ακολουθεί ως ισχυρή εναλλακτική. Η Weibull εμφανίζει καλή συμπεριφορά στο κύριο σώμα, αλλά σαφέστερη αδυναμία στην άνω ουρά, ενώ η κανονική κατανομή υπολείπεται σημαντικά.

Η εκτίμηση αυτή παραμένει προσωρινή και προϋποθέτει την επιβεβαίωση ότι όλα τα Q-Q διαγράμματα βασίζονται στο ίδιο δείγμα και στις ίδιες πιθανότητες ποσοστημορίων. Η τελική επιλογή θα προκύψει μετά την αριθμητική αξιολόγηση μέσω του στατιστικού χ^2 και τη σύνθεση όλων των γραφικών ευρημάτων.

4.7.5 Αριθμητική αξιολόγηση με βάση το στατιστικό χ^2

Μετά τη γραφική διερεύνηση της προσαρμογής, η σύγκριση των θεωρητικών κατανομών συμπληρώνεται με το στατιστικό χ^2 . Ο δείκτης συνοψίζει τη συνολική σταθμισμένη απόκλιση μεταξύ των παρατηρούμενων και των θεωρητικά αναμενόμενων συχνοτήτων στις κοινές κλάσεις συγκέντρωσης που χρησιμοποιήθηκαν στην ανάλυση.

Μικρότερη τιμή χ^2 σημαίνει ότι οι θεωρητικές συχνότητες βρίσκονται συνολικά πλησιέστερα στις εμπειρικές και, επομένως, υποδηλώνει καλύτερη σχετική αριθμητική προσαρμογή. Στην παρούσα εργασία, ωστόσο, το χ^2 χρησιμοποιείται κυρίως ως συγκριτικός δείκτης κατάταξης των τεσσάρων μοντέλων και όχι ως αυτοτελής έλεγχος αποδοχής ή απόρριψής τους. Για μια πλήρη στατιστική δοκιμή θα απαιτούνταν επιπλέον συνεκτίμηση των βαθμών ελευθερίας και της αντίστοιχης τιμής πιθανότητας.

Παράλληλα, ο δείκτης δεν αποκαλύπτει σε ποια ακριβώς περιοχή της εμπειρικής κατανομής εντοπίζονται οι αποκλίσεις. Ένα μοντέλο μπορεί να εμφανίζει χαμηλή συνολική τιμή χ^2 , αλλά να παρουσιάζει τοπική αδυναμία στις υψηλότερες συγκεντρώσεις ή στην άνω ουρά. Για τον λόγο αυτό, η αριθμητική αξιολόγηση ερμηνεύεται σε συνδυασμό με τα προηγούμενα γραφικά διαγνωστικά εργαλεία.

Κατανομή	χ^2	Θέση κατάταξης
Lognormal	192,77	1
Weibull	445,54	2
Gamma	480,65	3
Normal	1273,76	4

Πίνακας 4.7.5.1 Τιμές του στατιστικού χ^2 και σχετική κατάταξη των θεωρητικών κατανομών για τις ωριαίες συγκεντρώσεις PM_{10} στη Λειβαδιά το 2019.

Η λογαριθμοκανονική κατανομή παρουσιάζει τη μικρότερη τιμή χ^2 , ίση με 192,77. Το αποτέλεσμα δείχνει ότι, στις επιλεγμένες κλάσεις συγκέντρωσης, οι αναμενόμενες συχνότητες της βρίσκονται συνολικά πλησιέστερα στις παρατηρούμενες από ό,τι οι αντίστοιχες συχνότητες των υπόλοιπων μοντέλων. Η πρώτη θέση της είναι συμβατή κυρίως με την καλή της απόδοση στα ραβδογράμματα, όπου είχε εμφανίσει την περισσότερο ισορροπημένη αντιστοίχιση μεταξύ των κλάσεων.

Η Weibull ακολουθεί με $\chi^2 = 445,54$. Η τιμή αυτή είναι περίπου 2,3 φορές μεγαλύτερη από εκείνη της λογαριθμοκανονικής, γεγονός που δείχνει ότι η διαφορά μεταξύ των δύο μοντέλων δεν είναι αμελητέα. Παρότι η Weibull είχε παρουσιάσει ιδιαίτερα συνεκτική εικόνα στα διαγράμματα παρατηρούμενων έναντι αναμενόμενων συχνοτήτων και ελαφρύ πλεονέκτημα στα διαγράμματα υπολοίπων, η συνολική απόκλιση στις ευρείες κλάσεις συγκέντρωσης παραμένει σαφώς μεγαλύτερη.

Η Gamma καταλαμβάνει την τρίτη θέση με $\chi^2 = 480,65$. Η απόστασή της από τη Weibull είναι σχετικά μικρή, περίπου 35 μονάδες χ^2 , και δείχνει ότι οι δύο κατανομές παρουσιάζουν συγκρίσιμη συνολική αριθμητική επίδοση. Η Gamma υπολείπεται ελαφρά ως προς τη συμφωνία των συχνοτήτων, παρότι στο Q-Q διάγραμμα είχε εμφανίσει την ισχυρότερη οπτική αντιστοίχιση των θεωρητικών και εμπειρικών ποσοστημορίων.

Η κανονική κατανομή εμφανίζει τη μεγαλύτερη τιμή, $\chi^2 = 1.273,76$. Η τιμή αυτή είναι πολλαπλάσια της αντίστοιχης της λογαριθμοκανονικής και επιβεβαιώνει αριθμητικά τις μεγάλες αποκλίσεις που είχαν ήδη παρατηρηθεί στα γραφικά διαγνωστικά εργαλεία. Η συμμετρική μορφή της δεν μπορεί να αποδώσει αποτελεσματικά τη συγκέντρωση της εμπειρικής μάζας στις χαμηλομεσαίες τιμές και την ταυτόχρονη ανάπτυξη της δεξιάς ουράς.

Τα αριθμητικά αποτελέσματα δεν ταυτίζονται πλήρως με την κατάταξη όλων των επιμέρους γραφικών κριτηρίων. Στα ραβδογράμματα η λογαριθμοκανονική είχε παρουσιάσει την πληρέστερη συνολική ισορροπία, ενώ στα observed-expected διαγράμματα και στα residual plots η Weibull εμφάνισε περισσότερο συνεκτική δομή. Αντίστοιχα, το Q-Q διάγραμμα ανέδειξε τη Gamma ως την κατανομή με την καλύτερη περίπου αντιστοίχιση ποσοστημορίων στο μεγαλύτερο μέρος του απεικονιζόμενου εύρους.

Η διαφοροποίηση αυτή δεν αποτελεί ασυνέπεια της ανάλυσης, αλλά αποτυπώνει τη διαφορετική λειτουργία των διαγνωστικών εργαλείων. Το χ^2 δίνει μεγαλύτερη έμφαση στη συνολική συμφωνία των συχνοτήτων ανά κλάση, ενώ τα Q-Q διαγράμματα εξετάζουν άμεσα το μέγεθος των ποσοστημορίων και είναι περισσότερο ευαίσθητα στη συμπεριφορά της άνω ουράς. Τα διαγράμματα υπολοίπων, από την άλλη πλευρά, αποκαλύπτουν τη θέση, την κατεύθυνση και τη διάρκεια των τοπικών αποκλίσεων.

Υπό αυτή την έννοια, η αριθμητική υπεροχή της λογαριθμοκανονικής σημαίνει ότι επιτυγχάνει την καλύτερη συνολική κατανομή των συχνοτήτων στις προκαθορισμένες κλάσεις. Δεν σημαίνει, όμως, ότι υπερέχει σε κάθε επιμέρους τμήμα της εμπειρικής κατανομής. Το Q-Q διάγραμμά της έδειξε συστηματική υπερεκτίμηση σε μέρος των μεσαίων ποσοστημορίων και υποεκτίμηση των πλέον υψηλών παρατηρήσεων. Επομένως, η προσαρμογή της είναι η καλύτερη συγκριτικά, αλλά όχι πλήρης.

Η δεύτερη θέση της Weibull αντανακλά επίσης μια σύνθετη εικόνα. Το μοντέλο εμφανίζει σταθερή σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων και

περιορισμένη δομή υπολοίπων μετά το αρχικό τμήμα, αλλά υποεκτιμά αισθητά την ανοικτή κλάση άνω των $50 \mu\text{g}/\text{m}^3$ και τα υψηλότερα εμπειρικά ποσοστημόρια. Η Gamma, αντίθετα, αποδίδει καλύτερα τα περισσότερα ποσοστημόρια, αλλά παρουσιάζει μεγαλύτερες αποκλίσεις στις συνολικές συχνότητες των κλάσεων.

Συνεπώς, ο Πίνακας 4.7.5.1 δεν λειτουργεί απλώς ως κατάλογος αριθμητικών τιμών, αλλά ως σημείο σύνθεσης των προηγούμενων ευρημάτων. Η λογαριθμοκανονική αποκτά σαφές πλεονέκτημα ως προς τη συνολική αριθμητική συμφωνία, η Weibull και η Gamma σχηματίζουν μια δεύτερη ομάδα με αρκετά κοντινή επίδοση, ενώ η κανονική κατανομή απομακρύνεται σημαντικά από τις τρεις ασύμμετρες επιλογές.

Η αριθμητική αξιολόγηση μετατοπίζει, επομένως, το συνολικό βάρος της επιλογής προς τη λογαριθμοκανονική κατανομή. Η τελική απόφαση, ωστόσο, πρέπει να αναγνωρίζει τόσο την υπεροχή της στο χ^2 όσο και τα επιμέρους πλεονεκτήματα της Weibull στη δομή των συχνοτήτων και της Gamma στην αντιστοίχιση των ποσοστημορίων. Η σύνθεση αυτή αποτελεί το αντικείμενο της επόμενης υποενότητας, όπου θα επιλεγεί η καταλληλότερη θεωρητική κατανομή για τη Λειβαδιά το 2019.

4.7.6 Επιλογή καταλληλότερης κατανομής για τη Λειβαδιά το 2019

Η τελική επιλογή για τη Λειβαδιά προκύπτει από μια σύνθετη, αλλά αρκετά καθαρή, εικόνα. Τα επιμέρους διαγνωστικά εργαλεία δεν ανέδειξαν πάντοτε το ίδιο μοντέλο ως ισχυρότερο, όμως η λογαριθμοκανονική κατανομή παρουσίασε την πιο ισορροπημένη συνολική επίδοση και τη σαφέστερη αριθμητική υπεροχή.

Στη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων ανά κλάση, η λογαριθμοκανονική απέδωσε καλύτερα τη γενική μορφή της εμπειρικής κατανομής. Ιδιαίτερα σημαντική ήταν η καλή της συμφωνία στην κλάση $10\text{-}20 \mu\text{g}/\text{m}^3$, όπου συγκεντρώνεται το μεγαλύτερο πλήθος παρατηρήσεων. Παρότι υποεκτίμησε την κλάση $20\text{-}30 \mu\text{g}/\text{m}^3$ και υπερεκτίμησε τις δύο ανώτερες κλάσεις, οι αποκλίσεις της ήταν συνολικά περισσότερο ισορροπημένες από εκείνες των υπόλοιπων κατανομών.

Η εικόνα των διαγνωστικών διαγραμμάτων observed-expected ήταν πιο σύνθετη. Εκεί, η Weibull παρουσίασε την πιο συνεκτική περίπου γραμμική σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων, ενώ η λογαριθμοκανονική εμφάνισε μεγαλύτερη

διασπορά στις υψηλότερες αναμενόμενες συχνότητες. Το αποτέλεσμα αυτό έδειξε ότι η Weibull αποδίδει αποτελεσματικά τη γενική σχέση των επιμέρους συχνοτήτων, χωρίς όμως να υπερέχει αντίστοιχα σε όλα τα υπόλοιπα κριτήρια.

Παρόμοια διαφοροποίηση προέκυψε και από τα διαγράμματα υπολοίπων. Η Weibull εμφάνισε μικρότερο αρχικό εύρος αποκλίσεων και ταχύτερη σταθεροποίηση γύρω από τη γραμμή μηδενός, ενώ η λογαριθμοκανονική παρουσίασε μεγαλύτερες αρχικές διακυμάνσεις και μια μικρή παρατεταμένη ζώνη αρνητικών υπολοίπων. Ωστόσο, και στις δύο κατανομές οι αποκλίσεις περιορίστηκαν γρήγορα και στο μεγαλύτερο μέρος της ταξινομημένης σειράς τα υπόλοιπα παρέμειναν πολύ κοντά στο μηδέν.

Τα Q-Q διαγράμματα ανέδειξαν διαφορετικό πλεονέκτημα. Η Gamma παρουσίασε την καλύτερη συνολική ευθυγράμμιση με τη διαγώνιο στο μεγαλύτερο μέρος των ποσοστημορίων, ενώ η λογαριθμοκανονική ακολούθησε με επίσης καλή συμπεριφορά. Η Weibull εμφάνισε εντονότερη ανοδική απόκλιση στο ανώτερο τμήμα, στοιχείο που υποδηλώνει μεγαλύτερη υποεκτίμηση των υψηλών συγκεντρώσεων.

Καθοριστικό βάρος στην τελική επιλογή έχει η αριθμητική αξιολόγηση. Η λογαριθμοκανονική κατέγραψε τη μικρότερη τιμή του στατιστικού χ^2 : $\chi^2=192,77$ έναντι 445,54 για τη Weibull, 480,65 για τη Gamma και 1.273,76 για την κανονική κατανομή. Η διαφορά αυτή δείχνει ότι η λογαριθμοκανονική επιτυγχάνει αισθητά μικρότερη συνολική απόκλιση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων στις εξεταζόμενες κλάσεις.

Η κανονική κατανομή εμφανίζει τη χαμηλότερη καταλληλότητα. Οι αποκλίσεις της είναι μεγάλες και συστηματικές στα ραβδογράμματα, στα observed-expected διαγράμματα, στα residual plots και στα Q-Q διαγράμματα. Η πολύ υψηλή τιμή χ^2 επιβεβαιώνει ότι η συμμετρική της μορφή δεν μπορεί να αποδώσει αποτελεσματικά τη θετική ασυμμετρία και τη δεξιά ουρά της εμπειρικής κατανομής.

Η Gamma αποτελεί αξιόλογη επιλογή ως προς την αντιστοίχιση των ποσοστημορίων, αλλά υστερεί στη συνολική κατανομή των συχνοτήτων. Η Weibull παρουσιάζει ισχυρή γραφική συμπεριφορά στα observed-expected διαγράμματα και στα υπόλοιπα, αλλά η αριθμητική της επίδοση είναι σαφώς ασθενέστερη από εκείνη της λογαριθμοκανονικής και το Q-Q διάγραμμα αναδεικνύει μεγαλύτερη αδυναμία στην άνω ουρά.

Η λογαριθμοκανονική δεν υπερέχει σε κάθε μεμονωμένο διάγραμμα. Υπερέχει, όμως, στη συνολική σύνθεση των αποτελεσμάτων: παρουσιάζει την καλύτερη συμφωνία ανά κλάση, διατηρεί ανταγωνιστική συμπεριφορά στα υπόλοιπα γραφικά κριτήρια και καταγράφει με

σαφή διαφορά τη χαμηλότερη τιμή χ^2 . Για τον λόγο αυτό, επιλέγεται ως η καταλληλότερη από τις τέσσερις εξεταζόμενες κατανομές για τη Λειβαδιά το 2019.

Η επιλογή αυτή πρέπει να ερμηνευθεί συγκριτικά και όχι ως ένδειξη πλήρους προσαρμογής. Στα υψηλότερα ποσοστημόρια παρατηρείται ανοδική απόκλιση, γεγονός που δείχνει ότι η λογαριθμοκανονική υποεκτιμά ορισμένες από τις μεγαλύτερες συγκεντρώσεις. Οι ακραίες τιμές παραμένουν, επομένως, απαιτητικό τμήμα όχι μόνο για τη λογαριθμοκανονική, αλλά και για όλα τα εξεταζόμενα μοντέλα.

Το αποτέλεσμα της Λειβαδιάς συμφωνεί με εκείνα της Άμφισσας και της Λαμίας, όπου είχε επίσης επιλεγεί η λογαριθμοκανονική κατανομή. Η επανάληψη αυτή αποτελεί ισχυρή εμπειρική ένδειξη ότι η συγκεκριμένη κατανομή μπορεί να περιγράφει αποτελεσματικά αρκετές από τις εξεταζόμενες χρονοσειρές PM_{10} . Δεν θεμελιώνει, όμως, έναν καθολικό κανόνα, όπως έδειξε η διαφορετική επιλογή της Weibull για τη Χαλκίδα.

Συνεπώς, το βασικό συμπέρασμα δεν είναι ότι η λογαριθμοκανονική αποτελεί πάντοτε το καταλληλότερο μοντέλο, αλλά ότι στη Λειβαδιά επιτυγχάνει την καλύτερη συνολική ισορροπία μεταξύ συμφωνίας των συχνοτήτων, συμπεριφοράς των αποκλίσεων και αριθμητικής προσαρμογής.

4.8 Περίπτωση Καρπενησίου: γραφική απεικόνιση της πληρότητας των δεδομένων

Η ένταξη μιας ετήσιας χρονοσειράς στη διαδικασία προσαρμογής θεωρητικών κατανομών προϋποθέτει επαρκή πληρότητα και χρονική συνέχεια των ωριαίων μετρήσεων. Η απλή ύπαρξη διαθέσιμων τιμών δεν αρκεί, καθώς μεγάλα ή παρατεταμένα κενά μπορούν να αλλοιώσουν την εμπειρική κατανομή των συγκεντρώσεων PM_{10} και, κατ' επέκταση, να επηρεάσουν την εκτίμηση των παραμέτρων και την αξιολόγηση της ποιότητας προσαρμογής. Για τον λόγο αυτό, πριν από τη στατιστική ανάλυση πραγματοποιήθηκε έλεγχος όλων των διαθέσιμων ετών του σταθμού του Καρπενησίου. Η αξιολόγηση βασίστηκε στον αριθμό των έγκυρων ωριαίων παρατηρήσεων, στο ποσοστό πληρότητας σε σχέση με τον αναμενόμενο αριθμό ωρών του έτους και στη χρονική κατανομή των διαθέσιμων μετρήσεων. Εφαρμόστηκαν τα ίδια ακριβώς κριτήρια που χρησιμοποιήθηκαν και για τους υπόλοιπους σταθμούς, ώστε η επιλογή των δεδομένων να παραμείνει ενιαία

και αναπαραγώγιμη. Ως εσωτερικό κριτήριο ένταξης στην τελική διαδικασία προσαρμογής ορίστηκε ποσοστό πληρότητας τουλάχιστον 90%, σε συνδυασμό με ικανοποιητική χρονική συνέχεια. Η δεύτερη προϋπόθεση είναι ουσιώδης, επειδή δύο χρονοσειρές με παρόμοιο ποσοστό πληρότητας μπορεί να διαφέρουν σημαντικά ως προς την ποιότητά τους: σε μία περίπτωση οι ελλείψεις μπορεί να είναι διάσπαρτες, ενώ σε άλλη να σχηματίζουν μεγάλα συνεχόμενα χρονικά κενά.

Από τον έλεγχο του Καρπενησίου δεν προέκυψε κάποιο έτος που να ικανοποιεί ταυτόχρονα τις απαιτήσεις πληρότητας και χρονικής συνέχειας. Ως εκ τούτου, δεν σχηματίστηκε ετήσιο σύνολο δεδομένων κατάλληλο για την εφαρμογή των κατανομών Normal, Gamma, Weibull και Lognormal. Η συνέχιση της ανάλυσης με ελλιπή δεδομένα θα μπορούσε να οδηγήσει σε παραπλανητική εικόνα της συχνοτικής δομής και ιδίως σε υποεκπροσώπηση περιόδων με διαφορετική εποχική ή επεισοδιακή συμπεριφορά. Ο αποκλεισμός του σταθμού δεν αποτελεί, επομένως, αυθαίρετη επιλογή ούτε διαφορετική μεθοδολογική μεταχείριση. Αντιθέτως, αποτελεί άμεσο αποτέλεσμα της εφαρμογής του ίδιου ελέγχου ποιότητας που χρησιμοποιήθηκε για όλες τις εξεταζόμενες πόλεις. Η διατήρηση κοινών κριτηρίων είναι απαραίτητη, ώστε οι συγκρίσεις μεταξύ των σταθμών να βασίζονται σε χρονοσειρές αντίστοιχης αξιοπιστίας.

Η γραφική απεικόνιση της πληρότητας λειτουργεί ως τεκμήριο της απόφασης αυτής. Επιτρέπει την άμεση αναγνώριση των ετών με μειωμένη διαθεσιμότητα και αποκαλύπτει αν οι ελλείψεις κατανέμονται ομοιόμορφα ή συγκεντρώνονται σε συγκεκριμένες χρονικές περιόδους. Έτσι, ο αναγνώστης μπορεί να διαπιστώσει ότι η μη εφαρμογή θεωρητικών κατανομών στο Καρπενήσι προκύπτει από τα χαρακτηριστικά των διαθέσιμων δεδομένων και όχι από υποκειμενική επιλογή.

Το Καρπενήσι αντιμετωπίζεται, συνεπώς, ως ειδική περίπτωση ελέγχου καταλληλότητας του δείγματος και όχι ως περίπτωση αποτυχημένης προσαρμογής. Η διάκριση είναι σημαντική: δεν διαπιστώθηκε ότι κάποια θεωρητική κατανομή δεν περιγράφει ικανοποιητικά τις συγκεντρώσεις της πόλης, αλλά ότι τα διαθέσιμα δεδομένα δεν επιτρέπουν αξιόπιστη διερεύνηση του συγκεκριμένου ερωτήματος.

Η παρουσίαση της περίπτωσης αυτής ενισχύει τη μεθοδολογική διαφάνεια της εργασίας. Τεκμηριώνει ότι η στατιστική προσαρμογή εφαρμόστηκε μόνο όπου υπήρχαν επαρκείς προϋποθέσεις αντιπροσωπευτικότητας και ότι η απουσία αποτελεσμάτων για το Καρπενήσι αποτελεί εύρημα του ελέγχου ποιότητας των δεδομένων και όχι παράλειψη της ανάλυσης.

4.8.1 Έλεγχος πληρότητας των δεδομένων Καρπενησίου

Ο έλεγχος πληρότητας των δεδομένων του Καρπενησίου προηγήθηκε οποιασδήποτε προσπάθειας προσαρμογής θεωρητικών κατανομών. Η διαδικασία αυτή ήταν απαραίτητη, καθώς η αξιοπιστία της κατανομικής ανάλυσης εξαρτάται από την ύπαρξη επαρκούς αριθμού έγκυρων ωριαίων παρατηρήσεων και από την ικανοποιητική κάλυψη του εξεταζόμενου έτους. Για τον σκοπό αυτό εξετάστηκαν όλες οι διαθέσιμες ετήσιες χρονοσειρές της περιόδου 2017-2024. Για κάθε έτος υπολογίστηκαν ο θεωρητικά αναμενόμενος αριθμός ωριαίων παρατηρήσεων, ο αριθμός των έγκυρων τιμών και το αντίστοιχο ποσοστό πληρότητας. Ως εσωτερικό κριτήριο ένταξης στην τελική στατιστική ανάλυση είχε οριστεί ποσοστό πληρότητας τουλάχιστον 90%.

Έτος	Ημέρες	Θεωρητικές ώρες	Έγκυρες ώρες	Πληρότητα (%)	Κατάσταση
2017	365	8760	3755	42,87	Μη αναλύσιμο
2018	365	8760	6923	79,03	Μη αναλύσιμο
2019	365	8760	6740	76,94	Μη αναλύσιμο
2020	366	8784	4860	55,33	Μη αναλύσιμο
2021	365	8760	0	0,00	Μη αναλύσιμο
2022	365	8760	0	0,00	Μη αναλύσιμο
2023	365	8760	0	0,00	Μη αναλύσιμο
2024	366	8784	0	0,00	Μη αναλύσιμο

Πίνακας 4.8.1. Έλεγχος πληρότητας των ωριαίων συγκεντρώσεων PM₁₀ στον σταθμό του Καρπενησίου για την περίοδο 2017-2024.

Τα αποτελέσματα δείχνουν ότι κανένα από τα εξεταζόμενα έτη δεν ικανοποιεί το όριο πληρότητας του 90%. Η καλύτερη διαθέσιμη χρονοσειρά είναι εκείνη του 2018, με 6.923

έγκυρες ωριαίες παρατηρήσεις και πληρότητα 79,03%. Ακολουθεί το 2019 με 6.740 παρατηρήσεις και πληρότητα 76,94%. Παρότι τα δύο αυτά έτη διαθέτουν σαφώς περισσότερα δεδομένα από τα υπόλοιπα, υπολείπονται του προκαθορισμένου ορίου κατά περίπου 11 και 13 ποσοστιαίες μονάδες αντίστοιχα.

Η απόκλιση αυτή δεν θεωρήθηκε αμελητέα. Στο 2018 απουσιάζουν 1.837 ωριαίες παρατηρήσεις, ενώ στο 2019 οι ελλείπουσες ώρες ανέρχονται σε 2.020. Επομένως, περίπου το ένα πέμπτο έως το ένα τέταρτο κάθε ετήσιας χρονοσειράς δεν καλύπτεται από έγκυρες μετρήσεις. Χωρίς περαιτέρω τεκμηρίωση ότι τα κενά είναι τυχαία και ομοιόμορφα κατανεμημένα στον χρόνο, δεν μπορεί να θεωρηθεί ασφαλές ότι οι διαθέσιμες παρατηρήσεις αναπαριστούν πλήρως την ετήσια συμπεριφορά των συγκεντρώσεων.

Ακόμη χαμηλότερη πληρότητα εμφανίζεται το 2017, για το οποίο είναι διαθέσιμες 3.755 από τις 8.760 αναμενόμενες ώρες. Το αντίστοιχο ποσοστό, 42,87%, σημαίνει ότι απουσιάζει περισσότερο από το μισό της ετήσιας χρονοσειράς. Ανάλογη εικόνα παρουσιάζεται το 2020, με 4.860 έγκυρες τιμές σε σύνολο 8.784 θεωρητικών ωρών και πληρότητα 55,33%. Μετά το 2020 παρατηρείται πλήρης διακοπή της διαθεσιμότητας έγκυρων μετρήσεων. Για τα έτη 2021-2024 δεν καταγράφεται καμία έγκυρη ωριαία τιμή και η πληρότητα είναι μηδενική. Τα έτη αυτά αποκλείονται, επομένως, εξ ορισμού από οποιαδήποτε περιγραφική ή κατανομική ανάλυση.

Η ανεπάρκεια της ετήσιας κάλυψης δεν αποτελεί μόνο ζήτημα μεγέθους δείγματος. Αν τα κενά συγκεντρώνονται σε συγκεκριμένους μήνες ή εποχές, ενδέχεται να απουσιάζουν περίοδοι με διαφορετική μετεωρολογική και ατμοσφαιρική συμπεριφορά. Στην περίπτωση αυτή, η εμπειρική κατανομή θα μπορούσε να υπεραντιπροσωπεύει ορισμένα επίπεδα συγκέντρωσης και να υποαντιπροσωπεύει άλλα. Ιδιαίτερα οι σπάνιες υψηλές τιμές είναι πιθανό να μην καταγραφούν επαρκώς, με αποτέλεσμα να αλλοιωθεί η μορφή της δεξιάς ουράς.

Για τον λόγο αυτό δεν επιλέχθηκε απλώς το έτος με τη μεγαλύτερη διαθέσιμη πληρότητα. Η εφαρμογή θεωρητικών κατανομών στο 2018 ή στο 2019, μόνο επειδή αποτελούν τις λιγότερο ελλειπείς χρονοσειρές του σταθμού, θα παραβίαζε το κοινό κριτήριο που εφαρμόστηκε στις υπόλοιπες πόλεις και θα μείωνε τη συγκρισιμότητα των αποτελεσμάτων.

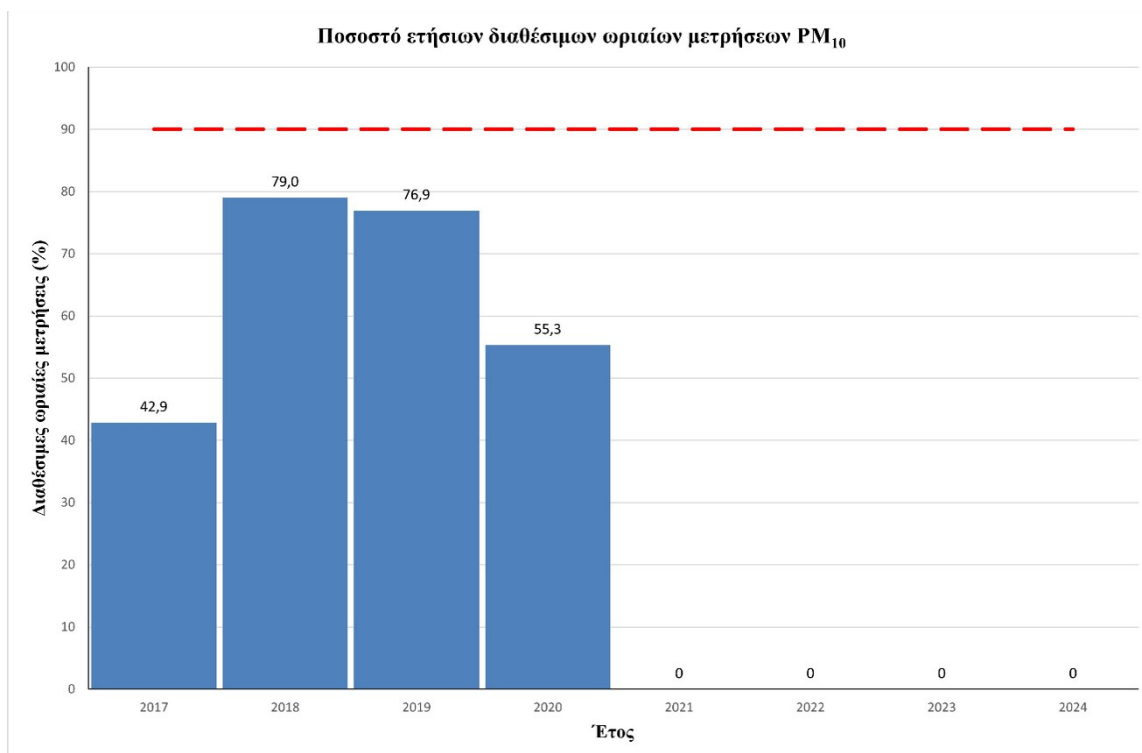
Συνεπώς, καμία ετήσια χρονοσειρά του Καρπενησίου δεν εντάχθηκε στη διαδικασία προσαρμογής των κατανομών Normal, Gamma, Weibull και Lognormal. Η απόφαση αυτή δεν αποτελεί επιλεκτικό αποκλεισμό του σταθμού, αλλά άμεσο αποτέλεσμα του προκαθορισμένου ελέγχου ποιότητας. Το Καρπενήσι διατηρείται στην ανάλυση ως τεκμηριωμένη περίπτωση ανεπαρκούς διαθεσιμότητας δεδομένων, αναδεικνύοντας ότι η μη εφαρμογή ενός στατιστικού μοντέλου μπορεί να αποτελεί μεθοδολογικά ορθότερη επιλογή από την προσαρμογή του σε μη επαρκώς αντιπροσωπευτική χρονοσειρά.

4.8.2 Γραφική απεικόνιση της πληρότητας ανά έτος

Για τη συμπληρωματική οπτική τεκμηρίωση του ελέγχου πληρότητας, τα ποσοστά διαθεσιμότητας των ωριαίων μετρήσεων PM_{10} στον σταθμό του Καρπενησίου παρουσιάζονται γραφικά ανά έτος. Η γραφική αυτή απεικόνιση επιτρέπει την άμεση αποτίμηση της επάρκειας των δεδομένων και διευκολύνει τη σύγκριση κάθε ετήσιας χρονοσειράς με το εσωτερικό όριο αποδοχής του 90%, το οποίο χρησιμοποιήθηκε στην παρούσα εργασία ως κριτήριο ένταξης στην κατανομική ανάλυση.

Σε αντίθεση με τον πίνακα, ο οποίος παρέχει αναλυτικά τις αριθμητικές τιμές της πληρότητας, το διάγραμμα αναδεικνύει πιο καθαρά τη συνολική εικόνα της διαθεσιμότητας των δεδομένων διαχρονικά. Με τον τρόπο αυτό καθίσταται ευκολότερα αντιληπτό όχι μόνο ποια έτη εμφανίζουν υψηλότερη σχετική πληρότητα, αλλά και πόσο σημαντική παραμένει η απόκλισή τους από το απαιτούμενο κατώφλι αποδοχής.

Ενώ ο πίνακας παρέχει την ακριβή αριθμητική πληροφορία, το διάγραμμα αναδεικνύει οπτικά τη συνολική εικόνα της ανεπάρκειας των δεδομένων και διευκολύνει τη σύγκριση μεταξύ των ετών.



Σχήμα 4.8.2. Ποσοστό διαθέσιμων ωριαίων μετρήσεων PM_{10} ανά έτος στον σταθμό Καρπενησίου και σύγκριση με το εσωτερικό όριο πληρότητας 90%.

Από το Σχήμα 4.8.2 προκύπτει με σαφήνεια ότι κανένα από τα εξεταζόμενα έτη δεν προσεγγίζει το απαιτούμενο επίπεδο πληρότητας. Οι υψηλότερες τιμές καταγράφονται το 2018 και το 2019, με ποσοστά 79,03% και 76,94% αντίστοιχα, τα οποία, αν και αισθητά υψηλότερα από εκείνα των λοιπών ετών, εξακολουθούν να υπολείπονται του ορίου αποδοχής. Το 2020 παρουσιάζει περαιτέρω μείωση της πληρότητας, ενώ το 2017 εμφανίζει ακόμη χαμηλότερη διαθεσιμότητα δεδομένων, γεγονός που υποδηλώνει ήδη από τα πρώτα έτη σημαντική ανεπάρκεια της χρονοσειράς.

Ιδιαίτερα χαρακτηριστική είναι η εικόνα μετά το 2020, καθώς από το 2021 έως και το 2024 η πληρότητα μηδενίζεται. Η πλήρης απουσία έγκυρων ωριαίων παρατηρήσεων στα έτη αυτά δεν αφήνει κανένα περιθώριο για στατιστική επεξεργασία ή για εφαρμογή θεωρητικών κατανομών. Η απότομη αυτή μετάβαση από μερικώς διαθέσιμες χρονοσειρές σε πλήρη έλλειψη δεδομένων αποτυπώνεται εξαιρετικά καθαρά στο διάγραμμα και ενισχύει την τεκμηρίωση του αποκλεισμού του σταθμού από την κύρια αναλυτική διαδικασία.

Συνολικά, η γραφική απεικόνιση επιβεβαιώνει και ενισχύει τα ευρήματα του Πίνακα 4.8.1. Το Καρπενήσι δεν διαθέτει καμία ετήσια χρονοσειρά που να ικανοποιεί το κριτήριο πληρότητας και, συνεπώς, δεν μπορεί να ενταχθεί με μεθοδολογική ασφάλεια στη διαδικασία προσαρμογής θεωρητικών κατανομών. Η παρουσίαση αυτή είναι ιδιαίτερα σημαντική για τη διαφάνεια της ανάλυσης, καθώς δείχνει ότι ο αποκλεισμός του σταθμού δεν αποτελεί αυθαίρετη επιλογή, αλλά άμεση συνέπεια της εφαρμογής ενιαίων και σαφώς ορισμένων ποιοτικών κριτηρίων.

4.8.3 Μεθοδολογική αιτιολόγηση αποκλεισμού

Η περίπτωση του Καρπενησίου αναδεικνύει ότι ο έλεγχος ποιότητας των δεδομένων δεν αποτελεί τυπικό προκαταρκτικό βήμα, αλλά ουσιαστική προϋπόθεση για την αξιόπιστη εφαρμογή στατιστικών μοντέλων. Παρότι ο σταθμός εξετάστηκε με την ίδια διαδικασία που εφαρμόστηκε και στις υπόλοιπες πόλεις, κανένα από τα διαθέσιμα έτη δεν ικανοποίησε το εσωτερικό κριτήριο πληρότητας και χρονικής συνέχειας που είχε τεθεί στη μεθοδολογία.

Για τον λόγο αυτό, δεν εφαρμόστηκαν στο Καρπενήσι οι κατανομές Normal, Gamma, Weibull και Lognormal. Η απόφαση αυτή δεν αποτελεί παράλειψη ούτε διαφορετική μεταχείριση του σταθμού, αλλά συνέπεια της ενιαίας διαδικασίας επιλογής του αναλυτικού δείγματος. Η χρήση ελλιπών χρονοσειρών θα μπορούσε να οδηγήσει σε μη αντιπροσωπευτική εμπειρική κατανομή, με αποτέλεσμα να επηρεαστούν τόσο η εκτίμηση των παραμέτρων όσο και η συγκριτική αξιολόγηση των θεωρητικών μοντέλων.

Η σύγκριση με τις υπόλοιπες πόλεις καθιστά σαφέστερο τον ρόλο του αρχικού ελέγχου. Στην Αμφισσα, στη Λαμία, στη Χαλκίδα και στη Λειβαδιά εντοπίστηκαν ετήσιες χρονοσειρές που πληρούσαν τις προϋποθέσεις ένταξης και μπορούσαν να υποστηρίξουν τη διαδικασία προσαρμογής. Στο Καρπενήσι, αντίθετα, η ίδια ακριβώς διαδικασία κατέληξε σε αρνητικό αποτέλεσμα. Επομένως, η τελική σύνθεση του δείγματος δεν καθορίστηκε απλώς από την ύπαρξη σταθμών ή μεμονωμένων μετρήσεων, αλλά από την επάρκεια και την ποιότητα κάθε ετήσιας χρονοσειράς.

Η επιλογή αυτή ενισχύει τη συγκρισιμότητα των αποτελεσμάτων. Η ένταξη του Καρπενησίου με χαμηλότερα κριτήρια πληρότητας θα δημιουργούσε άνιση μεθοδολογική

μεταχείριση και θα καθιστούσε τις μεταξύ πόλεων συγκρίσεις λιγότερο αξιόπιστες. Αντίθετα, η διατήρηση του ίδιου ορίου και των ίδιων ποιοτικών προϋποθέσεων για όλους τους σταθμούς εξασφαλίζει ότι οι τελικές αναλύσεις βασίζονται σε δεδομένα αντίστοιχης επάρκειας.

Παράλληλα, η ξεχωριστή παρουσίαση του Καρπενησίου ενισχύει τη διαφάνεια της εργασίας. Αν ο σταθμός απουσίαζε απλώς από τις ενότητες προσαρμογής, χωρίς αιτιολόγηση, θα μπορούσε να δημιουργηθεί η εντύπωση ότι δεν εξετάστηκε ή ότι αποκλείστηκε αυθαίρετα. Η παράθεση του πίνακα πληρότητας και της γραφικής απεικόνισης αποδεικνύει ότι ο σταθμός αξιολογήθηκε κανονικά, αλλά δεν διέθετε ετήσια χρονοσειρά κατάλληλη για ασφαλή στατιστική αξιοποίηση.

Το Καρπενήσι δεν πρέπει, επομένως, να ερμηνευθεί ως περίπτωση αποτυχημένης προσαρμογής. Δεν διαπιστώθηκε ότι κάποια θεωρητική κατανομή δεν μπορεί να περιγράψει τις συγκεντρώσεις PM_{10} της πόλης. Διαπιστώθηκε ότι τα διαθέσιμα δεδομένα δεν επαρκούν για να εξεταστεί αξιόπιστα αυτό το ερώτημα. Η διάκριση αυτή είναι μεθοδολογικά ουσιώδης, επειδή αποτρέπει την εξαγωγή συμπερασμάτων που δεν υποστηρίζονται επαρκώς από το διαθέσιμο δείγμα.

Συνολικά, η ενότητα 4.8 τεκμηριώνει ότι η μη εφαρμογή στατιστικής μοντελοποίησης μπορεί να αποτελεί εξίσου σημαντική επιστημονική απόφαση με την επιλογή ενός κατάλληλου μοντέλου. Στην περίπτωση του Καρπενησίου, η αποχή από την προσαρμογή θεωρητικών κατανομών προστατεύει την εγκυρότητα της ανάλυσης και επιβεβαιώνει ότι τα τελικά αποτελέσματα του κεφαλαίου βασίζονται αποκλειστικά σε χρονοσειρές που πληρούν σαφή, ενιαία και αναπαραγώγιμα κριτήρια ποιότητας.

4.9 Συγκριτική αξιολόγηση των θεωρητικών κατανομών

Η ολοκλήρωση της προσαρμογής των κατανομών Normal, Gamma, Weibull και Lognormal στις αναλύσιμες χρονοσειρές επιτρέπει πλέον τη συνολική σύνθεση των αποτελεσμάτων. Στις προηγούμενες ενότητες κάθε πόλη εξετάστηκε αυτοτελώς, με βάση τη συμφωνία παρατηρούμενων και αναμενόμενων συχνοτήτων, τα διαγνωστικά διαγράμματα observed-expected, τα διαγράμματα υπολοίπων, τα Q-Q plots και την αριθμητική αξιολόγηση μέσω του στατιστικού χ^2 .

Στην παρούσα ενότητα το ενδιαφέρον μετατοπίζεται από την επιμέρους αξιολόγηση στη συγκριτική αποτίμηση των τεσσάρων θεωρητικών μοντέλων. Στόχος δεν είναι η επανάληψη των αποτελεσμάτων κάθε πόλης, αλλά η ανάδειξη των γενικών προτύπων που προέκυψαν από την εφαρμογή της ίδιας μεθοδολογίας σε διαφορετικές ετήσιες χρονοσειρές PM_{10} .

Η σύγκριση αυτή επιτρέπει να εξεταστεί αν κάποια θεωρητική κατανομή παρουσιάζει σταθερά καλύτερη συμπεριφορά ή αν η καταλληλότητά της εξαρτάται από τα ιδιαίτερα χαρακτηριστικά κάθε χρονοσειράς. Το ερώτημα αυτό είναι ουσιώδες, καθώς οι συγκεντρώσεις PM_{10} στις εξεταζόμενες πόλεις εμφανίζουν κοινά χαρακτηριστικά, όπως μη αρνητικές τιμές, θετική ασυμμετρία και παρουσία δεξιάς ουράς, χωρίς όμως να έχουν την ίδια ακριβώς κατανομική δομή. Οι χρονοσειρές διαφοροποιούνται ως προς τη μέση τιμή, τη διασπορά, την ασυμμετρία, την κύρτωση και την ένταση των υψηλών παρατηρήσεων. Οι διαφορές αυτές επηρεάζουν άμεσα την απόδοση των θεωρητικών μοντέλων. Μια κατανομή μπορεί να αποδίδει αποτελεσματικά το κύριο σώμα των δεδομένων, αλλά να δυσκολεύεται στην άνω ουρά, ενώ μια άλλη μπορεί να προσαρμόζεται καλύτερα στα ποσοστημόρια χωρίς να αναπαράγει με την ίδια ακρίβεια τις συχνότητες των επιμέρους κλάσεων.

Για τον λόγο αυτό, η συγκριτική αξιολόγηση δεν στηρίζεται αποκλειστικά στη χαμηλότερη τιμή χ^2 . Το στατιστικό χ^2 παρέχει ένα ενιαίο αριθμητικό μέτρο της συνολικής απόκλισης μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων, αλλά δεν δείχνει πού ακριβώς εντοπίζονται οι αποκλίσεις. Τα γραφικά διαγνωστικά εργαλεία συμπληρώνουν την αριθμητική εικόνα, αποκαλύπτοντας τη δομή των σφαλμάτων, τη συμφωνία στο κύριο σώμα και τη συμπεριφορά στα υψηλότερα ποσοστημόρια.

Η σύνθεση των αποτελεσμάτων έχει, επομένως, διπλό στόχο. Αφενός, εξετάζει τη συνολική συμπεριφορά κάθε θεωρητικής κατανομής στις τέσσερις αναλύσιμες πόλεις. Αφετέρου, διερευνά κατά πόσο οι διαφορετικές επιλογές μοντέλου μπορούν να συνδεθούν με τις ιδιαίτερες κατανομικές ιδιότητες κάθε χρονοσειράς.

Ιδιαίτερη σημασία έχει και η περίπτωση του Καρπενησίου. Ο σταθμός δεν συμμετέχει στη συγκριτική αξιολόγηση των θεωρητικών κατανομών, επειδή δεν εντοπίστηκε ετήσια χρονοσειρά που να πληροί τα κριτήρια πληρότητας και χρονικής συνέχειας. Η εξαίρεσή του διασφαλίζει ότι η σύγκριση βασίζεται αποκλειστικά σε σύνολα δεδομένων επαρκούς ποιότητας και αποτρέπει τη δημιουργία μη ισότιμων ή παραπλανητικών συγκρίσεων.

Η ενότητα 4.9 ολοκληρώνει, συνεπώς, τη μετάβαση από τις επιμέρους εφαρμογές στη συνολική ερμηνεία των αποτελεσμάτων. Μέσα από αυτή τη σύνθεση θα αναδειχθούν τόσο οι κατανομές που εμφανίζουν συχνότερα ευνοϊκή συμπεριφορά όσο και οι περιπτώσεις στις οποίες η τελική επιλογή διαφοροποιείται. Έτσι καθίσταται δυνατή μια περισσότερο τεκμηριωμένη απάντηση στο βασικό ερευνητικό ερώτημα της εργασίας: αν υπάρχει ένα θεωρητικό μοντέλο που περιγράφει γενικά τις ωριαίες συγκεντρώσεις PM_{10} ή αν η καταλληλότερη κατανομή πρέπει να προσδιορίζεται χωριστά για κάθε πόλη και ετήσια χρονοσειρά.

4.9.1 Συγκριτική παρουσίαση των τιμών χ^2 και των τελικών επιλογών

Ο Πίνακας 4.9 συγκεντρώνει τις τιμές του στατιστικού χ^2 για τις τέσσερις θεωρητικές κατανομές που εφαρμόστηκαν στις αναλύσιμες χρονοσειρές PM_{10} και παρουσιάζει την τελική επιλογή για κάθε πόλη. Με τον τρόπο αυτό καθίσταται δυνατή τόσο η ενδοαστική σύγκριση των υποψήφιων μοντέλων όσο και η συνολική αποτίμηση της συμπεριφοράς τους στις διαφορετικές χρονοσειρές.

Πόλη	Έτος	Normal χ^2	Gamma χ^2	Weibull χ^2	Lognormal χ^2	Τελική επιλογή
Αμφισσα	2018	1634,93	466,61	476,42	96,26	Lognormal
Λαμία	2017	1568,23	574,16	346,07	227,85	Lognormal
Λειβαδιά	2019	1273,76	480,65	445,54	192,77	Lognormal
Χαλκίδα	2017	2062,81	591,19	46,09	234,53	Weibull

Πίνακας 4.9 Συγκριτική αξιολόγηση των θεωρητικών κατανομών στις αναλύσιμες χρονοσειρές PM_{10}

Η πρώτη βασική διαπίστωση είναι η σαφής και επαναλαμβανόμενη αδυναμία της κανονικής κατανομής. Σε όλες τις πόλεις παρουσιάζει τη μεγαλύτερη τιμή χ^2 , με πολύ μεγάλη απόσταση από τις τρεις ασύμμετρες κατανομές. Το αποτέλεσμα αυτό επιβεβαιώνει ότι η συμμετρική μορφή της δεν ανταποκρίνεται ικανοποιητικά στη θετικά ασύμμετρη δομή των ωριαίων συγκεντρώσεων PM_{10} .

Αντίθετα, οι Gamma, Weibull και Lognormal εμφανίζουν αισθητά χαμηλότερες τιμές και, επομένως, καλύτερη συνολική αριθμητική προσαρμογή. Ωστόσο, η συμπεριφορά τους δεν είναι ίδια σε όλες τις χρονοσειρές. Η Gamma βελτιώνει συστηματικά την εικόνα σε σχέση με τη Normal, χωρίς όμως να αναδεικνύεται ως τελική επιλογή σε κάποια πόλη. Η Weibull παρουσιάζει πιο μεταβλητή απόδοση, ενώ η Lognormal εμφανίζει τη μεγαλύτερη συνολική συνέπεια.

Η λογαριθμοκανονική κατανομή επιλέγεται στην Άμφισσα, στη Λαμία και στη Λειβαδιά. Στις τρεις αυτές περιπτώσεις καταγράφει τη χαμηλότερη τιμή χ^2 και συνδυάζει την αριθμητική υπεροχή με γενικά ισχυρή εικόνα στα γραφικά διαγνωστικά εργαλεία. Η επανάληψη της επιλογής της σε τρεις διαφορετικές χρονοσειρές δείχνει ότι αποτελεί ιδιαίτερα αποτελεσματικό θεωρητικό πρότυπο για δεδομένα με υψηλή συγκέντρωση παρατηρήσεων στις χαμηλές και μεσαίες τιμές και με σταδιακά αραιότερη δεξιά ουρά.

Η υπεροχή της, ωστόσο, δεν έχει την ίδια ένταση σε όλες τις πόλεις. Στην Άμφισσα η τιμή χ^2 της Lognormal, 96,26, είναι πολύ χαμηλότερη από εκείνες της Gamma και της Weibull, οι οποίες βρίσκονται κοντά μεταξύ τους. Στη Λαμία η απόσταση από τη Weibull είναι μικρότερη, αλλά παραμένει σαφής. Αντίστοιχα, στη Λειβαδιά η Lognormal υπερέχει αισθητά τόσο της Weibull όσο και της Gamma, παρά τα επιμέρους πλεονεκτήματα των δύο τελευταίων σε ορισμένα γραφικά κριτήρια.

Η Χαλκίδα διαφοροποιείται έντονα από το γενικό αυτό πρότυπο. Η Weibull παρουσιάζει $\chi^2 = 46,09$, τιμή πολύ χαμηλότερη από την αντίστοιχη της Lognormal, $\chi^2 = 234,53$. Η αριθμητική αυτή απόσταση δεν είναι οριακή και δείχνει ότι η συγκεκριμένη χρονοσειρά διαθέτει διαφορετική συχνοτική δομή, η οποία αποδίδεται αποτελεσματικότερα από τη Weibull. Η Χαλκίδα αποτελεί, επομένως, την κύρια εξαίρεση στην επικράτηση της λογαριθμοκανονικής κατανομής.

Ιδιαίτερο ενδιαφέρον παρουσιάζει και η σχετική θέση των Weibull και Gamma. Στην Άμφισσα οι τιμές τους είναι σχεδόν ισοδύναμες, με μικρό πλεονέκτημα της Gamma. Στη Λαμία η Weibull υπερέχει καθαρότερα, ενώ στη Λειβαδιά οι δύο κατανομές εμφανίζουν

σχετικά κοντινή αριθμητική επίδοση. Στη Χαλκίδα, αντίθετα, η Weibull απομακρύνεται εντυπωσιακά από τη Gamma και κατακτά με σαφήνεια την πρώτη θέση. Η μεταβλητότητα αυτή δείχνει ότι η απόδοση της Weibull εξαρτάται περισσότερο από τα ιδιαίτερα χαρακτηριστικά της εκάστοτε χρονοσειράς.

Η συνολική εικόνα δεν υποστηρίζει, επομένως, την ύπαρξη μιας κατανομής που υπερέχει καθολικά. Η Lognormal αποτελεί το επικρατέστερο μοντέλο, επειδή επιλέγεται στις περισσότερες περιπτώσεις και εμφανίζει τη μεγαλύτερη σταθερότητα. Ωστόσο, η σαφής υπεροχή της Weibull στη Χαλκίδα δείχνει ότι η τελική επιλογή πρέπει να παραμένει εξαρτώμενη από τη μορφή της εμπειρικής κατανομής και να μην προκύπτει αυτοματοποιημένα από ένα γενικό πρότυπο.

Ο Πίνακας 4.9 λειτουργεί, επομένως, όχι μόνο ως σύνοψη αριθμητικών αποτελεσμάτων, αλλά και ως ένδειξη της ετερογένειας των εξεταζόμενων χρονοσειρών. Από τη μία πλευρά, αναδεικνύει τη συστηματική αδυναμία της Normal και τη γενική καταλληλότητα των ασύμμετρων μοντέλων. Από την άλλη, δείχνει ότι ακόμη και μεταξύ των ασύμμετρων κατανομών η σχετική επίδοση μεταβάλλεται ουσιαστικά από πόλη σε πόλη.

Συνολικά, η λογαριθμοκανονική κατανομή μπορεί να θεωρηθεί το βασικό θεωρητικό πρότυπο της παρούσας ανάλυσης, χωρίς όμως να αποτελεί καθολική λύση. Η περίπτωση της Χαλκίδας επιβεβαιώνει ότι η καταλληλότερη κατανομή πρέπει να προσδιορίζεται χωριστά για κάθε ετήσια χρονοσειρά, με βάση τη συνδυαστική ερμηνεία των αριθμητικών και γραφικών διαγνωστικών κριτηρίων.

4.9.2 Συνολική ερμηνεία της απόδοσης των κατανομών

Η συγκριτική αξιολόγηση αναδεικνύει αρχικά τη συστηματική αδυναμία της κανονικής κατανομής να περιγράψει ικανοποιητικά τις εμπειρικές κατανομές των ωριαίων συγκεντρώσεων PM_{10} . Η Normal κατέγραψε την υψηλότερη τιμή χ^2 σε όλες τις αναλύσιμες χρονοσειρές και παρουσίασε τις εντονότερες αποκλίσεις στα γραφικά διαγνωστικά εργαλεία. Η συμπεριφορά αυτή συνδέεται με τη συμμετρική μορφή της, η οποία δεν ανταποκρίνεται επαρκώς σε δεδομένα μη αρνητικά, θετικά ασύμμετρα και με εκτεταμένη δεξιά ουρά.

Η αδυναμία της Normal δεν περιορίστηκε σε ένα μόνο στάδιο της ανάλυσης. Στη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων εμφάνισε συστηματικές αστοχίες στις βασικές κλάσεις συγκέντρωσης. Στα observed-expected διαγράμματα τα σημεία παρουσίασαν μεγάλη διασπορά και οργανωμένη απομάκρυνση από τη διαγώνιο αναφοράς, ενώ στα residual plots καταγράφηκαν έντονα κυματοειδή πρότυπα. Αντίστοιχα, τα Q-Q plots ανέδειξαν σημαντική απόκλιση, ιδίως στα υψηλά ποσοστημόρια. Συνεπώς, η χαμηλή απόδοσή της προκύπτει από το σύνολο των διαγνωστικών ενδείξεων και όχι μόνο από τις τιμές του χ^2 .

Το δεύτερο βασικό εύρημα αφορά τη σαφή υπεροχή των τριών ασύμμετρων κατανομών. Οι Gamma, Weibull και Lognormal είναι καταλληλότερες για μη αρνητικά δεδομένα και μπορούν να αποδώσουν διαφορετικούς βαθμούς θετικής ασυμμετρίας. Για τον λόγο αυτό παρουσίασαν σημαντικά χαμηλότερες τιμές χ^2 και καλύτερη γραφική συμπεριφορά από τη Normal σε όλες τις πόλεις.

Η Gamma εμφάνισε σταθερή βελτίωση έναντι της κανονικής, χωρίς όμως να αναδειχθεί ως τελική επιλογή σε κάποια χρονοσειρά. Σε αρκετές περιπτώσεις απέδωσε ικανοποιητικά σημαντικό μέρος της κύριας εμπειρικής μάζας και παρουσίασε καλή ευθυγράμμιση σε τμήματα των Q-Q διαγραμμάτων. Ωστόσο, οι αποκλίσεις της στις χαμηλές ή υψηλές κλάσεις και η λιγότερο ομοιόμορφη δομή των υπολοίπων περιόρισαν τη συνολική της επίδοση. Η Gamma μπορεί, επομένως, να θεωρηθεί αξιόλογο εναλλακτικό μοντέλο, αλλά όχι το περισσότερο συνεπές στις συγκεκριμένες εφαρμογές.

Η Weibull παρουσίασε περισσότερο εξαρτώμενη από τη χρονοσειρά συμπεριφορά. Στη Λαμία και στη Λειβαδιά εμφάνισε καλή αριθμητική και γραφική απόδοση, χωρίς να υπερβεί συνολικά τη Lognormal. Στην Άμφισσα η επίδοσή της ήταν παρόμοια με εκείνη της Gamma και σαφώς ασθενέστερη από τη λογαριθμοκανονική. Στη Χαλκίδα, αντίθετα, αναδείχθηκε ως η καταλληλότερη κατανομή, με ιδιαίτερα χαμηλή τιμή χ^2 και ισχυρή συμφωνία στις συχνότητες και στα υπόλοιπα.

Η περίπτωση της Χαλκίδας δείχνει ότι η Weibull μπορεί να αποδώσει εξαιρετικά συγκεκριμένες εμπειρικές δομές, ακόμη και όταν δεν αποτελεί την κυρίαρχη επιλογή στο σύνολο των πόλεων. Η απόδοσή της φαίνεται να είναι ιδιαίτερα ισχυρή όταν η κατανομή της κύριας μάζας των παρατηρήσεων προσαρμόζεται αποτελεσματικά στη μορφή της, παρά τις πιθανές αστοχίες στην άνω ουρά. Επομένως, η Weibull δεν λειτουργεί ως γενικά σταθερό πρότυπο, αλλά ως μοντέλο που μπορεί να παρουσιάσει πολύ μεγάλη υπεροχή σε ειδικές περιπτώσεις.

Η Lognormal εμφάνισε τη μεγαλύτερη συνολική συνέπεια. Επιλέχθηκε στην Αμφισσα, στη Λαμία και στη Λειβαδιά και κατέγραψε στις τρεις αυτές πόλεις τη μικρότερη τιμή χ^2 . Η επαναλαμβανόμενη υπεροχή της δείχνει ότι μπορεί να αποδώσει αποτελεσματικά χρονοσειρές στις οποίες η πλειονότητα των παρατηρήσεων συγκεντρώνεται σε χαμηλές και μεσαίες τιμές, ενώ οι υψηλές συγκεντρώσεις εμφανίζονται με μικρότερη συχνότητα και δημιουργούν δεξιά ουρά.

Το βασικό πλεονέκτημα της λογαριθμοκανονικής δεν είναι ότι αποδίδει τέλεια κάθε επιμέρους τμήμα της κατανομής, αλλά ότι παρουσιάζει τη σταθερότερη ισορροπία μεταξύ των διαφορετικών κριτηρίων. Στις περισσότερες περιπτώσεις συνδύασε χαμηλή συνολική αριθμητική απόκλιση, καλή συμφωνία στις βασικές κλάσεις και ικανοποιητική συμπεριφορά στα διαγνωστικά διαγράμματα. Παρότι στα Q-Q plots εμφανίστηκαν αποκλίσεις στα υψηλότερα ποσοστημόρια, οι αστοχίες αυτές δεν αντέστρεψαν τη συνολικά ευνοϊκή της εικόνα.

Η σύγκριση αναδεικνύει επίσης ότι η έννοια της «καλύτερης προσαρμογής» δεν είναι μονοδιάστατη. Η κατανομή με τη χαμηλότερη τιμή χ^2 δεν είναι αναγκαστικά εκείνη που εμφανίζει την καλύτερη ευθυγράμμιση στα Q-Q plots ή τη μικρότερη τοπική απόκλιση σε κάθε κλάση. Τα διαφορετικά εργαλεία αξιολογούν διαφορετικές ιδιότητες: το χ^2 τη συνολική συμφωνία των συχνοτήτων, τα residual plots τη δομή των σφαλμάτων και τα Q-Q plots την αντιστοίχιση των ποσοστημορίων και της ουράς.

Υπό αυτό το πρίσμα, η επικράτηση της Lognormal σε τρεις πόλεις δεν πρέπει να ερμηνευθεί ως καθολικός κανόνας για όλες τις χρονοσειρές PM_{10} . Αντίθετα, αποτελεί ισχυρό εμπειρικό πρότυπο μέσα στο συγκεκριμένο σύνολο δεδομένων. Η σαφής διαφοροποίηση της Χαλκίδας επιβεβαιώνει ότι η μορφή της εμπειρικής κατανομής μπορεί να μεταβάλλεται ουσιαστικά από πόλη σε πόλη και να ευνοεί διαφορετικό θεωρητικό μοντέλο.

Συνολικά, τα αποτελέσματα οδηγούν σε τρία κεντρικά συμπεράσματα: η Normal παρουσιάζει σταθερά τη χαμηλότερη καταλληλότητα, οι ασύμμετρες κατανομές αποτελούν σαφώς αποτελεσματικότερες επιλογές και η Lognormal εμφανίζει τη μεγαλύτερη συνολική συνέπεια χωρίς να υπερέχει καθολικά. Η τελική επιλογή πρέπει, επομένως, να πραγματοποιείται χωριστά για κάθε χρονοσειρά και να βασίζεται στη σύνθεση αριθμητικών και γραφικών ενδείξεων.

4.9.3 Διαφοροποίηση της Χαλκίδας και σημασία της πόλης

Η Χαλκίδα αποτελεί τη μοναδική αναλύσιμη περίπτωση στην οποία η λογαριθμοκανονική κατανομή δεν αναδείχθηκε ως καταλληλότερο μοντέλο. Η επικράτηση της Weibull διαφοροποιεί ουσιαστικά τη συγκεκριμένη χρονοσειρά από εκείνες της Άμφισσας, της Λαμίας και της Λειβαδιάς και δείχνει ότι η συχνότερη υπεροχή ενός θεωρητικού μοντέλου δεν συνεπάγεται καθολική εφαρμογή του.

Η διαφοροποίηση αυτή συνδέεται με την ιδιαίτερη εμπειρική εικόνα της Χαλκίδας. Η πόλη παρουσίασε την υψηλότερη μέση συγκέντρωση, τη μεγαλύτερη τυπική απόκλιση και τη μεγαλύτερη μέγιστη τιμή μεταξύ των τεσσάρων αναλύσιμων χρονοσειρών. Παράλληλα, καταγράφηκαν πολύ υψηλές τιμές ασυμμετρίας και κύρτωσης, οι οποίες φανερώνουν έντονη συγκέντρωση της εμπειρικής μάζας σε χαμηλότερες και μεσαίες τιμές, σε συνδυασμό με περιορισμένο αριθμό εξαιρετικά υψηλών παρατηρήσεων. Επομένως, η μορφή της κατανομής της δεν διαφοροποιείται μόνο ποσοτικά, αλλά και ως προς τον τρόπο με τον οποίο κατανέμονται οι συχνότητες μεταξύ του κύριου σώματος και της δεξιάς ουράς.

Η Weibull ανταποκρίθηκε ιδιαίτερα αποτελεσματικά στη συχνοτική αυτή δομή. Στα ραβδογράμματα παρουσίασε την πιο ισορροπημένη αντιστοίχιση στις περισσότερες κλάσεις, ενώ στα observed-expected διαγράμματα εμφάνισε την περισσότερο συνεκτική σχέση μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων. Αντίστοιχα, στα residual plots οι μεγαλύτερες αστοχίες περιορίστηκαν στο αρχικό τμήμα και τα υπόλοιπα σταθεροποιήθηκαν γρήγορα γύρω από το μηδέν.

Η υπεροχή της αποτυπώθηκε με ακόμη μεγαλύτερη σαφήνεια στην αριθμητική αξιολόγηση. Η Weibull κατέγραψε στη Χαλκίδα: $\chi^2=46,09$

έναντι 234,53 για τη λογαριθμοκανονική, 591,19 για τη Gamma και 2.062,81 για την κανονική κατανομή. Η διαφορά από τη δεύτερη καλύτερη επίδοση είναι μεγάλη και δείχνει ότι η επιλογή της Weibull δεν προέκυψε από μια οριακή αριθμητική υπεροχή, αλλά από σαφώς μικρότερη συνολική απόκλιση στις εξεταζόμενες κλάσεις.

Η εικόνα αυτή συνοδεύεται, ωστόσο, από μια ουσιαστική επιφύλαξη. Στο Q-Q διάγραμμα η Weibull παρουσίασε έντονη ανοδική απόκλιση στα υψηλότερα ποσοστημόρια, γεγονός που υποδηλώνει υποεκτίμηση των μεγαλύτερων συγκεντρώσεων. Συνεπώς, η καλή συνολική συμφωνία των συχνοτήτων δεν σημαίνει ότι το μοντέλο αναπαράγει πλήρως την ακραία δεξιά ουρά. Η επιλογή της βασίζεται στην καλύτερη συνολική ισορροπία μεταξύ των διαθέσιμων κριτηρίων και όχι σε υπεροχή σε κάθε επιμέρους πτυχή της προσαρμογής.

Ακριβώς για τον λόγο αυτό, η Χαλκίδα έχει ιδιαίτερη σημασία για την ερμηνεία των συνολικών αποτελεσμάτων. Αν η κατανομή επιλεγόταν μηχανικά με βάση τη συχνότερη επικράτηση στις υπόλοιπες πόλεις, η λογαριθμοκανονική θα εφαρμοζόταν και εδώ. Μια τέτοια επιλογή, όμως, θα αγνοούσε τόσο τη σαφή αριθμητική υπεροχή της Weibull όσο και την καλύτερη συμπεριφορά της στα περισσότερα γραφικά κριτήρια που σχετίζονται με τις συχνότητες.

Η περίπτωση της Χαλκίδας λειτουργεί, επομένως, ως έλεγχος της ίδιας της μεθοδολογικής προσέγγισης. Δείχνει ότι η διαδικασία αξιολόγησης ήταν ικανή να αναγνωρίσει μια πραγματική διαφοροποίηση και δεν οδήγησε εκ των προτέρων στην επιλογή του ίδιου μοντέλου για όλες τις πόλεις. Η ύπαρξη διαφορετικής τελικής επιλογής ενισχύει, αντί να αποδυναμώνει, την αξιοπιστία της ανάλυσης, επειδή αποδεικνύει ότι τα κριτήρια εφαρμόστηκαν ουσιαστικά και όχι τυπικά.

Παράλληλα, τα αποτελέσματα παρέχουν ένδειξη χωρικής ετερογένειας στη στατιστική συμπεριφορά των PM_{10} . Η έννοια αυτή αφορά τη διαφορετική μορφή των εμπειρικών κατανομών και δεν πρέπει, με βάση τα δεδομένα της παρούσας ανάλυσης, να μετατρέπεται σε βέβαιο συμπέρασμα για τους μηχανισμούς που την προκάλεσαν. Η διερεύνηση πιθανών διαφορών στις πηγές εκπομπών, στις μετεωρολογικές συνθήκες ή στην εποχικότητα θα απαιτούσε πρόσθετες μεταβλητές και ειδικότερη ανάλυση.

Επομένως, η Χαλκίδα δεν αποτελεί απλώς μια εξαίρεση από την επικράτηση της λογαριθμοκανονικής κατανομής. Αποτελεί την περίπτωση που αναδεικνύει σαφέστερα το βασικό μεθοδολογικό συμπέρασμα της εργασίας: κοινά χαρακτηριστικά, όπως η μη αρνητικότητα, η θετική ασυμμετρία και η δεξιά ουρά, δεν αρκούν για να επιβάλουν την ίδια θεωρητική κατανομή σε όλες τις χρονοσειρές. Η τελική επιλογή πρέπει να πραγματοποιείται χωριστά για κάθε πόλη και έτος, με βάση τη σύνθεση της εμπειρικής μορφής, των γραφικών διαγνωστικών και της αριθμητικής αξιολόγησης.

4.9.4 Συμπέρασμα συγκριτικής αξιολόγησης

Συνολικά, η συγκριτική αξιολόγηση των τεσσάρων θεωρητικών κατανομών καταδεικνύει ότι οι εξεταζόμενες ετήσιες χρονοσειρές των ωριαίων συγκεντρώσεων PM_{10} χαρακτηρίζονται από θετική ασυμμετρία, παρουσία υψηλών και ακραίων τιμών και σαφή

απόκλιση από την κανονικότητα. Για τον λόγο αυτό, η κανονική κατανομή δεν αποδείχθηκε κατάλληλη για την περιγραφή των εμπειρικών κατανομών, καθώς η συμμετρική της μορφή δεν απέδωσε ικανοποιητικά ούτε το κύριο σώμα των δεδομένων ούτε τη δεξιά ουρά.

Η αδυναμία της Normal επιβεβαιώθηκε σε όλα σχεδόν τα διαγνωστικά στάδια. Στη σύγκριση παρατηρούμενων και αναμενόμενων συχνοτήτων παρουσίασε μεγάλες και συστηματικές αποκλίσεις, στα observed-expected διαγράμματα τα σημεία απομακρύνθηκαν αισθητά από τη γραμμή ιδανικής συμφωνίας, ενώ στα διαγράμματα υπολοίπων εμφανίστηκαν οργανωμένα κυματοειδή πρότυπα. Αντίστοιχα, τα Q-Q plots ανέδειξαν έντονη ασυμφωνία στα υψηλότερα ποσοστημόρια. Οι πολύ υψηλές τιμές του στατιστικού χ^2 σε όλες τις πόλεις επιβεβαίωσαν αριθμητικά τη χαμηλή σχετική καταλληλότητά της.

Αντίθετα, οι ασύμμετρες κατανομές παρουσίασαν σαφώς καλύτερη προσαρμογή. Η Gamma, η Weibull και η λογαριθμοκανονική είναι περισσότερο συμβατές με μη αρνητικά και θετικά ασύμμετρα δεδομένα, γεγονός που αντανακλάται τόσο στις χαμηλότερες τιμές χ^2 όσο και στη βελτιωμένη εικόνα των διαγνωστικών γραφημάτων. Ωστόσο, η απόδοσή τους δεν ήταν ίδια σε όλες τις χρονοσειρές.

Η λογαριθμοκανονική αποτέλεσε την επικρατέστερη θεωρητική προσέγγιση, καθώς επιλέχθηκε στην Άμφισσα, στη Λαμία και στη Λειβαδιά. Στις τρεις αυτές περιπτώσεις κατέγραψε τη χαμηλότερη τιμή χ^2 και παρουσίασε γενικά ισχυρή συμπεριφορά στα περισσότερα γραφικά κριτήρια. Η επαναλαμβανόμενη επιλογή της δείχνει ότι αποτελεί αποτελεσματικό μοντέλο για χρονοσειρές στις οποίες η πλειονότητα των παρατηρήσεων συγκεντρώνεται σε χαμηλές και μεσαίες τιμές, ενώ οι υψηλότερες συγκεντρώσεις εμφανίζονται με μικρότερη συχνότητα και δημιουργούν δεξιά ουρά.

Η υπεροχή της Lognormal δεν ήταν απόλυτη σε κάθε επιμέρους διάγραμμα. Σε ορισμένες πόλεις η Weibull ή η Gamma παρουσίασαν καλύτερη εικόνα σε συγκεκριμένα διαγνωστικά εργαλεία, όπως στα διαγράμματα υπολοίπων ή στα Q-Q plots. Παρ' όλα αυτά, η λογαριθμοκανονική εμφάνισε τη μεγαλύτερη συνολική ισορροπία μεταξύ αριθμητικής προσαρμογής, συμφωνίας των συχνοτήτων και συμπεριφοράς των ποσοστημορίων.

Η Weibull αποτέλεσε τη σημαντικότερη εναλλακτική και αναδείχθηκε ως η καταλληλότερη κατανομή για τη Χαλκίδα. Στη συγκεκριμένη χρονοσειρά παρουσίασε ιδιαίτερα χαμηλή τιμή χ^2 και σαφή υπεροχή στα διαγράμματα συχνοτήτων και υπολοίπων.

Το αποτέλεσμα αυτό δείχνει ότι η εμπειρική δομή της Χαλκίδας αποδίδεται αποτελεσματικότερα από τη Weibull, παρά το γεγονός ότι η Lognormal επικράτησε στις υπόλοιπες τρεις πόλεις.

Η περίπτωση της Χαλκίδας είναι ουσιαστική για την ερμηνεία των συνολικών αποτελεσμάτων. Αποδεικνύει ότι η συχνότερη επιλογή μιας κατανομής δεν συνεπάγεται καθολική καταλληλότητα. Παρότι όλες οι χρονοσειρές παρουσιάζουν θετική ασυμμετρία και δεξιά ουρά, διαφοροποιούνται ως προς τη διασπορά, την κύρτωση, την ένταση των υψηλών παρατηρήσεων και την κατανομή της κύριας εμπειρικής μάζας. Οι διαφορές αυτές μπορούν να μεταβάλλουν ουσιαστικά τη σχετική απόδοση των θεωρητικών μοντέλων.

Η Gamma παρουσίασε σταθερή βελτίωση έναντι της Normal, χωρίς όμως να αναδειχθεί ως τελική επιλογή σε κάποια πόλη. Σε αρκετές περιπτώσεις απέδωσε ικανοποιητικά σημαντικό μέρος των ποσοστημορίων ή της κύριας μάζας των παρατηρήσεων, αλλά εμφάνισε λιγότερο ομοιόμορφη συμπεριφορά στις συχνότητες και στα υπόλοιπα. Επομένως, αποτελεί αξιόλογη εναλλακτική ασύμμετρη κατανομή, αλλά όχι την περισσότερο συνεπή για τα συγκεκριμένα δεδομένα.

Το συνολικό συμπέρασμα είναι ότι δεν υπάρχει μία μοναδική θεωρητική κατανομή που να μπορεί να θεωρηθεί καθολικά βέλτιστη για όλες τις χρονοσειρές PM₁₀. Η Lognormal παρουσιάζει τη μεγαλύτερη συνολική συνέπεια και μπορεί να θεωρηθεί το επικρατέστερο μοντέλο της παρούσας ανάλυσης. Ωστόσο, η διαφοροποίηση της Χαλκίδας δείχνει ότι η τελική επιλογή πρέπει να πραγματοποιείται χωριστά για κάθε πόλη και έτος.

Η πολυκριτηριακή προσέγγιση ενισχύει την αξιοπιστία της αξιολόγησης, καθώς αποφεύγει τη μηχανική επιλογή μοντέλου με βάση έναν μόνο δείκτη. Το στατιστικό χ^2 συνοψίζει τη συνολική απόκλιση των συχνοτήτων, τα observed-expected διαγράμματα αποτυπώνουν τη γενική σχέση μεταξύ εμπειρικών και θεωρητικών τιμών, τα residual plots αναδεικνύουν τη δομή και τη συστηματικότητα των σφαλμάτων και τα Q-Q plots εξετάζουν την αντιστοίχιση των ποσοστημορίων, ιδίως προς την άνω ουρά. Η σύνθεση αυτών των εργαλείων παρέχει πληρέστερη εικόνα από οποιοδήποτε μεμονωμένο κριτήριο. Επομένως, η λογαριθμοκανονική κατανομή αποτελεί την επικρατέστερη θεωρητική προσέγγιση για τις περισσότερες αναλύσιμες χρονοσειρές της παρούσας εργασίας, ενώ η Weibull αναδεικνύεται ως σημαντική εναλλακτική και ως η καταλληλότερη επιλογή για τη Χαλκίδα. Το βασικό μεθοδολογικό συμπέρασμα είναι ότι η επιλογή θεωρητικής

κατανομής πρέπει να αντιμετωπίζεται ως διαδικασία αξιολόγησης ανά χρονοσειρά και όχι ως καθολική εφαρμογή ενός ενιαίου μοντέλου.

Τέλος, τα συμπεράσματα αφορούν τις συγκεκριμένες πόλεις και τα επιλεγμένα έτη που πληρούσαν τα κριτήρια πληρότητας και χρονικής συνέχειας. Δεν πρέπει, επομένως, να γενικεύονται αυτομάτως στη μακροχρόνια κατανομική συμπεριφορά των συγκεντρώσεων PM_{10} σε ολόκληρη τη Στερεά Ελλάδα. Μια ευρύτερη γενίκευση θα απαιτούσε περισσότερες πλήρεις ετήσιες χρονοσειρές και επανάληψη της ίδιας διαδικασίας σε διαφορετικά έτη. Εντός, όμως, του διαθέσιμου αναλυτικού πλαισίου, η Lognormal προσφέρει τη μεγαλύτερη συνολική συνέπεια, ενώ η Weibull τεκμηριώνεται ως η σημαντικότερη εναλλακτική όταν η εμπειρική μορφή των δεδομένων το απαιτεί.

5. Συμπεράσματα

Στο παρόν κεφάλαιο συντίθενται τα βασικά συμπεράσματα της διπλωματικής εργασίας, όπως προέκυψαν από τη στατιστική διερεύνηση των ωριαίων συγκεντρώσεων PM_{10} στις εξεταζόμενες πόλεις της Στερεάς Ελλάδας. Η παρουσίαση δεν επαναλαμβάνει αναλυτικά τα αποτελέσματα του προηγούμενου κεφαλαίου, αλλά τα ερμηνεύει σε σχέση με τον αρχικό σκοπό της έρευνας, τα επιμέρους ερευνητικά ερωτήματα και τη μεθοδολογική προσέγγιση που εφαρμόστηκε.

Το ενδιαφέρον επικεντρώνεται στη μορφή των εμπειρικών κατανομών των συγκεντρώσεων PM_{10} και στη συγκριτική καταλληλότητα των θεωρητικών μοντέλων που εξετάστηκαν. Παράλληλα, αξιολογείται η σημασία της πληρότητας και της χρονικής συνέχειας των δεδομένων, καθώς και ο ρόλος χαρακτηριστικών όπως η θετική ασυμμετρία, η αυξημένη κύρτωση και η παρουσία σπάνιων υψηλών τιμών στη διαδικασία επιλογής της καταλληλότερης κατανομής.

Ιδιαίτερη έμφαση δίνεται στη συνδυαστική χρήση αριθμητικών και γραφικών διαγνωστικών κριτηρίων. Η επιλογή ενός θεωρητικού μοντέλου δεν αντιμετωπίστηκε ως μηχανική διαδικασία που βασίζεται σε έναν μόνο δείκτη, αλλά ως σύνθεση διαφορετικών μορφών πληροφορίας σχετικά με τη συμφωνία των συχνοτήτων, τη δομή των αποκλίσεων και την αντιστοίχιση των ποσοστημορίων.

Αρχικά παρουσιάζεται συνοπτικά η ερευνητική διαδικασία που ακολουθήθηκε. Στη συνέχεια διατυπώνονται τα βασικά ευρήματα και απαντώνται τα ερευνητικά ερωτήματα της εργασίας. Τέλος, αξιολογείται η μεθοδολογική συνεισφορά της προσέγγισης, επισημαίνονται οι βασικοί περιορισμοί και προτείνονται κατευθύνσεις για την περαιτέρω διερεύνηση της κατανομικής συμπεριφοράς των συγκεντρώσεων PM_{10} .

5.1 Σύνοψη της εργασίας

Η παρούσα διπλωματική εργασία διερεύνησε τη στατιστική συμπεριφορά των ωριαίων συγκεντρώσεων PM_{10} σε επιλεγμένες πόλεις της Στερεάς Ελλάδας, με βασικό στόχο τον προσδιορισμό της θεωρητικής κατανομής πιθανότητας που προσεγγίζει καταλληλότερα την εμπειρική μορφή των διαθέσιμων δεδομένων. Η επιλογή των PM_{10} ως αντικειμένου μελέτης συνδέεται τόσο με τη σημασία τους για την αξιολόγηση της ποιότητας του αέρα όσο και με την ιδιαίτερη στατιστική συμπεριφορά των συγκεντρώσεών τους, οι οποίες

συχνά χαρακτηρίζονται από μη αρνητικές τιμές, θετική ασυμμετρία και περιορισμένο αριθμό πολύ υψηλών παρατηρήσεων.

Η έρευνα δεν περιορίστηκε στη σύγκριση μέσων τιμών ή άλλων συνοπτικών στατιστικών δεικτών. Επικεντρώθηκε στη συνολική κατανομική δομή των δεδομένων, ώστε να εξεταστούν χαρακτηριστικά όπως η θέση της κύριας εμπειρικής μάζας, η διασπορά, η ασυμμετρία, η κύρτωση και η μορφή της δεξιάς ουράς. Η προσέγγιση αυτή παρείχε πληρέστερη εικόνα από εκείνη που θα προέκυπτε αποκλειστικά από τους δείκτες κεντρικής τάσης και μεταβλητότητας.

Η ανάλυση βασίστηκε σε ωριαίες μετρήσεις από σταθμούς παρακολούθησης της ποιότητας του αέρα στην Άμφισσα, στη Λαμία, στη Λειβαδιά, στη Χαλκίδα και στο Καρπενήσι. Πριν από οποιαδήποτε κατανομική προσαρμογή πραγματοποιήθηκε έλεγχος του αριθμού των έγκυρων παρατηρήσεων, του ποσοστού πληρότητας και της χρονικής συνέχειας κάθε διαθέσιμης ετήσιας χρονοσειράς. Το στάδιο αυτό αποτέλεσε ουσιαστικό μέρος της μεθοδολογίας, καθώς η εφαρμογή θεωρητικών κατανομών σε ελλιπή ή χρονικά ασυνεχή δεδομένα θα μπορούσε να οδηγήσει σε μη αντιπροσωπευτική εμπειρική εικόνα.

Ως εσωτερικό κριτήριο ένταξης στην τελική ανάλυση χρησιμοποιήθηκε ποσοστό πληρότητας τουλάχιστον 90%, σε συνδυασμό με ικανοποιητική χρονική συνέχεια. Με βάση τα κριτήρια αυτά επιλέχθηκαν οι χρονοσειρές της Άμφισσας για το 2018, της Λαμίας για το 2017, της Λειβαδιάς για το 2019 και της Χαλκίδας για το 2017. Για το Καρπενήσι δεν εντοπίστηκε έτος που να πληροί τις απαιτούμενες προϋποθέσεις και, επομένως, ο σταθμός δεν συμπεριλήφθηκε στη διαδικασία προσαρμογής. Η περίπτωση αυτή παρουσιάστηκε ξεχωριστά, ώστε η μη ένταξή της να τεκμηριωθεί ως αποτέλεσμα του ίδιου ελέγχου ποιότητας που εφαρμόστηκε σε όλες τις πόλεις.

Στις τέσσερις αναλύσιμες χρονοσειρές εξετάστηκαν οι κατανομές Normal, Gamma, Weibull και Lognormal. Η Normal χρησιμοποιήθηκε ως συμμετρικό μοντέλο αναφοράς, ενώ οι Gamma, Weibull και Lognormal επιλέχθηκαν ως ασύμμετρες κατανομές κατάλληλες για μη αρνητικές μεταβλητές. Η σύγκρισή τους επέτρεψε να διερευνηθεί όχι μόνο αν η κανονική κατανομή είναι ανεπαρκής, αλλά και ποιο από τα ασύμμετρα μοντέλα προσαρμόζεται αποτελεσματικότερα στην ιδιαίτερη δομή κάθε χρονοσειράς.

Η αξιολόγηση πραγματοποιήθηκε σε διαδοχικά και συμπληρωματικά στάδια. Αρχικά συγκρίθηκαν οι παρατηρούμενες και οι θεωρητικά αναμενόμενες συχνότητες στις προκαθορισμένες κλάσεις συγκέντρωσης. Στη συνέχεια εξετάστηκαν τα διαγνωστικά διαγράμματα observed-expected, τα διαγράμματα υπολοίπων και τα Q-Q διαγράμματα. Η

αριθμητική αξιολόγηση συμπληρώθηκε με το στατιστικό χ^2 , το οποίο χρησιμοποιήθηκε ως συγκριτικός δείκτης της συνολικής απόκλισης μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων.

Κάθε διαγνωστικό εργαλείο αξιολόγησε διαφορετική πλευρά της προσαρμογής. Η σύγκριση ανά κλάση ανέδειξε τον τρόπο με τον οποίο κάθε μοντέλο κατανέμει τη θεωρητική πιθανότητα στις βασικές περιοχές συγκέντρωσης. Τα observed-expected διαγράμματα αποτύπωσαν τη συνολική σχέση μεταξύ εμπειρικών και θεωρητικών συχνοτήτων, ενώ τα residual plots κατέγραψαν τη θέση, την κατεύθυνση και τη διάρκεια των αποκλίσεων. Τα Q-Q διαγράμματα επέτρεψαν την αξιολόγηση της αντιστοίχισης των θεωρητικών και εμπειρικών ποσοστημορίων, με ιδιαίτερη έμφαση στο ανώτερο τμήμα της κατανομής.

Η χρήση πολλαπλών διαγνωστικών κριτηρίων κρίθηκε απαραίτητη, επειδή μια θεωρητική κατανομή μπορεί να εμφανίζει καλή συνολική συμφωνία στις συχνότητες, αλλά να αποκλίνει στα υψηλότερα ποσοστημόρια ή σε συγκεκριμένες περιοχές της εμπειρικής κατανομής. Για τον λόγο αυτό, η τελική επιλογή δεν στηρίχθηκε αποκλειστικά στη χαμηλότερη τιμή χ^2 , αλλά στη σύνθεση της αριθμητικής κατάταξης με τα γραφικά ευρήματα.

Συνολικά, η ερευνητική διαδικασία οργανώθηκε σε τρία κύρια επίπεδα: έλεγχο ποιότητας και επιλογή κατάλληλων χρονοσειρών, περιγραφική και εμπειρική διερεύνηση των δεδομένων και συγκριτική προσαρμογή θεωρητικών κατανομών. Η διαδοχική αυτή προσέγγιση επέτρεψε την αξιολόγηση της κατανομικής συμπεριφοράς των PM_{10} με ενιαία μεθοδολογία, χωρίς να παραβλέπονται οι διαφοροποιήσεις που εμφανίζονται μεταξύ των πόλεων και των επιλεγμένων ετών.

5.2 Βασικά ευρήματα της ανάλυσης

Η ανάλυση ανέδειξε, καταρχάς, ότι η ποιότητα των πρωτογενών δεδομένων αποτελεί αναγκαία προϋπόθεση για την αξιόπιστη εφαρμογή θεωρητικών κατανομών. Ο έλεγχος πληρότητας και χρονικής συνέχειας οδήγησε στην επιλογή τεσσάρων ετήσιων χρονοσειρών: της Άμφισσας για το 2018, της Λαμίας για το 2017, της Λειβαδιάς για το 2019 και της Χαλκίδας για το 2017. Αντίθετα, για το Καρπενήσι δεν εντοπίστηκε έτος που

να ικανοποιεί το προκαθορισμένο όριο πληρότητας και τις απαιτήσεις χρονικής κάλυψης. Η περίπτωση αυτή επιβεβαιώνει ότι η λειτουργία ενός σταθμού και η ύπαρξη ορισμένων μετρήσεων δεν εξασφαλίζουν από μόνες τους την καταλληλότητα των δεδομένων για ετήσια κατανομική ανάλυση.

Η περιγραφική διερεύνηση έδειξε ότι και οι τέσσερις αναλύσιμες χρονοσειρές παρουσιάζουν σαφή θετική ασυμμετρία. Σε όλες τις πόλεις η μέση τιμή ήταν μεγαλύτερη από τη διάμεσο, ενώ οι θετικές τιμές ασυμμετρίας και η αυξημένη κύρτωση επιβεβαίωσαν την παρουσία δεξιάς ουράς και περιορισμένου αριθμού πολύ υψηλών παρατηρήσεων. Επομένως, οι συγκεντρώσεις δεν κατανέμονται συμμετρικά γύρω από μία κεντρική τιμή, αλλά συγκεντρώνονται κυρίως σε χαμηλά και μεσαία επίπεδα, με σταδιακά μικρότερη συχνότητα προς τις υψηλότερες τιμές.

Οι διαφοροποιήσεις μεταξύ των πόλεων ήταν επίσης ουσιαστικές. Η Άμφισσα, η Λαμία και η Λειβαδιά εμφάνισαν σχετικά κοντινές μέσες τιμές, αλλά διέφεραν ως προς τη μεταβλητότητα και την ένταση της δεξιάς ουράς. Η Χαλκίδα ξεχώρισε σαφώς, καθώς παρουσίασε την υψηλότερη μέση συγκέντρωση, τη μεγαλύτερη τυπική απόκλιση, τη μεγαλύτερη μέγιστη τιμή και τις υψηλότερες τιμές ασυμμετρίας και κύρτωσης. Η εικόνα αυτή δείχνει ότι, παρά τα κοινά βασικά χαρακτηριστικά, οι εμπειρικές κατανομές δεν έχουν ταυτόσημη μορφή.

Η κανονική κατανομή παρουσίασε τη χαμηλότερη ποιότητα προσαρμογής σε όλες τις αναλύσιμες χρονοσειρές. Η αδυναμία της αποτυπώθηκε τόσο στις υψηλές τιμές του στατιστικού χ^2 όσο και στα γραφικά διαγνωστικά εργαλεία. Η συμμετρική της μορφή δεν απέδωσε ικανοποιητικά τη συγκέντρωση των παρατηρήσεων στις χαμηλομεσαίες τιμές, τη θετική ασυμμετρία και την εκτεταμένη δεξιά ουρά. Το αποτέλεσμα αυτό ήταν σταθερό σε όλες τις πόλεις και δεν εξαρτήθηκε από ένα μόνο κριτήριο αξιολόγησης.

Οι Gamma, Weibull και Lognormal παρουσίασαν αισθητά καλύτερη συμπεριφορά, επιβεβαιώνοντας τη μεγαλύτερη καταλληλότητα των ασύμμετρων μοντέλων για μη αρνητικά δεδομένα αυτού του τύπου. Ωστόσο, κανένα από τα τρία δεν υπερείχε σε κάθε διαγνωστικό εργαλείο και σε κάθε πόλη. Σε αρκετές περιπτώσεις η κατανομή που παρουσίαζε την καλύτερη αντιστοίχιση των συχνοτήτων δεν ήταν εκείνη που απέδιδε καλύτερα τα υψηλότερα ποσοστημόρια. Η διαφοροποίηση αυτή κατέδειξε ότι η προσαρμογή αποτελεί πολυδιάστατη έννοια και δεν μπορεί να αποτιμηθεί πλήρως με έναν μόνο αριθμητικό ή γραφικό δείκτη.

Η λογαριθμοκανονική αναδείχθηκε ως η επικρατέστερη κατανομή, καθώς επιλέχθηκε για την Άμφισσα, τη Λαμία και τη Λειβαδιά. Στις τρεις αυτές χρονοσειρές κατέγραψε τη χαμηλότερη τιμή χ^2 και παρουσίασε συνολικά ισορροπημένη συμπεριφορά στα επιμέρους διαγνωστικά στάδια. Το βασικό της πλεονέκτημα δεν ήταν η τέλεια προσαρμογή σε κάθε τμήμα των δεδομένων, αλλά η μεγαλύτερη συνολική συνέπεια μεταξύ συμφωνίας των συχνοτήτων, δομής των αποκλίσεων και αντιστοίχισης των ποσοστημορίων.

Η Χαλκίδα αποτέλεσε τη σημαντικότερη διαφοροποίηση. Στη συγκεκριμένη χρονοσειρά η Weibull κατέγραψε με μεγάλη διαφορά τη χαμηλότερη τιμή χ^2 και παρουσίασε την πιο συνεκτική εικόνα στη σύγκριση των παρατηρούμενων και αναμενόμενων συχνοτήτων. Παρότι το Q-Q διάγραμμά της ανέδειξε αδυναμία στα υψηλότερα ποσοστημόρια, η συνολική σύνθεση των αποτελεσμάτων την κατέστησε την καταλληλότερη από τις τέσσερις εξεταζόμενες επιλογές για τη Χαλκίδα.

Ένα ακόμη σημαντικό εύρημα είναι ότι καμία από τις ασύμμετρες κατανομές δεν απέδωσε πλήρως τις πλέον ακραίες παρατηρήσεις σε όλες τις πόλεις. Τα Q-Q διαγράμματα εμφάνισαν συχνά αποκλίσεις στο ανώτερο τμήμα, ακόμη και όταν η συνολική συμφωνία των συχνοτήτων ήταν ικανοποιητική. Το γεγονός αυτό δείχνει ότι η περιγραφή της κύριας μάζας των δεδομένων και η αναπαράσταση της ακραίας δεξιάς ουράς αποτελούν διαφορετικές απαιτήσεις, οι οποίες δεν ικανοποιούνται πάντοτε στον ίδιο βαθμό από ένα ενιαίο θεωρητικό μοντέλο.

Συνολικά, τα ευρήματα υποστηρίζουν ότι η Lognormal αποτελεί το πιο σταθερό θεωρητικό πρότυπο για τις περισσότερες χρονοσειρές της παρούσας εργασίας, χωρίς όμως να συνιστά καθολική λύση. Η επικράτηση της Weibull στη Χαλκίδα καταδεικνύει ότι η επιλογή πρέπει να πραγματοποιείται χωριστά για κάθε πόλη και έτος, με βάση την εμπειρική μορφή των δεδομένων και τη συνδυαστική αξιολόγηση όλων των διαθέσιμων κριτηρίων.

5.3 Απάντηση στα ερευνητικά ερωτήματα

Η παρούσα εργασία οργανώθηκε γύρω από τέσσερα βασικά ερευνητικά ερωτήματα, τα οποία αφορούσαν την καταλληλότητα των διαθέσιμων χρονοσειρών, τα κύρια στατιστικά χαρακτηριστικά των συγκεντρώσεων PM_{10} , την επιλογή της θεωρητικής κατανομής με την

καλύτερη προσαρμογή και τη διερεύνηση διαφοροποιήσεων μεταξύ των εξεταζόμενων πόλεων. Τα αποτελέσματα της ανάλυσης επιτρέπουν την απάντηση των ερωτημάτων αυτών με συνεκτικό και τεκμηριωμένο τρόπο.

Ως προς την καταλληλότητα των διαθέσιμων χρονοσειρών, ο έλεγχος πληρότητας και χρονικής συνέχειας έδειξε ότι μόνο τέσσερις ετήσιες σειρές μπορούσαν να ενταχθούν με αξιοπιστία στην τελική στατιστική ανάλυση. Συγκεκριμένα, επιλέχθηκαν η Άμφισσα για το 2018, η Λαμία για το 2017, η Λειβαδιά για το 2019 και η Χαλκίδα για το 2017. Οι συγκεκριμένες χρονοσειρές πληρούσαν τα προκαθορισμένα κριτήρια ποιότητας και παρείχαν επαρκή βάση για την εφαρμογή των θεωρητικών κατανομών. Αντίθετα, το Καρπενήσι δεν εντάχθηκε στην κατανομική ανάλυση, καθώς κανένα από τα διαθέσιμα έτη δεν παρουσίαζε το απαιτούμενο επίπεδο πληρότητας. Το εύρημα αυτό επιβεβαιώνει ότι η στατιστική αξιοποίηση των δεδομένων προϋποθέτει συστηματικό έλεγχο ποιότητας και δεν μπορεί να βασίζεται αποκλειστικά στην ύπαρξη διαθέσιμων μετρήσεων.

Σε σχέση με τα βασικά στατιστικά χαρακτηριστικά των συγκεντρώσεων PM_{10} , οι αναλύσιμες χρονοσειρές παρουσίασαν κοινή γενική εικόνα θετικής ασυμμετρίας και αυξημένης κύρτωσης. Σε όλες τις πόλεις η μέση τιμή ήταν μεγαλύτερη από τη διάμεσο, γεγονός που δείχνει ότι η κατανομή επηρεάζεται από υψηλότερες παρατηρήσεις οι οποίες εμφανίζονται με μικρότερη συχνότητα. Παράλληλα, οι υψηλές τιμές ασυμμετρίας και κύρτωσης επιβεβαίωσαν ότι οι εμπειρικές κατανομές αποκλίνουν ουσιαστικά από τη συμμετρική μορφή. Οι περισσότερες παρατηρήσεις συγκεντρώθηκαν σε χαμηλές και μεσαίες τιμές, ενώ οι υψηλότερες συγκεντρώσεις σχημάτισαν δεξιά ουρά.

Παρά το κοινό αυτό πρότυπο, καταγράφηκαν ουσιαστικές διαφοροποιήσεις μεταξύ των πόλεων. Η Χαλκίδα παρουσίασε την υψηλότερη μέση συγκέντρωση, τη μεγαλύτερη διασπορά και τις εντονότερες τιμές ασυμμετρίας και κύρτωσης, ενώ η Λειβαδιά εμφάνισε επίσης ιδιαίτερα υψηλή κύρτωση και πολύ μεγάλη μέγιστη τιμή. Τα αποτελέσματα αυτά δείχνουν ότι οι χρονοσειρές δεν διαφέρουν μόνο ως προς το επίπεδο των συγκεντρώσεων, αλλά και ως προς τη συνολική μορφή της εμπειρικής κατανομής τους.

Ως προς την επιλογή της θεωρητικής κατανομής με την καλύτερη προσαρμογή, η λογαριθμοκανονική κατανομή αναδείχθηκε ως η επικρατέστερη επιλογή. Συγκεκριμένα, επιλέχθηκε για την Άμφισσα, τη Λαμία και τη Λειβαδιά, καθώς στις τρεις αυτές περιπτώσεις παρουσίασε τη μικρότερη τιμή του στατιστικού χ^2 και συνολικά ισχυρή συμπεριφορά στα γραφικά διαγνωστικά εργαλεία. Η επαναλαμβανόμενη υπεροχή της δείχνει ότι η Lognormal μπορεί να αποδώσει αποτελεσματικά χρονοσειρές στις οποίες η

κύρια μάζα των παρατηρήσεων συγκεντρώνεται σε χαμηλές και μεσαίες τιμές, ενώ οι υψηλότερες συγκεντρώσεις εμφανίζονται αραιότερα.

Η κανονική κατανομή, αντίθετα, δεν αποδείχθηκε κατάλληλη σε καμία από τις εξεταζόμενες χρονοσειρές. Σε όλες τις πόλεις εμφάνισε τις υψηλότερες τιμές χ^2 και τις μεγαλύτερες αποκλίσεις στα διαγνωστικά γραφήματα. Η συμμετρική της μορφή δεν ανταποκρίθηκε στη μη αρνητική και θετικά ασύμμετρη φύση των δεδομένων. Η Gamma παρουσίασε σαφή βελτίωση σε σχέση με τη Normal, χωρίς όμως να αναδειχθεί ως τελική επιλογή σε κάποια πόλη. Η Weibull εμφάνισε ισχυρή συμπεριφορά σε αρκετά διαγνωστικά στάδια και τελικά επικράτησε στη Χαλκίδα.

Η διαφορετική επιλογή για τη Χαλκίδα απαντά και στο ερώτημα αν η καταλληλότερη κατανομή διαφοροποιείται μεταξύ των πόλεων. Παρότι η Lognormal παρουσίασε τη μεγαλύτερη συνολική συνέπεια, η Weibull απέδωσε καλύτερα τη συγκεκριμένη χρονοσειρά και κατέγραψε σαφώς χαμηλότερη τιμή χ^2 από τις υπόλοιπες κατανομές. Το αποτέλεσμα αυτό δείχνει ότι η κατανομική συμπεριφορά των συγκεντρώσεων PM_{10} δεν είναι απολύτως ενιαία και ότι η τελική επιλογή εξαρτάται από τα ιδιαίτερα χαρακτηριστικά κάθε χρονοσειράς.

Η διαφοροποίηση της Χαλκίδας δεν πρέπει να ερμηνευθεί ως αντίφαση προς την επικράτηση της Lognormal στις άλλες τρεις πόλεις. Αντίθετα, επιβεβαιώνει ότι η επιλογή θεωρητικού μοντέλου δεν πρέπει να γίνεται μηχανικά με βάση τη συχνότερη επικράτηση μιας κατανομής. Η ίδια μεθοδολογική διαδικασία πρέπει να εφαρμόζεται χωριστά σε κάθε χρονοσειρά, ώστε να λαμβάνονται υπόψη η διασπορά, η ασυμμετρία, η κύρτωση, η θέση της κύριας εμπειρικής μάζας και η συμπεριφορά της δεξιάς ουράς.

Συνολικά, τα τέσσερα ερευνητικά ερωτήματα απαντήθηκαν μέσα από μια διαδοχική διαδικασία που ξεκίνησε από τον έλεγχο της ποιότητας των δεδομένων και ολοκληρώθηκε με τη συγκριτική αξιολόγηση των θεωρητικών κατανομών. Η ανάλυση έδειξε ότι οι κατάλληλες χρονοσειρές χαρακτηρίζονται από θετική ασυμμετρία και δεξιά ουρά, ότι οι ασύμμετρες κατανομές προσαρμόζονται σαφώς καλύτερα από την Normal και ότι η Lognormal αποτελεί την επικρατέστερη, αλλά όχι καθολικά βέλτιστη, θεωρητική προσέγγιση. Η τελική επιλογή πρέπει, επομένως, να πραγματοποιείται ανά πόλη και έτος, με βάση τη συνδυαστική ερμηνεία αριθμητικών και γραφικών διαγνωστικών κριτηρίων.

5.4 Μεθοδολογική αποτίμηση της προσέγγισης

Η μεθοδολογική προσέγγιση της εργασίας οργανώθηκε σε διαδοχικά στάδια, τα οποία περιέλαβαν τον έλεγχο της ποιότητας των δεδομένων, την περιγραφική διερεύνηση των χρονοσειρών, την εφαρμογή θεωρητικών κατανομών και τη συγκριτική αξιολόγηση της προσαρμογής τους. Η σειρά αυτή ήταν απαραίτητη, καθώς η αξιοπιστία της τελικής επιλογής εξαρτάται τόσο από την ποιότητα των αρχικών δεδομένων όσο και από την απόδοση των επιμέρους μοντέλων.

Ιδιαίτερη σημασία είχε ο έλεγχος πληρότητας και χρονικής συνέχειας. Η χρήση ελλιπών ή αποσπασματικών χρονοσειρών θα μπορούσε να οδηγήσει σε εμπειρικές κατανομές που δεν αποτυπώνουν ικανοποιητικά την ετήσια συμπεριφορά των συγκεντρώσεων PM_{10} . Για τον λόγο αυτό, η επιλογή των αναλύσιμων χρονοσειρών προηγήθηκε της κατανομικής προσαρμογής και βασίστηκε σε κοινά κριτήρια για όλες τις πόλεις.

Η περιγραφική στατιστική ανάλυση λειτούργησε ως ενδιάμεσο στάδιο, παρέχοντας μια πρώτη εικόνα της κεντρικής τάσης, της διασποράς, της ασυμμετρίας και της κύρτωσης. Οι δείκτες αυτοί δεν χρησιμοποιήθηκαν για την εκ των προτέρων επιλογή μιας συγκεκριμένης κατανομής, αλλά ως πλαίσιο για την ερμηνεία των αποτελεσμάτων της προσαρμογής.

Η αριθμητική αξιολόγηση βασίστηκε στο στατιστικό χ^2 , το οποίο χρησιμοποιήθηκε ως συγκριτικός δείκτης της συνολικής απόκλισης μεταξύ παρατηρούμενων και θεωρητικά αναμενόμενων συχνοτήτων. Το χ^2 ήταν χρήσιμο για την κατάταξη των υποψήφιων μοντέλων μέσα στην ίδια χρονοσειρά, χωρίς όμως να αντιμετωπιστεί ως αποκλειστικό κριτήριο. Η τιμή του δεν δείχνει από μόνη της σε ποιο τμήμα της εμπειρικής κατανομής εμφανίζονται οι μεγαλύτερες αποκλίσεις.

Για τον λόγο αυτό, η αριθμητική αξιολόγηση συμπληρώθηκε από γραφικά διαγνωστικά εργαλεία. Τα διαγράμματα παρατηρούμενων και αναμενόμενων συχνοτήτων ανέδειξαν την απόδοση των μοντέλων στις επιμέρους κλάσεις, τα observed-expected διαγράμματα τη γενική συμφωνία των συχνοτήτων και τα residual plots τη φορά και τη συστηματικότητα των σφαλμάτων. Τα Q-Q διαγράμματα επέτρεψαν την αξιολόγηση της συμφωνίας των ποσοστημορίων, ιδίως στο ανώτερο τμήμα της κατανομής.

Η συνδυαστική χρήση των παραπάνω εργαλείων αποτέλεσε βασικό πλεονέκτημα της προσέγγισης. Σε αρκετές περιπτώσεις τα γραφικά ευρήματα επιβεβαίωσαν την κατάταξη του χ^2 , ενώ σε άλλες ανέδειξαν αποκλίσεις στην άνω ουρά ή σε συγκεκριμένες κλάσεις

που δεν ήταν εμφανείς από τη συνολική αριθμητική τιμή. Έτσι, η τελική επιλογή δεν βασίστηκε μηχανικά σε έναν μόνο δείκτη.

Σημαντική ήταν επίσης η σύγκριση μιας συμμετρικής κατανομής, της Normal, με τρεις ασύμμετρες κατανομές, τις Gamma, Weibull και Lognormal. Η επιλογή αυτή επέτρεψε να αξιολογηθεί κατά πόσο η θετική ασυμμετρία και η μη αρνητικότητα των δεδομένων απαιτούν τη χρήση διαφορετικού θεωρητικού μοντέλου. Η σταθερά χαμηλή απόδοση της Normal επιβεβαίωσε ότι η συμμετρική προσέγγιση δεν ήταν κατάλληλη για τις συγκεκριμένες χρονοσειρές.

Η μεθοδολογία κρίνεται κατάλληλη για τον σκοπό της εργασίας, επειδή συνδύασε ενιαία κριτήρια εφαρμογής με αυτοτελή αξιολόγηση κάθε χρονοσειράς. Με τον τρόπο αυτό διατηρήθηκε η συγκρισιμότητα των αποτελεσμάτων, χωρίς να παραβλεφθούν οι διαφοροποιήσεις μεταξύ των πόλεων. Παράλληλα, η ανάλυση παρέμεινε εντός σαφών ορίων, καθώς αφορά την κατανομική μορφή των ωριαίων τιμών και όχι τη χρονική εξάρτηση, την εποχικότητα ή τους μηχανισμούς που διαμόρφωσαν τις συγκεντρώσεις.

5.5 Περιορισμοί της εργασίας

Παρά τη συστηματική μεθοδολογική διαδικασία που ακολουθήθηκε, η παρούσα εργασία υπόκειται σε ορισμένους περιορισμούς, οι οποίοι πρέπει να λαμβάνονται υπόψη κατά την ερμηνεία των αποτελεσμάτων. Οι περιορισμοί αυτοί δεν αναιρούν την αξία της ανάλυσης, αλλά καθορίζουν το εύρος μέσα στο οποίο μπορούν να γενικευθούν τα ευρήματα.

Ο βασικότερος περιορισμός αφορά τη διαθεσιμότητα και την πληρότητα των δεδομένων. Στην τελική κατανομική ανάλυση εντάχθηκαν μόνο οι ετήσιες χρονοσειρές που πληρούσαν τα προκαθορισμένα κριτήρια ποιότητας και χρονικής συνέχειας. Ως αποτέλεσμα, ο αριθμός των αναλύσιμων περιπτώσεων ήταν περιορισμένος και δεν κατέστη δυνατή η ισότιμη ένταξη όλων των σταθμών και ετών. Η περίπτωση του Καρπενησίου είναι ενδεικτική, καθώς η ανεπαρκής πληρότητα των διαθέσιμων μετρήσεων δεν επέτρεψε αξιόπιστη προσαρμογή θεωρητικών κατανομών.

Περιορισμός προκύπτει και από τη χρονική έκταση της μελέτης. Για κάθε πόλη αναλύθηκε ένα συγκεκριμένο έτος με επαρκή πληρότητα και όχι ένα ενιαίο πολυετές διάστημα. Επομένως, τα αποτελέσματα αποτυπώνουν τη συμπεριφορά των επιλεγμένων

ετήσιων χρονοσειρών, αλλά δεν επαρκούν για την εξαγωγή συμπερασμάτων σχετικά με μακροχρόνιες τάσεις, διαχρονικές μεταβολές ή τη σταθερότητα της καταλληλότερης κατανομής από έτος σε έτος.

Η εργασία επικεντρώθηκε αποκλειστικά στα PM_{10} και δεν περιέλαβε άλλους ατμοσφαιρικούς ρύπους, όπως $PM_{2.5}$, NO_2 , O_3 ή SO_2 . Η επιλογή αυτή ενίσχυσε τη συνοχή της ανάλυσης, αλλά δεν επιτρέπει ολοκληρωμένη αξιολόγηση της ποιότητας του αέρα στις εξεταζόμενες πόλεις, καθώς η ατμοσφαιρική ρύπανση διαμορφώνεται από τη συνδυασμένη παρουσία πολλών ρύπων.

Επιπλέον, στην ανάλυση δεν ενσωματώθηκαν μετεωρολογικές, τοπογραφικές ή εκπομπικές μεταβλητές. Παράγοντες όπως ο άνεμος, η θερμοκρασία, η υγρασία, η βροχόπτωση, η οικιακή θέρμανση, η κυκλοφορία και η μεταφορά σκόνης μπορούν να επηρεάσουν σημαντικά τις συγκεντρώσεις PM_{10} . Οι παράγοντες αυτοί εξετάστηκαν στο θεωρητικό μέρος, αλλά δεν αναλύθηκαν ποσοτικά. Κατά συνέπεια, τα αποτελέσματα περιγράφουν τη στατιστική μορφή των συγκεντρώσεων χωρίς να επιτρέπουν ασφαλή αιτιολογική απόδοση των διαφορών που παρατηρήθηκαν μεταξύ των πόλεων.

Ένας ακόμη περιορισμός συνδέεται με το σύνολο των θεωρητικών μοντέλων που εξετάστηκαν. Η σύγκριση περιορίστηκε στις Normal, Gamma, Weibull και Lognormal, οι οποίες αποτελούν ευρέως χρησιμοποιούμενες και ερμηνεύσιμες επιλογές για δεδομένα συγκεντρώσεων. Δεν αποκλείεται, ωστόσο, άλλες κατανομές, σύνθετα μοντέλα ή ειδικές μέθοδοι για ακραίες τιμές να αποδίδουν καλύτερα ορισμένες χρονοσειρές ή συγκεκριμένα τμήματα της δεξιάς ουράς.

Η χρήση προκαθορισμένων κλάσεων συχνότητας επηρεάζει επίσης την αριθμητική αξιολόγηση. Ιδιαίτερα η ανοικτή ανώτερη κλάση συγκεντρώνει παρατηρήσεις διαφορετικού μεγέθους σε μία κατηγορία και περιορίζει τη λεπτομερή διερεύνηση των υψηλότερων τιμών. Παρότι τα Q-Q διαγράμματα συμπλήρωσαν την εικόνα της άνω ουράς, η ανάλυση δεν είχε ως αποκλειστικό στόχο τη μοντελοποίηση ακραίων συγκεντρώσεων.

Τέλος, η μελέτη εξετάζει την κατανομική μορφή των ωριαίων τιμών και όχι τη χρονική τους εξάρτηση. Δεν αναλύθηκαν η αυτοσυσχέτιση, η εποχικότητα, οι ημερήσιοι ή εβδομαδιαίοι κύκλοι και η διάρκεια πιθανών περιόδων αυξημένων συγκεντρώσεων. Για τον λόγο αυτό, ο όρος χρονοσειρά αναφέρεται στην χρονικά διατεταγμένη προέλευση των δεδομένων, ενώ η εφαρμοζόμενη προσέγγιση αφορά κυρίως την οριακή κατανομή των παρατηρήσεων.

Οι παραπάνω περιορισμοί καθιστούν αναγκαία την προσεκτική γενίκευση των συμπερασμάτων. Η επικράτηση της Lognormal στις τρεις από τις τέσσερις αναλύσιμες χρονοσειρές και της Weibull στη Χαλκίδα αφορά τα συγκεκριμένα δεδομένα και τα επιλεγμένα έτη. Δεν συνιστά καθολικό κανόνα για όλες τις πόλεις, όλες τις χρονικές περιόδους ή το σύνολο της Στερεάς Ελλάδας.

5.6 Προτάσεις για μελλοντική έρευνα

Με βάση τα ευρήματα και τους περιορισμούς της παρούσας εργασίας, η μελλοντική έρευνα θα μπορούσε αρχικά να επεκτείνει την ανάλυση σε περισσότερα έτη και σταθμούς. Η αξιοποίηση πληρέστερων πολυετών χρονοσειρών θα επέτρεπε να εξεταστεί αν η καταλληλότερη κατανομή παραμένει σταθερή στον χρόνο ή μεταβάλλεται μεταξύ διαφορετικών περιόδων και περιοχών.

Ιδιαίτερα χρήσιμη θα ήταν και η ενσωμάτωση μετεωρολογικών μεταβλητών, όπως η θερμοκρασία, η υγρασία, η βροχόπτωση και τα χαρακτηριστικά του ανέμου. Η προσθήκη των παραμέτρων αυτών θα επέτρεπε τη σύνδεση της κατανομικής μορφής των συγκεντρώσεων PM_{10} με τις συνθήκες διασποράς, συσσώρευσης και απομάκρυνσης των σωματιδίων.

Παράλληλα, θα μπορούσαν να αξιοποιηθούν πληροφορίες για τοπικές πηγές και ειδικά γεγονότα, όπως η κυκλοφορία, η οικιακή θέρμανση, οι αγροτικές ή βιομηχανικές δραστηριότητες, οι πυρκαγιές και τα επεισόδια μεταφοράς σκόνης. Η προσέγγιση αυτή θα συνέβαλλε στην ερμηνεία των υψηλών παρατηρήσεων και της δεξιάς ουράς, χωρίς να περιορίζεται στην απλή περιγραφή της στατιστικής τους μορφής.

Μια ακόμη κατεύθυνση αφορά τη διερεύνηση και άλλων ρύπων, όπως τα $PM_{2.5}$, το NO_2 , το O_3 και το SO_2 . Η παράλληλη ανάλυση διαφορετικών ρύπων θα μπορούσε να αναδείξει ομοιότητες και διαφοροποιήσεις στη στατιστική τους συμπεριφορά και να προσφέρει πληρέστερη εικόνα της ποιότητας του αέρα.

Σε μεθοδολογικό επίπεδο, μπορούν να εξεταστούν πρόσθετες θεωρητικές κατανομές, μικτά μοντέλα και εξειδικευμένες μέθοδοι για την ανάλυση ακραίων τιμών. Ιδιαίτερο ενδιαφέρον παρουσιάζει η χωριστή μοντελοποίηση της κύριας μάζας των δεδομένων και της άνω ουράς, καθώς τα αποτελέσματα έδειξαν ότι ένα μοντέλο μπορεί να αποδίδει

ικανοποιητικά τις συχνότητες χωρίς να αναπαράγει εξίσου καλά τις υψηλότερες συγκεντρώσεις.

Τέλος, η μελλοντική ανάλυση θα μπορούσε να εξετάσει τη χρονική δομή των μετρήσεων, μέσω της εποχικότητας, της αυτοσυσχέτισης, των ημερήσιων κύκλων και της διάρκειας περιόδων αυξημένων συγκεντρώσεων. Με τον τρόπο αυτό, η κατανομική προσέγγιση θα μπορούσε να συνδυαστεί με τη δυναμική συμπεριφορά των PM_{10} και να προσφέρει πληρέστερη ερμηνεία της μεταβλητότητάς τους.

5.7 Τελικό συμπέρασμα

Η παρούσα διπλωματική εργασία διερεύνησε τη στατιστική συμπεριφορά των ωριαίων συγκεντρώσεων PM_{10} σε επιλεγμένες πόλεις της Στερεάς Ελλάδας και αξιολόγησε τέσσερις θεωρητικές κατανομές ως προς την ικανότητά τους να περιγράψουν την εμπειρική μορφή των δεδομένων. Η ανάλυση έδειξε ότι οι εξεταζόμενες χρονοσειρές χαρακτηρίζονται από μη αρνητικές τιμές, θετική ασυμμετρία, αυξημένη κύρτωση και παρουσία περιορισμένου αριθμού πολύ υψηλών παρατηρήσεων.

Τα χαρακτηριστικά αυτά δεν αποδόθηκαν ικανοποιητικά από την κανονική κατανομή, η οποία παρουσίασε σταθερά τη χαμηλότερη ποιότητα προσαρμογής. Αντίθετα, οι ασύμμετρες κατανομές ανταποκρίθηκαν αποτελεσματικότερα στη μορφή των δεδομένων, επιβεβαιώνοντας ότι η συμμετρία δεν αποτελεί κατάλληλη υπόθεση για τις συγκεκριμένες χρονοσειρές PM_{10} .

Η λογαριθμοκανονική κατανομή αναδείχθηκε ως η επικρατέστερη επιλογή, καθώς προσαρμόστηκε καταλλήλότερα στις χρονοσειρές της Άμφισσας, της Λαμίας και της Λειβαδιάς. Η Weibull επιλέχθηκε για τη Χαλκίδα, γεγονός που απέδειξε ότι η συχνότερη υπεροχή ενός μοντέλου δεν συνεπάγεται καθολική εφαρμογή του σε όλες τις περιπτώσεις. Η τελική επιλογή βασίστηκε στη συνδυαστική αξιολόγηση των παρατηρούμενων και αναμενόμενων συχνοτήτων, των διαγραμμάτων observed-expected, των υπολοίπων, των Q-Q διαγραμμάτων και του στατιστικού χ^2 . Η πολυκριτηριακή αυτή διαδικασία επέτρεψε να αξιολογηθεί όχι μόνο η γενική συμφωνία κάθε κατανομής, αλλά και η συμπεριφορά της σε διαφορετικά τμήματα της εμπειρικής κατανομής.

Το βασικό συμπέρασμα της εργασίας είναι ότι η Lognormal αποτελεί την πιο συνεπή θεωρητική προσέγγιση για τις περισσότερες αναλύσιμες χρονοσειρές, χωρίς όμως να είναι καθολικά βέλτιστη. Η διαφοροποίηση της Χαλκίδας επιβεβαιώνει ότι κάθε πόλη και έτος πρέπει να εξετάζονται αυτοτελώς, με κοινή μεθοδολογία αλλά χωρίς προκαθορισμένη επιλογή μοντέλου.

Τα ευρήματα αφορούν τα συγκεκριμένα διαθέσιμα δεδομένα και τα επιλεγμένα έτη και δεν πρέπει να γενικεύονται αυτόματα στη μακροχρόνια συμπεριφορά των PM_{10} σε ολόκληρη τη Στερεά Ελλάδα. Παρ' όλα αυτά, η εργασία παρέχει ένα συνεκτικό πλαίσιο ελέγχου ποιότητας, κατανομικής ανάλυσης και συγκριτικής αξιολόγησης, το οποίο μπορεί να αξιοποιηθεί και να επεκταθεί σε μελλοντικές μελέτες ποιότητας του αέρα.

Βιβλιογραφία

Aleksandropoulou, V., & Lazaridis, M. (2013). *Identification of the Influence of African Dust on PM₁₀ Concentrations at the Athens Air Quality Monitoring Network during the Period 2001-2010*. *Aerosol and Air Quality Research*, 13(5), 1492-1503. <https://doi.org/10.4209/aaqr.2012.12.0363>

Argyropoulos, G., Samara, C., Diapouli, E., Eleftheriadis, K., Papaoikonomou, K., & Kungolos, A. (2017). *Source apportionment of PM₁₀ and PM_{2.5} in major urban Greek agglomerations using a hybrid source-receptor modeling process*. *Science of The Total Environment*, 601-602, 906-917. <https://doi.org/10.1016/j.scitotenv.2017.05.088>

Bagcı, K. (2023). *Application of Statistical Distributions to PM₁₀ Concentrations: Van, Türkiye*. *Yüzüncü Yıl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 60, 87-95. <https://doi.org/10.53568/yyusbed.1214637>

Brook, R. D., Rajagopalan, S., Pope, C. A., III, Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Jr, Whitsel, L., & Kaufman, J. D. (2010). *Particulate Matter Air Pollution and Cardiovascular Disease*. *Circulation*, 121(21), 2331-2378. <https://doi.org/10.1161/cir.0b013e3181dbee1>

Copernicus. (2024, 13 June). *Saharan dust transport to impact air quality in Eastern Mediterranean*. Copernicus Atmosphere Monitoring Service. <https://atmosphere.copernicus.eu/saharan-dust-transport-impact-air-quality-eastern-mediterranean>

Diapouli, E., Manousakas, M., Vratolis, S., Vasilatou, V., Maggos, T., Saraga, D., Grigoratos, T., Argyropoulos, G., Voutsas, D., Samara, C., & Eleftheriadis, K. (2017). *Evolution of air pollution source contributions over one decade, derived by PM₁₀ and PM_{2.5} source apportionment in two metropolitan urban areas in Greece*. *Atmospheric Environment*, 164, 416-430. <https://doi.org/10.1016/j.atmosenv.2017.06.016>

Dimitriou, K., & Kassomenos, P. (2019). *Estimation of North African dust contribution on PM₁₀ episodes at four continental Greek cities*. *Ecological Indicators*, 106, 105530. <https://doi.org/10.1016/j.ecolind.2019.105530>

European Environment Agency (EEA). (2016). *Air pollution from emissions to exposure*. <https://www.eea.europa.eu/en/analysis/maps-and-charts/air-pollution-from-emissions-to-exposure>

European Environment Agency (EEA). (2019). *How air pollution affects our health*. <https://www.eea.europa.eu/en/topics/in-depth/air-pollution/how-it-affects-our-health>

European Environment Agency (EEA). (2021). *Sources and emissions of air pollutants in Europe*. In *Air quality in Europe 2021*. <https://www.eea.europa.eu/en/analysis/publications/air-quality-in-europe-2021/sources-and-emissions-of-air-pollutants-in-europe>

European Environment Agency (EEA). (2023). *Air pollution and children's health*. <https://www.eea.europa.eu/en/analysis/publications/air-pollution-and-childrens-health>

European Environment Agency (EEA). (2025, April 9). *Particulate matter - PM₁₀*. In Air quality status report 2025. <https://www.eea.europa.eu/en/analysis/publications/air-quality-status-report-2025/particulate-matter>

European Environment Agency (EEA). (2026, April 30). *Particulate matter - PM₁₀*. In Air quality status in Europe 2026. <https://www.eea.europa.eu/en/analysis/publications/air-quality-status-report-2026/particulate-matter-pm10>

European Parliament and Council of the European Union. (2008). *Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe*. Official Journal of the European Union <https://eur-lex.europa.eu/eli/dir/2008/50/oj/eng>

European Parliament and Council of the European Union. (2024). *Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on ambient air quality and cleaner air for Europe (recast)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/dir/2024/2881/oj/>

Giavis, G. M., Kambezidis, H. D., & Lykoudis, S. P. (2008). *Frequency distribution of particulate matter (PM₁₀) in urban environments*. International Journal of Environment and Pollution, Vol. 36, No.1-3, pp 99-109. <https://doi.org/10.1504/ijep.2009.021819>

Grivas, G., Chaloulakou, A., & Kassomenos, P. (2008). *An overview of the PM₁₀ pollution problem, in the Metropolitan Area of Athens, Greece*. Assessment of controlling factors and potential impact of long range transport. Science of The Total Environment, 389(1), 165-177. <https://doi.org/10.1016/j.scitotenv.2007.08.048>

Grivas, G., Cheristanidis, S., Chaloulakou, A., Koutrakis, P., & Mihalopoulos, N. (2018). *Elemental Composition and Source Apportionment of Fine and Coarse Particles at Traffic and Urban Background Locations in Athens, Greece*. Aerosol and Air Quality Research, 18(7), 1642-1659. <https://doi.org/10.4209/aaqr.2017.12.0567>

He, H., Schäfer, B., & Beck, C. (2022). *Spatial heterogeneity of air pollution statistics in Europe*. Scientific Reports, 12(1). <https://doi.org/10.1038/s41598-022-16109-2>

- Karagiannidis, A., Poupkou, A., Giannaros, T., Giannaros, C., Melas, D., & Argiriou, A. (2015). *The Air Quality of a Mediterranean Urban Environment Area and Its Relation to Major Meteorological Parameters*. Water, Air, & Soil Pollution, 226(1). <https://doi.org/10.1007/s11270-014-2239-8>
- Kassomenos, P. A., Paschalidou, A. K., Lykoudis, S., & Koletsis, I. (2014). *Temperature inversion characteristics in relation to synoptic circulation above Athens, Greece*. Environmental Monitoring and Assessment, 186(6), 3495-3502. <https://doi.org/10.1007/s10661-014-3632-x>
- Koçak, M., Mihalopoulos, N., & Kubilay, N. (2009). *Origin and source regions of PM₁₀ in the Eastern Mediterranean atmosphere*. Atmospheric Research, 92(4), 464-474. <https://doi.org/10.1016/j.atmosres.2009.01.005>
- Lu, H.-C. (2002). *The statistical characters of PM₁₀ concentration in Taiwan area*. Atmospheric Environment, 36(3), 491-502. [https://doi.org/10.1016/s1352-2310\(01\)00245-x](https://doi.org/10.1016/s1352-2310(01)00245-x)
- Lu, H.-C. & Fang, G.-C. (2003). *Predicting the exceedances of a critical PM10 concentration - a case study in Taiwan*. Atmospheric Environment, 37(25), 3491-3499. [https://doi.org/10.1016/s1352-2310\(03\)00453-9](https://doi.org/10.1016/s1352-2310(03)00453-9)
- Matthaios, V. N., Triantafyllou, A. G., & Koutrakis, P. (2016). *PM₁₀ episodes in Greece: Local sources versus long-range transport - observations and model simulations*. Journal of the Air & Waste Management Association, 67(1), 105-126. <https://doi.org/10.1080/10962247.2016.1231146>
- Meirelles, M. G., & Vasconcelos, H. C. (2025). *Physical-Statistical Characterization of PM₁₀ and PM_{2.5} Concentrations and Atmospheric Transport Events in the Azores During 2024*. Earth, 6(2), 54. <https://doi.org/10.3390/earth6020054>
- Mitsakou, C., Kallos, G., Papantoniou, N., Spyrou, C., Solomos, S., Astitha, M., & Housiadas, C. (2008). *Saharan dust levels in Greece and received inhalation doses*. Atmospheric Chemistry and Physics, 8(23), 7181-7192. <https://doi.org/10.5194/acp-8-7181-2008>
- Moustris, K. P., Petraki, E., Ntourou, K., Priniotakis, G., & Nikolopoulos, D. (2020). *Spatiotemporal Evaluation of PM10 Concentrations within the Greater Athens Area, Greece. Trends, Variability and Analysis of a 19 Years Data Series*. Environments, 7(10), 85. <https://doi.org/10.3390/environments7100085>
- Mukherjee, A. & Agrawal, M. (2017). *World air particulate matter: sources, distribution and health effects*. Environmental Chemistry Letters, 15(2), 283-309. <https://doi.org/10.1007/s10311-017-0611-9>

NIST/SEMATECH. (2012). *e-Handbook of Statistical Methods: Chi-Square Goodness-of-Fit Test*. National Institute of Standards and Technology. <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35f.htm>

Ntziachristos, L. & Boulter, P. (2023). Road transport: Automobile tyre and brake wear; road transport: Automobile road abrasion (NFR 1.A.3.b.vi-vii). In *EMEP/EEA air pollutant emission inventory guidebook 2023: Update 2023*. European Environment Agency.

Orellano, P., Reynoso, J., Quaranta, N., Bardach, A., & Ciapponi, A. (2020). *Short-term exposure to particulate matter (PM_{10} and $PM_{2.5}$), nitrogen dioxide (NO_2), and ozone (O_3) and all-cause and cause-specific mortality: Systematic review and meta-analysis*. *Environment International*, 142, 105876. <https://doi.org/10.1016/j.envint.2020.105876>

Papanastasiou, D. K. & Melas, D. (2010). *Application of PM_{10} 's Statistical Distribution to Air Quality Management - A Case Study in Central Greece*. 207(1-4), 115-122. [doi:10.1007/s11270-009-0123-8](https://doi.org/10.1007/s11270-009-0123-8)

Pateraki, St., Assimakopoulos, V. D., Maggos, Th., Fameli, K. M., Kotroni, V., & Vasiliakos, Ch. (2013). *Particulate matter pollution over a Mediterranean urban area*. *Science of The Total Environment*, 463-464, 508-524. <https://doi.org/10.1016/j.scitotenv.2013.05.076>

Plocoste, T., Calif, R., Euphrasie-Clotilde, L., & Brute, F.-N. (2020). *The statistical behavior of PM_{10} events over guadeloupean archipelago: Stationarity, modelling and extreme events*. *Atmospheric Research*, 241, 104956. <https://doi.org/10.1016/j.atmosres.2020.104956>

Pope, C. A., III, & Dockery, D. W. (2006). *Health Effects of Fine Particulate Air Pollution: Lines that Connect*. *Journal of the Air & Waste Management Association*, 56(6), 709-742. <https://doi.org/10.1080/10473289.2006.10464485>

Querol, X., Pey, J., Pandolfi, M., Alastuey, A., Cusack, M., Pérez, N., Moreno, T., Viana, M., Mihalopoulos, N., Kallos, G., & Kleanthous, S. (2009). African dust contributions to mean ambient PM_{10} mass-levels across the Mediterranean Basin. *Atmospheric Environment*, 43(28), 4266-4277. <https://doi.org/10.1016/j.atmosenv.2009.06.013>

Remoundaki, E., Bourliva, A., Kokkalis, P., Mamouri, R. E., Papayannis, A., Grigoratos, T., Samara, C., & Tsezos, M. (2011). *PM_{10} composition during an intense Saharan dust transport event over Athens (Greece)*. *Science of The Total Environment*, 409(20), 4361-4372. <https://doi.org/10.1016/j.scitotenv.2011.06.026>

Sansuddin, N., Ramli, N. A., Yahaya, A. S., Yusof, N. F. F. M., Ghazali, N. A., & Madhoun, W. A. A. (2010). *Statistical analysis of PM_{10} concentrations at different locations in Malaysia*. *Environmental Monitoring and Assessment*, 180 (1-4), 573-588. <https://doi.org/10.1007/s10661-010-1806-8>

Saxena, P., Sonwani, S. (2019). *Criteria Air Pollutants: Chemistry, Sources and Sinks*. In: *Criteria Air Pollutants and their Impact on Environmental Health*. Springer, Singapore. https://doi.org/10.1007/978-981-13-9992-3_2

Seinfeld, J.H. & Pandis, S.N. (2016). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons, Hoboken.

Tiwary, A., Williams, I. & Colls, J. (2019). *Air pollution measurement, modelling and mitigation*. CRC Press. Taylor & Francis Group

Tsiaousidis, D. T., Liora, N., Kontos, S., Poupkou, A., Akritidis, D., & Melas, D. (2023). *Evaluation of PM Chemical Composition in Thessaloniki, Greece Based on Air Quality Simulations*. Sustainability, 15(13), 10034. <https://doi.org/10.3390/su151310034>

U.S. Environmental Protection Agency (EPA). (2026a, May 5). *Health and environmental effects of particulate matter (PM)*. <https://www.epa.gov/pm-pollution/health-and-environmental-effects-particulate-matter-pm>

U.S. Environmental Protection Agency (EPA). (2026b, May 21). *Particulate Matter (PM) Basics*. <https://www.epa.gov/pm-pollution/particulate-matter-pm-basics>

Vardoulakis, S., & Kassomenos, P. (2008). Sources and factors affecting PM₁₀ levels in two European cities: Implications for local air quality management. *Atmospheric Environment*, 42(17), 3949-3963. <https://doi.org/10.1016/j.atmosenv.2006.12.021>

Viana, M., Pey, J., de Leeuw, F., Querol, X., Alastuey, A., dall'Osto, M., & Moreno, T. (2011). *Reporting on natural events in the EU Member States under Directive 2008/50/EC: Years 2008-2009* (ETC/ACM Technical Paper 2011/17). European Topic Centre on Air Pollution and Climate Change Mitigation.

Viana, M., Pey, J., Querol, X., Alastuey, A., de Leeuw, F., & Lükewille, A. (2014). Natural sources of atmospheric aerosols influencing air quality across Europe. *Science of The Total Environment*, 472, 825-833. <https://doi.org/10.1016/j.scitotenv.2013.11.140>

Wang X., Chen RJ, Chen BH & Kan HD., (2013). *Application of Statistical Distribution of PM₁₀ Concentration in Air Quality Management in 5 Representative Cities of China*. Biomedical and Environmental Sciences, 26(08), 638. <https://doi.org/10.3967/0895-3988.2013.08.002>

Wilks, D. S. (2019). *Statistical Methods in the Atmospheric Sciences* (4th ed.). Elsevier.

World Health Organization (WHO). (2021). *WHO global air quality guidelines: Particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization. <https://www.who.int/publications/i/item/9789240034228>

Ελληνική Στατιστική Αρχή (ΕΛΣΤΑΤ). (2026). *Λογαριασμοί εκπομπών αέριων ρύπων: Έτος 2023*. ΕΛΣΤΑΤ. <https://www.statistics.gr/el/statistics/-/publication/SOP08/->

Ζέρβας, Ε., & Σέμπος, Ι. (2024). *Αέρια ρύπανση και αέρια αντιρρυπαντική τεχνολογία* [Μεταπτυχιακό εγχειρίδιο]. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις. <https://dx.doi.org/10.57713/kallipos-1032>

Κατσαφάδος, Π., & Μαυροματίδης, Η. (2024). *Κλιματική αλλαγή* [Προπτυχιακό εγχειρίδιο]. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις. <https://dx.doi.org/10.57713/kallipos-983>

Κούγκολος, Α. (2022). *Περιβαλλοντική Μηχανική Ρύπανση & Προστασία Περιβάλλοντος*. Θεσσαλονίκη: Εκδόσεις Τζιόλα.

Κουϊμτζή, Θ., Φυτιάνου, Κ. & Σαμαρά-Κωνσταντίνου, Κ. (1998). *Χημεία περιβάλλοντος*. Θεσσαλονίκη: University studio press.

Ρεμουντάκη, Ε. (2003). Σωματίδια στην ατμόσφαιρα. Στο Ι. Ζιωμας & Ε. Ρεμουντάκη (Επιμ.), *Φυσικό Περιβάλλον και Ρύπανση* (σς. 177 - 265). Πάτρα: Ελληνικό Ανοικτό Πανεπιστήμιο

Υπουργείο Περιβάλλοντος και Ενέργειας (ΥΠΕΝ). (χ.χ.). *Δεδομένα μετρήσεων ατμοσφαιρικής ρύπανσης*. Εθνικό Δίκτυο Παρακολούθησης Ατμοσφαιρικής Ρύπανσης <https://ypen.gov.gr/perivallon/poiotita-tis-atmosfairas/dedomena-metriseon-atmosfairikis-rypansis>

Υπουργείο Περιβάλλοντος, Ενέργειας και Κλιματικής Αλλαγής (ΥΠΕΚΑ). (2011, 30 Μαρτίου). *Μέτρα για τη βελτίωση της ποιότητας της ατμόσφαιρας, σε συμμόρφωση με τις διατάξεις της οδηγίας 2008/50/ΕΚ «για την ποιότητα του ατμοσφαιρικού αέρα και καθαρότερο αέρα για την Ευρώπη» του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της Ευρωπαϊκής Ένωσης της 21ης Μαΐου 2008* (Αριθμ. Η.Π. 14122/549/Ε.103). Εφημερίδα της Κυβερνήσεως, Β'(488), 7111-7162. https://ypen.gov.gr/wp-content/uploads/legacy/Files/Perivallon/Poiotita%20Atmosfairas/Nomothesia/FEK488B_KYA14122.pdf

Υπουργείο Περιβάλλοντος, Ενέργειας και Κλιματικής Αλλαγής (ΥΠΕΚΑ). (2013, 23 Δεκεμβρίου). *Βραχυπρόθεσμα σχέδια δράσης για την αντιμετώπιση ατμοσφαιρικής ρύπανσης από αιωρούμενα σωματίδια* (ΚΥΑ οικ. 70601). Εφημερίδα της Κυβερνήσεως, Β'(3272), 46225-46228. https://ypen.gov.gr/wp-content/uploads/legacy/Files/Perivallon/PoiotitaAtmosfairas/Nomothesia/FEK3272B_KYA70601.pdf

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν.1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.