



**ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ**  
**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ**

**Διαχείριση και Τεχνολογία Ποιότητας (ΔΙΠ)**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

«Σύγκριση του παραδοσιακού Στατιστικού Ελέγχου Διεργασιών με μεθόδους Μηχανικής Μάθησης για τη Διαχείριση Ποιότητας: Μελέτη περίπτωσης σε ελληνική βιομηχανία.»

**Ιωάννης Σαράφης**

**Επιβλέπων:** Δημοσθένης Πολύζος

**Β' Αξιολογητής:** Σαπίδης Νικόλαος

ΘΕΣΣΑΛΟΝΙΚΗ  
ΣΕΠΤΕΜΒΡΙΟΣ, 2026

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο/η 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.

**ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ**  
**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ**

Διαχείριση και Τεχνολογία Ποιότητας (ΔΙΠ)

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

«Σύγκριση του παραδοσιακού Στατιστικού Ελέγχου Διεργασιών με μεθόδους Μηχανικής  
Μάθησης για τη Διαχείριση Ποιότητας: Μελέτη περίπτωσης σε ελληνική βιομηχανία.»

**Ιωάννης Σαράφης**

**ΕΠΙΤΡΟΠΗ ΚΡΙΣΗΣ**

Επιβλέπων: Δημοσθένης Πολύζος

Β' Αξιολογητής: Σαπίδης Νικόλαος

ΘΕΣΣΑΛΟΝΙΚΗ — ΣΕΠΤΕΜΒΡΙΟΣ 2026

## Περίληψη

Η παρούσα μεταπτυχιακή διπλωματική εργασία εξετάζει τη συγκριτική εφαρμογή παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών (Statistical Process Control – SPC) και μεθόδων Μηχανικής Μάθησης (Machine Learning – ML) στη Διαχείριση Ποιότητας, με εφαρμογή σε πραγματικά δεδομένα ηλεκτρικού ελέγχου φωτιστικών σωμάτων LED. Η βασική στόχευση της εργασίας είναι να διερευνήσει κατά πόσο οι προσεγγίσεις ML-SPC μπορούν να λειτουργήσουν συμπληρωματικά, παρέχοντας πρόσθετη πληροφορία για τη συμπεριφορά των διεργασιών.

Η εργασία βασίζεται σε εφαρμοσμένη μελέτη περίπτωσης και ακολουθεί ποσοτική αναλυτική προσέγγιση. Τα δεδομένα προέρχονται από τελικό ηλεκτρικό έλεγχο προϊόντων και περιλαμβάνουν ηλεκτρικές μετρήσεις, πληροφορίες ταυτοποίησης και τεχνικής διαμόρφωσης, καθώς και τελικό αποτέλεσμα ελέγχου. Λόγω της ύπαρξης διαφορετικών οικογενειών προϊόντων, ονομαστικών ισχύων και τεχνικών διαμορφώσεων, οι απόλυτες τιμές των μετρήσεων δεν είναι άμεσα συγκρίσιμες. Για τον λόγο αυτό, οι κρίσιμες μεταβλητές μετασχηματίστηκαν σε σχετικές αποκλίσεις από τις αντίστοιχες ονομαστικές τιμές.

Στο κομμάτι του Στατιστικού Ελέγχου Διεργασιών εφαρμόστηκαν διαγράμματα ατομικών τιμών και κινούμενου εύρους (I-MR) και διαγράμματα εκθετικά σταθμισμένου κινητού μέσου όρου (EWMA). Στο σκέλος της Μηχανικής Μάθησης εφαρμόστηκαν και συγκρίθηκαν μέθοδοι επιβλεπόμενης μάθησης (Supervised Learning) και μη επιβλεπόμενης μάθησης (Unsupervised Learning), όπως Ανάλυση Κύριων Συνιστωσών (PCA), ομαδοποίηση (Clustering), και Δάσος Απομόνωσης (Isolation Forest).

Τα αποτελέσματα κατέδειξαν ότι το ηλεκτρικό αποτύπωμα των προϊόντων περιέχει πληροφορία που μπορεί να αξιοποιηθεί για την αναγνώριση τεχνικών διαμορφώσεων και την ανίχνευση ασυνήθιστων ηλεκτρικών προτύπων. Παράλληλα, η χαμηλή επικάλυψη μεταξύ σημάτων SPC και ενδείξεων αποκλίσεων ML (Anomaly flags) υποστηρίζει τη λογική της συμπληρωματικότητας. Συνολικά, οι αλγόριθμοι Μηχανικής Μάθησης αξιολογούνται ως πιθανή επέκταση των συστημάτων που υιοθετούν τις αρχές του SPC και όχι ως άμεση αντικατάστασή τους.

**Λέξεις-κλειδιά:** Στατιστικός Έλεγχος Διεργασιών (SPC), Μηχανική Μάθηση, Διαχείριση Ποιότητας, Διαγράμματα Ελέγχου, Ανάλυση Κύριων Συνιστωσών (PCA), Ανίχνευση Αποκλινουσών Παρατηρήσεων, Πολυμεταβλητή Ανάλυση, Βιομηχανικά Δεδομένα

## Comparison of Traditional Statistical Process Control with Machine Learning Methods for Quality Management: A Case Study in Greek Industry

Ioannis Sarafis

### Abstract

This work examines the comparative application of traditional Statistical Process Control (SPC) methods and Machine Learning (ML) algorithms in Quality Management, using real industrial testing data from LED luminaires. The main objective of the study is to investigate whether ML-SPC approaches can operate supportively, providing additional information regarding process behaviour.

The dissertation is based on an applied case study and follows a quantitative analytical approach. The data originate from finished product testing and include electrical measurements, product identification and technical configuration information, as well as the final test result (Pass/No Pass). Due to the existence of different product families, nominal power levels and technical configurations, the absolute measurement values are not directly comparable. Hence, the critical variables were transformed into relative deviations from their corresponding nominal values.

In the SPC part of the analysis, Individuals–Moving Range Charts (I-MR) and Exponentially Weighted Moving Average Charts (EWMA) were applied. In the Machine Learning part, both Supervised Classification and Unsupervised Learning methods were used, such as Principal Component Analysis (PCA), Clustering, and Isolation Forests.

The results showed that the unique electrical fingerprint of the products contains information that can be used for the recognition of technical configurations and the detection of unusual electrical patterns. At the same time, the limited overlap between SPC alarms and ML anomaly flags supports the logic of complementarity between the two approaches. Overall, Machine Learning algorithms are evaluated as a potential extension of systems that adopt the principles of SPC, and not as a direct replacement for them.

**Keywords:** Statistical Process Control (SPC), Machine Learning, Quality Management, Control Charts, Principal Component Analysis (PCA), Anomaly Detection, Multivariate Analysis, Industrial Data

## Περιεχόμενα

|                                                                               |    |
|-------------------------------------------------------------------------------|----|
| Περίληψη .....                                                                | iv |
| Abstract .....                                                                | v  |
| Κεφάλαιο 1 - Εισαγωγή .....                                                   | 1  |
| 1.1 Δομή της Εργασίας .....                                                   | 2  |
| 1.2 Περιγραφή του προβλήματος.....                                            | 2  |
| 1.3 Σκοπός και στόχοι της έρευνας.....                                        | 3  |
| 1.4 Σύντομη Μεθοδολογική Προσέγγιση .....                                     | 4  |
| Κεφάλαιο 2 - Θεωρητικό Υπόβαθρο .....                                         | 5  |
| 2.1 Στατιστικός Έλεγχος Διεργασιών (SPC).....                                 | 5  |
| 2.1.1 Ιστορική εξέλιξη του Στατιστικού Ελέγχου Διεργασιών .....               | 6  |
| 2.1.2 Βασικές μέθοδοι και λειτουργία του Στατιστικού Ελέγχου Διεργασιών ..... | 7  |
| 2.1.3 Περιορισμοί και περιθώρια βελτίωσης του SPC .....                       | 9  |
| 2.2 Μηχανική Μάθηση στη Διαχείριση Ποιότητας .....                            | 11 |
| 2.2.1 Υπάρχουσα βιβλιογραφία σε ML-SPC.....                                   | 13 |
| 2.2. Επισκόπηση βασικών αλγορίθμων Μηχανικής Μάθησης.....                     | 15 |
| 2.3 Συγκριτική θεώρηση SPC και ML-SPC .....                                   | 18 |
| 2.4 Κριτήρια συγκριτικής αξιολόγησης SPC και ML-SPC .....                     | 21 |
| Κεφάλαιο 3 - Μεθοδολογία Έρευνας .....                                        | 25 |
| 3.1 Ερευνητικός σχεδιασμός και μεθοδολογική προσέγγιση .....                  | 25 |
| 3.2 Περιγραφή της μελέτης περίπτωσης.....                                     | 26 |
| 3.3 Περιγραφή δεδομένων και μεταβλητών .....                                  | 28 |
| 3.4 Προεπεξεργασία και προετοιμασία δεδομένων .....                           | 31 |
| 3.5 Εφαρμογή παραδοσιακού SPC .....                                           | 37 |
| 3.6 Ανάπτυξη μοντέλων Μηχανικής Μάθησης.....                                  | 41 |
| 3.6.1 Αναγνώριση Τεχνικής Διαμόρφωσης .....                                   | 44 |
| 3.6.2 Ανίχνευση ανωμαλιών και σύγκριση με SPC .....                           | 45 |

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| 3.7 Πλαίσιο σύγκρισης ML – SPC .....                                  | 46 |
| Κεφάλαιο 4 - Αποτελέσματα .....                                       | 50 |
| 4.1 Αποτελέσματα από την εφαρμογή Στατιστικού Ελέγχου Διεργασιών..... | 50 |
| 4.2 Αποτελέσματα ML .....                                             | 58 |
| 4.2.1 Περιγραφική εικόνα του ML dataset .....                         | 59 |
| 4.2.2 Ανάλυση Κύριων Συνιστωσών - PCA .....                           | 60 |
| 4.2.3 Αποτελέσματα Ταξινόμησης (Clustering).....                      | 62 |
| 4.2.4 Επιβλεπόμενη ταξινόμηση Driver Type και Configuration ID.....   | 63 |
| 4.2.5 Σημαντικότητα χαρακτηριστικών .....                             | 65 |
| 4.3 Συγκριτική αξιολόγηση μεταξύ SPC και ML .....                     | 66 |
| 4.4 Συνολική αξιολόγηση σκοπιμότητας υιοθέτησης ML.....               | 69 |
| Κεφάλαιο 5 – Συμπεράσματα, Περιορισμοί και Μελλοντική Έρευνα.....     | 71 |
| 5.1 Συνολική αποτίμηση της εργασίας .....                             | 71 |
| 5.2 Κύρια συμπεράσματα .....                                          | 72 |
| 5.3 Περιορισμοί της έρευνας .....                                     | 73 |
| 5.4 Προτάσεις για μελλοντική έρευνα .....                             | 74 |
| Βιβλιογραφικές Αναφορές .....                                         | 76 |
| ΠΑΡΑΡΤΗΜΑΤΑ .....                                                     | 78 |
| Παράρτημα Α – Διαγράμματα & Σχήματα.....                              | 78 |
| Παράρτημα Β – Κώδικας SPC & ML.....                                   | 85 |
| Επιβλεπόμενη Ταξινόμηση .....                                         | 88 |
| Παράρτημα Γ – Δείγμα Πρωτογενών Δεδομένων .....                       | 91 |

## Κατάλογος Εικόνων

|                                              |   |
|----------------------------------------------|---|
| Εικόνα 1 Είδη Διαγραμμάτων Ελέγχου SPC ..... | 8 |
|----------------------------------------------|---|

## Κατάλογος Διαγραμμάτων

|                                                                           |    |
|---------------------------------------------------------------------------|----|
| Διάγραμμα 1 Διαδικασία SPC .....                                          | 21 |
| Διάγραμμα 2 Διαδικασία ML-SPC .....                                       | 21 |
| Διάγραμμα 3 Μεθοδολογική Προσέγγιση Εργασίας .....                        | 25 |
| Διάγραμμα 4 Φάσεις Παραγωγικής Διαδικασίας Φωτιστικών Σωμάτων .....       | 27 |
| Διάγραμμα 5 Διαδικασία προεπεξεργασίας και προετοιμασίας δεδομένων .....  | 31 |
| Διάγραμμα 6 Διαγράμματα I-MR επιλεγμένης ομάδας .....                     | 52 |
| Διάγραμμα 7 Διάγραμμα Pareto αριθμού παρατηρήσεων ανά Configuration ..... | 59 |

## Κατάλογος Πινάκων

|                                                                                |    |
|--------------------------------------------------------------------------------|----|
| Πίνακας 1 Χρήσεις και διαφορές SPC με ML/AI-SPC .....                          | 21 |
| Πίνακας 2 Κατηγορίες Δεδομένων της Ανάλυσης .....                              | 29 |
| Πίνακας 3 Κρίσιμα όρια χαρακτηριστικών προϊόντος .....                         | 30 |
| Πίνακας 4 Διαδικασία καθαρισμού δεδομένων ποιοτικού ελέγχου .....              | 33 |
| Πίνακας 5 Στάδια Κατηγοριοποίησης προϊόντων .....                              | 34 |
| Πίνακας 6 Κατηγοριοποίηση δεδομένων ML .....                                   | 36 |
| Πίνακας 7 Σχέση Ορίων Ελέγχου SPC και Ορίων Ελέγχου Προδιαγραφών .....         | 40 |
| Πίνακας 8 Δείκτες αξιολόγησης SPC .....                                        | 41 |
| Πίνακας 9 Πιθανοί συνδυασμοί alarms .....                                      | 47 |
| Πίνακας 10 Μητρώο Καταστάσεων Συναγερμών (alarms) .....                        | 47 |
| Πίνακας 11 Επιλεγμένες ομάδες προϊόντων για εφαρμογή SPC .....                 | 51 |
| Πίνακας 12 Δείκτες αξιολόγησης SPC .....                                       | 57 |
| Πίνακας 13 Κατανομή παρατηρήσεων ανά διαμόρφωση .....                          | 59 |
| Πίνακας 14 Ανάλυση διακύμανσης PCA .....                                       | 60 |
| Πίνακας 15 Συντελεστές κύριων συνιστωσών PCA .....                             | 60 |
| Πίνακας 16 Σημαντικότητα παραγόντων ως προς την ταξινόμηση των προϊόντων ..... | 66 |
| Πίνακας 17 Δείκτες συμφωνίας Isolation Forest με SPC Alarms .....              | 67 |
| Πίνακας 18 Αριθμός Alarms SPC και Isolation Forest .....                       | 67 |

## Συντομογραφίες & Ακρωνύμια

|              |                                       |
|--------------|---------------------------------------|
| <b>AI</b>    | Artificial Intelligence               |
| <b>ARL</b>   | Average Run Length                    |
| <b>CUSUM</b> | Cumulative Sum                        |
| <b>EWMA</b>  | Exponentially Weighted Moving Average |
| <b>GMM</b>   | Gaussian Mixture Model                |
| <b>I-MR</b>  | Individuals–Moving Range              |
| <b>IF</b>    | Isolation Forest                      |
| <b>KNN</b>   | k-Nearest Neighbors                   |
| <b>ML</b>    | Machine Learning                      |
| <b>PCA</b>   | Principal Component Analysis          |
| <b>PF</b>    | Power Factor                          |
| <b>PO</b>    | Production Order                      |
| <b>RF</b>    | Random Forest                         |
| <b>SPC</b>   | Statistical Process Control           |
| <b>SVM</b>   | Support Vector Machine                |
| <b>THDU</b>  | Total Harmonic Distortion of Voltage  |
| <b>TQM</b>   | Total Quality Management              |

## Κεφάλαιο 1 - Εισαγωγή

Η Διαχείριση Ποιότητας αποτελεί πεδίο μείζονος σημασίας για τη σύγχρονη βιομηχανική δραστηριότητα, καθώς συνδέεται άμεσα με τη σταθερότητα των διεργασιών, τη συμμόρφωση των προϊόντων και τη δυνατότητα τεκμηριωμένης λήψης αποφάσεων. Στο σύγχρονο, ραγδαία εξελισσόμενο περιβάλλον όπου οι απαιτήσεις ποιότητας, αξιοπιστίας και ιχνηλασιμότητας αυξάνονται διαρκώς, τα δεδομένα που συλλέγονται κατά την παραγωγή και τον ποιοτικό έλεγχο αποκτούν ιδιαίτερη αξία. Η αξιοποίησή των δεδομένων Ποιότητας εξελίσσεται από την απλή τεκμηρίωση της συμμόρφωσης, στην καλύτερη κατανόηση της παραγωγικής διεργασίας και στον έγκαιρο εντοπισμό αποκλίσεων.

Ο Στατιστικός Έλεγχος Διεργασιών (Statistical Process Control – SPC) αποτελεί διαχρονικά σημαντικό εργαλείο παρακολούθησης της μεταβλητότητας στις βιομηχανικές διεργασίες. Με τη χρήση διαγραμμάτων ελέγχου, ο ΣΕΔ επιτρέπει την αξιολόγηση της στατιστικής σταθερότητας μίας διεργασίας και τον εντοπισμό πιθανών ειδικών αιτίων μεταβλητότητας. Ωστόσο, η εφαρμογή του σε σύγχρονα παραγωγικά περιβάλλοντα παρουσιάζει προκλήσεις, ιδιαίτερα σε περιπτώσεις όπου τα διαθέσιμα δεδομένα είναι πολυμεταβλητά, προέρχονται από διαφορετικά προϊόντα ή τεχνικές διαμορφώσεις και περιλαμβάνουν σύνθετες σχέσεις αλληλεπίδρασης μεταξύ μεταβλητών.

Παράλληλα, η εξέλιξη των υπολογιστικών μεθόδων και η αυξανόμενη διαθεσιμότητα βιομηχανικών δεδομένων έχουν δημιουργήσει τις κατάλληλες συνθήκες για την εφαρμογή πιο σύγχρονων εργαλείων και τεχνικών, όπως αυτές της Μηχανικής Μάθησης (Machine Learning – ML) στη Διαχείριση Ποιότητας. Οι μέθοδοι αυτές μπορούν να αξιοποιήσουν ταυτόχρονα περισσότερα χαρακτηριστικά, να αναγνωρίσουν πρότυπα στα δεδομένα και να εντοπίσουν πολυμεταβλητές ανωμαλίες, οι οποίες ενδέχεται να μην είναι άμεσα εμφανείς μέσω παραδοσιακών διαγραμμάτων ελέγχου. Για τον λόγο αυτό, τα τελευταία χρόνια αναπτύσσεται έντονο ερευνητικό ενδιαφέρον γύρω από τη σύνδεση του SPC με ML/AI προσεγγίσεις, στο πλαίσιο της ευρύτερης μετάβασης προς πιο δεδομενοκεντρικά συστήματα ποιότητας.

Η παρούσα εργασία εντάσσεται σε αυτό το πλαίσιο και εξετάζει τη συγκριτική εφαρμογή παραδοσιακών μεθόδων SPC και μεθόδων Μηχανικής Μάθησης σε πραγματικά δεδομένα προερχόμενα από βιομηχανική παραγωγή φωτιστικών σωμάτων LED. Η βασική στόχευση της

εργασίας δεν είναι να διερευνηθεί η δυνατότητα αντικατάστασης του SPC από τη Μηχανική Μάθηση, αλλά η διερεύνηση του κατά πόσο οι ML-SPC προσεγγίσεις μπορούν να λειτουργήσουν επικουρικά, παρέχοντας πρόσθετη πληροφορία για τη συμπεριφορά των διεργασιών. Πιο συγκεκριμένα, οι μέθοδοι SPC χρησιμοποιούνται ως στατιστικά θεμελιωμένο και ερμηνεύσιμο πλαίσιο αναφοράς για την παρακολούθηση της μεταβλητότητας της διεργασίας, ενώ οι μέθοδοι ML εξετάζονται ως πιθανή πολυμεταβλητή επέκταση της ανάλυσης. Με τον τρόπο αυτό, η εργασία επιδιώκει να αξιολογήσει αν το ηλεκτρικό αποτύπωμα των προϊόντων μπορεί να αξιοποιηθεί για την αναγνώριση τεχνικών διαφοροποιήσεων, την ανίχνευση ασυνήθιστων προτύπων και τη βελτίωση της πληροφόρησης που παρέχεται στο σύστημα ποιότητας.

## 1.1 Δομή της Εργασίας

Η παρούσα εργασία οργανώνεται σε πέντε κεφάλαια. Στο πρώτο κεφάλαιο παρουσιάζονται το γενικό πλαίσιο της έρευνας, το πρόβλημα που πραγματεύεται η εργασία, καθώς και οι σκοποί και οι στόχοι της. Το δεύτερο κεφάλαιο περιλαμβάνει τη βιβλιογραφική ανασκόπηση και το θεωρητικό υπόβαθρο, με έμφαση στον Στατιστικό Έλεγχο Διεργασιών, στη Μηχανική Μάθηση στη Διαχείριση Ποιότητας και τη μεταξύ τους σύνδεση.

Στο τρίτο κεφάλαιο παρουσιάζεται η μεθοδολογία της έρευνας. Πιο συγκεκριμένα, περιγράφονται η μελέτη περίπτωσης, η μορφή και το είδος των διαθέσιμων δεδομένων, τα στάδια προ επεξεργασίας των δεδομένων, η μεθοδολογία εφαρμογής των μεθόδων SPC, η ανάπτυξη των μοντέλων Μηχανικής Μάθησης και το πλαίσιο σύγκρισης των δύο προσεγγίσεων. Στο τέταρτο κεφάλαιο παρουσιάζονται τα αποτελέσματα της εφαρμογής των μεθόδων και διενεργείται συγκριτική αξιολόγηση SPC και ML-SPC. Τέλος, στο πέμπτο κεφάλαιο συνοψίζονται τα βασικά συμπεράσματα, διατυπώνονται οι περιορισμοί της έρευνας και οι προτάσεις για μελλοντική διερεύνηση.

## 1.2 Περιγραφή του προβλήματος

Οι σύγχρονες βιομηχανικές παραγωγικές διαδικασίες δημιουργούν αξιοσημείωτο όγκο δεδομένων από τα στάδια ελέγχου και παρακολούθησης της ποιότητας. Ωστόσο, τα δεδομένα αυτά αξιοποιούνται κατά βάση αποκλειστικά για την επιβεβαίωση της συμμόρφωσης των προϊόντων και όχι για συστηματική ανάλυση της συμπεριφοράς των διεργασιών. Η προσέγγιση

αυτή περιορίζει τη δυνατότητα έγκαιρης αναγνώρισης μεταβολών, αποκλίσεων ή εμφάνισης μη τυχαίων προτύπων που ενδέχεται να προηγούνται μιας μη συμμόρφωσης.

Στην παρούσα εργασία, το πρόβλημα αφορά την αξιοποίηση δεδομένων τελικού ηλεκτρικού ελέγχου από παραγωγική διαδικασία φωτιστικών σωμάτων LED. Τα δεδομένα αποτελούνται από ηλεκτρικές μετρήσεις, πληροφορίες ταυτοποίησης και τεχνικής διαμόρφωσης προϊόντος, καθώς και τελικό αποτέλεσμα ελέγχου (Pass/Fail). Η ύπαρξη διαφορετικών οικογενειών προϊόντων, ισχύων και τεχνικών διαμορφώσεων καθιστά την ανάλυση πιο σύνθετη, καθώς οι απόλυτες τιμές των μετρήσεων δεν είναι άμεσα συγκρίσιμες ανάμεσα σε διαφορετικά προϊόντα.

Παράλληλα, είναι ουσιώδης η διάκριση μεταξύ συμμόρφωσης προϊόντος και στατιστικής σταθερότητας διεργασίας. Το δυαδικό αποτέλεσμα Pass / No Pass εκφράζει τη συμμόρφωση με συγκεκριμένες προδιαγραφές, ενώ τα σήματα SPC αναφέρονται στη συμπεριφορά μίας διεργασίας σε σχέση με τη φυσική της μεταβλητότητα. Επομένως, το ερευνητικό ζήτημα δεν εξαντλείται στον εντοπισμό προϊόντων εκτός προδιαγραφών, αλλά αφορά πρωτίστως τη διερεύνηση του κατά πόσο οι παραδοσιακές μέθοδοι SPC και οι μέθοδοι Μηχανικής Μάθησης μπορούν να προσφέρουν συμπληρωματική πληροφόρηση για τη συμπεριφορά της υπό μελέτη διεργασίας.

### 1.3 Σκοπός και στόχοι της έρευνας

Σκοπός της παρούσας εργασίας είναι η συγκριτική αξιολόγηση παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών και μεθόδων Μηχανικής Μάθησης στη Διαχείριση Ποιότητας, με εφαρμογή σε πραγματικά δεδομένα ηλεκτρικού ελέγχου φωτιστικών σωμάτων LED. Η εργασία επιδιώκει να διερευνήσει κατά πόσο οι μέθοδοι ML μπορούν να ενισχύσουν ή να συμπληρώσουν την πληροφορία που παρέχουν τα παραδοσιακά εργαλεία SPC, ιδιαίτερα σε περιπτώσεις όπου η συμπεριφορά των διεργασιών επηρεάζεται από περισσότερες από μία μεταβλητές. Οι επιμέρους στόχοι της έρευνας αφορούν την ανάπτυξη και αξιολόγηση μεθόδων Μηχανικής Μάθησης για την αναγνώριση ηλεκτρικών προτύπων και τεχνικών διαμορφώσεων των προϊόντων, την εφαρμογή μεθόδων για την ανίχνευση ανωμαλιών για τη διερεύνηση πολυμεταβλητών αποκλίσεων, καθώς και την ανάδειξη των βασικών πλεονεκτημάτων, των περιορισμών και προϋποθέσεων εφαρμογής των ML-SPC προσεγγίσεων σε πραγματικό βιομηχανικό περιβάλλον. Τέλος, η εργασία αποσκοπεί να διερευνήσει τις απαιτήσεις και τις

προκλήσεις που προκύπτουν κατά τη χρήση μεθόδων Μηχανικής Μάθησης. Μέσα από αυτούς τους στόχους, η εργασία επιδιώκει να συμβάλει στην κατανόηση του τρόπου με τον οποίο τα παραδοσιακά εργαλεία SPC και οι σύγχρονες μέθοδοι Μηχανικής Μάθησης μπορούν να συνδυαστούν για την ενίσχυση της Διαχείρισης Ποιότητας.

## 1.4 Σύντομη Μεθοδολογική Προσέγγιση

Η εργασία βασίζεται σε εφαρμοσμένη μελέτη περίπτωσης και ακολουθεί ποσοτική αναλυτική προσέγγιση. Αρχικά, πραγματοποιείται θεωρητική παρουσίαση των βασικών αρχών του Στατιστικού Ελέγχου Διεργασιών και των μεθόδων Μηχανικής Μάθησης στη Διαχείριση Ποιότητας. Στη συνέχεια περιγράφονται τα στάδια της παραγωγικής διαδικασίας, τα δεδομένα τελικού ελέγχου και τα στάδια προεπεξεργασίας που απαιτούνται για τη δημιουργία συγκρίσιμων μεταβλητών και κατάλληλων σετ δεδομένων. Ακολούθως εφαρμόζονται μέθοδοι που παράγουν διαγράμματα SPC, όπως I-MR και EWMA, καθώς και μέθοδοι ML, όπως Ανάλυση Κύριων Συνιστωσών (*Principal Components Analysis – PCA*), ομαδοποίηση δεδομένων (*Clustering*), Επιβλεπόμενη Ταξινόμηση και Isolation Forest. Τα αποτελέσματα των δύο προσεγγίσεων συγκρίνονται με στόχο να αναδειχθούν τα πλεονεκτήματα, οι περιορισμοί και η πιθανή συνδυαστική τους αξία.

Η συμβολή αυτής της εργασίας έγκειται στην εφαρμογή και αξιολόγηση αυτών των μεθόδων σε πραγματικό περιβάλλον ελληνικής βιομηχανίας, όπου τα δεδομένα παρουσιάζουν χαρακτηριστικά που συναντώνται συχνά στην πράξη π.χ. διαφορετικές οικογένειες προϊόντων, πλήθος τεχνικών διαμορφώσεων και ανομοιόμορφη κατανομή παρατηρήσεων. Συνολικά, η παρούσα εργασία φιλοδοξεί να συμβάλει στη συζήτηση γύρω από τη μετάβαση από τα παραδοσιακά στατικά εργαλεία ελέγχου ποιότητας προς σύγχρονες σύνθετες, δεδομενοκεντρικές προσεγγίσεις. Μέσα από τη συγκεκριμένη μελέτη, αναδεικνύεται το γεγονός ότι η αξιοποίηση των δεδομένων ελέγχου μπορεί να χρησιμοποιηθεί πέρα από την επιβεβαίωση της συμμόρφωσης, στην συστηματική κατανόηση της παραγωγικής διεργασίας.

## Κεφάλαιο 2 - Θεωρητικό Υπόβαθρο

### 2.1 Στατιστικός Έλεγχος Διεργασιών (SPC)

Ο Στατιστικός Έλεγχος Διεργασιών (Statistical Process Control – SPC) αποτελεί μία από τις θεμελιώδεις μεθοδολογίες της Διαχείρισης Ποιότητας, με κύριο στόχο την παρακολούθηση, τον έλεγχο και τη βελτίωση της συμπεριφοράς των παραγωγικών διεργασιών με χρήση στατιστικής ανάλυσης δεδομένων. Η θεμελιώδης αρχή του SPC βασίζεται στην αναγνώριση και ποσοτικοποίηση της μεταβλητότητας που εμφανίζεται σε κάθε διεργασία, καθώς και στη διάκριση μεταξύ τυχαίων (common causes) και ειδικών (special causes) αιτίων μεταβολής, τα οποία επιδρούν με διαφορετικό τρόπο στη σταθερότητα μίας διεργασίας (Shewhart, 1931; Montgomery, 2019). Η διάκριση αυτή είναι ουσιώδης και αποτελεί κρίσιμο στοιχείο για τη λήψη αποφάσεων και την εφαρμογή κατάλληλων διορθωτικών ενεργειών.

Η ανάπτυξη του ΣΕΔ αποδίδεται στον Walter A. Shewhart, ο οποίος εισήγαγε την έννοια του στατιστικού ελέγχου στη Βιομηχανία, θέτοντας τις βάσεις για τη συστηματική παρακολούθηση της ποιότητας και την κατανόηση της μεταβλητότητας που διέπει τις διεργασίες (Shewhart, 1931). Στη σύγχρονη βιβλιογραφία, ο Στατιστικός Έλεγχος Διεργασιών αναγνωρίζεται ως θεμελιώδες εργαλείο των συστημάτων διαχείρισης ποιότητας, το οποίο συμβάλλει στη βελτίωση της απόδοσης των παραγωγικών συστημάτων (Montgomery, 2019; Escobar et al., 2025).

Στο πλαίσιο των παραγωγικών διαδικασιών, ο SPC λειτουργεί ως βασικός μηχανισμός υποστήριξης του ποιοτικού ελέγχου, ενσωματώνοντας πρακτικές συνεχούς παρακολούθησης και ανάλυσης δεδομένων διεργασίας. Ο ΣΕΔ εφαρμόζεται ευρέως σε περιβάλλοντα όπου απαιτείται υψηλός βαθμός αξιοπιστίας, ενώ αποτελεί αναπόσπαστο τμήμα σύγχρονων προσεγγίσεων βελτίωσης, όπως το Six Sigma (6σ) και το Lean Manufacturing (Montgomery, 2019; Boaventura et al., 2022).

Επιπλέον, ο SPC μπορεί να θεωρηθεί ως ένα δομημένο πλαίσιο παρακολούθησης διεργασιών (framework), το οποίο υποστηρίζει τη λήψη αποφάσεων σε συστήματα Ποιότητας, μέσω της συστηματικής συλλογής και στατιστικής ανάλυσης δεδομένων. Με την εφαρμογή του SPC, γίνεται δυνατή η αξιολόγηση της συμπεριφοράς μιας διεργασίας και η σύγκρισή της σε σχέση με την αναμενόμενη της απόδοση, συμβάλλοντας στον έγκαιρο εντοπισμό αποκλίσεων που

ενδέχεται να επηρεάσουν την ποιότητα του τελικού προϊόντος (Zhou et al., 2025; Kazmi et al., 2025). Η προσέγγιση αυτή ενισχύει τη δυνατότητα πρόληψης σφαλμάτων και βελτιώνει τη συνολική αποδοτικότητα των παραγωγικών διαδικασιών.

Στο σύγχρονο βιομηχανικό περιβάλλον, όπου η πολυπλοκότητα των διεργασιών και ο όγκος των διαθέσιμων δεδομένων αυξάνονται συνεχώς, ο SPC εξακολουθεί να αποτελεί τη μεθοδολογική βάση για την κατανόηση, την απεικόνιση και τον έλεγχο της μεταβλητότητας. Παράλληλα, η εξέλιξη των απαιτήσεων δημιουργεί το κατάλληλο εφελκυστικό για την ανάπτυξη και αξιοποίηση πιο προηγμένων τεχνικών ανάλυσης, όπως η ενσωμάτωση μεθόδων Μηχανικής Μάθησης (Zan et al., 2020; Hric & Sabahno, 2024).

### 2.1.1 Ιστορική εξέλιξη του Στατιστικού Ελέγχου Διεργασιών

Η ανάπτυξη του Στατιστικού Ελέγχου Διεργασιών (SPC) συνδέεται άμεσα με την ανάγκη συστηματικού ελέγχου της ποιότητας στη βιομηχανική παραγωγή κατά τις αρχές του 20ού αιώνα. Η αυξανόμενη πολυπλοκότητα των παραγωγικών διαδικασιών και η απαίτηση για σταθερή ποιότητα των παραγόμενων προϊόντων οδήγησαν στην ανάγκη για μεθόδους που να επιτρέπουν την παρακολούθηση και τον έλεγχο της μεταβλητότητας με ποσοτικό τρόπο. Σε αυτό τον άξονα, καθοριστική υπήρξε η συμβολή του Walter A. Shewhart, μηχανικού της Bell Laboratories κατά τη δεκαετία του 1920, ο οποίος εισήγαγε την έννοια του στατιστικού ελέγχου και ανέπτυξε τα πρώτα διαγράμματα ελέγχου, θέτοντας τα θεμέλια για τη διάκριση μεταξύ τυχαίων και ειδικών αιτίων μεταβλητότητας (Shewhart, 1931). Η προσέγγιση αυτή αποτέλεσε τη βάση για τη μετάβαση από τον εκ των υστέρων έλεγχο ποιότητας (inspection-based quality control) σε μία πιο προληπτική προσέγγιση της ποιότητας.

Στη συνέχεια, η εξέλιξη της Διαχείρισης Ποιότητας συνοδεύτηκε από σημαντικά ορόσημα που ενίσχυσαν τον ρόλο του SPC ως βασικού εργαλείου ελέγχου και βελτίωσης διεργασιών. Κατά τη δεκαετία του 1950, οι αρχές του Shewhart επεκτάθηκαν και διαδόθηκαν ευρέως μέσω του έργου του W. Edwards Deming και του Joseph M. Juran, συμβάλλοντας στην υιοθέτηση του SPC από τη ιαπωνική βιομηχανία και στη διαμόρφωση της φιλοσοφίας της συνεχούς βελτίωσης (Continuous Improvement) (Deming, 1986; Juran & Godfrey, 1999). Τις επόμενες δεκαετίες, ο ΣΕΔ ενσωματώθηκε σε ολοκληρωμένα συστήματα Διοίκησης Ολικής Ποιότητας (Total Quality Management – TQM), ενώ από τη δεκαετία του 1980 και έπειτα αποτελεί βασικό πυλώνα της μεθοδολογίας Six Sigma, με έμφαση στη μείωση της μεταβλητότητας και τη

βελτίωση της απόδοσης (Montgomery, 2019; Evans & Lindsay, 2020). Παράλληλα, η αυξανόμενη διαθεσιμότητα δεδομένων και η πολυπλοκότητα των σύγχρονων διεργασιών οδήγησαν στην ανάπτυξη πολυμεταβλητών προσεγγίσεων SPC, ικανών να λαμβάνουν υπόψη πολλαπλά χαρακτηριστικά ποιότητας ταυτόχρονα (Bersimis et al., 2007).

Στο σύγχρονο πλαίσιο της Βιομηχανίας 4.0 και της λεγόμενης Quality 4.0, ο Στατιστικός Έλεγχος Διεργασιών εξακολουθεί να αποτελεί θεμελιώδη μεθοδολογική βάση, πάνω στην οποία αναπτύσσονται νέες προσεγγίσεις, αξιοποιώντας προηγμένες τεχνικές ανάλυσης και τεχνητής νοημοσύνης (Escobar et al., 2025). Η ιστορική αυτή εξέλιξη αναδεικνύει τη μετάβαση από απλές στατιστικές τεχνικές σε ολοκληρωμένα συστήματα παρακολούθησης και βελτίωσης διεργασιών.

### 2.1.2 Βασικές μέθοδοι και λειτουργία του Στατιστικού Ελέγχου Διεργασιών

Ο Στατιστικός Έλεγχος Διεργασιών βασίζεται σε μία συστηματική διαδικασία παρακολούθησης της συμπεριφοράς μιας διεργασίας μέσω στατιστικών δεικτών και διαγραμμάτων ελέγχου. Η λειτουργική διαδικασία του SPC ξεκινά με τη συλλογή δεδομένων από τη διεργασία, τα οποία αντιπροσωπεύουν κρίσιμα χαρακτηριστικά ποιότητας. Στη συνέχεια, τα δεδομένα αυτά μετατρέπονται σε κατάλληλους στατιστικούς δείκτες, π.χ. μέσος όρος, εύρος ή διασπορά, και απεικονίζονται σε Διαγράμματα Ελέγχου (control charts). Η συνεχής παρακολούθηση των διαγραμμάτων επιτρέπει την αξιολόγηση της συμπεριφοράς της διεργασίας σε σχέση με προκαθορισμένα όρια ελέγχου (Control Limits), ενώ η ανίχνευση αποκλίσεων ή μη τυχαίων ακολουθιών οδηγεί στη λήψη αποφάσεων σχετικά με την κατάσταση της διεργασίας και την ανάγκη επέμβασης για την επαναφορά της διεργασίας υπό στατιστικό έλεγχο.

Τα διαγράμματα ελέγχου (ΔΕ) διακρίνονται σε δύο βασικές κατηγορίες, ανάλογα με τη φύση των δεδομένων που αναλύονται: τα διαγράμματα μεταβλητών (variables charts), τα οποία χρησιμοποιούνται για συνεχή δεδομένα, όπως μετρήσεις φυσικών μεγεθών και τα διαγράμματα ιδιοτήτων (attributes charts). Τα πρώτα εφαρμόζονται συνήθως σε συνεχή δεδομένα ενώ τα δεύτερα εφαρμόζονται σε διακριτά δεδομένα, όπως ο αριθμός ή το ποσοστό ελαττωματικών προϊόντων (Montgomery, 2019). Στα διαγράμματα μεταβλητών περιλαμβάνονται κυρίως τα διαγράμματα  $\bar{X}$ -R,  $\bar{X}$ -S και I-MR, τα οποία χρησιμοποιούνται για την παρακολούθηση του μέσου όρου και της διασποράς της διεργασίας. Αντίστοιχα, στα

διαγράμματα ιδιοτήτων ανήκουν τα διαγράμματα  $p$ ,  $np$ ,  $c$  και  $u$ , τα οποία βασίζονται σε κατανομές Bernoulli, διωνυμικές ή Poisson, ανάλογα με τη φύση των δεδομένων.



Εικόνα 1 Είδη Διαγραμμάτων Ελέγχου SPC

Πέρα από τα κλασικά διαγράμματα Shewhart, έχουν αναπτυχθεί και πιο προηγμένες μέθοδοι ελέγχου με αυξημένη ευαισθησία στην ανίχνευση μικρών και σταδιακών μεταβολών της διεργασίας. Τα διαγράμματα EWMA (Exponentially Weighted Moving Average) και CUSUM (Cumulative Sum) αποτελούν χαρακτηριστικά παραδείγματα, καθώς ενσωματώνουν πληροφορία από προηγούμενες παρατηρήσεις, επιτρέποντας την έγκαιρη ανίχνευση σταδιακών μεταβολών (Montgomery, 2019; Kazmi et al., 2025). Επιπλέον, η αυξανόμενη πολυπλοκότητα των σύγχρονων διεργασιών έχει οδηγήσει στην ανάπτυξη πολυμεταβλητών τεχνικών SPC, όπως το διάγραμμα Hotelling's  $T^2$ , το οποίο επιτρέπει την ταυτόχρονη παρακολούθηση πολλαπλών συσχετισμένων χαρακτηριστικών ποιότητας (Bersimis et al., 2007).

Η μαθηματική θεμελίωση των διαγραμμάτων ελέγχου βασίζεται στον υπολογισμό στατιστικών δεικτών και στον καθορισμό ορίων ελέγχου. Τα όρια αυτά ορίζονται συνήθως ως τρία τυπικά σφάλματα ( $\pm 3\sigma$ ) γύρω από τη μέση τιμή της διεργασίας και αποτελούν κριτήριο διάκρισης μεταξύ κανονικής και μη κανονικής λειτουργίας (Shewhart, 1931; Montgomery, 2019) ή

διεργασίας εντός ή εκτός στατιστικού ελέγχου. Η προσέγγιση αυτή προϋποθέτει, σε πολλές περιπτώσεις, την ικανοποίηση βασικών στατιστικών υποθέσεων, όπως η κανονικότητα της κατανομής και η ανεξαρτησία των παρατηρήσεων. Πέρα από την απλή υπέρβαση των ορίων ελέγχου, ιδιαίτερη σημασία έχει η αναγνώριση μη τυχαίων προτύπων στη συμπεριφορά των δεδομένων, όπως τάσεις, επαναλαμβανόμενα μοτίβα ή συνεχόμενες αποκλίσεις προς μία κατεύθυνση. Στο πλαίσιο αυτό, χρησιμοποιούνται κανόνες ερμηνείας, όπως οι κανόνες Western Electric, οι οποίοι συμβάλλουν στην έγκαιρη ανίχνευση αποκλίσεων ακόμη και όταν δεν υπάρχουν παρατηρήσεις εκτός ορίων ελέγχου.

Για την αξιολόγηση της απόδοσης των διαγραμμάτων ελέγχου χρησιμοποιείται ευρέως ο δείκτης Average Run Length (ARL), ο οποίος ορίζεται ως ο αναμενόμενος αριθμός δειγμάτων που απαιτούνται μέχρι την ανίχνευση μιας απόκλισης. Τυπικά, διακρίνονται δύο περιπτώσεις: το  $ARL_0$ , που αντιστοιχεί στην περίπτωση όπου η διεργασία βρίσκεται υπό στατιστικό έλεγχο, και το  $ARL_1$ , που αφορά την περίπτωση όπου έχει επέλθει μεταβολή στη διεργασία. Μαθηματικά, ο δείκτης ARL εκφράζεται ως:

$$ARL = \frac{1}{P(signal)}$$

,όπου  $P(signal)$  είναι η πιθανότητα εμφάνισης σήματος σε κάθε δείγμα.

Ο δείκτης αυτός αποτελεί βασικό μέτρο αξιολόγησης της ευαισθησίας και της αποτελεσματικότητας των διαγραμμάτων ελέγχου, καθώς συνδέεται άμεσα με την ταχύτητα ανίχνευσης αποκλίσεων.

Τέλος, οι μέθοδοι SPC εφαρμόζονται ευρέως σε βιομηχανικά περιβάλλοντα, τόσο σε επίπεδο ελέγχου παρτίδων όσο και σε συνθήκες συνεχούς παρακολούθησης της παραγωγικής διαδικασίας. Η χρήση τους εκτείνεται από τον ενδιάμεσο έλεγχο διεργασιών έως τον τελικό ποιοτικό έλεγχο προϊόντων, συμβάλλοντας στη μείωση της μεταβλητότητας, τη βελτίωση της ποιότητας και τη μείωση του κόστους κακής ποιότητας (Boaventura et al., 2022; Escobar et al., 2025).

### 2.1.3 Περιορισμοί και περιθώρια βελτίωσης του SPC

Οι κλασικές μέθοδοι Στατιστικού Ελέγχου Διεργασιών βασίζονται σε συγκεκριμένες στατιστικές και θεωρητικές παραδοχές, οι οποίες επηρεάζουν άμεσα την αποτελεσματικότητά

τους σε πραγματικές βιομηχανικές εφαρμογές. Τα περισσότερα παραδοσιακά διαγράμματα ελέγχου προϋποθέτουν ότι τα δεδομένα ακολουθούν κανονική κατανομή, ότι οι παρατηρήσεις είναι στατιστικά ανεξάρτητες και ότι η διεργασία παραμένει σταθερή κατά τη διάρκεια της παρακολούθησης (Montgomery, 2019). Ωστόσο, σε πολλές σύγχρονες διεργασίες οι υποθέσεις αυτές δεν ικανοποιούνται πλήρως, ιδιαίτερα όταν εμφανίζονται μη γραμμικές σχέσεις, αλληλεπιδράσεις μεταξύ μεταβλητών ή δυναμική μεταβολή των συνθηκών λειτουργίας. Ως αποτέλεσμα, τα παραδοσιακά SPC εργαλεία ενδέχεται να εμφανίζουν μειωμένη ευαισθησία στην ανίχνευση αποκλίσεων ή αυξημένη πιθανότητα ψευδών συναγερμών (Zan et al., 2020; Hric & Sabahno, 2024).

Επιπλέον, σημαντικός περιορισμός των μονομεταβλητών προσεγγίσεων SPC αποτελεί η αδυναμία αποτελεσματικής παρακολούθησης πολλαπλών συσχετισμένων χαρακτηριστικών ποιότητας ταυτόχρονα. Στις περισσότερες βιομηχανικές διεργασίες, οι μεταβλητές της παραγωγής παρουσιάζουν σημαντικό βαθμό αλληλεπίδρασης και συσχέτισης, γεγονός που δυσχεραίνει σημαντικά την ανεξάρτητη αξιολόγησή τους μέσω μεμονωμένων διαγραμμάτων ελέγχου. Παρότι οι πολυμεταβλητές τεχνικές SPC επιχειρούν να αντιμετωπίσουν το πρόβλημα αυτό, η αυξανόμενη πολυπλοκότητα των δεδομένων δημιουργεί πρόσθετες απαιτήσεις σε επίπεδο υπολογιστικής επεξεργασίας και ερμηνείας των αποτελεσμάτων (Bersimis et al., 2007; Biegel et al., 2022).

Παράλληλα, τα παραδοσιακά διαγράμματα ελέγχου παρουσιάζουν περιορισμούς στην ανίχνευση μικρών ή σταδιακών μεταβολών της διεργασίας. Η έγκαιρη ανίχνευση αποκλίσεων εξαρτάται σε μεγάλο βαθμό από την ευαισθησία του χρησιμοποιούμενου διαγράμματος, καθώς και από την ισορροπία μεταξύ ταχύτητας ανίχνευσης και πιθανότητας εμφάνισης ψευδών συναγερμών (false alarms) ή σφαλμάτων τύπου I. Παρότι τεχνικές όπως τα EWMA και CUSUM βελτιώνουν την ικανότητα ανίχνευσης μικρών μεταβολών, η διαδικασία αναγνώρισης μη τυχαίων προτύπων εξακολουθεί να βασίζεται σε προκαθορισμένους κανόνες, όπως οι Western Electric Rules, οι οποίοι δεν είναι πάντοτε επαρκείς για πολύπλοκα ή υψηλής διάστασης δεδομένα (Kazmi et al., 2025; Zhou et al., 2025).

Οι προκλήσεις αυτές εντείνονται στο σύγχρονο περιβάλλον της Βιομηχανίας 4.0 και των συστημάτων Smart Manufacturing, όπου οι παραγωγικές διεργασίες παράγουν συνεχώς μεγάλους όγκους πολυμεταβλητών δεδομένων σε πραγματικό χρόνο. Η αυξημένη

διαθεσιμότητα δεδομένων, η ανάγκη για online παρακολούθηση και η απαίτηση για ταχύτερη και ορθότερη λήψη αποφάσεων δημιουργούν περιορισμούς στις παραδοσιακές SPC προσεγγίσεις, οι οποίες σχεδιάστηκαν κυρίως για στατικότερα (Escobar et al., 2025) και λιγότερο πολύπλοκα περιβάλλοντα (Chou et al., 2020).

Η ανάγκη για προσαρμοστικότητα, αποτελεσματικότερη αναγνώριση προτύπων και αυτοματοποιημένη ανάλυση οδηγεί σταδιακά στην υιοθέτηση πιο δεδομενοκεντρικών (data-driven) προσεγγίσεων παρακολούθησης διεργασιών. Στο πλαίσιο αυτό, τα τελευταία χρόνια αναπτύσσονται προσεγγίσεις AI-SPC και ML-SPC, οι οποίες επιχειρούν να συνδυάσουν τη θεωρητική βάση του Στατιστικού Ελέγχου Διεργασιών με τεχνικές Μηχανικής Μάθησης και προηγμένης ανάλυσης δεδομένων. Οι προσεγγίσεις αυτές στοχεύουν στη βελτίωση της ικανότητας αναγνώρισης σύνθετων προτύπων, στην προσαρμογή σε δυναμικά περιβάλλοντα και στην ενίσχυση της προγνωστικής ικανότητας των συστημάτων ποιότητας (Bonaventura et al., 2022). Ως αποτέλεσμα, η Μηχανική Μάθηση εμφανίζεται πλέον ως μία πολλά υποσχόμενη κατεύθυνση για την εξέλιξη των σύγχρονων συστημάτων παρακολούθησης και ελέγχου διεργασιών.

## 2.2 Μηχανική Μάθηση στη Διαχείριση Ποιότητας

Η ραγδαία ανάπτυξη των τεχνολογιών Industry 4.0 και η συνεχής αύξηση του όγκου των διαθέσιμων παραγωγικών δεδομένων έχουν οδηγήσει στη σταδιακή ενσωμάτωση τεχνικών Μηχανικής Μάθησης (Machine Learning – ML) στα σύγχρονα βιομηχανικά συστήματα. Η μετάβαση από παραδοσιακές, αντιδραστικές πρακτικές ποιοτικού ελέγχου σε πιο προγνωστικές και δεδομενοκεντρικές προσεγγίσεις έχει ενισχύσει σημαντικά τον ρόλο της τεχνητής νοημοσύνης στη Διαχείριση Ποιότητας (Escobar et al., 2025). Στο πλαίσιο αυτό, η Μηχανική Μάθηση αξιοποιείται ολοένα και περισσότερο για την αυτοματοποιημένη ανάλυση δεδομένων, την παρακολούθηση διεργασιών και την ανίχνευση αποκλίσεων σε πραγματικό χρόνο (Zan et al., 2020).

Επιπλέον, η αυξανόμενη διαθεσιμότητα πολυδιάστατων βιομηχανικών δεδομένων (High-Dimensional Industrial Data) από αισθητήρες, συστήματα αυτοματισμού και διαδικασίες ελέγχου ποιότητας δημιουργεί νέες δυνατότητες αλλά και αυξημένες απαιτήσεις ως προς την επεξεργασία και αξιοποίηση των δεδομένων αυτών. Οι τεχνικές ML προσφέρουν τη δυνατότητα ανάλυσης μεγάλων όγκων δεδομένων και αναγνώρισης σύνθετων ή μη γραμμικών

προτύπων, τα οποία συχνά δεν είναι εύκολο να εντοπιστούν μέσω παραδοσιακών στατιστικών εργαλείων (Hric & Sabahno, 2024). Επιπλέον, οι δυνατότητες παρακολούθησης σε πραγματικό χρόνο (Real-Time Monitoring) και προσαρμοστικής παρακολούθησης (Adaptive Monitoring) ενισχύουν σημαντικά τη λειτουργία των συστημάτων ποιότητας σε δυναμικά βιομηχανικά περιβάλλοντα.

Όπως τονίζουν οι (Wang & Liu, 2024), στο πεδίο της Διαχείρισης Ποιότητας, η Μηχανική Μάθηση χρησιμοποιείται ως εργαλείο υποστήριξης διαδικασιών όπως:

- η πρόληψη ελαττωμάτων (Defect Prevention),
- η βελτιστοποίηση διεργασιών (Process Optimization),
- η μείωση του κόστους κακής ποιότητας και
- η υποστήριξη της λήψης αποφάσεων .

Η αξιοποίηση των δεδομένων παραγωγής (ή διεργασιών) επιτρέπει την ανάπτυξη πιο αποτελεσματικών συστημάτων διαχείρισης ποιότητας βασισμένων σε δεδομένα (Data-Driven Quality Management) και προγνωστικής διαχείρισης ποιότητας (Predictive Quality Management), τα οποία μπορούν να συμβάλουν στην έγκαιρη ανίχνευση αποκλίσεων και στη βελτίωση της συνολικής αποδοτικότητας της παραγωγής.

Οι εφαρμογές της Μηχανικής Μάθησης στη Διαχείριση Ποιότητας περιλαμβάνουν ένα ευρύ φάσμα λειτουργιών, όπως η ταξινόμηση ελαττωμάτων (Classification), η πρόβλεψη μεταβλητών διεργασίας (Regression), η ομαδοποίηση δεδομένων (Clustering) και η ανίχνευση ανωμαλιών (Anomaly Detection) (Biegel et al., 2022). Οι τεχνικές αυτές χρησιμοποιούνται σε εφαρμογές παρακολούθησης διεργασιών (Process Monitoring), ανίχνευσης σφαλμάτων (Fault Detection) και προγνωστικής συντήρησης (Predictive Maintenance), ιδιαίτερα σε σύνθετα βιομηχανικά περιβάλλοντα όπου η ποσότητα και η πολυπλοκότητα των δεδομένων υπερβαίνουν τις δυνατότητες των παραδοσιακών εργαλείων ελέγχου (Zan et al., 2020).

Η εμφάνιση προσεγγίσεων «ενισχυμένου» SPC μέσω τεχνητής νοημοσύνης (AI-Enhanced SPC) και SPC βασισμένου σε Μηχανική Μάθηση (ML-SPC) αποτελεί φυσική εξέλιξη της προσπάθειας ενίσχυσης των παραδοσιακών συστημάτων Στατιστικού Ελέγχου Διεργασιών. Οι προσεγγίσεις αυτές επιχειρούν να συνδυάσουν τη θεωρητική βάση του SPC με τεχνικές Μηχανικής Μάθησης, επιτρέποντας πιο προσαρμοστική παρακολούθηση διεργασιών. Σε

αντίθεση με τα παραδοσιακά συστήματα που βασίζονται κυρίως σε στατικά όρια ελέγχου, τα διαγράμματα ελέγχου βασισμένα σε ML (ML-Based Control Charts), ως συστήματα ευφούς παρακολούθησης διεργασιών, αξιοποιούν προηγμένες αλγοριθμικές τεχνικές αναγνώρισης προτύπων και ανίχνευσης αποκλίσεων, βελτιώνοντας τη δυνατότητα παρακολούθησης πολύπλοκων διεργασιών (Boaventura et al., 2022; Kazmi et al., 2025).

Παρά τις σημαντικές δυνατότητες των τεχνικών ML, η εφαρμογή τους σε βιομηχανικά περιβάλλοντα συνοδεύεται από σημαντικές τεχνικές και οργανωτικές απαιτήσεις. Η ανάπτυξη αποτελεσματικών μοντέλων προϋποθέτει κατάλληλες διαδικασίες συλλογής, διαχείρισης και προεπεξεργασίας δεδομένων, διαδικασίες επιλογής και μετασχηματισμού χαρακτηριστικών, καθώς και επαρκή εκπαίδευση και ορθή παραμετροποίηση των μοντέλων. Παράλληλα, παράγοντες όπως η ύπαρξη θορύβου στα δεδομένα, η ανισορροπία μεταξύ κλάσεων (imbalanced classes) και οι αυξημένες υπολογιστικές απαιτήσεις μπορούν να επηρεάσουν σημαντικά την απόδοση των μοντέλων και τη δυνατότητα εφαρμογής τους σε πραγματικά βιομηχανικά περιβάλλοντα (Chou et al., 2020; Biegel et al., 2022). Ως αποτέλεσμα, η επιτυχής ενσωμάτωση της Μηχανικής Μάθησης στη Διαχείριση Ποιότητας προϋποθέτει όχι μόνο τεχνολογική υποστήριξη και υπολογιστική ισχύ, αλλά και κατάλληλη προσαρμογή των διαδικασιών συλλογής και αξιοποίησης δεδομένων.

### 2.2.1 Υπάρχουσα βιβλιογραφία σε ML-SPC

Η αυξανόμενη πολυπλοκότητα των σύγχρονων βιομηχανικών διεργασιών και η συνεχής αύξηση του όγκου των διαθέσιμων παραγωγικών δεδομένων έχουν στρέψει τους ερευνητές στη διεθνή βιβλιογραφία στην ανάπτυξη νέων προσεγγίσεων παρακολούθησης διεργασιών, οι οποίες συνδυάζουν τις αρχές του Στατιστικού Ελέγχου Διεργασιών με τεχνικές Μηχανικής Μάθησης (ML). Η μετάβαση από τα παραδοσιακά στατιστικά μοντέλα προς πιο προσαρμοστικά και δεδομενοκεντρικά συστήματα παρακολούθησης αποτελεί μία από τις βασικότερες τάσεις της σύγχρονης έρευνας στον τομέα της Διαχείρισης Ποιότητας (Escobar et al., 2025). Οι προσεγγίσεις αυτές στοχεύουν κυρίως στη βελτίωση της ικανότητας ανίχνευσης αποκλίσεων, στη διαχείριση πολυμεταβλητών δεδομένων και στην υποστήριξη παρακολούθησης διεργασιών σε πραγματικό χρόνο.

Σημαντικό μέρος της σύγχρονης βιβλιογραφίας επικεντρώνεται στην ανάπτυξη διαγραμμάτων ελέγχου βασισμένων σε τεχνικές Μηχανικής Μάθησης (ML-Based Control Charts). Οι Wang

και Liu (2024) παρουσίασαν ένα πολυμεταβλητό διάγραμμα ελέγχου βασισμένο στον αλγόριθμο Isolation Forest, με στόχο τη βελτίωση της ανίχνευσης ανωμαλιών σε πολύπλοκα βιομηχανικά δεδομένα. Αντίστοιχα, οι Kazmi και Noor-ul-Amin (2025) πρότειναν μία υβριδική προσέγγιση συνδυασμού διαγραμμάτων EWMA με τεχνικές ML για πολυμεταβλητή παρακολούθηση διεργασιών, επιτυγχάνοντας αυξημένη ευαισθησία στην ανίχνευση μικρών μεταβολών. Επιπλέον, οι Hric και Sabahno (2024) ανέπτυξαν διαγράμματα ελέγχου βασισμένα σε μοντέλα Γενικευμένων Γραμμικών Μοντέλων (Generalized Linear Models – GLM) με ενσωμάτωση τεχνικών Μηχανικής Μάθησης, προτείνοντας πιο ευέλικτες δομές ελέγχου για σύνθετες διεργασίες.

Παράλληλα, στη βιβλιογραφία παρουσιάζεται αυξανόμενο ενδιαφέρον σχετικά με την αξιοποίηση τεχνικών Βαθιάς Μάθησης (Deep Learning) και πολυμεταβλητής παρακολούθησης διεργασιών. Οι Zan et al. (2020) πρότειναν ένα πλαίσιο ευφυούς SPC βασισμένο σε Deep Learning, με στόχο την αποτελεσματικότερη αναγνώριση πολύπλοκων προτύπων συμπεριφοράς διεργασιών. Αντίστοιχα, οι Biegel et al. (2022) εφάρμοσαν τεχνικές Deep Learning για πολυμεταβλητό στατιστικό έλεγχο σε διεργασία διαμόρφωσης μεταλλικών φύλλων, αναδεικνύοντας τις δυνατότητες διαχείρισης πολυδιάστατων βιομηχανικών δεδομένων. Επιπλέον, οι Zhou et al. (2025) παρουσίασαν προσεγγίσεις παρακολούθησης χωροχρονικών ροών δεδομένων μέσω Διαγραμμάτων Ελέγχου ML, επεκτείνοντας την εφαρμογή του ML-SPC σε ακόμη πιο σύνθετα περιβάλλοντα δεδομένων.

Τα αποτελέσματα που αναφέρονται στη σχετική βιβλιογραφία υποδεικνύουν ότι οι ML-SPC προσεγγίσεις μπορούν να προσφέρουν σημαντικά πλεονεκτήματα σε σχέση με τα παραδοσιακά συστήματα SPC. Μεταξύ αυτών περιλαμβάνονται η βελτιωμένη ικανότητα αναγνώρισης σύνθετων ή μη γραμμικών προτύπων, η αυξημένη ευαισθησία στην ανίχνευση αποκλίσεων, η μείωση ψευδών συναγερμών και η δυνατότητα προσαρμοστικής παρακολούθησης διεργασιών (Boaventura et al., 2022; Wang & Liu, 2024). Παράλληλα, οι τεχνικές αυτές συνδέονται με ευρύτερους στόχους της Διαχείρισης Ποιότητας, όπως η βελτίωση της αποδοτικότητας της παραγωγής, η υποστήριξη της λήψης αποφάσεων και η μείωση του κόστους κακής ποιότητας (Escobar et al., 2025).

Παρά το αυξημένο ερευνητικό ενδιαφέρον, η βιβλιογραφία εξακολουθεί να αναδεικνύει ορισμένους περιορισμούς και ερευνητικά κενά. Πολλές προσεγγίσεις διερεύνησης -

αξιολόγησης που πραγματεύονται ML-SPC αξιολογούνται κυρίως σε θεωρητικά ή περιορισμένα περιβάλλοντα δεδομένων, ενώ παραμένει περιορισμένος ο αριθμός μελετών που εξετάζουν την εφαρμογή τους σε πραγματικές βιομηχανικές διεργασίες μεγάλης κλίμακας (Chou et al., 2020). Επιπλέον, ζητήματα όπως η αυξημένη υπολογιστική πολυπλοκότητα, η ανάγκη κατάλληλης παραμετροποίησης των μοντέλων και η δυσκολία γενίκευσης των αποτελεσμάτων εξακολουθούν να αποτελούν αντικείμενο ερευνητικής διερεύνησης (Biegel et al., 2022; Hric & Sabahno, 2024). Στο πλαίσιο αυτό, προκύπτει η ανάγκη για περαιτέρω συγκριτική διερεύνηση μεταξύ παραδοσιακών μεθόδων SPC και προσεγγίσεων ML-SPC, με αξιοποίηση πραγματικών βιομηχανικών δεδομένων και έμφαση τόσο στην τεχνική απόδοση των μοντέλων όσο και στη συνολική συμβολή τους στη βελτίωση των συστημάτων Διαχείρισης Ποιότητας.

## 2.2. Επισκόπηση βασικών αλγορίθμων Μηχανικής Μάθησης

Η εφαρμογή της Μηχανικής Μάθησης στη Διαχείριση Ποιότητας μπορεί να προσεγγιστεί μέσα από διαφορετικές κατηγορίες αλγορίθμων, ανάλογα με τη φύση των διαθέσιμων δεδομένων και τον στόχο που τίθεται στην ανάλυση. Στο πλαίσιο της παρούσας εργασίας, ιδιαίτερο ενδιαφέρον παρουσιάζουν τόσο οι μέθοδοι επιβλεπόμενης μάθησης (Supervised Learning), οι οποίες αξιοποιούν γνωστές ετικέτες εξόδου, όσο και οι μέθοδοι μη επιβλεπόμενης μάθησης (Unsupervised Learning), οι οποίες στοχεύουν στην αναγνώριση δομών, προτύπων ή ανωμαλιών χωρίς προκαθορισμένη ταξινόμηση. Η διάκριση αυτή είναι σημαντική για εφαρμογές SPC–ML, καθώς οι επιβλεπόμενες μέθοδοι μπορούν να χρησιμοποιηθούν για πρόβλεψη κατηγοριών ποιότητας, ενώ οι μη επιβλεπόμενες μέθοδοι είναι ιδιαίτερα χρήσιμες για διερευνητική ανάλυση και ανίχνευση ασυνήθιστης συμπεριφοράς διεργασιών (Chou et al., 2020; Boaventura et al., 2022).

Στην κατηγορία της επιβλεπόμενης μάθησης, η Λογιστική Παλινδρόμηση (Logistic Regression) αποτελεί μία από τις απλούστερες και πιο εύκολες στην ερμηνεία μεθόδους δυαδικής ταξινόμησης. Χρησιμοποιείται για την εκτίμηση της πιθανότητας μία παρατήρηση να ανήκει σε συγκεκριμένη κατηγορία και μπορεί να λειτουργήσει ως βασικό μοντέλο αναφοράς σε προβλήματα αποδοχής ή απόρριψης προϊόντων (Pass/Fail). Παρότι βασίζεται σε σχετικά απλές παραδοχές, η σαφήνεια και η ευκολία ερμηνείας της την καθιστούν χρήσιμη ως αρχικό σημείο σύγκρισης έναντι πιο σύνθετων αλγορίθμων (Cox, 1958).

Μια επίσης απλή μέθοδος ταξινόμησης είναι ο αλγόριθμος των  $k$ -Πλησιέστερων Γειτόνων ( $k$ -Nearest Neighbors – KNN). Ο KNN βασίζεται στην υπόθεση ότι παρατηρήσεις με παρόμοια χαρακτηριστικά είναι πιθανό να ανήκουν στην ίδια κατηγορία. Η ταξινόμηση πραγματοποιείται με βάση την εγγύτητα μιας νέας παρατήρησης προς τις ήδη γνωστές παρατηρήσεις του συνόλου εκπαίδευσης. Η μέθοδος είναι εύκολη στην κατανόηση και μπορεί να χρησιμοποιηθεί ως απλό μοντέλο αναφοράς, ωστόσο είναι ιδιαίτερα ευαίσθητη στην κλίμακα των μεταβλητών και στη διάσταση των δεδομένων, γεγονός που καθιστά απαραίτητη την κατάλληλη κανονικοποίηση (Cover & Hart, 1967).

Τα Δέντρα Απόφασης (Decision Trees) αποτελούν μία πιο ευέλικτη προσέγγιση επιβλεπόμενης μάθησης, καθώς βασίζονται στη διαδοχική διάσπαση των δεδομένων με βάση κανόνες απόφασης. Η δομή τους επιτρέπει την αποτύπωση μη γραμμικών σχέσεων και αλληλεπιδράσεων μεταξύ μεταβλητών, ενώ παράλληλα παρέχει σχετικά υψηλή ερμηνευσιμότητα. Για τον λόγο αυτό, τα Δέντρα Απόφασης είναι ιδιαίτερα χρήσιμα σε περιβάλλοντα ποιότητας, όπου η κατανόηση των παραγόντων που οδηγούν σε μία συγκεκριμένη ταξινόμηση είναι συχνά εξίσου σημαντική με την ακρίβεια του μοντέλου. Ωστόσο, η μέθοδος μπορεί να παρουσιάσει υπερπροσαρμογή (Overfitting), ιδιαίτερα όταν το δέντρο είναι πολύ βαθύ ή όταν δεν εφαρμόζονται κατάλληλοι περιορισμοί πολυπλοκότητας (Quinlan, 1986).

Τα Τυχαία Δάση (Random Forests) επεκτείνουν τη λογική των Δέντρων Απόφασης (Decision Trees), συνδυάζοντας μεγάλο αριθμό δέντρων σε ένα σύνολο μοντέλων. Η τεχνική αυτή μειώνει την πιθανότητα υπερπροσαρμογής και βελτιώνει τη σταθερότητα της πρόβλεψης, καθώς η τελική απόφαση προκύπτει από τον συνδυασμό των επιμέρους δέντρων. Τα Random Forests είναι κατάλληλα για προβλήματα ταξινόμησης στη Διαχείριση Ποιότητας, καθώς μπορούν να χειριστούν μη γραμμικές σχέσεις, αλληλεπιδράσεις μεταξύ μεταβλητών και δεδομένα με σύνθετη δομή. Επιπλέον, παρέχουν δυνατότητα εκτίμησης της σχετικής σημαντικότητας των μεταβλητών, στοιχείο ιδιαίτερα χρήσιμο σε συστήματα Διαχείρισης Ποιότητας, όπου απαιτείται κατανόηση των παραγόντων που επηρεάζουν περισσότερο την τελική ταξινόμηση (Breiman, 2001; Boaventura et al., 2022).

Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines – SVM) αποτελούν πιο σύνθετη μέθοδο επιβλεπόμενης ταξινόμησης, η οποία στοχεύει στον προσδιορισμό του

βέλτιστου ορίου διαχωρισμού μεταξύ κατηγοριών. Μέσω της χρήσης συναρτήσεων πυρήνα (Kernel Functions), οι SVM μπορούν να χειριστούν και μη γραμμικά προβλήματα ταξινόμησης, μετασχηματίζοντας τα δεδομένα σε χώρους υψηλότερης διάστασης. Η μέθοδος είναι ιδιαίτερα ισχυρή σε προβλήματα όπου οι κλάσεις μπορούν να διαχωριστούν με σαφές όριο απόφασης, αλλά απαιτεί προσεκτική παραμετροποίηση και κατάλληλη κλιμάκωση των δεδομένων (Cortes & Vapnik, 1995; Zan et al., 2020).

Στην κατηγορία της μη επιβλεπόμενης μάθησης, η Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis – PCA) αποτελεί βασική τεχνική μείωσης διαστάσεων και διερευνητικής ανάλυσης πολυμεταβλητών δεδομένων. Η PCA μετασχηματίζει ένα σύνολο συσχετισμένων μεταβλητών σε ένα μικρότερο σύνολο νέων, ασυσχέτιστων μεταβλητών, οι οποίες ονομάζονται κύριες συνιστώσες. Με τον τρόπο αυτό, καθίσταται δυνατή η απλοποίηση πολύπλοκων συνόλων δεδομένων, η οπτικοποίηση της δομής τους και η αναγνώριση πιθανών ασυνήθιστων παρατηρήσεων. Στο πλαίσιο της παρακολούθησης διεργασιών, η PCA έχει ιδιαίτερη αξία, καθώς μπορεί να υποστηρίξει την ανάλυση πολυμεταβλητών δεδομένων και τη διερεύνηση συσχετίσεων μεταξύ χαρακτηριστικών ποιότητας (Hotelling, 1933; Biegel et al., 2022).

Η Ομαδοποίηση k-Means (k-Means Clustering) αποτελεί μία από τις πιο διαδεδομένες μεθόδους ομαδοποίησης δεδομένων. Η μέθοδος διαχωρίζει τις παρατηρήσεις σε ομάδες με βάση την ομοιότητα των χαρακτηριστικών τους, επιδιώκοντας την ελαχιστοποίηση της απόστασης των παρατηρήσεων από το κέντρο της ομάδας στην οποία ανήκουν. Στη Διαχείριση Ποιότητας, η ομαδοποίηση μπορεί να χρησιμοποιηθεί για την αναγνώριση διαφορετικής συμπεριφοράς διεργασίας, την ανάλυση πιθανών προτύπων και τη διερεύνηση αποκλίσεων που δεν είναι εμφανείς μέσω απλών στατιστικών δεικτών. Ωστόσο, η απόδοση της μεθόδου επηρεάζεται από την επιλογή του αριθμού των ομάδων (k) και από την κλίμακα των μεταβλητών (Hartigan & Wong, 1979; Zan et al., 2020).

Ο Αλγόριθμος Απομόνωσης Δασών (Isolation Forest) αποτελεί σύγχρονη τεχνική ανίχνευσης ανωμαλιών (Anomaly Detection), η οποία βασίζεται στην αρχή ότι οι ασυνήθιστες παρατηρήσεις μπορούν να απομονωθούν ευκολότερα από τις κανονικές. Σε αντίθεση με άλλες μεθόδους που επιχειρούν να μοντελοποιήσουν την κανονική συμπεριφορά των δεδομένων, ο Isolation Forest επικεντρώνεται στον εντοπισμό παρατηρήσεων που αποκλίνουν από τη γενική

δομή του συνόλου δεδομένων. Η μέθοδος έχει ιδιαίτερη συνάφεια με εφαρμογές ML-SPC, καθώς μπορεί να χρησιμοποιηθεί για πολυμεταβλητή ανίχνευση αποκλίσεων και για την ανάπτυξη διαγραμμάτων ελέγχου βασισμένων σε βαθμολογίες ανωμαλίας (Anomaly Scores) (Liu et al., 2012; Wang & Liu, 2024; Perdakis et al., 2026).

Τέλος, οι Αυτόματοι Κωδικοποιητές (Autoencoders) και γενικότερα οι μέθοδοι Βαθιάς Μάθησης (Deep Learning) αποτελούν πιο προηγμένες προσεγγίσεις για την ανάλυση πολύπλοκων και πολυδιάστατων δεδομένων. Οι Autoencoders εκπαιδεύονται ώστε να αναπαράγουν τα δεδομένα εισόδου στην έξοδο, συμπιέζοντας την πληροφορία σε μία ενδιάμεση αναπαράσταση. Σε εφαρμογές ανίχνευσης ανωμαλιών, μεγάλες αποκλίσεις μεταξύ αρχικής και ανακατασκευασμένης παρατήρησης μπορούν να υποδηλώνουν ασυνήθιστη συμπεριφορά. Παρότι οι τεχνικές αυτές προσφέρουν σημαντικές δυνατότητες αναγνώρισης σύνθετων και μη γραμμικών προτύπων, συνοδεύονται από αυξημένες απαιτήσεις σε δεδομένα, υπολογιστικούς πόρους και παραμετροποίηση (Zan et al., 2020; Biegel et al., 2022).

Η επιλογή των κατάλληλων αλγορίθμων σε μία εφαρμογή SPC–ML δεν εξαρτάται αποκλειστικά από την προγνωστική τους απόδοση, αλλά και από τη φύση του προβλήματος, τη δομή των δεδομένων και τη δυνατότητα πρακτικής εφαρμογής τους. Για τον σκοπό αυτό, σημαντικά κριτήρια επιλογής αποτελούν η καταλληλότητα για δυαδική ταξινόμηση (π.χ. Pass/Fail), η ικανότητα διαχείρισης πολυμεταβλητών δεδομένων, η δυνατότητα αναγνώρισης μη γραμμικών σχέσεων, η ευαισθησία σε διαδικασίες κανονικοποίησης και η δυνατότητα σύγκρισης με παραδοσιακές μεθόδους SPC (Boaventura et al., 2022). Παράλληλα, σε περιβάλλοντα Διαχείρισης Ποιότητας, ιδιαίτερη σημασία έχει η πρακτικότητα της εφαρμογής, η αξιοπιστία της απόδοσης και η δυνατότητα αξιοποίησης των αποτελεσμάτων για λήψη αποφάσεων (Chou et al., 2020; Kazmi & Noor-ul-Amin, 2025).

### 2.3 Συγκριτική θεώρηση SPC και ML-SPC

Η σύγκριση μεταξύ του παραδοσιακού Στατιστικού Ελέγχου Διεργασιών και των προσεγγίσεων SPC βασισμένων σε Μηχανική Μάθηση (ML-SPC) δεν πρέπει να αντιμετωπίζεται ως αντιπαράθεση δύο ανεξάρτητων μεθοδολογιών. Αντίθετα, οι δύο προσεγγίσεις μπορούν να θεωρηθούν ως διαφορετικά πλαίσια ανάλυσης και ελέγχου της μεταβλητότητας των διεργασιών. Ο παραδοσιακός SPC στηρίζεται σε στατιστικά μεγέθη, προκαθορισμένα όρια ελέγχου και κανόνες ερμηνείας, με στόχο τη διάκριση μεταξύ τυχαίας

και ειδικής μεταβλητότητας (random cause and special cause variation). Αντίστοιχα, οι μέθοδοι ML-SPC επιχειρούν να επεκτείνουν αυτή τη λογική, αξιοποιώντας αλγορίθμους που μπορούν να αναγνωρίζουν πιο σύνθετα, πολυμεταβλητά ή μη γραμμικά πρότυπα στα δεδομένα της διεργασίας (Boaventura et al., 2022; Wang & Liu, 2024).

Η βασική διαφοροποίηση μεταξύ των δύο προσεγγίσεων εντοπίζεται στον τρόπο αντίληψης και επεξεργασίας της μεταβλητότητας. Ο παραδοσιακός SPC εξετάζει τη μεταβλητότητα κυρίως μέσω στατιστικών παραμέτρων, όπως ο μέσος όρος, η διασπορά και η απόκλιση από τα όρια ελέγχου. Η προσέγγιση αυτή είναι ιδιαίτερα αποτελεσματική όταν η διεργασία είναι σχετικά σταθερή, οι βασικές στατιστικές υποθέσεις ικανοποιούνται και τα χαρακτηριστικά ποιότητας μπορούν να παρακολουθούνται με σαφή και προκαθορισμένο τρόπο (Montgomery, 2019). Αντίθετα, οι μέθοδοι ML-SPC μπορούν να αξιοποιήσουν μεγαλύτερο αριθμό μεταβλητών και να εντοπίσουν σχέσεις που δεν εκφράζονται εύκολα μέσω κλασικών στατιστικών δεικτών, όπως αλληλεπιδράσεις μεταξύ χαρακτηριστικών, μη γραμμικές συμπεριφορές ή σύνθετα πρότυπα αποκλίσεων (Zan et al., 2020; Hric & Sabahno, 2024).

Το παραδοσιακό SPC υπερέχει κυρίως ως προς την απλότητα, τη διαφάνεια και την ερμηνευσιμότητα. Τα διαγράμματα ελέγχου είναι ευρέως κατανοητά από στελέχη ποιότητας και παραγωγής, ενώ η χρήση τους συνδέεται άμεσα με καθιερωμένες πρακτικές βιομηχανικού ελέγχου. Επιπλέον, τα στατιστικά όρια και οι κανόνες ερμηνείας παρέχουν ένα σαφές και τυποποιημένο πλαίσιο λήψης αποφάσεων, το οποίο μπορεί να εφαρμοστεί χωρίς ιδιαίτερα υψηλές υπολογιστικές απαιτήσεις. Για τον λόγο αυτό, ο SPC εξακολουθεί να αποτελεί βασική μεθοδολογία παρακολούθησης διεργασιών και εργαλείο αναφοράς σε συστήματα Διαχείρισης Ποιότητας (Shewhart, 1931; Montgomery, 2019).

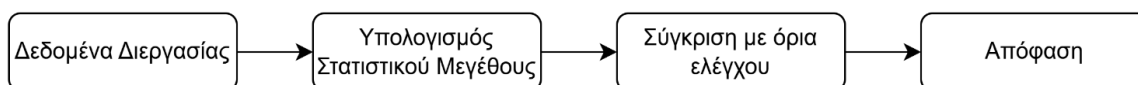
Από την άλλη πλευρά, οι προσεγγίσεις ML-SPC εμφανίζουν πλεονεκτήματα σε περιβάλλοντα όπου τα δεδομένα είναι πολυδιάστατα, οι σχέσεις μεταξύ μεταβλητών είναι σύνθετες και η παραδοσιακή ανάλυση ενδέχεται να μην επαρκεί. Μέθοδοι όπως ο Αλγόριθμος Απομόνωσης Δασών (Isolation Forest), τα υβριδικά διαγράμματα EWMA βασισμένα σε Μηχανική Μάθηση και τα μοντέλα Βαθιάς Μάθησης (Deep Learning) έχουν χρησιμοποιηθεί για την ενίσχυση της ανίχνευσης ανωμαλιών και της πολυμεταβλητής παρακολούθησης διεργασιών (Biegel et al., 2022; Kazmi & Noor-ul-Amin, 2025; Perdikis et al., 2026). Η δυνατότητα των αλγορίθμων αυτών να παράγουν βαθμολογίες ανωμαλίας (Anomaly Scores), πιθανότητες ταξινόμησης ή

πιο σύνθετες αναπαραστάσεις των δεδομένων τους καθιστά χρήσιμους σε περιβάλλοντα όπου απαιτείται μεγαλύτερη προσαρμοστικότητα.

Παρόλα αυτά, αξίζει να σημειώσουμε πως τα μοντέλα ML-SPC δεν παρουσιάζουν μόνο πλεονεκτήματα. Τα μοντέλα Μηχανικής Μάθησης απαιτούν κατάλληλη προεπεξεργασία δεδομένων, επιλογή χαρακτηριστικών, παραμετροποίηση και επαρκή ποσότητα αξιόπιστων δεδομένων, ποσότητα που εξαρτάται και από την πολυπλοκότητα τους, τη διεργασία και την εμπιστοσύνη των αποτελεσμάτων που απαιτούνται από την εφαρμογή. Επιπλέον, η πολυπλοκότητα των μοντέλων μπορεί να δυσκολέψει την άμεση ερμηνεία των αποτελεσμάτων από τα στελέχη ποιότητας, ειδικά όταν τα μοντέλα λειτουργούν ως πιο σύνθετα συστήματα ταξινόμησης ή ανίχνευσης ανωμαλιών. Αντίθετα, ο παραδοσιακός SPC παρέχει πιο άμεση και κατανοητή απεικόνιση της διεργασιακής συμπεριφοράς, γεγονός που τον καθιστά ιδιαίτερα χρήσιμο για καθημερινή παρακολούθηση και επιχειρησιακή λήψη αποφάσεων (Chou et al., 2020; Biegel et al., 2022).

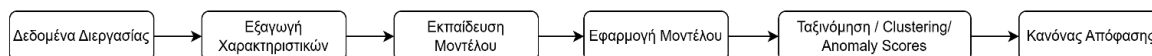
Συνεπώς, η πιο παραγωγική θεώρηση δεν είναι ότι η Μηχανική Μάθηση αντικαθιστά τον Στατιστικό Έλεγχο Διεργασιών, αλλά ότι μπορεί να λειτουργήσει συμπληρωματικά προς αυτόν. Ο SPC παρέχει τη στατιστική και επιχειρησιακή βάση για την παρακολούθηση της διεργασίας, ενώ οι μέθοδοι ML μπορούν να ενισχύσουν την ικανότητα αναγνώρισης σύνθετων προτύπων, την ανίχνευση αποκλίσεων και την προγνωστική υποστήριξη αποφάσεων. Η συνδυαστική αξιοποίηση των δύο προσεγγίσεων δημιουργεί ένα πλαίσιο ενισχυμένου SPC μέσω τεχνητής νοημοσύνης (AI-Enhanced SPC), το οποίο διατηρεί τη λογική του ελέγχου διεργασιών, αλλά την επεκτείνει σε περιβάλλοντα μεγαλύτερης πολυπλοκότητας και αυξημένων απαιτήσεων δεδομένων (Boaventura et al., 2022; Escobar et al., 2025).

Σε επίπεδο λειτουργικής λογικής, η σύγκριση μπορεί να αποτυπωθεί ως μετάβαση από ένα στατιστικά καθορισμένο σύστημα αποφάσεων σε ένα περισσότερο δεδομενοκεντρικό και προσαρμοστικό σύστημα παρακολούθησης. Στον παραδοσιακό SPC, η διαδικασία ακολουθεί συνήθως τη ροή:



Διάγραμμα 1 Διαδικασία SPC

Αντίστοιχα, σε ένα πλαίσιο ML-SPC, η διαδικασία μπορεί να περιλαμβάνει:



**Διάγραμμα 2 Διαδικασία ML-SPC**

Η διαφορά αυτή δεν αναιρεί τη βασική λογική του SPC, αλλά την επεκτείνει μέσω της αξιοποίησης πιο σύνθετων μηχανισμών μάθησης από δεδομένα. Οι βασικές/χρήσεις και διαφορές συνοψίζονται στον πίνακα 1.

**Πίνακας 1 Χρήσεις και διαφορές SPC με ML/AI-SPC**

| Παραδοσιακό SPC                  | ML-SPC / AI-Enhanced SPC                      |
|----------------------------------|-----------------------------------------------|
| Δεδομένα διεργασίας              | Δεδομένα διεργασίας                           |
| Υπολογισμός στατιστικού μεγέθους | Εξαγωγή / μετασχηματισμός χαρακτηριστικών     |
| Στατικά όρια ελέγχου             | Μοντέλο Μηχανικής Μάθησης                     |
| Σήμα / μη σήμα ελέγχου           | Πιθανότητα ταξινόμησης ή βαθμολογία ανωμαλίας |
| Διορθωτική ενέργεια              | Προληπτική ή διορθωτική υποστήριξη απόφασης   |

Με βάση τα παραπάνω, η συγκριτική θεώρηση SPC και ML-SPC δεν περιορίζεται στο ποια μέθοδος εμφανίζει καλύτερη ακρίβεια, αλλά αφορά τη συνολική αξιολόγηση της καταλληλότητας κάθε προσέγγισης ανάλογα με τη φύση της διεργασίας, τη δομή των δεδομένων, την απαιτούμενη ευαισθησία ανίχνευσης και τη δυνατότητα πρακτικής εφαρμογής σε βιομηχανικό περιβάλλον. Η προσέγγιση αυτή οδηγεί στην ανάγκη καθορισμού συγκεκριμένων κριτηρίων αξιολόγησης, τα οποία παρουσιάζονται στο επόμενο υποκεφάλαιο.

## 2.4 Κριτήρια συγκριτικής αξιολόγησης SPC και ML-SPC

Η συγκριτική αξιολόγηση του παραδοσιακού Στατιστικού Ελέγχου Διεργασιών και των προσεγγίσεων ML-SPC απαιτεί τον καθορισμό κοινών και αντικειμενικών κριτηρίων. Η σύγκριση δεν μπορεί να περιοριστεί αποκλειστικά στην ακρίβεια ενός μοντέλου ή στην εμφάνιση σημάτων ελέγχου, αλλά πρέπει να εξετάζει συνολικά την τεχνική απόδοση, την επιχειρησιακή χρησιμότητα και τη δυνατότητα εφαρμογής κάθε προσέγγισης σε πραγματικό βιομηχανικό περιβάλλον. Για τον λόγο αυτό, η αξιολόγηση οφείλει να περιλαμβάνει τόσο στατιστικούς και προγνωστικούς δείκτες όσο και κριτήρια που σχετίζονται με τη λειτουργία του συστήματος ποιότητας (Montgomery, 2019; Chou et al., 2020).

### Ικανότητα Ανίχνευσης Αποκλίσεων

Βασικότερο κριτήριο της συγκριτικής αξιολόγησης αποτελεί η ικανότητα ανίχνευσης αποκλίσεων. Στο παραδοσιακό SPC, η ανίχνευση βασίζεται κυρίως στην υπέρβαση των ορίων ελέγχου ή στην εμφάνιση μη τυχαίων προτύπων στη συμπεριφορά των δεδομένων (π.χ. General Electric Rules). Αντίστοιχα, στις προσεγγίσεις ML-SPC, η ανίχνευση μπορεί να βασίζεται σε πιθανότητες ταξινόμησης, βαθμολογίες ανωμαλίας (Anomaly Scores) ή προβλέψεις του μοντέλου. Η αξιολόγηση της ικανότητας ανίχνευσης είναι κρίσιμη, καθώς συνδέεται άμεσα με την έγκαιρη αναγνώριση ειδικών αιτιών μεταβλητότητας και πιθανών προβλημάτων ποιότητας - π.χ. απόκλιση από μέση τιμή, τάση κλπ. (Wang & Liu, 2024; Kazmi & Noor-ul-Amin, 2025).

### Ευαισθησία σε μικρές ή σταδιακές μεταβολές

Δεύτερο σημαντικό κριτήριο είναι η ευαισθησία των τεχνικών σε μικρές ή σταδιακές μεταβολές της διεργασίας. Τα κλασικά διαγράμματα Shewhart είναι αποτελεσματικά στην ανίχνευση σχετικά μεγάλων μετατοπίσεων, αλλά ενδέχεται να παρουσιάζουν καθυστέρηση όταν οι μεταβολές είναι μικρές ή εξελίσσονται σταδιακά. Για τον λόγο αυτό έχουν αναπτυχθεί τεχνικές όπως τα διαγράμματα EWMA και CUSUM, τα οποία αυξάνουν την ευαισθησία μέσω της αξιοποίησης πληροφορίας  $n$  (ιστορικότητα) από προηγούμενες παρατηρήσεις. Αντίστοιχα, τα ML-SPC μοντέλα μπορούν να αξιολογηθούν ως προς την ικανότητά τους να αναγνωρίζουν λεπτές αλλαγές στη δομή των δεδομένων ή σε πολυμεταβλητά πρότυπα συμπεριφοράς (Montgomery, 2019; Zhou et al., 2025).

### Ψευδείς Συναγερμοί (False Alarms)

Τρίτο κριτήριο αποτελεί η συχνότητα ψευδών συναγερμών. Ένα σύστημα παρακολούθησης διεργασιών δεν αρκεί να είναι ευαίσθητο· πρέπει παράλληλα να αποφεύγει την υπερβολική παραγωγή σημάτων που δεν αντιστοιχούν σε πραγματική απόκλιση (Σφάλμα τύπου I). Οι ψευδείς συναγερμοί μπορούν να οδηγήσουν σε περιττούς ελέγχους, διακοπές παραγωγής, αυξημένο φόρτο εργασίας και μειωμένη εμπιστοσύνη στο σύστημα παρακολούθησης. Επομένως, η σύγκριση μεταξύ SPC και ML-SPC πρέπει να εξετάζει την ισορροπία μεταξύ ευαισθησίας και σταθερότητας αποφάσεων, ιδιαίτερα σε εφαρμογές όπου η πρακτική

χρησιμότητα των σημάτων είναι εξίσου σημαντική με την τεχνική ακρίβεια (Boaventura et al., 2022; Perdikis et al., 2026).

### Χρόνος Ανίχνευσης Μεταβολής (ARL)

Για την αξιολόγηση ενός συστήματος εντοπισμού/ ανίχνευσης μεταβολών, ιδιαίτερη σημασία έχει ο χρόνος (ή αντίστοιχα το τεμάχιο) ανίχνευσης της μεταβολής. Στο πλαίσιο του SPC, ένας καθιερωμένος δείκτης αξιολόγησης είναι το Average Run Length (ARL), το οποίο εκφράζει τον αναμενόμενο αριθμό δειγμάτων μέχρι την εμφάνιση σήματος (alarm). Στην ουσία, μία χαμηλή τιμή του ARL σε διεργασία εκτός στατιστικού ελέγχου υποδηλώνει ταχύτερη ανίχνευση μεταβολής, ενώ υψηλότερο ARL σε διεργασία εντός στατιστικού ελέγχου συνδέεται με χαμηλότερη πιθανότητα ψευδούς συναγερμού (Σφάλμα τύπου I). Σε αντίστοιχο πνεύμα, οι προσεγγίσεις ML-SPC μπορούν να αξιολογηθούν ως προς τη χρονική καθυστέρηση ανίχνευσης, δηλαδή το διάστημα που απαιτείται μέχρι το μοντέλο να αναγνωρίσει μία απόκλιση ή αυξημένη πιθανότητα αστοχίας (Montgomery, 2019; Kazmi & Noor-ul-Amin, 2025). Για τις επιβλεπόμενες μεθόδους Μηχανικής Μάθησης (ML), ιδιαίτερο ρόλο έχουν οι δείκτες προγνωστικής απόδοσης. Σε προβλήματα δυαδικής ταξινόμησης, όπως η διάκριση αποδεκτών και μη αποδεκτών προϊόντων, μπορούν να χρησιμοποιηθούν δείκτες όπως η ακρίβεια (Accuracy), η ευαισθησία ή ανάκληση (Sensitivity / Recall), η ειδικότητα (Specificity), η ακρίβεια θετικής πρόβλεψης (Precision) και η καμπύλη ROC. Η χρήση περισσότερων του ενός δεικτών είναι απαραίτητη, ιδιαίτερα όταν οι κλάσεις δεν είναι ισορροπημένες, καθώς η συνολική ακρίβεια μπορεί να δώσει παραπλανητική εικόνα της πραγματικής απόδοσης του μοντέλου (Chou et al., 2020; Biegel et al., 2022).

### Πολυμεταβλητές ή Μη γραμμικές σχέσεις

Ένα ακόμη κρίσιμο κριτήριο είναι η ικανότητα διαχείρισης πολυμεταβλητών και μη γραμμικών σχέσεων. Τα παραδοσιακά διαγράμματα ελέγχου λειτουργούν αποτελεσματικά όταν τα χαρακτηριστικά ποιότητας είναι λίγα, σαφώς ορισμένα και μπορούν να παρακολουθούνται ανεξάρτητα ή μέσω καθιερωμένων πολυμεταβλητών στατιστικών μεθόδων. Ωστόσο, σε σύγχρονες βιομηχανικές εφαρμογές, οι μεταβλητές μπορεί να εμφανίζουν συσχετίσεις, αλληλεπιδράσεις και μη γραμμικές σχέσεις. Οι τεχνικές Μηχανικής Μάθησης προσφέρουν τη δυνατότητα αναγνώρισης αυτών των σύνθετων προτύπων, γεγονός

που αποτελεί σημαντικό πλεονέκτημα σε περιβάλλοντα υψηλής πολυπλοκότητας (Zan et al., 2020; Hric & Sabahno, 2024).

Παράλληλα, πέραν των αμιγώς τεχνικών κριτηρίων πρέπει ακόμη να λαμβάνεται υπόψη η δυνατότητα εφαρμογής και οι προκλήσεις που εμφανίζονται σε βιομηχανικό περιβάλλον. Η πρακτική αξιοποίηση μιας μεθόδου εξαρτάται από παράγοντες όπως:

- η διαθεσιμότητα και ποιότητα των διαθέσιμων δεδομένων,
- η ανάγκη προεπεξεργασίας (καθαρισμός δεδομένων, κατηγοριοποίηση κλπ),
- η ευκολία παραμετροποίησης,
- οι υπολογιστικές απαιτήσεις και
- η δυνατότητα ένταξης στα υφιστάμενα συστήματα ποιότητας.

Το παραδοσιακό SPC έχει το πλεονέκτημα της απλότητας και της καθιερωμένης χρήσης, ενώ οι μέθοδοι ML-SPC απαιτούν μεγαλύτερη τεχνική υποδομή και εξειδίκευση. Επομένως, η επιλογή μεθόδου πρέπει να βασίζεται όχι μόνο στην απόδοση, αλλά και στη δυνατότητα σταθερής και βιώσιμης εφαρμογής της στην πράξη (Escobar et al., 2025; Chou et al., 2020).

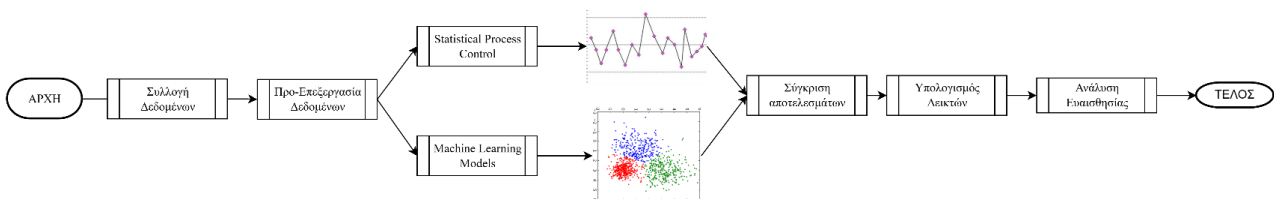
Συνολικά, τα ανωτέρω κριτήρια διαμορφώνουν ένα ολοκληρωμένο πλαίσιο αξιολόγησης για τη σύγκριση του παραδοσιακού SPC και των προσεγγίσεων ML-SPC. Το πλαίσιο αυτό επιτρέπει την εξέταση όχι μόνο της στατιστικής ή προγνωστικής απόδοσης των μεθόδων, αλλά και της επιχειρησιακής τους χρησιμότητας, της τεχνικής τους εφικτότητας και της πιθανής συμβολής τους στη μείωση του κόστους και στη βελτίωση της βιομηχανικής ποιότητας.

## Κεφάλαιο 3 - Μεθοδολογία Έρευνας

### 3.1 Ερευνητικός σχεδιασμός και μεθοδολογική προσέγγιση

Στην παρούσα εργασία υιοθετείται εφαρμοσμένη ερευνητική προσέγγιση, βασισμένη σε ποσοτική ανάλυση πραγματικών δεδομένων από βιομηχανικό περιβάλλον. Ο εφαρμοσμένος χαρακτήρας της συνδέεται με την αξιολόγηση μεθόδων που μπορούν να υποστηρίξουν πρακτικές αποφάσεις στη Διαχείριση Ποιότητας, ενώ η μελέτη περίπτωσης επιτρέπει την εξέταση των μεθόδων υπό πραγματικές συνθήκες λειτουργίας. Με τον τρόπο αυτό, στην ανάλυση λαμβάνεται υπόψη τόσο το συγκεκριμένο παραγωγικό πλαίσιο όσο και τη μορφή των διαθέσιμων δεδομένων (Yin, 2018).

Η μεθοδολογία που θα ακολουθήσουμε αναπτύσσεται σε διαδοχικά στάδια. Αρχικά, παρουσιάζεται το επιχειρηματικό πλαίσιο και η υπό εξέταση διεργασία. Στη συνέχεια, περιγράφονται τα δεδομένα, οι μεταβλητές και οι περιορισμοί που σχετίζονται με τη χρήση αυτών. Ακολουθεί η περιγραφή της διαδικασίας προ-επεξεργασίας των δεδομένων, ώστε να διασφαλιστεί η καταλληλότητα τους για στατιστική ανάλυση. Στη συνέχεια, εφαρμόζουμε τις παραδοσιακές τεχνικές Στατιστικού Ελέγχου Ποιότητας (SPC), αναπτύσσονται τα Διαγράμματα Ελέγχου και τα μοντέλα Μηχανικής Μάθησης. Τα αποτελέσματα των δύο προσεγγίσεων συγκρίνονται βάσει κριτηρίων που αναλύονται στο 3.7..



Διάγραμμα 3 Μεθοδολογική Προσέγγιση Εργασίας

Ο ερευνητικός σχεδιασμός της εργασίας στηρίζεται στην εφαρμογή δύο διαφορετικών προσεγγίσεων στα ίδια δεδομένα: αφενός των παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών (I-MR, EWMA κλπ.) και αφετέρου μεθόδων Μηχανικής Μάθησης (ML). Με τον τρόπο αυτό, η σύγκριση δεν πραγματοποιείται σε θεωρητικό επίπεδο, αλλά μέσω εφαρμογής και αξιολόγησης των μεθόδων σε κοινή βάση δεδομένων. Η προσέγγιση αυτή επιτρέπει την πιο αντικειμενική διερεύνηση της σχετικής απόδοσης των δύο μεθοδολογιών και της

δυνατότητας ενίσχυσης του παραδοσιακού SPC μέσω ML-SPC προσεγγίσεων (Chou et al., 2020; Boaventura et al., 2022).

Ο σκοπός της μεθοδολογικής προσέγγισης δεν περιορίζεται στην ανάπτυξη ενός προγνωστικού μοντέλου ή στην εφαρμογή διαγραμμάτων ελέγχου στα δεδομένα, αλλά στην πραγματοποίηση συνολικής αξιολόγησης της σκοπιμότητας υιοθέτησης τεχνικών Μηχανικής Μάθησης σε ένα υφιστάμενο βιομηχανικό σύστημα ποιότητας. Για αυτό τον λόγο, στην αξιολόγηση ενσωματώνονται τόσο τεχνικές και λειτουργικές, όσο και οικονομικές παραμέτρους. Η σύνδεση με τη λογική της προγνωστικής διαχείρισης ποιότητας και της Quality 4.0 καθιστά τη μεθοδολογία κατάλληλη για τη διερεύνηση του βαθμού στον οποίο οι σύγχρονες μέθοδοι ανάλυσης δεδομένων μπορούν να συμπληρώσουν τα παραδοσιακά εργαλεία ελέγχου διεργασιών (Escobar et al., 2025)

Η σύγκριση των δύο προσεγγίσεων πραγματοποιείται με βάση κριτήρια που συνδέονται τόσο με τη στατιστική και προγνωστική απόδοση όσο και με τη βιομηχανική χρησιμότητα των μεθόδων.

Σε αυτό το σημείο, είναι ωφέλιμο να τονίσουμε ότι η προσέγγιση επιλέχθηκε με βάση τη δομή και τις ιδιαιτερότητες της συγκεκριμένης εφαρμογής και των δεδομένων που χρησιμοποιούνται. Συνεπώς, αυτή η εργασία σε καμία περίπτωση δεν αποσκοπεί στην εξαγωγή γενικευμένων συμπερασμάτων που αφορούν άλλες διεργασίες ή άλλους τύπους δεδομένων. Πάρα ταύτα, η συγκεκριμένη προσέγγιση μπορεί να προσφέρει μία κατεύθυνση με τεκμηριωμένη αξιολόγηση εφαρμογής μεθόδων SPC και ML-SPC σε ένα πραγματικό Σύστημα Διαχείρισης Ποιότητας (ΣΔΠ).

### 3.2 Περιγραφή της μελέτης περίπτωσης

Η μελέτη περίπτωσης που πραγματεύεται η παρούσα εργασία αφορά ανώνυμη ελληνική βιομηχανική επιχείρηση, η οποία δραστηριοποιείται στο αντικείμενο της παραγωγής φωτιστικών σωμάτων LED. Η συγκεκριμένη εταιρεία επιλέχθηκε λόγω της ύπαρξης συστηματικά καταγεγραμμένων δεδομένων ποιοτικού ελέγχου προϊόντων με διάφορες μεταβλητές, τα οποία θα αξιοποιήσουμε για τη συγκριτική εφαρμογή μεθόδων Στατιστικού Ελέγχου Διεργασιών (SPC) και μεθόδων Μηχανικής Μάθησης. Για λόγους εμπιστευτικότητας,

κρίθηκε αναγκαίο τα δεδομένα που χρησιμοποιούνται και παρουσιάζονται να ανωνυμοποιούνται, χωρίς ωστόσο να επηρεαστεί η τεχνική εγκυρότητα της ανάλυσης.

Το αντικείμενο της επιχείρησης είναι η κατασκευή και εμπορία φωτιστικών σωμάτων LED, τα οποία ανήκουν σε τρεις βασικές οικογένειες προϊόντων (έστω A, B, C). Εντός των οικογενειών αυτών υφίστανται διαφορετικά μοντέλα (διαχωρίζονται βάσει ισχύος) και τεχνικές διαμορφώσεις, οι οποίες διαφοροποιούνται ως προς διάφορα χαρακτηριστικά του προϊόντος. Η ύπαρξη πολλαπλών διαμορφώσεων αποτελεί σημαντική παράμετρο της μελέτης, καθώς επηρεάζει τόσο τα αποδεκτά όρια λειτουργίας όσο και τον τρόπο με τον οποίο πρέπει να αξιολογούνται και να συγκρίνονται οι μετρήσεις. Για τον λόγο αυτό, η πληροφορία που αφορά το μοντέλο ή τη διαμόρφωση του προϊόντος λαμβάνεται υπόψη στα επόμενα στάδια της ανάλυσης.

Η παραγωγική διαδικασία των φωτιστικών σωμάτων LED παρουσιάζεται περιγραφικά στη συνέχεια. Σε γενικό επίπεδο, η ροή περιλαμβάνει την προετοιμασία του σώματος, τη συναρμολόγηση της οπτικής μονάδας, τη συναρμολόγηση του καπακιού, τη συναρμολόγηση σώματος και οπτικής μονάδας, τη συναρμολόγηση σώματος και καπακιού, τον ηλεκτρικό έλεγχο, τον έλεγχο γήρανσης και, τέλος, τη συσκευασία και παλετοποίηση των έτοιμων προϊόντων. Η συνολική ροή της διεργασίας παρουσιάζεται διαγραμματικά στο σχετικό διάγραμμα ροής παραγωγής, το οποίο αποτυπώνει τη θέση του ηλεκτρικού ελέγχου μέσα στην αλληλουχία των φάσεων παραγωγής.

### Ροή Παραγωγικής Διαδικασίας



Διάγραμμα 4 Φάσεις Παραγωγικής Διαδικασίας Φωτιστικών Σωμάτων

Ο τελικός ηλεκτρικός έλεγχος αποτελεί κρίσιμο στάδιο της παραγωγικής διαδικασίας και εφαρμόζεται στο σύνολο(100%) των παραγόμενων προϊόντων που πρόκειται να διατεθούν. Τα

δεδομένα των αποτελεσμάτων του ελέγχου έχουν πλήρη ιχνηλασιμότητα, δηλαδή κάθε προϊόν, ελέγχεται και καταχωρείται με το αναγνωριστικό του (Serial – ID). Το αποτέλεσμα του ηλεκτρικού ελέγχου τελικών προϊόντων αποτελεί την μέτρηση που καθορίζει εάν τα τελικά προϊόντα είναι άρτια ώστε να διατεθούν στην αγορά, πρόκειται δηλαδή για δυαδικό Pass/No Pass. Η απόφαση αυτή βασίζεται σε μεταβλητά όρια, τα οποία διαφοροποιούνται ποσοστιαία ανάλογα με την ισχύ ή τη διαμόρφωση του προϊόντος. Με τον τρόπο αυτό, ο ηλεκτρικός έλεγχος λειτουργεί ως μηχανισμός επιβεβαίωσης της συμμόρφωσης του προϊόντος πριν από την τελική αποδέσμευση.

Τα δεδομένα της μελέτης περίπτωσης που επιλέχθηκε εμφανίζουν κάποια χαρακτηριστικά κατάλληλα για να διεξάγουμε σύγκριση μεταξύ των παραδοσιακών μεθόδων SPC και ML-SPC, καθώς συνδυάζει πραγματικά δεδομένα ποιοτικού ελέγχου, διαφορετικές οικογένειες προϊόντων, διαφορετικές τεχνικές διαμορφώσεις μεταξύ προϊόντων και σαφή κατηγοριοποίηση των αποτελεσμάτων του ελέγχου (data labels). Η χρονική περίοδος συλλογής των παρατηρήσεων καλύπτει τα τελευταία 5 έτη, παρέχοντας επαρκή αριθμό μετρήσεων για την αξιολόγηση της διεργασίας. Μέσα από τη μελέτη περίπτωσης επιχειρούμε να διερευνήσουμε τη σταθερότητα της παραγωγικής διαδικασίας, την εμφάνιση αποκλίσεων, της ικανότητας πρόβλεψης ή ταξινόμησης αποτελεσμάτων ελέγχου.

### 3.3 Περιγραφή δεδομένων και μεταβλητών

Τα δεδομένα της μελέτης προέρχονται από το σύστημα τελικού ηλεκτρικού ελέγχου, δηλαδή αποτελούν δεδομένα ποιοτικού ελέγχου, τα οποία συλλέγονται κατά την παραγωγική διαδικασία και όχι στο πλαίσιο εργαστηριακών μετρήσεων ή πειραματικού σχεδιασμού. Η χρονική περίοδος συλλογής καλύπτει 5 έτη λειτουργίας, παρέχοντας σημαντικό πλήθος εγγραφών.

Τα δεδομένα αφορούν προϊόντα που ελέγχονται στο τελικό στάδιο της ΠΔ και περιλαμβάνουν πληροφορίες όπως μετρήσεις ηλεκτρικών χαρακτηριστικών, πληροφορίες ταυτοποίησης και ιχνηλασιμότητας του προϊόντος και κατηγοριοποίηση του αποτελέσματος ελέγχου ως προς τη συμμόρφωση του προϊόντος.

Στον Πίνακα 2 παρουσιάζονται συνοπτικά οι κατηγορίες των δεδομένων που συλλέγονται και θα χρησιμοποιηθούν στα πλαίσια της ανάλυσης.

**Πίνακας 2 Κατηγορίες Δεδομένων της Ανάλυσης**

| Κατηγορία μεταβλητής           | Μεταβλητή                                                                    | Τύπος                   | Σύμβολο / Μ.<br>Μ. | Ρόλος στην ανάλυση                             |
|--------------------------------|------------------------------------------------------------------------------|-------------------------|--------------------|------------------------------------------------|
| Ταυτοποίηση/<br>ιχνηλασιμότητα | Serial ID                                                                    | Κατηγορική /<br>ID      | -                  | Σύνδεση μέτρησης με<br>συγκεκριμένο προϊόν     |
| Χρόνος/Ακολουθία<br>Ελέγχου    | Timestamp - Ημερομηνία & ώρα<br>ελέγχου                                      | Χρονική                 | -                  | Ανάλυση παραγωγικής<br>ακολουθίας              |
| Κατηγορία Προϊόντος            | Οικογένεια, μοντέλο (ισχύς),<br>διαμόρφωση                                   | Κατηγορική              | -                  | Ομαδοποίηση                                    |
| Ηλεκτρικές μετρήσεις           | Ισχύς (Wattage)                                                              | Συνεχής                 | W                  | Έλεγχος συμβατότητας με<br>ονομαστική ισχύ     |
| Ηλεκτρικές μετρήσεις           | Τάση (Voltage)                                                               | Συνεχής                 | V                  | Συνθήκη λειτουργίας / είσοδος<br>μοντέλου      |
| Ηλεκτρικές μετρήσεις           | Ρεύμα εισόδου (Input Current)                                                | Συνεχής                 | A                  | Συσχέτιση με ισχύ και ηλεκτρική<br>συμπεριφορά |
| Ηλεκτρικές μετρήσεις           | Συντελεστής ισχύος (Power Factor)                                            | Συνεχής                 | Αδιάστατο / -      | Έλεγχος αποδοτικότητας /<br>συμμόρφωσης        |
| Ηλεκτρικές μετρήσεις           | Ολική αρμονική παραμόρφωση<br>τάσης (THDU%)                                  | Συνεχής                 | %                  | Δείκτης ποιότητας ηλεκτρικής<br>συμπεριφοράς   |
| Όρια ελέγχου                   | Όρια προδιαγραφών (SL),<br>κανονιστικά όρια,<br>όρια ελέγχου διεργασίας (CL) | Αριθμητική /<br>κανόνας | -                  | Βάση αξιολόγησης συμμόρφωσης                   |
| Αποτέλεσμα ελέγχου             | Pass / No Pass                                                               | Δυαδική<br>κατηγορική   | 0/1                | Μεταβλητή εξόδου / ετικέτα<br>ταξινόμησης      |

Ουσιαστικά, η βασική μονάδα παρατήρησης είναι ο τελικός ηλεκτρικός έλεγχος ενός φωτιστικού σώματος, δηλαδή κάθε μέτρηση (datapoint) αντιστοιχεί σε ένα συγκεκριμένο προϊόν, μία χρονική στιγμή ελέγχου, ένα συγκεκριμένο μοντέλο, μία τεχνική διαμόρφωση και μοναδικό αποτέλεσμα ελέγχου. Το αποτέλεσμα προκύπτει ως Pass/No Pass, και χρησιμοποιείται για την αξιολόγηση της συμμόρφωσης του προϊόντος σύμφωνα με τις προδιαγραφές του πελάτη, αλλά και τις κανονιστικές/νομοθετικές απαιτήσεις. Η ένδειξη Pass αποδίδεται σε προϊόντα οι τιμές των οποίων ικανοποιούν τις απαιτούμενες προδιαγραφές και τα καθορισμένα όρια λειτουργίας, ενώ η ένδειξη No Pass αντιστοιχεί σε προϊόν για το οποίο τουλάχιστον μία από τις ελεγχόμενες παραμέτρους δεν βρίσκεται εντός των αποδεκτών ορίων. Οι έλεγχοι πραγματοποιούνται σειριακά σε αυτόματη ακολουθία από κατάλληλο πολύ-όργανο, το οποίο διακριβώνεται στα προτεινόμενα από τον κατασκευαστή και τα πρότυπα Ποιότητας διαστήματα. τα δεδομένα της μελέτης αποτελούν διαδοχικές μετρήσεις τελικού ελέγχου παραγωγής και όχι επαναλαμβανόμενες μετρήσεις του ίδιου προϊόντος στον χρόνο. Η διάκριση

αυτή είναι ουσιώδης, καθώς επηρεάζει τον τρόπο εφαρμογής τόσο των μεθόδων SPC όσο και των μεθόδων

Μηχανικής Μάθησης, στην περίπτωση της επιβλεπόμενης μάθησης (Supervised Learning), όπου το αντιμετωπίζουμε ως πρόβλημα ταξινόμησης.

Επιπλέον, είναι σημαντικό σε αυτό το σημείο να τονίσουμε ότι η ύπαρξη πολλαπλών (3) τριών οικογενειών προϊόντων και διάφορων τεχνικών διαμορφώσεων απαιτεί ειδική αντιμετώπιση. Εξαιτίας της διαφορετικής ισχύος των μοντέλων και των διαμορφώσεων (π.χ. Τύπος τροφοδοτικού, αριθμός φωτοдиодων LED), οι ηλεκτρικές μετρήσεις δεν είναι άμεσα συγκρίσιμες μεταξύ διαφορετικών προϊόντων. Η ύπαρξη διαφορετικών οικογενειών προϊόντων, μοντέλων και συγκεκριμένων διαμορφώσεων ανά μοντέλο, είναι κρίσιμη τόσο για την εφαρμογή του Στατιστικού Ελέγχου Διεργασιών όσο και για τη διαδικασία της μοντελοποίησης ML. Στο πλαίσιο του SPC, διαφοροποιείται ο τρόπος καθορισμού των ορίων και υποομάδων, ενώ στη Μηχανική Μάθηση μεταβάλλεται η μορφή των χαρακτηριστικών εισόδου και η δυνατότητα επέκτασης των μοντέλων μεταξύ διαφορετικών προϊόντων. Για τον λόγο αυτό, στη συνέχεια εφαρμόζονται διάφορες τεχνικές για την ομαδοποίηση ανά οικογένεια ή μοντέλο, την κανονικοποίηση των μεταβλητών και τις μεταβολές των ορίων ελέγχου.

Όσο ν' αφορά τα όρια των διαγραμμάτων ελέγχου, αυτά διαφοροποιούνται ανάλογα με την ισχύ ή την τεχνική διαμόρφωση. Συνεπώς, η αξιολόγηση των αποτελεσμάτων πρέπει να πραγματοποιείται λαμβάνοντας υπόψη αυτές τις διαφοροποιήσεις. Οι κρίσιμες τιμές των ορίων που λαμβάνονται υπόψη συνοψίζονται ως εξής:

**Πίνακας 3 Κρίσιμα όρια χαρακτηριστικών προϊόντος**

| Είδος Ορίου                  | Σύμβολο  | Χρήση στην ανάλυση                             |
|------------------------------|----------|------------------------------------------------|
| Όρια ελέγχου διεργασίας      | LCL, UCL | Έλεγχος φυσικής μεταβλητότητας της διεργασίας  |
| Όρια προδιαγραφών            | LSL, USL | Έλεγχος συμφωνίας με προδιαγραφές              |
| Νομοθετικά/ κανονιστικά όρια | LL, UL   | Έλεγχος συμμόρφωσης με κανονιστικές απαιτήσεις |

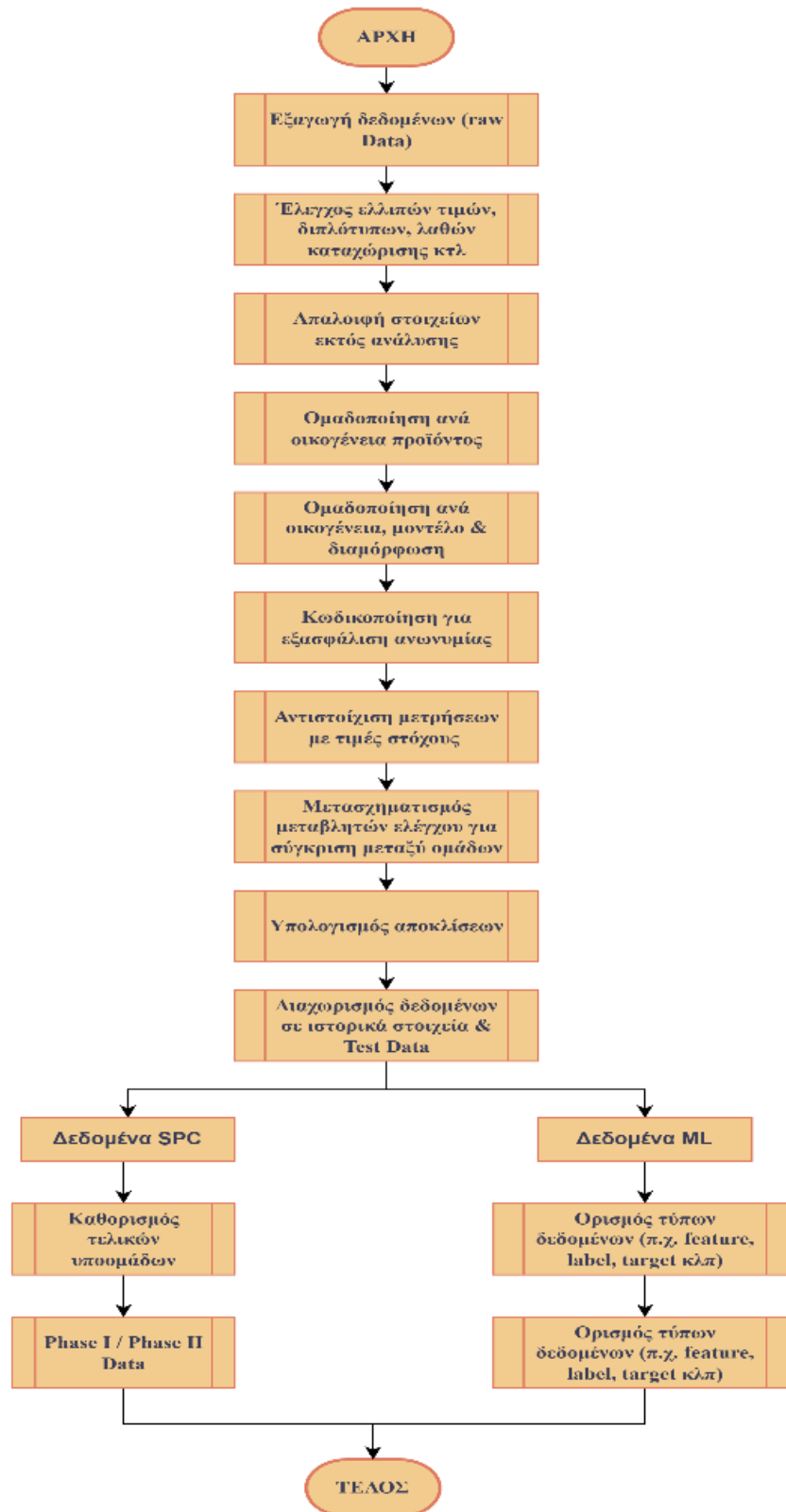
Τελικά, ο σχεδιασμός της μεθοδολογίας οφείλει να λαμβάνει υπόψη και συγκεκριμένους περιορισμούς που προκύπτουν τόσο από τη μορφή των δεδομένων, όσο και από τις ιδιαιτερότητες της Παραγωγικής Διαδικασίας. Πρώτον, τα δεδομένα λαμβάνονται μόνο κατά

το στάδιο του τελικού ηλεκτρικού ελέγχου (1 σημείο δειγματοληψίας). Ως αποτέλεσμα, δεν αξιολογούνται παράγοντες που προκαλούν διασπορά και εμφανίζονται σε προηγούμενο στάδιο της ΠΔ. Επιπλέον, όπως αναφέραμε και νωρίτερα, η ύπαρξη διαφορετικών οικογενειών, μοντέλων και μεταβλητών ορίων ελέγχου απαιτεί αρκετά στάδια προετοιμασίας των δεδομένων, ώστε να μην συγκρίνονται ανόμοια προϊόντα και οδηγούμαστε σε λανθασμένη ερμηνεία των αποτελεσμάτων. Τέλος, δεδομένου ότι η φυσική μεταβλητότητα της διεργασίας είναι σημαντικά μικρότερη από τα όρια των προδιαγραφών (Specification Limits – SL), δημιουργείται σχετική ανισορροπία μεταξύ εγγραφών Pass και No Pass, γεγονός το οποίο μας περιορίζει σημαντικά στην εφαρμογή της ταξινόμησης των μοντέλων Μηχανικής Μάθησης.

### 3.4 Προεπεξεργασία και προετοιμασία δεδομένων

Η προ-επεξεργασία των διαθέσιμων δεδομένων αποτελεί ένα από τα κυριότερα στάδια εφαρμογής της μεθοδολογίας, καθώς οι κατάλληλοι μετασχηματισμοί των δεδομένων και οι παραδοχές που λαμβάνονται σε αυτό το στάδιο επηρεάζουν σημαντικά την ποιότητα των αποτελεσμάτων της ανάλυσης. Για τον λόγο αυτό, η προεπεξεργασία δεν αφορά αποκλειστικά στον εντοπισμό και τη διόρθωση πιθανών ελλείψεων ή σφαλμάτων στα δεδομένα, αλλά περιλαμβάνει και τον μετασχηματισμό των μετρήσεων σε μορφή κατάλληλη για συγκριτική ανάλυση. Θεμελιώδης επιλογή της μεθοδολογίας που εφαρμόζεται είναι η χρήση σχετικών (%) μεταβολών και αποκλίσεων, ώστε μετρήσεις από προϊόντα διαφορετικής ισχύος και τεχνικής διαμόρφωσης να μπορούν να συγκριθούν και να αξιολογηθούν σε κοινή βάση.

Η διαδικασία προετοιμασίας αποτελείται από αρκετά στάδια και στοχεύει στη δημιουργία δύο κατάλληλα διαμορφωμένων συνόλων δεδομένων, ενός συνόλου για την εφαρμογή των παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών και ενός συνόλου για την ανάπτυξη και αξιολόγηση μοντέλων Μηχανικής Μάθησης. Τονίζουμε ότι τα δύο dataset προκύπτουν από την ίδια βάση δεδομένων του ηλεκτρικού ελέγχου. Η διαφοροποίηση κρίνεται απαραίτητη, καθώς οι 2 μέθοδοι έχουν διαφορετικές απαιτήσεις. Η διαδικασία προεπεξεργασίας και προετοιμασίας των δεδομένων παρουσιάζεται συνοπτικά στο Διάγραμμα 5.



Διάγραμμα 5 Διαδικασία προεπεξεργασίας και προετοιμασίας δεδομένων

### Καθαρισμός & αρχική επεξεργασία:

Αρχικά, εξάγουμε τα δεδομένα από τη βάση και κάνουμε τους απαραίτητους μετασχηματισμούς που μας επιτρέπουν να τα επεξεργαστούμε. Στη συνέχεια, ελέγχουμε την πληρότητα και τη δομή των διαθέσιμων δεδομένων, δηλαδή την ύπαρξη κενών τιμών, διπλότυπων και μη λογικών τιμών (π.χ. λανθασμένες καταχωρίσεις). Σε αυτό το σημείο, είναι σημαντικό να διαπιστώσουμε ότι τα Serial ID των φωτιστικών είναι μοναδικά και κάθε εγγραφή αντιστοιχεί σε μοναδικό προϊόν. Σε περίπτωση που υπάρχουν διπλότυπα, ερευνάται η αιτία (π.χ. λόγω επανακατεργασίας μετά από αποτυχημένο έλεγχο, λάθος χειριστή, σφάλμα καταγραφής), ώστε η εγγραφή να κατηγοριοποιηθεί κατάλληλα. Παράλληλα, ελέγχεται η ορθότητα των τύπων δεδομένων, ώστε ιδιαίτερα οι αριθμητικές μεταβλητές να μπορούν να χρησιμοποιηθούν για υπολογισμούς και οι χρονικές μεταβλητές να μας καθοδηγούν αναφορικά με την παραγωγική ακολουθία.

Οι βασικές ενέργειες που περιγράφονται, παρουσιάζονται συνοπτικά στον Πίνακα 4.

**Πίνακας 4 Διαδικασία καθαρισμού δεδομένων ποιοτικού ελέγχου**

| Έλεγχος                  | Σκοπός                                                   | Ενέργεια                                |
|--------------------------|----------------------------------------------------------|-----------------------------------------|
| Πληρότητα βασικών πεδίων | Επιβεβαίωση ύπαρξης των απαραίτητων μεταβλητών           | Έλεγχος και σήμανση ελλειπών εγγραφών   |
| Κενές τιμές              | Εντοπισμός μη διαθέσιμων μετρήσεων                       | Εξαίρεση ή διερεύνηση ανά περίπτωση     |
| Τύποι δεδομένων          | Διασφάλιση σωστής μορφής αριθμητικών και χρονικών πεδίων | Μετατροπή τύπων / καθαρισμός            |
| Serial ID                | Έλεγχος μοναδικότητας και ιχνηλασιμότητας                | Εντοπισμός διπλοτύπων ή επανελέγχων     |
| Timestamp                | Διατήρηση χρονικής σειράς παραγωγής                      | Ταξινόμηση με βάση τη χρονική ακολουθία |
| Pass / No Pass           | Έλεγχος ύπαρξης ετικέτας αποτελέσματος                   | Κωδικοποίηση σε δυαδική μορφή (0/1)     |
| Μη ρεαλιστικές τιμές     | Εντοπισμός πιθανών σφαλμάτων μέτρησης ή καταγραφής       | Σήμανση και τεχνική αξιολόγηση          |

### Ομαδοποίηση:

Αφού ολοκληρωθεί ο αρχικός καθαρισμός και έλεγχος των δεδομένων, ακολουθεί η ομαδοποίηση των εγγραφών με βάση τα τεχνικά χαρακτηριστικά του προϊόντος. Η ομαδοποίηση αυτή πραγματοποιείται αρχικά σε επίπεδο οικογένειας προϊόντος, στη συνέχεια

ανά μοντέλο (W) και τέλος, όπου είναι εφικτό, ανά τεχνική διαμόρφωση. Σε κάθε επίπεδο εξετάζεται το πλήθος των διαθέσιμων εγγραφών, ώστε να αξιολογηθεί αν η ομάδα μπορεί να χρησιμοποιηθεί για ξεχωριστή SPC ή ML ανάλυση. Με αυτό τον τρόπο, εξαλείφεται ο κίνδυνος να συγκρίνουμε προϊόντα χωρίς προηγούμενη ομαδοποίηση ή κανονικοποίηση, το οποίο θα μπορούσε να οδηγήσει σε λανθασμένη ερμηνεία της μεταβλητότητας. Για να γίνει απολύτως κατανοητό σε ποια χαρακτηριστικά οφείλεται η διαφοροποίηση, παρουσιάζουμε τον Πίνακα 5.

Πίνακας 5 Στάδια Κατηγοριοποίησης προϊόντων

| Επίπεδο ομαδοποίησης | Περιγραφή                                                                                                           | Χρήση στη μεθοδολογία               |
|----------------------|---------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| Οικογένεια προϊόντος | Βασική κατηγορία προϊόντος                                                                                          | Γενική ανάλυση και αρχική σύγκριση  |
| Μοντέλο / ισχύς      | Διαφοροποίηση βάσει ονομαστικής ισχύος<br>(π.χ. Οικογένεια A, 40W) " Μοντέλο A40                                    | Τεχνικά ομοιογενέστερη ανάλυση      |
| Τεχνική Διαμόρφωση   | Συγκεκριμένη τεχνική διαμόρφωση<br>(Π.χ. Οικογένεια A, 40W, Driver P, 30 LED " Διαμόρφωση A40P30 ≠ A40T30 ≠ A40P40) | Ανάλυση όπου υπάρχει επαρκές πλήθος |

Σε διαμορφώσεις με περιορισμένο αριθμό εγγραφών αξιολογείται η καταλληλότητα ένταξής τους ως ξεχωριστή κατηγορία ανάλογα με την επιμέρους ανάλυση. Για την εφαρμογή Στατιστικού Ελέγχου Διεργασιών, χρειαζόμαστε επαρκές πλήθος δεδομένων για να εκτιμήσουμε τη φυσική μεταβλητότητα της διεργασίας, ενώ αντίστοιχα αρκετές εγγραφές απαιτούνται και για την εκπαίδευση (training) των ML μοντέλων.

#### Μετασχηματισμός σε σχετικά (relative) μεγέθη:

Για να αντιμετωπίσουμε την πολυπλοκότητα λόγω οικογενειών, μοντέλων και διαμορφώσεων, οι κρίσιμες μεταβλητές μετασχηματίζονται σε σχετικές μεταβολές και ποσοστιαίες αποκλίσεις, βάση της προδιαγραφόμενης ή ονομαστικής τιμής.

Ενδεικτικά, για μία μετρούμενη τιμή ( $x_i$ ) και αντίστοιχη ονομαστική τιμή ( $x_{nom}$ ) η απόκλιση και η ποσοστιαία απόκλιση μπορούν να υπολογιστούν ως εξής:

$$\text{Απόκλιση} = x_i - x_{nom}$$

$$\text{Ποσοστιαία Απόκλιση}\% = \frac{x_i - x_{nom}}{x_{nom}} * 100$$

Είναι προφανές ότι η ποσοστιαία απόκλιση είναι μέγεθος απαλλαγμένο από την επίδραση της κλίμακας της μετρούμενης τιμής, άρα ουσιαστικά αυτός ο μετασχηματισμός τυποποιεί τις αποκλίσεις ώστε να μπορούν να συγκριθούν και προϊόντα διαφορετικής ισχύος(μοντέλα). Οι σχετικές αποκλίσεις μπορούν να χρησιμοποιηθούν τόσο στην εφαρμογή ΣΕΔ όσο και στη δημιουργία μεταβλητών εισόδου για τα μοντέλα Μηχανικής Μάθησης. Μετά τον μετασχηματισμό, η ελεγχόμενη τιμή δεν είναι η μέτρηση του ηλεκτρικού ελέγχου, αλλά η ποσοστιαία απόκλιση από την τιμή στόχο.

#### Χρονικός Διαχωρισμός:

Τόσο στον Στατιστικό Έλεγχο Διεργασιών, όσο και στους αλγορίθμους Machine Learning, απαιτείται διαχωρισμός των δεδομένων σε 2 κατηγορίες. Στην περίπτωση του SPC, τα πρώτα δεδομένα που συλλέγονται, και πάντα με την υπόθεση ότι η διεργασία βρίσκεται εντός στατιστικού ελέγχου, χρησιμοποιούνται για τον προσδιορισμό των χαρακτηριστικών και των τιμών των Διαγραμμάτων Ελέγχου (Phase I). Αντίστοιχα, για ML το πρώτο τμήμα των δεδομένων (training dataset) χρησιμοποιείται για την εκπαίδευση των μοντέλων, ώστε να μπορούμε να βγάλουμε συμπεράσματα για το επόμενο τμήμα (test dataset). Σε πολλές εφαρμογές Μηχανικής Μάθησης, ο διαχωρισμός μεταξύ των 2 κατηγοριών δεδομένων είναι τυχαίος. Στο πλαίσιο αυτής της εργασίας, εμείς θα ακολουθήσουμε διαφορετική προσέγγιση κατηγοριοποιώντας τα δεδομένα χρονικά. Για λόγους συνέπειας, ο χρονικός διαχωρισμός των δεδομένων είναι αντίστοιχος μεταξύ των δύο προσεγγίσεων. Με τον τρόπο αυτό, η σύγκριση πραγματοποιείται σε κοινή χρονική βάση και περιορίζεται ο κίνδυνος διαρροής πληροφορίας.

#### Δεδομένα SPC:

Το σύνολο δεδομένων για την εφαρμογή των μεθόδων Στατιστικού Ελέγχου Διεργασιών διαμορφώνεται με στόχο την παρακολούθηση της φυσικής μεταβλητότητας της διεργασίας. Η εφαρμογή των διαγραμμάτων SPC πραγματοποιείται σε μετασχηματισμένες μεταβλητές ποσοστιαίας απόκλισης από τις τιμές στόχους, ώστε τα υπολογιζόμενα όρια ελέγχου (LCL, UCL) να σχετίζονται με τη συμπεριφορά της διεργασίας και όχι τις τεχνικές διαφοροποιήσεις μεταξύ προϊόντων. Σε αυτό το στάδιο εξετάζεται η επάρκεια του πλήθους παρατηρήσεων ανά ομάδα, ώστε να αποφασιστεί σε ποιο επίπεδο είναι δόκιμο να πραγματοποιηθεί η ανάλυση.

Με βάση το πλήθος και τη δομή των δεδομένων, η ανάλυση πραγματοποιείται σε επίπεδο μοντέλου και, σε κάποιες περιπτώσεις, σε επίπεδο τεχνικής διαμόρφωσης. Σε τεχνικές διαμορφώσεις μικρών ομάδων, οι εγγραφές δεν διαχωρίζονται και χρησιμοποιούνται για τον προσδιορισμό των Ορίων Ελέγχου των μοντέλων. Η διαδικασία κατασκευής των Δ.Ε. παρουσιάζεται αναλυτικά στο 3.5.

### Δεδομένα ML:

Το σύνολο δεδομένων για τη Μηχανική Μάθηση διαμορφώνεται ως πίνακας χαρακτηριστικών εισόδου και μεταβλητής στόχου. Τα χαρακτηριστικά εισόδου περιλαμβάνουν ηλεκτρικές μετρήσεις (ισχύος, PF, V, I), σχετικές αποκλίσεις από τις τιμές στόχους, καθώς και πληροφορίες για την ταξινόμηση του προϊόντος, όπως η οικογένεια, το μοντέλο και η διαμόρφωση. Η μεταβλητή εξόδου - στόχος είναι το αποτέλεσμα Αποδοχής/Απόρριψης ενός προϊόντος.

Η επιλογή των χαρακτηριστικών εισόδου γίνεται με κριτήριο τη δυνατότητα ανίχνευσης αποκλίσεων αποφεύγοντας τη χρήση πεδίων που μας οδηγούν σε διαρροή πληροφορίας. Ως διαρροή πληροφορίας ορίζεται η χρήση μεταβλητών εισόδου που περιέχουν άμεσα ή έμμεσα ενδείξεις για το τελικό αποτέλεσμα αποδοχής ή απόρριψης. Η εισαγωγή τέτοιων μεταβλητών σε αλγόριθμο ML οδηγεί το μοντέλο να αναπαράγει τον υφιστάμενο κανόνα απόφασης αντί να διερευνήσει τα δεδομένα και να εξάγει χρήσιμα πρότυπα - σχέσεις. Για τον λόγο αυτό, τα πεδία που αποτελούν άμεσα αποτέλεσμα της τελικής απόφασης ελέγχου δεν χρησιμοποιούνται ως μεταβλητές εισόδου. Ο Πίνακας 6 περιέχει συνοπτικές πληροφορίες για τα δεδομένα που χρησιμοποιούνται στο ML.

**Πίνακας 6 Κατηγοριοποίηση δεδομένων ML**

| Στοιχείο dataset ML                     | Περιγραφή                                                                                                      |
|-----------------------------------------|----------------------------------------------------------------------------------------------------------------|
| Μεταβλητές Εισόδου                      | Ηλεκτρικές μετρήσεις<br>Σχετική απόκλιση ( $\Delta P$ %)<br>Πληροφορίες Προϊόντος<br>Ποσοστό ελαττωματικών (p) |
| Στόχος – Μεταβλητή Εξόδου               | Απόρριψη / Αποδοχή                                                                                             |
| Χωρισμός μεταβλητών εκπαίδευσης/ελέγχου | Χρονικός διαχωρισμός                                                                                           |
| Δείκτες απόδοσης                        | Alarms, % ελαττωματικών που ανιχνεύθηκαν, ευαισθησία                                                           |

Όπως και με την περίπτωση του SPC, ο διαχωρισμός εκπαίδευσης και αξιολόγησης πραγματοποιείται χρονικά. Τα παλαιότερα δεδομένα χρησιμοποιούνται για την εκπαίδευση των μοντέλων, ενώ τα νεότερα δεδομένα χρησιμοποιούνται για την αξιολόγηση της απόδοσής τους.

Η ανισορροπία μεταξύ αποδεκτών (Pass) και απορριπτέων (No Pass) εγγραφών αποτελεί μείζον αλλά αναμενόμενο ζήτημα, καθώς σε μία πραγματική παραγωγική διαδικασία το πλήθος των συμμορφούμενων προϊόντων είναι σημαντικά μεγαλύτερο από το πλήθος των μη συμμορφούμενων (p). Η συγκεκριμένη ανισορροπία επιδρά κυρίως στην ανάπτυξη και αξιολόγηση επιβλεπόμενων μοντέλων ταξινόμησης (supervised learning models), δηλαδή μοντέλων στα οποία δίνουμε εκ των προτέρων το αποτέλεσμα του ελέγχου (σε αυτή την περίπτωση P/NP), ενώ πρέπει να λαμβάνεται υπόψη και κατά την ερμηνεία μη επιβλεπόμενων μεθόδων (unsupervised learning), όπως ανίχνευση ανωμαλιών. Το κυριότερο ζητούμενο στο στάδιο της αρχικής επεξεργασίας είναι η καταγραφή και ποσοτικοποίηση της ανισορροπίας. Οι τεχνικές που εφαρμόζονται αναλύονται στο υποκεφάλαιο 3.6.

### 3.5 Εφαρμογή παραδοσιακού SPC

Η εφαρμογή των παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών διενεργείται στο σύνολο δεδομένων που έχουμε ήδη επεξεργαστεί. Ο κύριος στόχος της ανάλυσης είναι η αξιολόγηση της σταθερότητας της διεργασίας, ο έγκαιρος εντοπισμός αποκλίσεων και η δημιουργία ενός σημείου αναφοράς για τη μετέπειτα σύγκριση με τις μεθόδους Μηχανικής Μάθησης. Σε αυτό το σημείο, καλούμαστε να επιλέξουμε τα κατάλληλα Διαγράμματα Ελέγχου (ΔΕ) για τις μεταβλητές που επιθυμούμε να ελέγξουμε.

Με γνώμονα το ότι κάθε εγγραφή αντιστοιχεί σε έναν τελικό ηλεκτρικό έλεγχο ενός προϊόντος και όχι σε επαναλαμβανόμενες μετρήσεις του ίδιου τεμαχίου, η βασική εφαρμογή του ΣΕΔ οφείλει να βασίζεται σε διαγράμματα μεμονωμένων παρατηρήσεων. Τα καταλληλότερα διαγράμματα είναι τα I-MR ή ΔΕ μεμονωμένων παρατηρήσεων και κινούμενου εύρους. Τα διαγράμματα αυτά είναι κατάλληλα για τη διαδοχική παρακολούθηση μεμονωμένων παρατηρήσεων, ιδιαίτερα όταν δεν υπάρχουν φυσικά σχηματισμένες υποομάδες. Σημειώνεται εδώ ότι εφόσον έχουμε ομοιογενείς ομάδες σε ορθή χρονική ακολουθία, θα μπορούσαμε να ομαδοποιήσουμε τις παρατηρήσεις και να χρησιμοποιήσουμε τις υπο-ομάδες για να

παράγουμε διαγράμματα μέσου όρου – εύρους  $\bar{X}$ -R ή  $\bar{X}$ -S. Στο πλαίσιο της παρούσας εργασίας, αυτό δεν είναι θεμιτό δεδομένου ότι οι μετρήσεις που έχουμε δεν πληρούν σε κάθε περίπτωση αυτές τις προϋποθέσεις (π.χ. έχουμε διαδοχικά δεδομένα ίδιου μοντέλου αλλά διαφορετικής μέρας, έχουμε διαδοχικές μετρήσεις διαφορετικών μοντέλων).

Συνεπώς, εφεξής θα εργαστούμε με διαγράμματα ελέγχου I-MR που εφαρμόζονται σε σχετικές αποκλίσεις κρίσιμων ηλεκτρικών μεταβλητών, όπως η ισχύς(P), ο συντελεστής ισχύος(PF) και η αρμονική παραμόρφωση(THDu%), όπου αυτό είναι μεθοδολογικά εφικτό. Το διάγραμμα μεμονωμένων παρατηρήσεων παρακολουθεί την εξέλιξη κάθε μέτρησης στη χρονική ή παραγωγική ακολουθία, ενώ το διάγραμμα κινούμενου εύρους παρακολουθεί τη μεταβολή μεταξύ διαδοχικών παρατηρήσεων.

Η εφαρμογή των διαγραμμάτων SPC πραγματοποιείται σε επίπεδο ομοιογενών ομάδων δεδομένων, όπως περιγράψαμε παραπάνω. Δεν εφαρμόζεται ενιαία ανάλυση αδιακρίτως σε όλα τα προϊόντα, καθώς αυτό θα οδηγούσε σε σύγκριση ανόμοιων τεχνικών διαμορφώσεων. Αντίθετα, εξετάζεται η δυνατότητα εφαρμογής αρχικά σε επίπεδο οικογένειας προϊόντος, στη συνέχεια σε επίπεδο μοντέλου και τέλος σε επίπεδο τεχνικής διαμόρφωσης, ανάλογα με το διαθέσιμο πλήθος παρατηρήσεων και τον βαθμό τεχνικής ομοιογένειας.

### Phase I:

Στην παρούσα εργασία, δεν θα χρησιμοποιήσουμε προηγουμένως καθορισμένα όρια Phase I, αλλά θα χρησιμοποιήσουμε αναδρομική προσέγγιση. Συγκεκριμένα, ένα αρχικό χρονικό τμήμα των διαθέσιμων δεδομένων αξιοποιείται για την εκτίμηση της φυσικής μεταβλητότητας της διεργασίας και τον υπολογισμό των ορίων ελέγχου. Τα μεταγενέστερα δεδομένα χρησιμοποιούνται στη συνέχεια για την αξιολόγηση της σταθερότητας της διεργασίας σε λογική Phase II.

Στο Phase I υπολογίζουμε την κεντρική γραμμή(KΓ) και τα όρια ελέγχου(OE) για κάθε μεταβλητή και ομάδα που εξετάζεται. Συνήθως, στο Phase I προτιμούμε να μην συμπεριλαμβάνονται στον υπολογισμό των KΓ και OE σημεία εκτός ορίων. Σε περίπτωση που εντοπιστούν εμφανή σημεία εκτός ελέγχου κατά την αρχική εκτίμηση, αυτά εξετάζονται ως πιθανές ειδικές αιτίες. Εφόσον υπάρχει τεχνική τεκμηρίωση ότι πρόκειται για ειδική αιτία ή σφάλμα καταγραφής, μπορεί να πραγματοποιηθεί επανυπολογισμός των ορίων ελέγχου. Σε

αντίθετη περίπτωση, τα σημεία διατηρούνται και η επίδρασή τους λαμβάνεται υπόψη στην ερμηνεία.

Για την περίπτωση των I-MR charts, το διάγραμμα μεμονωμένων παρατηρήσεων (I) βασίζεται στο μέσο επίπεδο των παρατηρήσεων και στη μεταβλητότητα που εκτιμάται από το κινούμενο εύρος (MR). Ενδεικτικά, για το Individuals Chart:

$$K\Gamma = \bar{X}$$

$$UCL = \bar{X} + \frac{3 * \overline{MR}}{d_2}$$

$$LCL = \bar{X} - \frac{3 * \overline{MR}}{d_2}$$

,  $\bar{X}$  όπου είναι ο μέσος των μεμονωμένων παρατηρήσεων της Phase I περιόδου,  $\overline{MR}$  είναι το μέσο κινούμενο εύρος και  $d_2$  είναι σταθερά που εξαρτάται από το μέγεθος του κινούμενου εύρους. Για κινούμενο εύρος δύο διαδοχικών παρατηρήσεων, η τιμή του  $d_2$  είναι 1,128.

Είναι σημαντικό να τονίσουμε ότι η προσέγγιση αυτή επιτρέπει τον υπολογισμό ορίων ελέγχου που προκύπτουν από τη φυσική μεταβλητότητα της διεργασίας και όχι από τα όρια προδιαγραφών, τα οποία δεν έχουν σχέση με τη διεργασία, αλλά με τις απαιτήσεις του πελάτη και της νομοθεσίας. Με τον τρόπο αυτό διατηρείται σαφής διάκριση μεταξύ της στατιστικής σταθερότητας της διεργασίας και της συμμόρφωσης του προϊόντος ως προς τις τεχνικές προδιαγραφές.

### Phase II:

Αφού έχουμε υπολογίσει τα όρια στο Phase I, τα ΔΕ εφαρμόζονται στα μεταγενέστερα δεδομένα (Phase II). Τα δεδομένα της 2ης φάσης δεν χρησιμοποιούνται για τον υπολογισμό των ορίων ελέγχου, αλλά μόνο ως αξιολόγηση της σταθερότητας της διεργασίας, ουσιαστικά αξιοποιούνται για να εξεταστεί εάν εμφανίζονται σημεία και πρότυπα που υποδηλώνουν πιθανή ειδική αιτία μεταβλητότητας.

Κατά τη Φάση II, καταγράφονται τα σημεία που υπερβαίνουν τα όρια ελέγχου, η χρονική τους διαδοχή, η μεταβλητή στην οποία εμφανίζονται και το αντίστοιχο προϊόν. Ύστερα εξετάζεται αν η ύπαρξη αυτών των αποκλίσεων σηματοδοτεί ότι η διεργασία έχει αρχίσει να παρεκκλίνει

του στατιστικού ελέγχου ή αν πρόκειται για False Alarm. Η πληροφορία αυτή αξιοποιείται στη συνέχεια για τη σύγκριση με τα αποτελέσματα των ML-SPC προσεγγίσεων.

Η ερμηνεία των διαγραμμάτων SPC δεν περιορίζεται αποκλειστικά στον εντοπισμό σημείων εκτός των ορίων ελέγχου. Εκτός από την υπέρβαση των LCL ή UCL, αξιολογείται και η εμφάνιση μη τυχαίων προτύπων στη συμπεριφορά των δεδομένων (π.χ. Western Electric Rules), όπως συνεχόμενα σημεία από την ίδια πλευρά της κεντρικής γραμμής, αυξητικές ή πτωτικές τάσεις, απότομες μεταβολές ή ασυνήθιστη συγκέντρωση σημείων κοντά στα όρια ελέγχου.

Πέρα από τα Διαγράμματα Ελέγχου I-MR, μπορούν επίσης να εξεταστούν άλλες προσεγγίσεις SPC. Για παράδειγμα, τα διαγράμματα EWMA μπορούν να χρησιμοποιηθούν για την ανίχνευση μικρών ή σταδιακών μετατοπίσεων στη διεργασία, καθώς ενσωματώνουν την πληροφορία από προηγούμενες μετρήσεις και είναι πιο ευαίσθητα σε μικρές μεταβολές σε σχέση άλλα ΔΕ.

Τα αποτελέσματα της SPC ανάλυσης θα συγκριθούν στη συνέχεια με τις μεθόδους Μηχανικής Μάθησης. Συνεπώς, από κάθε εφαρμογή SPC εξάγονται συγκεκριμένοι δείκτες, οι οποίοι αποτυπώνουν την ικανότητα των διαγραμμάτων να εντοπίζουν αποκλίσεις και να υποστηρίζουν τη λήψη αποφάσεων μέσα στο Σύστημα Διαχείρισης Ποιότητας. Σε αυτό το σημείο, είναι σημαντικό να τονίσουμε ότι το SPC αξιολογείται κυρίως ως εργαλείο παρακολούθησης της μεταβλητότητας διεργασίας και δεν δύναται να συγχέεται με το Pass / No Pass, τα οποία αποτελούν σημεία εκτός προδιαγραφών. Επομένως, η σχέση μεταξύ SPC signals και Pass / No Pass δεν πρέπει να αντιμετωπίζεται ως απλή σχέση πρόβλεψης, αλλά ως ένδειξη του κατά πόσο οι στατιστικές αποκλίσεις της διεργασίας συνδέονται με ταυτόχρονες ή μεταγενέστερες μη συμμορφώσεις προϊόντων.

Διακρίνουμε λοιπόν τις παρακάτω περιπτώσεις:

**Πίνακας 7 Σχέση Ορίων Ελέγχου SPC και Ορίων Ελέγχου Προδιαγραφών**

| Σενάριο | Εντός SPC control limits; | Εντός specification limits; | Ερμηνεία                                               |
|---------|---------------------------|-----------------------------|--------------------------------------------------------|
| 1       | Ναι                       | Ναι                         | Σταθερή διεργασία και συμμορφούμενο προϊόν             |
| 2       | Όχι                       | Ναι                         | Διεργασιακή απόκλιση, αλλά ακόμη συμμορφούμενο προϊόν  |
| 3       | Ναι                       | Όχι                         | Προϊόν εκτός προδιαγραφών χωρίς προηγούμενο SPC signal |
| 4       | Όχι                       | Όχι                         | Απόκλιση διεργασίας και μη συμμορφούμενο προϊόν        |

Με βάση τα παραπάνω, οι δείκτες που μπορούμε να εξάγουμε και να αξιοποιήσουμε για την αξιολόγηση είναι:

**Πίνακας 8 Δείκτες αξιολόγησης SPC**

| Δείκτης / έξοδος                  | Υπολογισμός                                           | Ερμηνεία                                            |
|-----------------------------------|-------------------------------------------------------|-----------------------------------------------------|
| Αριθμός SPC signals               | Πλήθος σημείων ή patterns που ενεργοποιούν κανόνα SPC | Συχνότητα ενδείξεων στατιστικής αστάθειας           |
| Signal rate                       | # SPC signals / σύνολο παρατηρήσεων                   | Ένταση ενεργοποίησης του SPC                        |
| No Pass rate                      | # No Pass / σύνολο παρατηρήσεων                       | Συχνότητα μη συμμορφώσεων βάσει προδιαγραφών        |
| Συνύπαρξη signal και No Pass      | Cross-tab SPC signal με Pass / No Pass                | Σχέση στατιστικής αστάθειας με συμμόρφωση προϊόντος |
| Αριθμός τεμαχίων πριν από No Pass | Απόσταση προηγούμενου SPC signal από επόμενο No Pass  | Πιθανή προειδοποιητική ικανότητα                    |
| Empirical run length              | Παρατηρήσεις μεταξύ διαδοχικών SPC signals            | Συχνότητα εμφάνισης σημάτων SPC                     |

Οι δείκτες αυτοί δεν ερμηνεύονται μεμονωμένα, αλλά σε συνδυασμό με το τεχνικό πλαίσιο κάθε προϊόντος και το αντίστοιχο επίπεδο ανάλυσης. Η ύπαρξη ενός συναγερμού από το SPC δεν σημαίνει απαραίτητα ότι το προϊόν είναι εκτός προδιαγραφών, αλλά ότι η διεργασία εμφανίζει συμπεριφορά που πιθανόν να δικαιολογεί αλλαγή κατάστασης από τον στατιστικό έλεγχο.

### 3.6 Ανάπτυξη μοντέλων Μηχανικής Μάθησης

Η εφαρμογή των μεθόδων Μηχανικής Μάθησης πραγματοποιείται ως συμπληρωματικό στάδιο της ανάλυσης SPC, με στόχο τη διερεύνηση πολυμεταβλητών προτύπων στα δεδομένα τελικού ηλεκτρικού ελέγχου. Σε αντίθεση με τα κλασσικά διαγράμματα ελέγχου, τα οποία

εφαρμόζονται κυρίως σε μία κρίσιμη μεταβλητή κάθε φορά, το υπολογιστικό πλαίσιο ML σχεδιάστηκε ώστε να αξιοποιεί ταυτόχρονα περισσότερα ηλεκτρικά χαρακτηριστικά κάθε τεμαχίου. Με τον τρόπο αυτό εξετάζεται αν οι συνδυασμοί ηλεκτρικών μετρήσεων μπορούν να αναδείξουν τεχνικές διαφοροποιήσεις, αλλαγές στη συμπεριφορά της διεργασίας ή πολυμεταβλητές ανωμαλίες που δεν είναι απαραίτητα εμφανείς από μία μόνο μεταβλητή. Η διαδικασία περιλαμβάνει διάφορα στάδια προεπεξεργασίας, επιλογής χαρακτηριστικών, προτυποποίησης, εφαρμογής των αλγορίθμων Μηχανικής Μάθησης και αξιολόγηση των αποτελεσμάτων ιδιαίτερα σε σχέση με τα σήματα του SPC.

Στην παρούσα εργασία, το τμήμα που αφορά τη Μηχανική Μάθηση καταρτίστηκε σε δύο διακριτές πλην συμπληρωματικές προσεγγίσεις. Η μία προσέγγιση αφορά την αναγνώριση του ηλεκτρικού αποτυπώματος διαφορετικών τεχνικών διαμορφώσεων. Ο στόχος είναι να εξεταστεί αν τα δεδομένα που προκύπτουν από τις μετρήσεις του τελικού ελέγχου παρέχουν αρκετή πληροφορία ώστε να γίνει εμφανής ο διαχωρισμός μεταξύ προϊόντων ίδιας οικογένειας και ισχύος, αλλά με διαφορετικούς drivers, πλακέτες LED ή συνδυασμούς τεχνικών ρυθμίσεων. Η δεύτερη προσέγγιση αφορά την ανίχνευση πολυμεταβλητών ανωμαλιών και τη σύγκριση των αποτελεσμάτων των αλγορίθμων Μηχανικής Μάθησης με τα alarms που προκύπτουν από τις κλασσικές μεθόδους SPC. Με τον τρόπο αυτό, το ML δεν χρησιμοποιείται μόνο ως αλγόριθμος ταξινόμησης Pass/No Pass, αλλά ως ισχυρό εργαλείο διερεύνησης της ηλεκτρικής υπογραφής με πολλές προεκτάσεις στην παρακολούθηση της παραγωγικής διαδικασίας.

Αρχικά, από το σύνολο των δεδομένων επιλέγουμε συγκεκριμένο υποσύνολο προϊόντων για την εφαρμογή των μεθόδων ML. Η κυριότερη αιτία είναι ότι οι αλγόριθμοι Μηχανικής Μάθησης έχουν συγκεκριμένες απαιτήσεις σε δεδομένα, όχι μόνο στο πλήθος αλλά και στην κατανομή. Συνεπώς, η επιλογή δεν βασίζεται αποκλειστικά στο συνολικό πλήθος των παρατηρήσεων, αλλά και στην ύπαρξη επαρκούς ποικιλίας τεχνικών διαμορφώσεων. Το υποσύνολο που χρησιμοποιούμε ως case study αφορά την ομάδα προϊόντων Α περίπου 60 W, η οποία περιλαμβάνει επιμέρους ισχύς 60,5 W, 61,5 W, 62,5 W και 63,5 W. Η ομαδοποίηση αυτή κρίθηκε αποδεκτή, αφού τα προϊόντα ανήκουν στην ίδια οικογένεια ισχύος, ενώ η χρήση κανονικοποιημένων μεταβλητών επιτρέπει την αποφυγή τεχνητού διαχωρισμού λόγω μικρών διαφορών στην αναμενόμενη ισχύ.

Για κάθε εγγραφή του συνόλου δεδομένων δημιουργούμε μοναδικό αναγνωριστικό διαμόρφωσης. Το Configuration ID σχηματίζεται από τον συνδυασμό βασικών τεχνικών μεταβλητών, όπως το μοντέλο, ο τύπος driver και η πλακέτα LED. Η χρήση του Configuration ID είναι απαραίτητη επειδή επιτρέπει την ομαδοποίηση των παρατηρήσεων που αντιστοιχούν σε ομοιογενείς τεχνικές συνθήκες. Για το τμήμα της Μη Επιβλεπόμενης Μάθησης (unsupervised), αποκλείουμε το Configuration ID, αφού ουσιαστικά το ερευνητικό ερώτημα είναι αν τα μοντέλα ML μπορούν να διακρίνουν το Config\_ID χωρίς αυτή την πληροφορία. Αντίθετα, το Config\_ID χρησιμοποιήθηκε ως μεταβλητή ερμηνείας μετά την εφαρμογή των αλγορίθμων ML, ώστε να εξεταστεί αν οι ομάδες που προκύπτουν από τα δεδομένα αντιστοιχούν σε πραγματικές τεχνικές διαφοροποιήσεις.

Το σύνολο των χαρακτηριστικών που χρησιμοποιήθηκε για την ανάπτυξη των αλγορίθμων ML περιορίστηκε σε αριθμητικές μεταβλητές που περιγράφουν το ηλεκτρικό αποτύπωμα του τεμαχίου. Συγκεκριμένα, χρησιμοποιήθηκαν η ΔΡ%, ο συντελεστής ισχύος, η ολική αρμονική παραμόρφωση, η τάση ελέγχου και ένας παράγωγος δείκτης ρεύματος ανά Watt (I/W). Ο δείκτης I/W χρησιμοποιήθηκε ώστε το ρεύμα εισόδου να εκφραστεί σε σχετική μορφή ως προς την απορροφούμενη ισχύ. Αντίθετα, δεν χρησιμοποιήθηκαν ως είσοδοι μεταβλητές όπως το αποτέλεσμα Pass/No Pass, το Configuration ID, ο τύπος τροφοδοτικού, η πλακέτα LED, το Pnom ή τα SPC alarms, καθώς η χρήση τους θα μπορούσε να οδηγήσει σε διαρροή δεδομένων που θα οδηγούσε το μοντέλο να συσχετίσει τις ανωμαλίες με αυτές τις μεταβλητές. Οι μεταβλητές αυτές διατηρήθηκαν μόνο για την αξιολόγηση και ερμηνεία των αποτελεσμάτων.

Μετά την επιλογή των χαρακτηριστικών, παρατηρήσεις με ελλιπή δεδομένα σε αριθμητικές μεταβλητές εξαιρέθηκαν από το υπολογιστικό dataset του ML, καθώς οι αλγόριθμοι PCA, clustering, Isolation Forest και supervised classification απαιτούν πλήρες αριθμητικό διάνυσμα χαρακτηριστικών για κάθε παρατήρηση. Στη συνέχεια, οι αριθμητικές μεταβλητές κανονικοποιήθηκαν με χρήση z-score, σύμφωνα με τη σχέση:

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

όπου  $x$  είναι η αρχική τιμή,  $\mu$  ο μέσος όρος της μεταβλητής και  $\sigma$  η τυπική απόκλιση.

Η τυποποίηση είναι μεγάλης σημασίας, καθώς οι μεταβλητές εκφράζονται σε διαφορετικές μονάδες και έχουν διαφορές στα αριθμητικά εύρη.

### 3.6.1 Αναγνώριση Τεχνικής Διαμόρφωσης

Η πρώτη προσέγγιση ML έχει ως στόχο να διερευνηθεί αν η ηλεκτρική «υπογραφή» των τεμαχίων, δηλαδή ο συνδυασμός των ηλεκτρικών χαρακτηριστικών τους είναι επαρκής για την αναγνώριση διαφορετικών τεχνικών configurations. Για τον σκοπό αυτό εφαρμόστηκαν τόσο unsupervised όσο και supervised μέθοδοι. Οι Μη επιβλεπόμενες μέθοδοι χρησιμοποιήθηκαν για να διαπιστωθεί αν τα δεδομένα σχηματίζουν φυσικές ομάδες χωρίς να δοθούν στον αλγόριθμο οι πραγματικές ετικέτες configuration, ενώ οι supervised μέθοδοι χρησιμοποιήθηκαν για να εξεταστεί αν τα ηλεκτρικά χαρακτηριστικά μπορούν να ταξινομήσουν σωστά τον τύπο driver ή το Configuration ID.

Αρχικά εφαρμόστηκε η μέθοδος της Ανάλυσης Κύριων Συνιστωσών (Principal Component Analysis - PCA) για τη μείωση της διάστασης των δεδομένων και την απεικόνιση της πολυμεταβλητής δομής τους. Από τα κανονικοποιημένα χαρακτηριστικά υπολογίστηκαν οι κύριες συνιστώσες, το ποσοστό διασποράς που καθεμία τους εξηγεί, καθώς και τα loadings (συντελεστές) των αρχικών μεταβλητών. Τα PCA scores επιτρέπουν την οπτική προβολή των παρατηρήσεων σε τρεις διαστάσεις, ενώ τα PCA loadings χρησιμοποιήθηκαν για να ερμηνευθεί ποια ηλεκτρικά χαρακτηριστικά είναι πιο κρίσιμα για τον διαχωρισμό των παρατηρήσεων. Για παράδειγμα, μία κύρια συνιστώσα (PC) που επηρεάζεται κυρίως από τον συντελεστή ισχύος και την αρμονική παραμόρφωση μπορεί να ερμηνευθεί ως άξονας διαφοροποίησης της συμπεριφοράς του driver.

Στη συνέχεια εφαρμόστηκαν μέθοδοι ταξινόμησης (clustering). Ο αλγόριθμος K-means χρησιμοποιείται ως κύρια μέθοδος ομαδοποίησης, διαχωρίζοντας τις παρατηρήσεις σε προκαθορισμένο αριθμό κλάσεων με βάση την επιμέρους απόστασή τους από τα κέντρα της κάθε ομάδας. Επειδή ο k-means υποθέτει μόνο κυκλικές κλάσεις, χρησιμοποιούμε συμπληρωματικά και το Gaussian Mixture Model, το οποίο υποθέτει ότι τα δεδομένα προέρχονται από ένα μίγμα πολυδιάστατων κανονικών κατανομών και αποδίδει πιθανότητες συμμετοχής κάθε παρατήρησης σε κάθε κλάση. Τα clusters που προκύπτουν συγκρίνονται με τα πραγματικά Configuration IDs και τους τύπους driver, ώστε να αξιολογηθεί κατά πόσο οι κλάσεις που προκύπτουν από τα ηλεκτρικά χαρακτηριστικά αντιστοιχούν σε πραγματικές τεχνικές διαφοροποιήσεις.

Σε συνέχεια της εφαρμογής Μη Επιβλεπόμενων μεθόδων, εφαρμόζουμε και επιβλεπόμενη Ταξινόμηση για την αναγνώριση του driver και του Configuration ID από το ηλεκτρικό αποτύπωμα. Ως μεταβλητές εισόδου χρησιμοποιούμε αποκλειστικά τις αριθμητικές ηλεκτρικές μεταβλητές (ΔΡ, V, I, PF, THDu), ενώ μεταβλητές στόχους ορίζουμε αφενός τον τύπο driver και συμπληρωματικά σε 2<sup>η</sup> φάση το Configuration ID. Για την επιβλεπόμενη ταξινόμηση αξιοποιούμε ευρέως χρησιμοποιούμενα μοντέλα, όπως Logistic Regression και Random Forest. Τα μοντέλα αξιολογούνται με δείκτες όπως η ακρίβεια (accuracy), ισορροπημένη ακρίβεια (balanced accuracy) και macro F1-score, καθώς και με πίνακες σύγχυσης. Ιδιαίτερη έμφαση δόθηκε στο balanced accuracy και στο macro F1-score, επειδή οι τεχνικές διαμορφώσεις δεν έχουν όλα τον ίδιο αριθμό παρατηρήσεων. Η αξιολόγηση πραγματοποιήθηκε με δύο τρόπους. Αρχικά, χρησιμοποιήθηκε stratified random split (τυχαίος διαχωρισμός σε test και training set), ώστε τόσο το training set όσο και το test set να περιλαμβάνουν δεδομένα από όλες τις κλάσεις. Στη συνέχεια εφαρμόζουμε αυστηρότερη αξιολόγηση, στην οποία ορισμένα Production Orders κρατήθηκαν εσκεμμένα εκτός εκπαίδευσης για να χρησιμοποιηθούν ως test set. Η δεύτερη αξιολόγηση θεωρείται πιο απαιτητική, καθώς εξετάζει αν ο αλγόριθμος δύναται να γενικεύσει σε διαφορετικά POs και όχι μόνο σε τυχαία τεμάχια από παρόμοιες συνθήκες παραγωγής. Η σύγκριση ανάμεσα στις 2 προσεγγίσεις μας επιτρέπει να εκτιμήσουμε αν η αναγνώριση των Configurations βασίζεται σε γενικεύσιμη ηλεκτρική υπογραφή ή επηρεάζεται σημαντικά από batch effects.

### 3.6.2 Ανίχνευση ανωμαλιών και σύγκριση με SPC

Η δεύτερη προσέγγιση ML έχει ως κύριο στόχο την ανίχνευση πολυμεταβλητών ανωμαλιών και τη σύγκριση των αποτελεσμάτων των μοντέλων Μηχανικής Μάθησης με τα alarms που προκύπτουν από τις κλασσικές μεθόδους SPC. Για τον σκοπό αυτό χρησιμοποιούμε κυρίως τον αλγόριθμο Isolation Forest (IF). Η μέθοδος που χρησιμοποιείται βασίζεται στη λογική ότι οι «ανώμαλες» παρατηρήσεις είναι πιο εύκολα να απομονωθούν από τις κανονικές, επειδή βρίσκονται σε πιο αραιές ή πιο απομακρυσμένες περιοχές του πολυδιάστατου χώρου χαρακτηριστικών.

Ο αλγόριθμος Isolation Forest εφαρμόστηκε αρχικά σε συνολικό επίπεδο για το επιλεγμένο μοντέλο A60, με σκοπό να εντοπιστούν τεμάχια με «ασυνήθιστο» ηλεκτρικό προφίλ σε σχέση με το σύνολο δεδομένων. Στη συνέχεια, το IF εφαρμόστηκε ανά τεχνική διαμόρφωση. Η

δεύτερη προσέγγιση προσομοιάζει περισσότερο στη λογική που χρησιμοποιήσαμε και στο Phase I SPC, αφού κάθε τεχνική διαμόρφωση αντιμετωπίζεται σαν ξεχωριστός πληθυσμός με δική του φυσιολογική συμπεριφορά. Συνεπώς, κάθε τεμάχιο - datapoint αξιολογείται σε σχέση μόνο με το δικό του τεχνικό configuration και όχι με όλα τα υπόλοιπα προϊόντα.

Τα αποτελέσματα του Isolation Forest εκφράζονται μέσω anomaly scores και anomaly flags (0/1). Το ποσοστό contamination χρησιμοποιείται ως παραδοχή για το αναμενόμενο ποσοστό «ασυνήθιστων» παρατηρήσεων και απαιτεί μεγάλη προσοχή, καθώς επηρεάζει την ευαισθησία του μοντέλου, το πλήθος δηλαδή των τεμαχίων που χαρακτηρίζονται ως anomalies.

Με τη δομή αυτή, το ML σκέλος της εργασίας οργανώνεται σε δύο αλληλοσυμπληρούμενα επίπεδα: αφενός ως εργαλείο αναγνώρισης ηλεκτρικών υπογραφών που σχετίζονται με διαφορετικά τεχνικά configurations και αφετέρου ως εργαλείο ανίχνευσης πολυμεταβλητών ανωμαλιών σε σύγκριση με τα κλασσικά SPC alarms. Η προσέγγιση αυτή επιτρέπει την αξιολόγηση της Μηχανικής Μάθησης όχι ως αντικατάσταση του SPC, αλλά ως πιθανή πολυμεταβλητή επέκταση ενός συστήματος στατιστικής παρακολούθησης διεργασιών.

### 3.7 Πλαίσιο σύγκρισης ML – SPC

Η σύγκριση μεταξύ των παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών και των μεθόδων Μηχανικής Μάθησης πραγματοποιείται με στόχο να διερευνηθεί αν οι ML-SPC προσεγγίσεις μπορούν να λειτουργήσουν συμπληρωματικά προς το SPC στην παρακολούθηση της παραγωγικής διεργασίας. Στα πλαίσια αυτού του έργου, ο ΣΔΕ και η Μηχανική Μάθηση δεν αντιμετωπίζονται ως πλήρως ισοδύναμες μέθοδοι, αλλά ως συμπληρωματικές προσεγγίσεις που απαντούν παραπλήσια αλλά διαφορετικά ερωτήματα. Το SPC εξετάζει κυρίως εάν μία διεργασία αποκλίνει από τη στατιστικά αναμενόμενη συμπεριφορά της, ενώ οι μέθοδοι Μηχανικής Μάθησης αξιολογούν ταυτόχρονα τον συνδυασμό πολλαπλών χαρακτηριστικών και μπορούν να αναδείξουν πολυμεταβλητές ανωμαλίες ή ηλεκτρικά πρότυπα που συνήθως δεν είναι εμφανή σε μονομεταβλητά διαγράμματα ελέγχου.

Στο πλαίσιο της παρούσας εργασίας, η σύγκριση επικεντρώνεται στη 2<sup>η</sup> προσέγγιση που αξιοποιήσαμε στο προηγούμενο υποκεφάλαιο, δηλαδή κυρίως στη σχέση μεταξύ των σημάτων SPC και των ενδείξεων ανωμαλίας που προκύπτουν από τα unsupervised μοντέλα Μηχανικής Μάθησης. Τα Διαγράμματα Ελέγχου του SPC έχουν ως έξοδο alarms που προκύπτουν από

μεταβολές στη μεταβλητή σχετικής απόκλισης ισχύος. Αντίστοιχα, τα ML anomaly flags προκύπτουν από την εφαρμογή μεθόδων όπως ο Isolation Forest, στο πολυμεταβλητό σύνολο των ηλεκτρικών χαρακτηριστικών των προϊόντων.

Για να καταστεί δυνατή αυτή η σύγκριση, δημιουργούμε ένα συνδυαστικό δείκτη SPC alarm, ο οποίος ενεργοποιείται όταν εμφανίζεται σήμα είτε στο I-chart είτε στο MR-chart (OR). Με τον τρόπο αυτό, κάθε παρατήρηση χαρακτηρίζεται δυαδικά (0/1) ως προς την ύπαρξη alarm στο SPC διάγραμμα, ανεξαρτήτως αν είναι στο I, στο MR ή και στα δύο. Αντίστοιχα, από το μοντέλο Μηχανικής Μάθησης προκύπτει δυαδική ένδειξη anomaly flag, η οποία δείχνει αν μία παρατήρηση έχει χαρακτηριστεί ως πολυμεταβλητή ανωμαλία. Τα παραπάνω συνοψίζονται στον Πίνακα 9.

**Πίνακας 9 Πιθανοί συνδυασμοί alarms**

| Ένδειξη             | Πηγή                                     | Ερμηνεία                                                 |
|---------------------|------------------------------------------|----------------------------------------------------------|
| SPC alarm = 1       | I-chart ή MR-chart                       | Στατιστική ένδειξη απόκλισης από τη φυσική μεταβλητότητα |
| SPC alarm = 0       | I-chart και MR-chart                     | Μη ύπαρξη σήματος SPC                                    |
| ML anomaly flag = 1 | Isolation Forest ή αντίστοιχο ML μοντέλο | Πολυμεταβλητή ένδειξη ασυνήθιστου ηλεκτρικού προφίλ      |
| ML anomaly flag = 0 | Isolation Forest ή αντίστοιχο ML μοντέλο | Μη ανίχνευση πολυμεταβλητής ανωμαλίας                    |

Η αξιολόγηση της σχέσης μεταξύ των προσεγγίσεων ML και SPC πραγματοποιείται με δημιουργία πινάκων συμφωνίας μεταξύ SPC alarms και ML anomaly flags. Οι πίνακες αυτοί αποτυπώνουν σε ποιες περιπτώσεις οι δύο μέθοδοι συμφωνούν και σε ποιες περιπτώσεις οι έξοδοι διαφοροποιούνται. Η συμφωνία μπορεί να αφορά είτε ταυτόχρονη αναγνώριση απόκλισης είτε ταυτόχρονη απουσία ένδειξης. Αντίστοιχα, η ασυμφωνία μπορεί να αφορά περιπτώσεις όπου μόνο το SPC δίνει σήμα ή περιπτώσεις όπου μόνο το ML χαρακτηρίζει την παρατήρηση ως ανωμαλία. Το μητρώο καταστάσεων παρουσιάζεται παρακάτω:

**Πίνακας 10 Μητρώο Καταστάσεων Συναγερμών (alarms)**

|               | ML anomaly = 1 | ML anomaly = 0 |
|---------------|----------------|----------------|
| SPC alarm = 1 | 1              | 0              |
| SPC alarm = 0 | 0              | 1              |

Στο πλαίσιο αυτό μπορούν να υπολογιστούν δείκτες σύγκρισης όπως

- Precision,
- Recall και
- F1-score,

με την προϋπόθεση ότι τα SPC alarms χρησιμοποιούνται ως στατιστικό σημείο αναφοράς και όχι ως απόλυτη αλήθεια. Η διάκριση αυτή είναι σημαντική, διότι ένα SPC alarm δεν σημαίνει απαραίτητα ότι το προϊόν είναι μη συμμορφούμενο ή ότι υπάρχει πραγματικό ελάττωμα. Αντίθετα, σημαίνει ότι η συγκεκριμένη παρατήρηση ή η μεταβολή μεταξύ διαδοχικών παρατηρήσεων αποκλίνει από τη φυσική μεταβλητότητα που εκτιμούμε για τη διεργασία.

Οι δείκτες συμφωνίας υπολογίζονται ως εξής:

$$Precision = \frac{MLanomalies\text{που συμφωνούν με SPC}}{\text{αριθμός } MLanomalies}$$

$$Recall = \frac{MLanomalies\text{που συμφωνούν με SPC}}{\text{αριθμός SPC alarms}}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Η ερμηνεία των δεικτών αυτών γίνεται με προσοχή, καθώς δεν αξιολογείται η ικανότητα του ML να προβλέψει μία αντικειμενική ετικέτα αστοχίας, αλλά ο βαθμός συμφωνίας του με ένα στατιστικό σημείο αναφοράς (το alarm του SPC).

Η περίπτωση όπου και οι δύο μέθοδοι εντοπίζουν απόκλιση θεωρείται ιδιαίτερα σημαντική, καθώς υποδηλώνει ότι η παρατήρηση παρουσιάζει ταυτόχρονα μονομεταβλητή στατιστική απόκλιση και πολυμεταβλητή ανωμαλία. Αντίθετα, όταν εμφανίζεται μόνο SPC alarm, είναι πιθανό η απόκλιση να αφορά κυρίως τη μεταβλητή που παρακολουθείται από το διάγραμμα ελέγχου, χωρίς το συνολικό ηλεκτρικό προφίλ να διαφοροποιείται επαρκώς ώστε να χαρακτηριστεί ανωμαλία από το ML. Αντίστοιχα, όταν εμφανίζεται μόνο ML anomaly flag, είναι πιθανό η παρατήρηση να μην αποκλίνει έντονα σε μία μεμονωμένη μεταβλητή, αλλά να παρουσιάζει ασυνήθιστο συνδυασμό ηλεκτρικών χαρακτηριστικών.

Η ερμηνεία των ασυμφωνιών μεταξύ SPC και ML αποτελεί βασικό μέρος της συγκριτικής αξιολόγησης. Οι ασυμφωνίες δεν θεωρούνται απαραίτητα σφάλμα μίας από τις δύο μεθόδους, αλλά πιθανή ένδειξη ότι οι μέθοδοι εντοπίζουν διαφορετικές πτυχές της διεργασιακής συμπεριφοράς. Το SPC βασίζεται σε στατιστικά όρια ελέγχου και είναι ιδιαίτερα ερμηνεύσιμο, όμως παρακολουθεί συνήθως μία κρίσιμη μεταβλητή κάθε φορά. Αντίθετα, τα ML μοντέλα μπορούν να αξιοποιήσουν ταυτόχρονα περισσότερες μεταβλητές, αλλά η ερμηνεία των αποτελεσμάτων τους απαιτεί μεγαλύτερη προσοχή.

Παράλληλα, πρέπει να γίνει σαφής διάκριση μεταξύ των SPC alarms, των ML anomaly flags και του τελικού αποτελέσματος Pass / No Pass. Το Pass / No Pass αποτελεί ένδειξη συμμόρφωσης του προϊόντος με τις τεχνικές ή κανονιστικές προδιαγραφές, ενώ τα SPC alarms και τα ML anomaly flags αποτελούν ενδείξεις μεταβολής ή ασυνήθιστης συμπεριφοράς της διεργασίας. Συνεπώς, η παρούσα σύγκριση δεν αντιμετωπίζει το Pass / No Pass ως μοναδικό κριτήριο αξιολόγησης της διεργασιακής σταθερότητας, ούτε θεωρεί ότι κάθε ένδειξη ανωμαλίας πρέπει απαραίτητα να αντιστοιχεί σε μη συμμορφούμενο προϊόν.

Η σύγκριση πραγματοποιείται σε κοινή χρονική βάση, ώστε να διατηρηθεί η μεθοδολογική συνέπεια μεταξύ των δύο προσεγγίσεων. Τα δεδομένα που χρησιμοποιούνται για τον υπολογισμό των ορίων SPC αντιστοιχούν στην περίοδο Phase I, ενώ τα μεταγενέστερα δεδομένα χρησιμοποιούνται για την παρακολούθηση Phase II. Αντίστοιχα, στα μοντέλα Μηχανικής Μάθησης εφαρμόζεται χρονικός διαχωρισμός, ώστε η εκπαίδευση να βασίζεται σε παλαιότερα δεδομένα και η αξιολόγηση σε νεότερες παρατηρήσεις. Με τον τρόπο αυτό περιορίζεται ο κίνδυνος διαρροής πληροφορίας και προσομοιώνεται καλύτερα η πραγματική εφαρμογή ενός συστήματος παρακολούθησης σε παραγωγικό περιβάλλον.

Συνολικά, το πλαίσιο σύγκρισης SPC και ML-SPC βασίζεται στη λογική της συμπληρωματικότητας. Το SPC παρέχει ένα απλό, στατιστικά θεμελιωμένο και ερμηνεύσιμο πλαίσιο παρακολούθησης της διεργασίας. Οι μέθοδοι Μηχανικής Μάθησης, από την άλλη πλευρά, μπορούν να λειτουργήσουν ως πολυμεταβλητή επέκταση της ανάλυσης, εντοπίζοντας συνδυασμούς χαρακτηριστικών που ενδεχομένως δεν είναι εμφανείς σε μονομεταβλητά διαγράμματα ελέγχου. Η αξιολόγηση, επομένως, δεν αποσκοπεί στην απόδειξη ότι η μία μέθοδος αντικαθιστά την άλλη, αλλά στη διερεύνηση του βαθμού στον οποίο η χρήση ML μπορεί να προσθέσει πληροφορία και να ενισχύσει το υφιστάμενο πλαίσιο SPC.

## Κεφάλαιο 4 - Αποτελέσματα

Στο παρόν κεφάλαιο παρουσιάζονται τα αποτελέσματα που προέκυψαν από την εφαρμογή της μεθοδολογίας που περιεγράφηκε στο Κεφάλαιο 3. Η ανάλυση οργανώνεται σε στάδια, ξεκινώντας από την εφαρμογή των παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών (SPC) στις επιλεγμένες ομάδες προϊόντων και συνεχίζοντας με την αξιολόγηση και παρουσίαση των αποτελεσμάτων των μεθόδων Μηχανικής Μάθησης. Βασικός στόχος του κεφαλαίου είναι να αποτυπωθεί η συμπεριφορά της διεργασίας, να εντοπιστούν τυχόν αποκλίσεις ή μη τυχαία πρότυπα και να δημιουργηθεί η βάση για τη συγκριτική αξιολόγηση μεταξύ SPC και ML-SPC.

### 4.1 Αποτελέσματα από την εφαρμογή Στατιστικού Ελέγχου Διεργασιών

Λόγω του μεγάλου αριθμού δεδομένων (~ 30.000 εγγραφές) και της πολυπλοκότητας τους, παρουσιάζουμε την εφαρμογή των μεθόδων Στατιστικού Ελέγχου Διεργασιών σε επιλεγμένες και αντιπροσωπευτικές ομάδες προϊόντων, οι οποίες κρίθηκαν κατάλληλες λόγω του πλήθους διαθέσιμων δεδομένων, της αντιπροσωπευτικότητάς τους ως προς τις βασικές οικογένειες προϊόντων και της διαφορετικής συμπεριφοράς που παρουσιάζουν ως προς τη μεταβλητότητα και τη συμμόρφωση. Η ανάλυση βασίζεται στη μεταβλητή σχετικής απόκλισης ισχύος ( $\Delta P\%$ ), η οποία υπολογίζεται ως ποσοστιαία απόκλιση της μετρούμενης ισχύος από την τιμή στόχο του αντίστοιχου προϊόντος.

$$\Delta P\% = \frac{P_i - P_{nom}}{P_{nom}} * 100$$

Για την αξιολόγηση της συμμόρφωσης ως προς την ισχύ ( $P$ ) χρησιμοποιήθηκαν όρια προδιαγραφών (Specification Limits –  $SL$ ). Τα όρια αυτά τίθενται στο επίπεδο του  $\pm 10\%$  από την ονομαστική  $P_{nom}$  και δεν ταυτίζονται με τα όρια ελέγχου των διαγραμμάτων SPC, τα οποία υπολογίζονται χρησιμοποιώντας τη φυσική μεταβλητότητα της διεργασίας. Στον Πίνακα 11 παρουσιάζονται συνοπτικά οι ομάδες προϊόντων που αξιοποιήθηκαν στην ανάλυση SPC, καθώς και βασικοί δείκτες περιγραφικής στατιστικής, συμμόρφωσης και διεργασιακής ικανότητας. Οι ομάδες επιλέχθηκαν με τέτοιο τρόπο, ώστε να καλύπτουν διαφορετικές οικογένειες προϊόντων και διαφορετικές ονομαστικές ισχύεις, με στόχο να αποτυπωθεί ένα αντιπροσωπευτικό εύρος συμπεριφορών.

**Πίνακας 11 Επιλεγμένες ομάδες προϊόντων για εφαρμογή SPC**

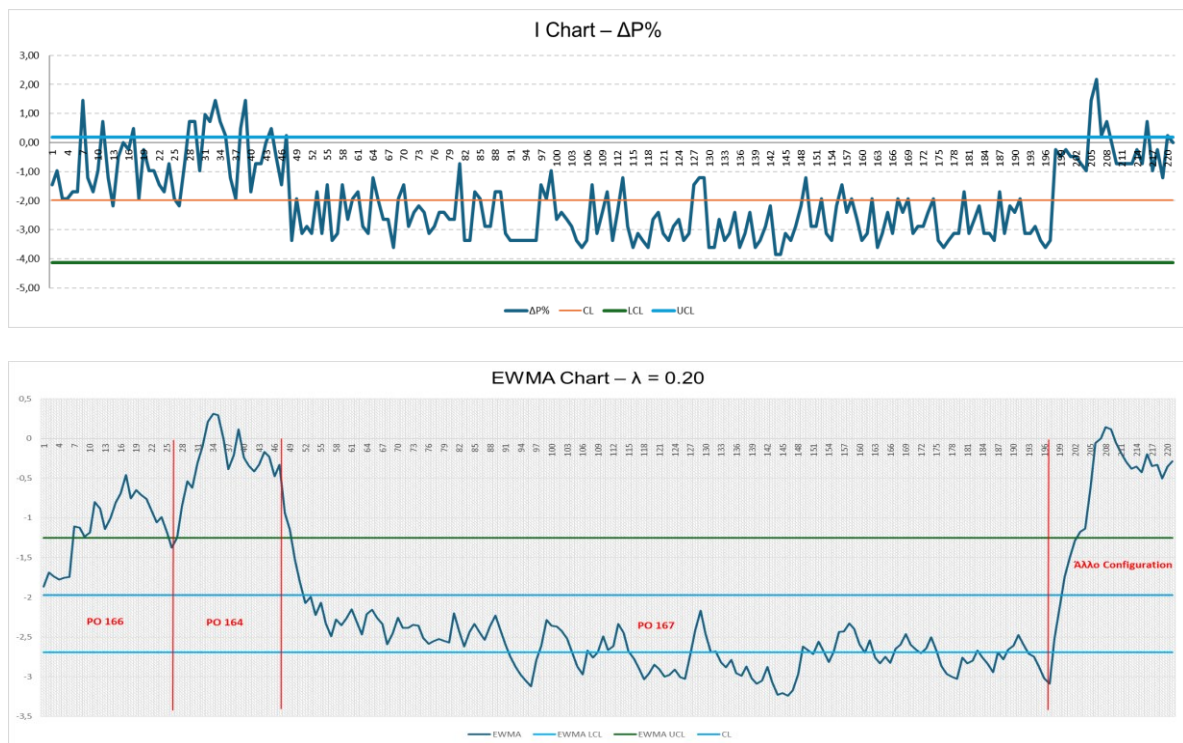
| Ομάδα προϊόντος | Πλήθος εγγραφών | Reject rate (%) | Εκτός $\pm 10\%$ | SL | Μέση τιμή $\Delta P\%$ | Τυπική απόκλιση $\Delta P\%$ | Cp   | Cpk  | I-chart alarms | MR alarms | EWMA alarms ( $\lambda=0.2$ ) |
|-----------------|-----------------|-----------------|------------------|----|------------------------|------------------------------|------|------|----------------|-----------|-------------------------------|
| <b>A30.5</b>    | 6.310           | 0,82            | 109              |    | 0,65                   | 5,17                         | 2,94 | 2,75 | 132            | 136       | 1.478                         |
| <b>A60.5</b>    | 5.072           | 0,02            | 3                |    | 0,03                   | 1,07                         | 5,49 | 5,47 | 240            | 47        | 2.756                         |
| <b>A41.5</b>    | 756             | 0,00            | 3                |    | -0,31                  | 2,64                         | 3,76 | 3,64 | 5              | 10        | 37                            |
| <b>B12.5</b>    | 4.252           | 0,07            | 3                |    | 0,09                   | 1,26                         | 2,78 | 2,75 | 16             | 39        | 248                           |
| <b>B48.5</b>    | 2.585           | 0,00            | 1                |    | 0,26                   | 1,37                         | 3,46 | 3,37 | 48             | 35        | 542                           |
| <b>C220</b>     | 867             | 0,00            | 0                |    | -0,18                  | 1,41                         | 4,09 | 4,01 | 98             | 7         | 335                           |
| <b>C182.5</b>   | 89              | 17,98           | 16               |    | -0,09                  | 0,58                         | 0,91 | 0,90 | 17             | 27        | 73                            |

Από τα αποτελέσματα του Πίνακα 11 προκύπτουν διάφορα ενδιαφέροντα συμπεράσματα. Αρχικά, παρατηρούμε ότι οι περισσότερες ομάδες προϊόντων εμφανίζουν χαμηλά ποσοστά απόρριψης (Reject Rate %) και μικρή μέση απόκλιση από την ονομαστική ισχύ. Παρόλα αυτά, η συμπεριφορά των μοντέλων παρουσιάζει σημαντικές διαφοροποιήσεις. Ορισμένες ομάδες εμφανίζουν υψηλό ποσοστό συμμόρφωσης ως προς τα όρια προδιαγραφών (SL), αλλά παράλληλα παράγουν αυξημένο αριθμό SPC alarms, γεγονός που αποδεικνύει ότι η στατιστική σταθερότητα της διεργασίας και η συμμόρφωση με τις προδιαγραφές δεν ταυτίζονται. Η παρατήρηση αυτή γίνεται πιο εμφανής από τις τιμές των δεικτών Cp και Cpk, τους δείκτες δυνατότητας της διεργασίας, οι οποίοι σε όλες τις ομάδες πλην της τελευταίας είναι σημαντικά μεγαλύτεροι της μονάδας. Η ομάδα C182.5 αποτελεί την πλέον αδύναμη περίπτωση ως προς τη διεργασιακή ικανότητα, καθώς εμφανίζει  $Cp = 0,91$  και  $Cpk = 0,90$ , οι οποίες υποδηλώνουν ότι η φυσική μεταβλητότητα της διεργασίας ( $\pm 3\sigma$ ) είναι μεγαλύτερη από το εύρος των προδιαγραφών. Το αποτέλεσμα αυτό συνάδει με το αυξημένο reject rate % που παρατηρείται για τη συγκεκριμένη ομάδα. Η ομάδα αυτή αποτελεί ειδική περίπτωση, καθώς συνδυάζει μικρότερο πλήθος εγγραφών (89) με αυξημένη συχνότητα μη συμμορφώσεων. Για τον λόγο αυτό, τα αποτελέσματά της πρέπει να ερμηνευθούν με προσοχή, καθώς το μικρό μέγεθος δείγματος περιορίζει την ασφαλή εξαγωγή συμπερασμάτων.

Από τα αποτελέσματα παρατηρείται επίσης ότι ο αριθμός των SPC signals διαφέρει σημαντικά μεταξύ των ομάδων. Η ομάδα A30.5 εμφανίζει 132 I-chart signals και 136 MR-chart signals, γεγονός που καταδεικνύει αυξημένη μεταβλητότητα ή πιθανή ύπαρξη διακριτών περιόδων με διαφορετική συμπεριφορά (π.χ. διαφορετικά POs). Η παρατήρηση αυτή ενισχύεται από τη

συγκριτικά υψηλή συνολική τυπική απόκλιση της  $\Delta P\%$ , η οποία ανέρχεται σε 5,17%. Την ίδια συμπεριφορά βλέπουμε και διαγραμματικά στο I-chart επιλεγμένων εντολών παραγωγής (παρτίδων) της συγκεκριμένης ομάδας.

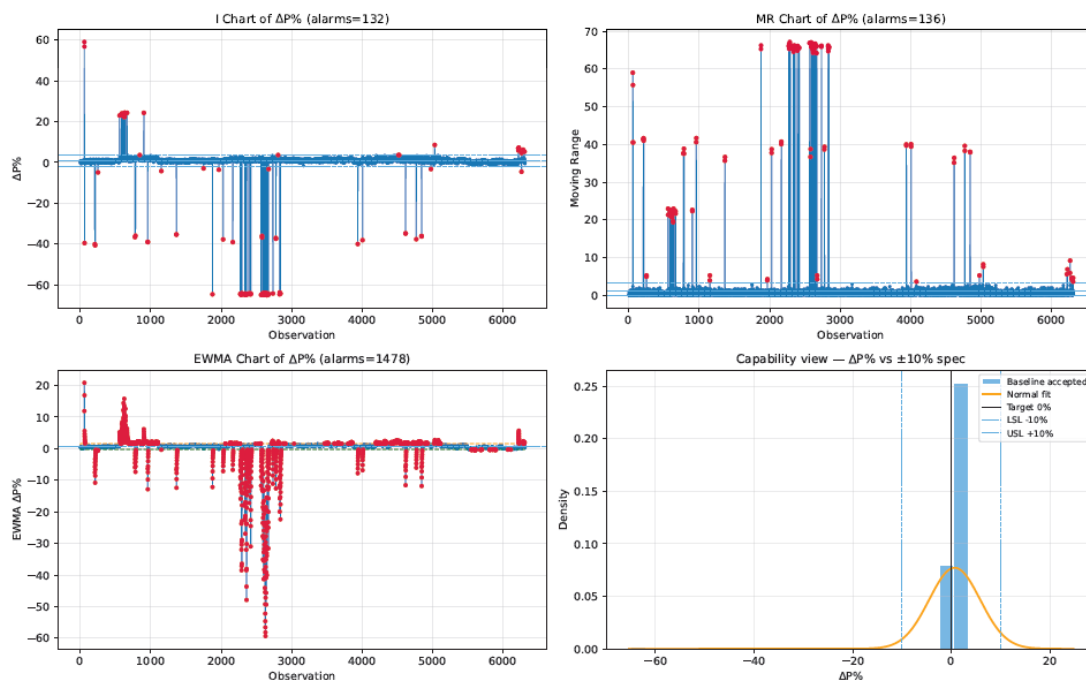
Διάγραμμα 6 Διαγράμματα I-MR επιλεγμένης ομάδας



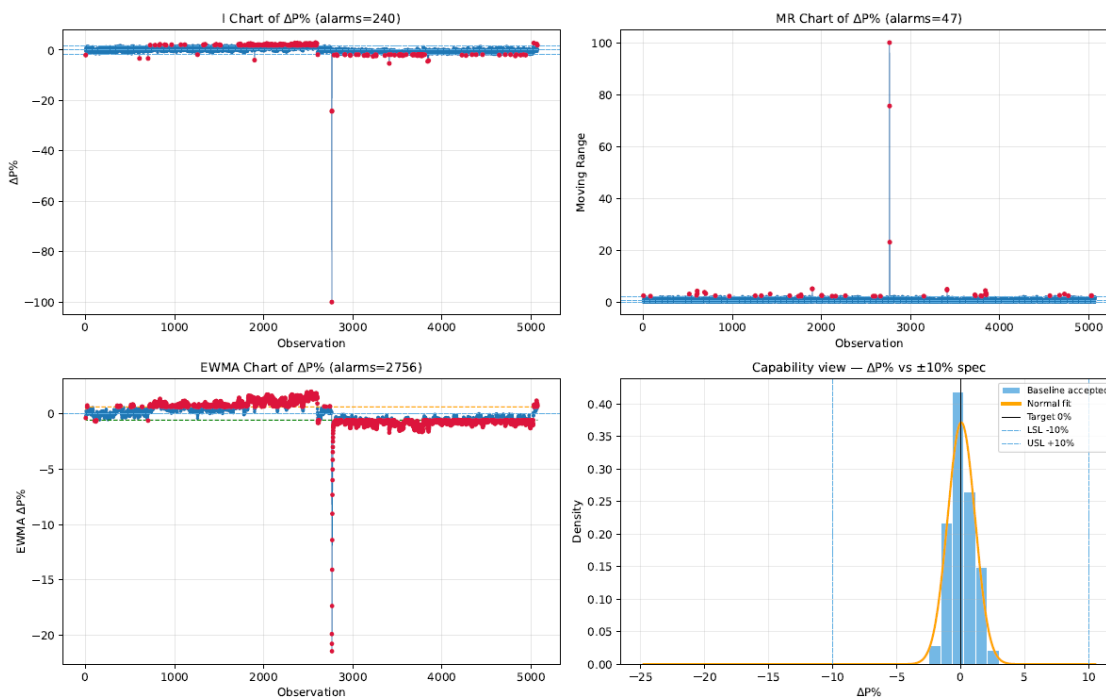
Η διαφοροποίηση αυτή μεταξύ των εντολών παραγωγής (PO) είναι εντόνως ορατή και στο διάγραμμα EWMA. Εδώ υπεισέρχεται και η διαφοροποίηση μεταξύ τεχνικών διαμορφώσεων των ίδιων μοντέλων, την οποία θα εξετάσουμε στην περίπτωση του μοντέλου A30.5.

Αντίθετα, η ομάδα B12.5 εμφανίζει περιορισμένο αριθμό I-chart signals, παρά το μεγάλο πλήθος εγγραφών. Η συμπεριφορά αυτή υποδηλώνει σταθερότερη διεργασία ως προς τη συγκεκριμένη μεταβλητή. Παρόμοια εικόνα εμφανίζεται και στην ομάδα A41.5W, όπου παρατηρούνται μόνο 5 I-chart signals και 10 MR-chart signals. Τα διαγράμματα ελέγχου I-MR εφαρμόστηκαν στις επιλεγμένες ομάδες προϊόντων με στόχο την παρακολούθηση της συμπεριφοράς στην παραγωγική ροή. Τα I-chart alarms αντιστοιχούν σε παρατηρήσεις που υπερβαίνουν τα στατιστικά όρια ελέγχου, ενώ τα MR-chart alarms αντιστοιχούν σε μεγάλες μεταβολές μεταξύ διαδοχικών παρατηρήσεων. Τα SPC διαγράμματα παρουσιάζονται παρακάτω.

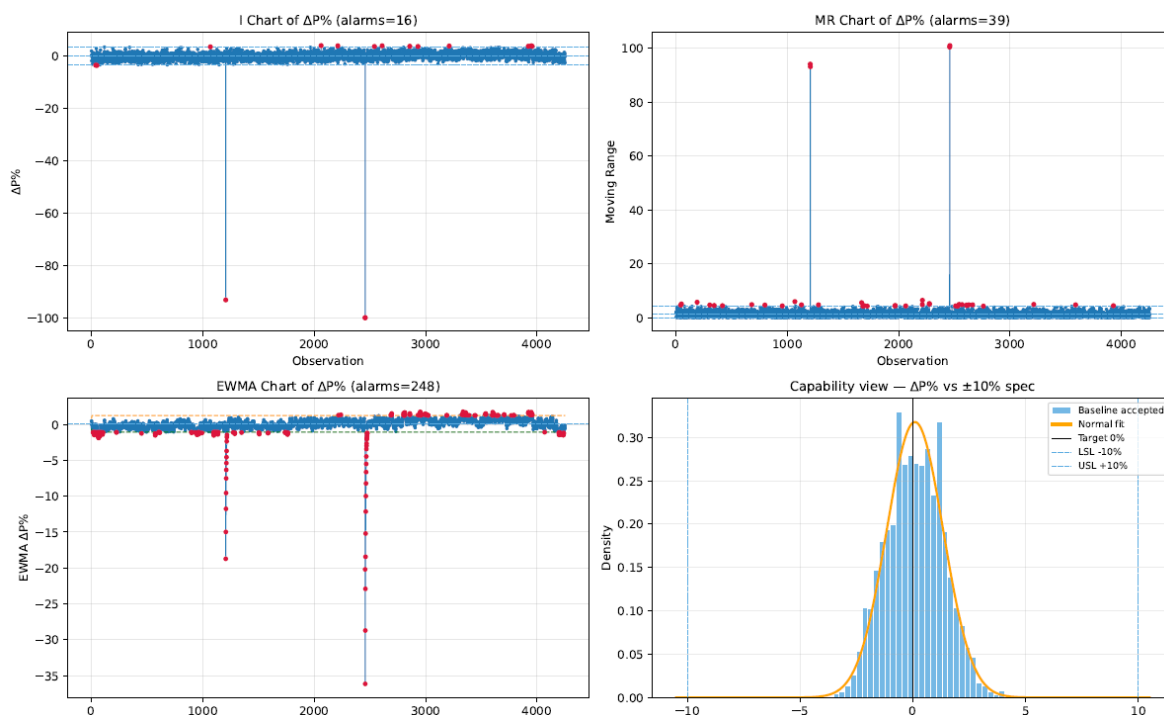
### Ομάδα Α30.5



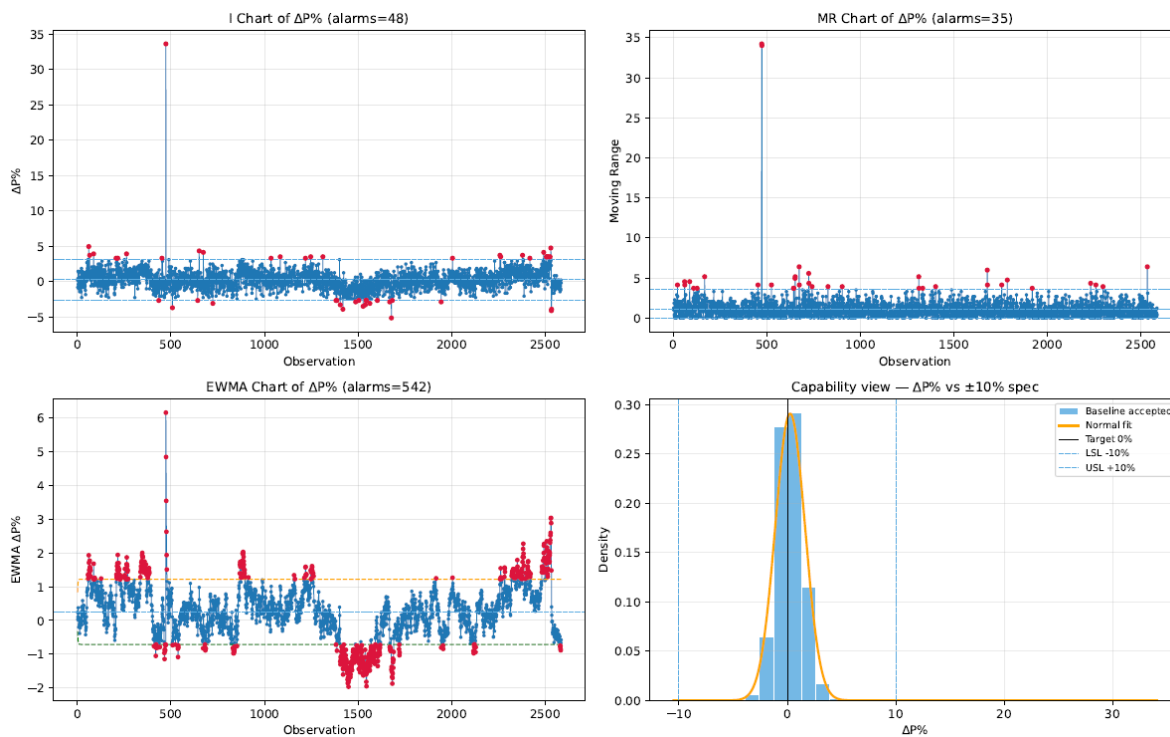
### Ομάδα Α60.5



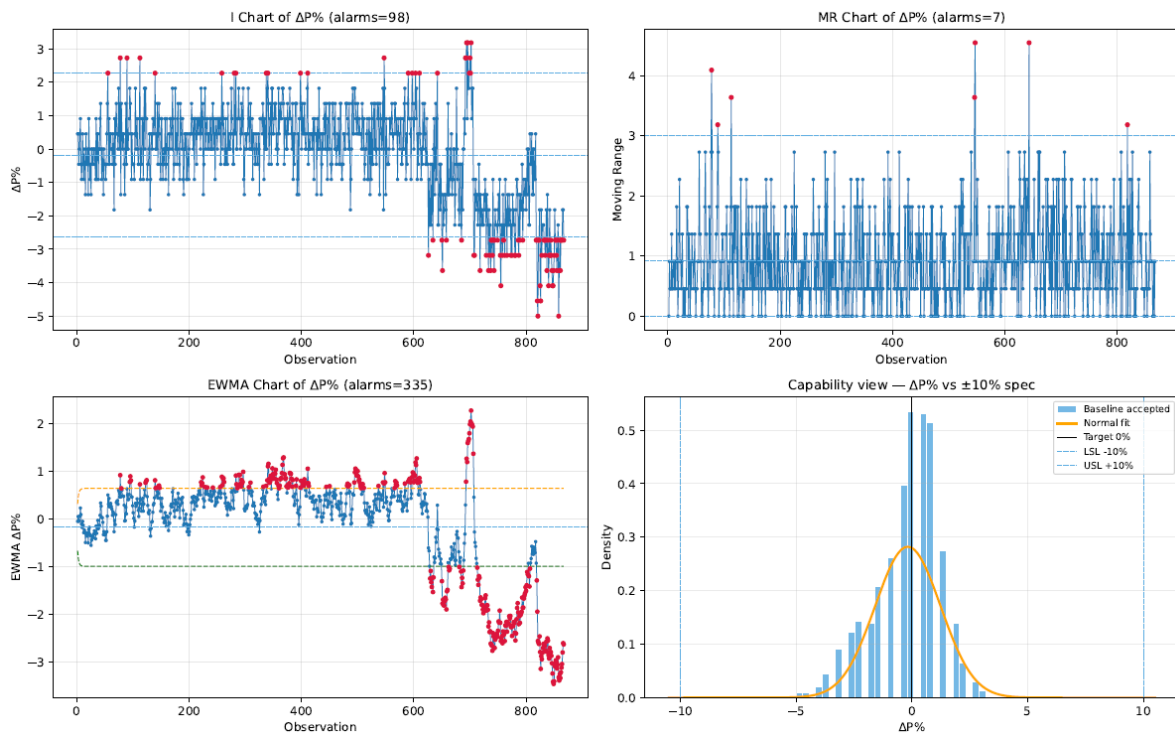
### Ομάδα B12.5



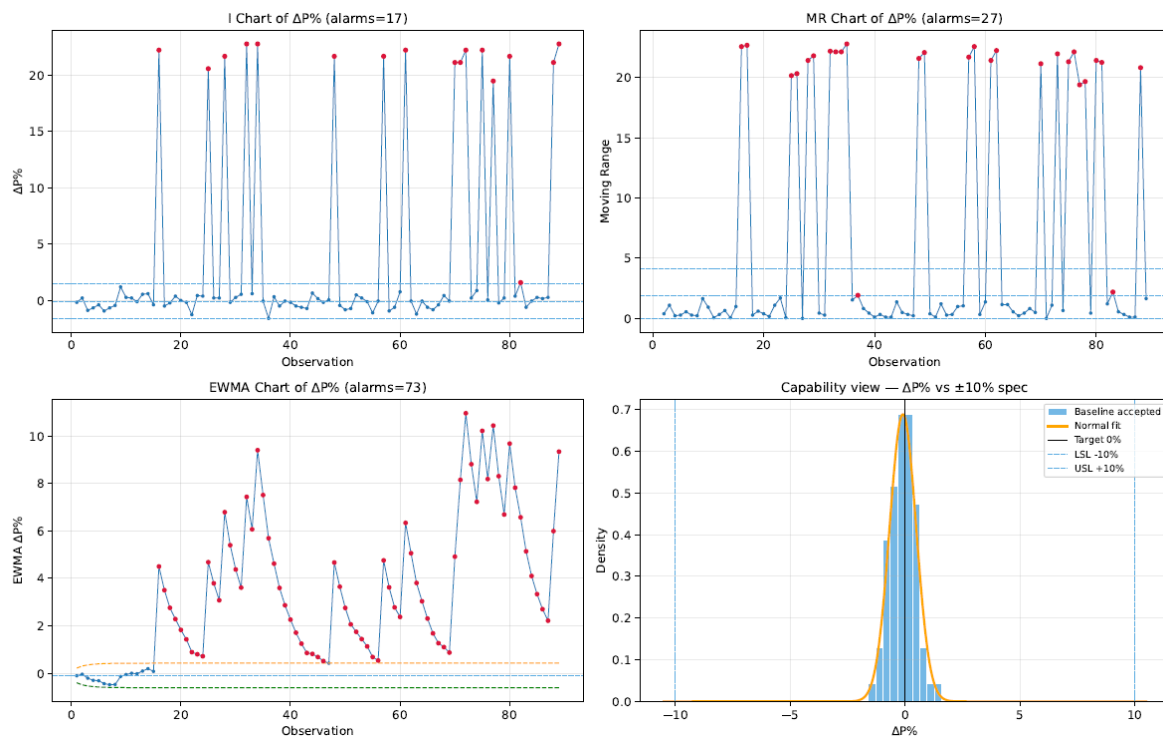
### Ομάδα B48.5



### Ομάδα C220



### Ομάδα C182.5



Τα διαγράμματα EWMA, λόγω της μεγαλύτερης ευαισθησίας τους σε μικρές και σταδιακές μετατοπίσεις, εμφανίζουν σε αρκετές περιπτώσεις σημαντικά μεγαλύτερο αριθμό alarms. Ενδεικτικά, η ομάδα A60.5 παρουσιάζει 2.756 EWMA alarms (>50% των παρατηρήσεων), παρά το χαμηλό reject rate και τις ελάχιστες μετρήσεις εκτός προδιαγραφών. Το αποτέλεσμα αυτό δείχνει ότι η EWMA προσέγγιση μπορεί να εντοπίζει μικρές μεταβολές της διεργασίας, οι οποίες δεν οδηγούν απαραίτητα σε μη συμμόρφωση προϊόντος. Για τον λόγο αυτό, τα EWMA alarms πρέπει να αξιολογούνται ως δείκτες ευαισθησίας της διεργασίας και όχι ως άμεσες ενδείξεις απόρριψης προϊόντος.

Ένα κρίσιμο στοιχείο της ανάλυσης είναι ότι τα SPC signals και το αποτέλεσμα Pass / No Pass δεν έχουν την ίδια έννοια. Όπως εξηγήσαμε και στο Κεφάλαιο 3, ένα προϊόν μπορεί να εμφανίζει SPC signal χωρίς να είναι απορριπτέο σύμφωνα με τις προδιαγραφές. Στην περίπτωση αυτή, το SPC λειτουργεί ως μηχανισμός έγκαιρης προειδοποίησης, καθώς τείνει να εντοπίζει μεταβολές στη διεργασία προτού αυτή οδηγήσει απαραίτητα σε μη συμμόρφωση. Αντίστοιχα, είναι δυνατόν να εμφανιστεί No Pass χωρίς προηγούμενο SPC alarm, ιδιαίτερα όταν η απόκλιση από τα SL οφείλεται σε μεμονωμένο περιστατικό ή σε παράμετρο που δεν αποτυπώνεται πλήρως από το συγκεκριμένο διάγραμμα ελέγχου. Η διάκριση αυτή είναι εμφανής σε ομάδες όπως η A60.5 και η C220, όπου παρατηρούνται πολύ χαμηλά ή μηδενικά reject rates, αλλά ταυτόχρονα εμφανίζονται SPC signals. Αντίθετα, στην ομάδα C182.5 η αυξημένη συχνότητα μη συμμορφώσεων συνοδεύεται και από αυξημένο αριθμό SPC signals, υποδηλώνοντας πιο έντονη σύνδεση μεταξύ διεργασιακής αστάθειας και μη συμμόρφωσης.

Για την αξιολόγηση του SPC, πέρα από τον αριθμό σημάτων που εμφανίζονται στα διαγράμματα ελέγχου, κρίθηκε απαραίτητος ο υπολογισμός κανονικοποιημένων δεικτών, ώστε να είναι δυνατή η σύγκριση μεταξύ ομάδων με διαφορετικό πλήθος εγγραφών. Έτσι, υπολογίστηκαν οι ρυθμοί εμφάνισης σημάτων στα I-chart, MR-chart και EWMA chart, αλλά και οι αντίστοιχοι δείκτες συμμόρφωσης ως προς τα όρια προδιαγραφών και το τελικό αποτέλεσμα ελέγχου.

Οι βασικοί δείκτες υπολογίζονται ως εξής:

$$\text{I-Alarm Rate} = \frac{N_{I\text{-signals}}}{N_{\text{total}}} \times 100$$

$$\text{MR-Alarm Rate} = \frac{N_{\text{MR-signals}}}{N_{\text{total}}} \times 100$$

$$\text{EWMA-Alarm Rate} = \frac{N_{\text{EWMA-signals}}}{N_{\text{total}}} \times 100$$

$$\text{Ποσοστό εκτός προδιαγραφών} = \frac{N_{\text{SpecOut}}}{N_{\text{total}}} \times 100$$

$$\text{Ποσοστό απορρίψεων} = \frac{N_{\text{Rejected}}}{N_{\text{total}}} \times 100$$

Επιπλέον, υπολογίζεται ένας εμπειρικός δείκτης Run Length μεταξύ σημάτων, ο οποίος προσεγγίζει τον μέσο αριθμό παρατηρήσεων ανά εμφάνιση σήματος (ARL):

$$\text{Empirical Run Length} \approx \frac{N_{\text{total}}}{N_{\text{signals}}}$$

Ο δείκτης αυτός δεν μπορεί να ταυτιστεί με το θεωρητικό ARL, αφού δεν διαχωρίζει τις καταστάσεις εντός και εκτός ελέγχου. Ωστόσο, παρέχει μια πιο πρακτική εικόνα για τη συχνότητα εμφάνισης σημάτων σε κάθε ομάδα προϊόντων. Αν θέλαμε να υπολογίσουμε το πραγματικό ARL, θα έπρεπε να παράγουμε «τεχνητά» σήματα – μετρήσεις εκτός ελέγχου (π.χ. να απέχουν >3σ από το μέσο ή με βάση τους κανόνες WE) και να παρατηρήσουμε τότε τα διαγράμματα SPC θα έδιναν alarm signal.

Πίνακας 12 Δείκτες αξιολόγησης SPC

| ΟΜΑΔΑ<br>ΠΡΟΪΟΝΤΟΣ | I-SIGNAL<br>RATE (%) | MR-SIGNAL<br>RATE (%) | EWMA SIGNAL<br>RATE (%) | SPEC-OUT<br>RATE (%) | REJECT<br>RATE<br>(%) | EMPIRICAL RUN<br>LENGTH |
|--------------------|----------------------|-----------------------|-------------------------|----------------------|-----------------------|-------------------------|
| <b>A30.5</b>       | 2,09                 | 2,16                  | 23,42                   | 1,73                 | 0,82                  | 47,80                   |
| <b>A60.5</b>       | 4,73                 | 0,93                  | 54,34                   | 0,06                 | 0,02                  | 21,13                   |
| <b>A41.5</b>       | 0,66                 | 1,32                  | 4,89                    | 0,40                 | 0,00                  | 151,20                  |
| <b>B12.5</b>       | 0,38                 | 0,92                  | 5,83                    | 0,07                 | 0,07                  | 265,75                  |
| <b>B48.5</b>       | 1,86                 | 1,35                  | 20,97                   | 0,04                 | 0,00                  | 53,85                   |
| <b>C220</b>        | 11,30                | 0,81                  | 38,64                   | 0,00                 | 0,00                  | 8,85                    |
| <b>C182.5</b>      | 19,10                | 30,34                 | 82,02                   | 17,98                | 17,98                 | 5,24                    |

Από τον υπολογισμό των δεικτών προκύπτει ότι η αξιολόγηση των αποτελεσμάτων διαφοροποιείται σημαντικά όταν κάνουμε σύγκριση λαμβάνοντας υπόψη και τα μεγέθη των

δειγμάτων. Για παράδειγμα, η ομάδα A30.5 εμφανίζει υψηλό απόλυτο αριθμό I-chart alarms, όμως ο αντίστοιχος I-signal rate ανέρχεται σε μόλις 2,09%. Τονίζουμε επίσης, ότι εδώ η σύγκριση μεταξύ signal rates, spec-out rates και reject rates χρησιμοποιείται για να διερευνηθεί η σχέση μεταξύ διεργασιακής μεταβλητότητας και τελικής συμμόρφωσης και όχι για να θεωρηθεί το SPC ως άμεση διαδικασία ταξινόμησης αποδοχής ή απόρριψης.

Συμπερασματικά, η εφαρμογή των παραδοσιακών μεθόδων SPC σε διάφορες ομάδες προϊόντων ανέδειξε ότι η συμμόρφωση με τις προδιαγραφές και η σταθερότητα της διεργασίας αποτελούν δύο διαφορετικές αλλά αλληλοσυνδεόμενες διαστάσεις της ποιότητας. Οι περισσότερες ομάδες προϊόντων που επιλέχθηκαν εμφανίζουν χαμηλά ποσοστά απόρριψης και ικανοποιητική ικανότητα διεργασίας ως προς τη μεταβλητή ΔΡ%. Το γεγονός αυτό σε καμία περίπτωση δεν σημαίνει αυτόματα ότι η διεργασία βρίσκεται εντός στατιστικού ελέγχου. Αντιθέτως, η ύπαρξη I-MR και EWMA alarms καταδεικνύει ότι η διεργασία μπορεί να εμφανίζει μεταβολές ή μη τυχαία πρότυπα ακόμη και αν τα προϊόντα παραμένουν εντός προδιαγραφών.

Τα αποτελέσματα της εφαρμογής SPC δείχνουν ότι ο ΣΕΔ μπορεί να προσφέρει χρήσιμη πληροφορία για τη συμπεριφορά διάφορων ομάδων προϊόντων και να υποστηρίξει την περαιτέρω στοχευμένη διερεύνηση αποκλίσεων. Τα αποτελέσματα αυτά αποτελούν το κύριο σημείο αναφοράς για τη μετέπειτα σύγκριση με τις μεθόδους Μηχανικής Μάθησης. Ενώ ο Στατιστικός Έλεγχος Διεργασιών παρέχει σαφές πλαίσιο παρακολούθησης των διεργασιών, η ύπαρξη πολυμεταβλητότητας, διαφορετικών τεχνικών διαμορφώσεων και πιθανών μη γραμμικών σχέσεων δημιουργεί ενδιαφέρουσες προεκτάσεις σχετικά με την αξιολόγηση ML-SPC προσεγγίσεων στο επόμενο στάδιο της ανάλυσης.

## 4.2 Αποτελέσματα ML

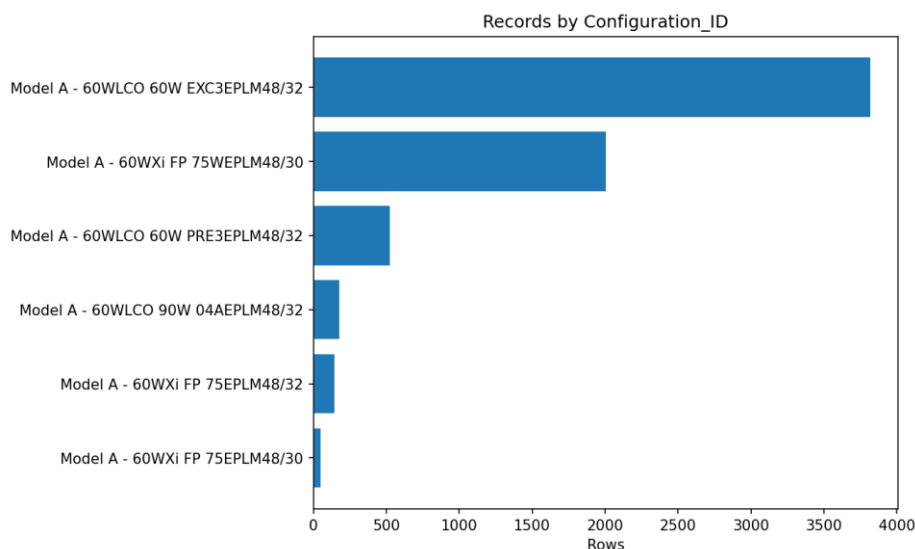
Η εφαρμογή των μεθόδων Μηχανικής Μάθησης πραγματοποιήθηκε στο επιλεγμένο υποσύνολο προϊόντων της οικογένειας A σε φωτιστικά ισχύος 60 W. Η επιλογή της συγκεκριμένης ομάδας έγινε λόγω του μεγάλου πλήθους διαθέσιμων εγγραφών και της ύπαρξης διαφορετικών τεχνικών διαμορφώσεων (configurations driver/πλακέτας LED), οι οποίες επιτρέπουν τη διερεύνηση του κατά πόσο το ηλεκτρικό αποτύπωμα του τελικού ελέγχου περιέχει πληροφορία ικανή να διαχωρίσει διαφορετικά configurations.

Το τελικό υπολογιστικό σύνολο δεδομένων περιλάμβανε συνολικά 6.692 εγγραφές με πλήρη διαθέσιμα ηλεκτρικά χαρακτηριστικά για την ανάλυση ML. Τα χαρακτηριστικά που χρησιμοποιήθηκαν ήταν η σχετική απόκλιση ισχύος  $\Delta P\%$ , ο συντελεστής ισχύος, η ολική αρμονική παραμόρφωση, η τάση ελέγχου και ο δείκτης ρεύματος ανά Watt που υπολογίστηκε.

Η υπολογιστική υλοποίηση πραγματοποιήθηκε σε γλώσσα προγραμματισμού Python. Η διαχείριση και επεξεργασία των δεδομένων έγινε με τη βιβλιοθήκη pandas, οι αριθμητικοί υπολογισμοί με τη numpy, ενώ οι αλγόριθμοι PCA, K-means, GMM, Isolation Forest και Επιβλεπόμενης Μάθησης υλοποιήθηκαν με τη βιβλιοθήκη scikit-learn, η οποία τους περιέχει. Για την αναπαραγωγικότητα των αποτελεσμάτων χρησιμοποιήθηκε σταθερό random state, ενώ όλα τα ενδιάμεσα και τελικά αποτελέσματα εξήχθησαν σε αρχεία MS Excel και γραφήματα για ευκολία στη διαχείριση και την παρουσίαση.

#### 4.2.1 Περιγραφική εικόνα του ML dataset

Το σύνολο δεδομένων που χρησιμοποιήσαμε για την εφαρμογή της Μηχανικής Μάθησης περιλαμβάνει έξι διαφορετικές τεχνικές διαμορφώσεις και πέντε διαφορετικούς τύπους driver. Η κατανομή των κλάσεων δεν είναι ισορροπημένη, καθώς δύο configurations συγκεντρώνουν η μερίδα του λέοντος των παρατηρήσεων. Συγκεκριμένα, η διαμόρφωση με driver LCO 60W EXC3 και πλακέτα EPLM48/32 αντιπροσωπεύει περίπου το 57% των εγγραφών, ενώ το configuration με driver Xi FP 75W και πλακέτα EPLM48/30 περίπου το 30%.



Διάγραμμα 7 Διάγραμμα Pareto αριθμού παρατηρήσεων ανά Configuration

**Πίνακας 13 Κατανομή παρατηρήσεων ανά διαμόρφωση**

| Διαμόρφωση                           | Πλήθος εγγραφών | Ποσοστό |
|--------------------------------------|-----------------|---------|
| Model A - 60W LCO 60W EXC3 EPLM48/32 | 3.818           | 57,05%  |
| Model A - 60W Xi FP 75W EPLM48/30    | 2.002           | 29,92%  |
| Model A - 60W LCO 60W PRE3 EPLM48/32 | 503             | 7,52%   |
| Model A - 60W LCO 90W 04A EPLM48/32  | 177             | 2,64%   |
| Model A - 60W Xi FP 75 EPLM48/32     | 142             | 2,12%   |
| Model A - 60W Xi FP 75 EPLM48/30     | 50              | 0,75%   |

Η ανισορροπία αυτή λήφθηκε υπόψη κατά την αξιολόγηση των επιβλεπόμενων μοντέλων και χρησιμοποιήσαμε δείκτες όπως το balanced accuracy και το macro F1-score, οι οποίοι είναι καταλληλότεροι από την απλή ακρίβεια όταν οι κλάσεις δεν έχουν παρόμοιο πλήθος παρατηρήσεων, για να εξαλείψουμε κατά το δυνατόν τις επιδράσεις του μεγέθους των κυρίαρχων κλάσεων.

#### 4.2.2 Ανάλυση Κύριων Συνιστωσών - PCA

Η Ανάλυση Κύριων Συνιστωσών εφαρμόστηκε στα κανονικοποιημένα χαρακτηριστικά με κύριο στόχο τη διερεύνηση της πολυμεταβλητής δομής των δεδομένων. Τα αποτελέσματα αποκάλυψαν ότι οι τρεις πρώτες κύριες συνιστώσες κυριαρχούν και εξηγούν συνολικά το 86,10% της μεταβλητότητας των μετρήσεων. Τα ποσοστά διακύμανσης που εξηγεί κάθε συνιστώσα παρουσιάζονται στον Πίνακα 14.

**Πίνακας 14 Ανάλυση διακύμανσης PCA**

| Κύρια συνιστώσα | Ποσοστό διακύμανσης | Αθροιστικό ποσοστό |
|-----------------|---------------------|--------------------|
| PC1             | 36,67%              | 36,67%             |
| PC2             | 27,92%              | 64,60%             |
| PC3             | 21,50%              | 86,10%             |
| PC4             | 13,78%              | 99,88%             |
| PC5             | 0,12%               | 100,00%            |

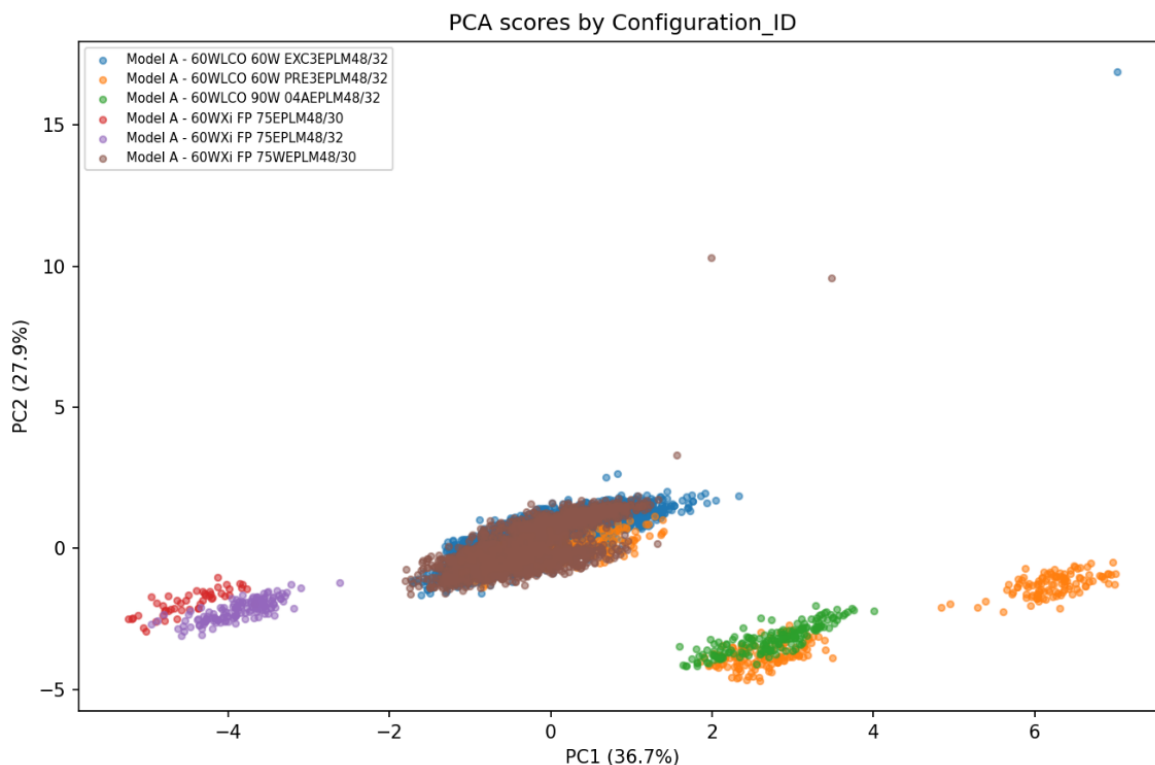
Η κύρια συνιστώσα PC1 επηρεάζεται προπάντων από τον δείκτη ρεύματος ανά Watt και την τάση ελέγχου, ενώ η PC2 συνδέεται περισσότερο με τον συντελεστή ισχύος, την τάση και τη σχετική απόκλιση ισχύος. Η PC3 επηρεάζεται έντονα από την ολική αρμονική παραμόρφωση και τη σχετική απόκλιση ισχύος. Τα αποτελέσματα αυτά δείχνουν ότι η διαφοροποίηση μεταξύ

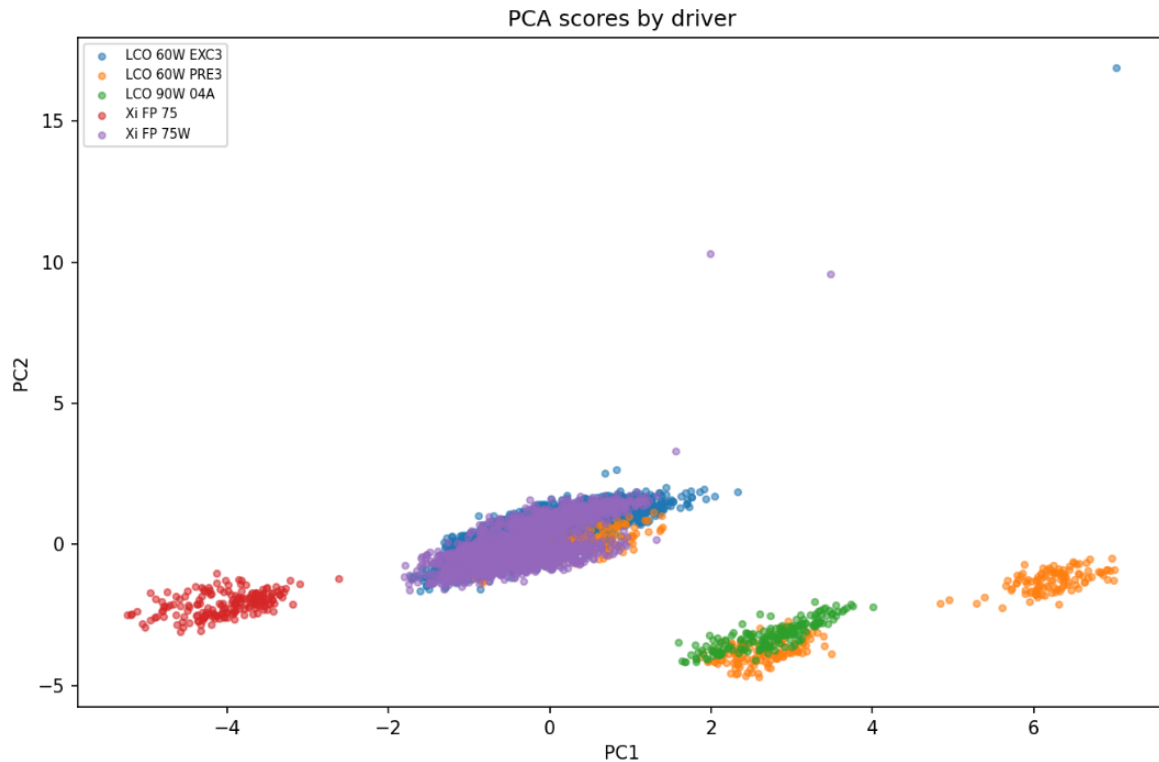
παρατηρήσεων δεν εξαρτάται αποκλειστικά από μία ηλεκτρική μεταβλητή, αλλά από τον συνδυασμό περισσότερων χαρακτηριστικών.

Πίνακας 15 Συντελεστές κύριων συνιστωσών PCA

| Χαρακτηριστικό | PC1    | PC2    | PC3    | PC4    | PC5    |
|----------------|--------|--------|--------|--------|--------|
| ΔΡ%            | -0,147 | -0,485 | 0,558  | 0,657  | -0,004 |
| PF             | -0,375 | 0,685  | 0,184  | 0,268  | 0,534  |
| THDu %         | 0,213  | 0,121  | 0,802  | -0,544 | -0,009 |
| Voltage        | -0,524 | -0,509 | -0,070 | -0,430 | 0,526  |
| Ρεύμα ανά W    | 0,720  | -0,149 | -0,079 | 0,122  | 0,662  |

Αυτή η συμπεριφορά είναι εμφανής και στα διαγράμματα PCA.

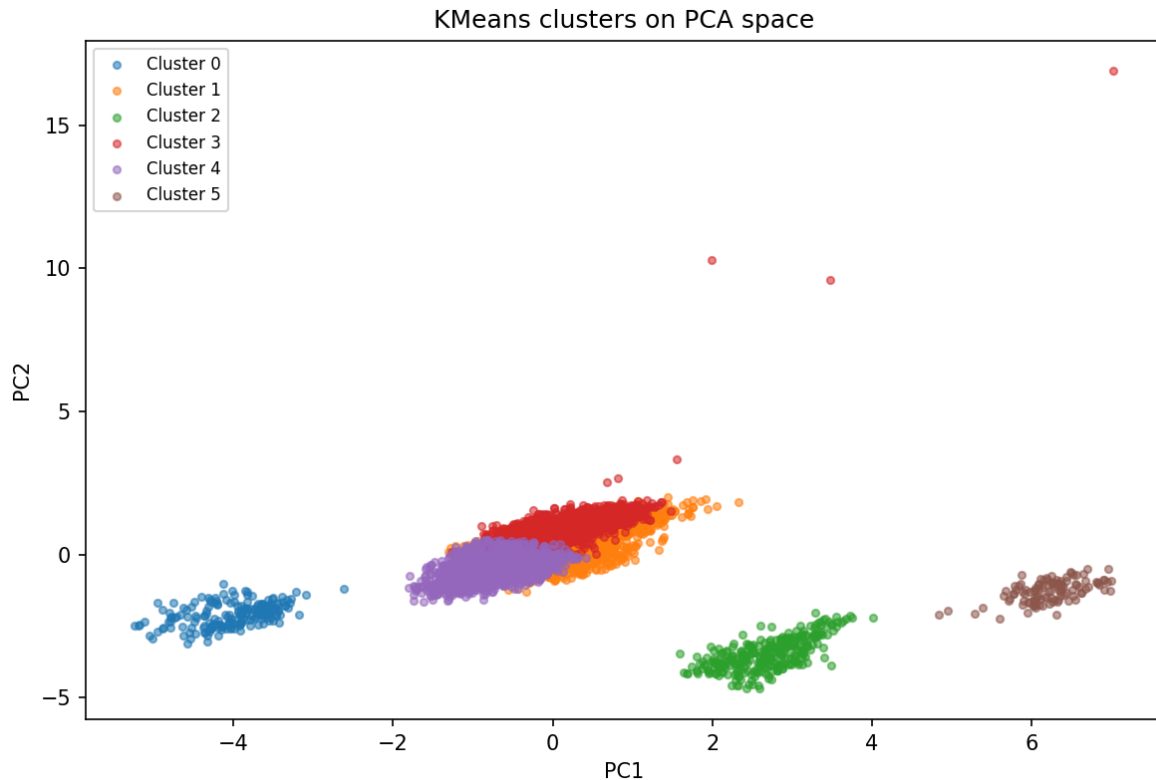




Τα αποτελέσματα της PCA δείχνουν να υποστηρίζουν την αρχική υπόθεση ότι το τελικό ηλεκτρικό αποτύπωμα περιέχει πολυμεταβλητή πληροφορία, η οποία μπορεί να αξιοποιηθεί για τη διερεύνηση τεχνικών διαφοροποιήσεων μεταξύ configurations.

#### 4.2.3 Αποτελέσματα Ταξινόμησης (Clustering)

Για τη διερεύνηση ύπαρξης φυσικών ομάδων στα ηλεκτρικά δεδομένα εφαρμόστηκαν οι αλγόριθμοι K-means και Gaussian Mixture Model. Στον K-means χρησιμοποιήθηκε αριθμός κλάσεων ίσος με το πλήθος των configurations  $k=6$ , ενώ στο GMM εξετάστηκαν διαφορετικοί αριθμοί components και επιλέχθηκε λύση με οκτώ components, βάσει του χαμηλότερου BIC.



Τα αποτελέσματα δείχνουν ότι κάποιες από τις τεχνικές διαμορφώσεις αναγνωρίζονται καθαρά από τα ηλεκτρικά χαρακτηριστικά τους. Για παράδειγμα, οι διαμορφώσεις με driver Xi FP 75 (μπλε) και οι διαμορφώσεις με driver LCO 90W 04A (πράσινο) συγκεντρώνονται σε διακριτά clusters. Αντίθετα, οι δύο πολυπληθέστερες διαμορφώσεις, δηλαδή LCO 60W EXC3 και Xi FP 75W, κατανέμονται σε περισσότερα από ένα clusters, γεγονός που υποδηλώνει εσωτερική διαφοροποίηση μεταξύ των ομάδων ή πιθανή επίδραση άλλων παραγόντων στην παραγωγική διαδικασία πέρα από το ίδιο το configuration. Η εικόνα αυτή δείχνει ότι τα ηλεκτρικά χαρακτηριστικά διαθέτουν ικανότητα διαχωρισμού ορισμένων τεχνικών διαμορφώσεων, αλλά όχι πλήρη και απόλυτο διαχωρισμό όλων των configurations. Το clustering επομένως λειτουργεί περισσότερο ως διερευνητικό εργαλείο αναγνώρισης προτύπων παρά ως πλήρης μηχανισμός ταξινόμησης.

#### 4.2.4 Επιβλεπόμενη ταξινόμηση Driver Type και Configuration ID

Στη συνέχεια εφαρμόζουμε μοντέλα ταξινόμησης Επιβλεπόμενης Μάθησης με στόχο την προσέγγιση του τύπου driver και, δευτερευόντως, του Configuration ID από το ηλεκτρικό αποτύπωμα του τελικού ηλεκτρικού ελέγχου.

Τα μοντέλα που αξιολογήθηκαν ήταν:

1. Λογιστική Παλινδρόμηση - Logistic Regression
2. Τυχαία Δάση - Random Forest
3. Εξαιρετικά Τυχαιοποιημένα Δέντρα - Extra Trees

με χρήση ισορροπημένων βαρών λόγω της ανισορροπίας των κλάσεων.

Η αξιολόγηση πραγματοποιήθηκε με δύο σενάρια. Στο 1<sup>ο</sup> σενάριο χρησιμοποιήσαμε stratified random split 70/30, έτσι ώστε τόσο το σετ εκπαίδευσης όσο και το σετ ελέγχου να περιλαμβάνουν παρατηρήσεις από όλες τις κλάσεις. Στο 2<sup>ο</sup> σενάριο χρησιμοποιήσαμε PO-based holdout, δηλαδή συγκεκριμένες Εντολές Παραγωγής (PO's) κρατήθηκαν εκτός σετ εκπαίδευσης και χρησιμοποιήθηκαν αποκλειστικά για αξιολόγηση για να διαπιστώσουμε αν το μοντέλο μπορεί να προβλέπει και δεδομένα πέρα από αυτά που έχει εκπαιδευτεί.

#### Stratified Random Split:

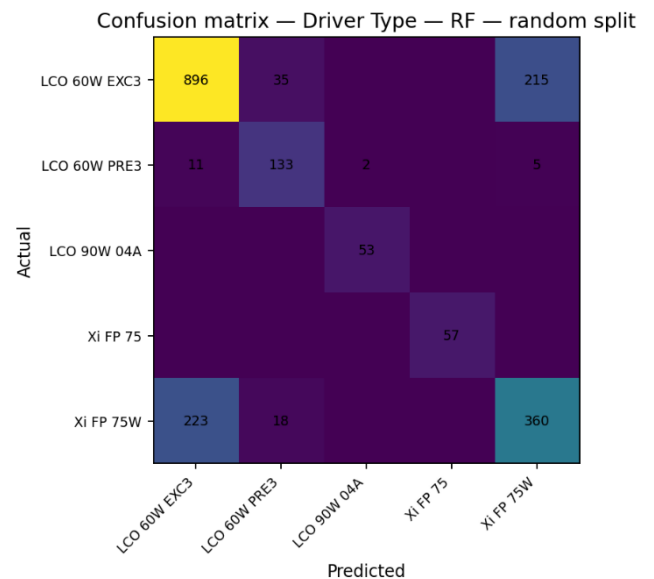
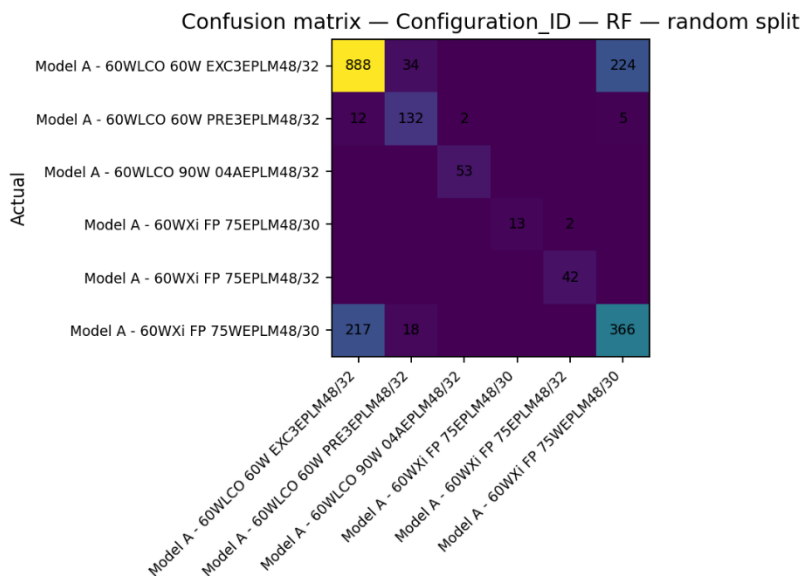
| Στόχος           | Διαχωρισμός             | Καλύτερο μοντέλο | Accuracy | Balanced accuracy | Macro F1 |
|------------------|-------------------------|------------------|----------|-------------------|----------|
| Τύπος Driver     | Stratified random 70/30 | Random Forest    | 74,65%   | 85,23%            | 83,36%   |
| Configuration ID | Stratified random 70/30 | Random Forest    | 74,40%   | 85,41%            | 84,53%   |

Τα αποτελέσματα αυτά δείχνουν ότι τα ηλεκτρικά χαρακτηριστικά όντως περιέχουν πληροφορία που μπορεί να χρησιμοποιηθεί για την αναγνώριση τεχνικών διαμορφώσεων. Η υψηλή balanced accuracy και το υψηλό macro F1-score καταδεικνύουν ότι η απόδοση δεν περιορίζεται αποκλειστικά στις πολυπληθείς κλάσεις, αλλά παραμένει σε ικανοποιητικό επίπεδο και για τις μικρότερες κατηγορίες.

#### PO – Based Holdout:

| Στόχος           | Split            | Καλύτερο / ενδεικτικό μοντέλο     | Accuracy | Balanced accuracy | Macro F1 |
|------------------|------------------|-----------------------------------|----------|-------------------|----------|
| Driver Type      | PO-based holdout | Logistic Regression / Tree models | ~27,3%   | ~49,9%            | ~29,7%   |
| Configuration ID | PO-based holdout | Logistic Regression               | 43,84%   | 26,88%            | 23,31%   |

Εδώ έχουμε πολύ διαφορετική εικόνα.



Διάγραμμα 8 Διαγράμματα Σύγκρισης Τεχνικών διαμορφώσεων

Η απόδοση των προβλέψεων στο PO-based holdout υποδηλώνει ότι η αναγνώριση των configurations δεν μπορεί να γενικευτεί πλήρως σε νέα Production Orders. Ουσιαστικά, αυτό σημαίνει ότι η ηλεκτρική υπογραφή υπάρχει, αλλά επηρεάζεται πιθανότατα από PO effects, διαφοροποιήσεις παραγωγικών συνθηκών ή από τη στενή σύνδεση ορισμένων διαμορφώσεων με συγκεκριμένα Production Orders. Συνεπώς, τα αποτελέσματα του random split πρέπει να ερμηνευθούν με προσοχή και ως ένδειξη ύπαρξης της διαχωριστικής πληροφορίας στα δεδομένα, ενώ τα αποτελέσματα του PO-based holdout ως αυστηρότερη αξιολόγηση.

#### 4.2.5 Σημαντικότητα χαρακτηριστικών

Η ανάλυση σημαντικότητας των ηλεκτρικών χαρακτηριστικών ως αποτέλεσμα από τα μοντέλα δάσους έδειξε ότι ο δείκτης ρεύματος ανά Watt, ο συντελεστής ισχύος και η τάση ελέγχου αποτελούν πιο σημαντικούς παράγοντες για την ταξινόμηση των τεχνικών διαμορφώσεων.

Για την ταξινόμηση του Driver Type με Random Forest σε stratified random split, η υψηλότερη σημαντικότητα εμφανίστηκε στον δείκτη Current per W, ακολουθούμενη από τον Power Factor και την τάση (V).

Πίνακας 16 Σημαντικότητα παραγόντων ως προς την ταξινόμηση των προϊόντων

| Χαρακτηριστικό | Σημαντικότητα Τυχαίων Δασών για τύπο Driver |
|----------------|---------------------------------------------|
|----------------|---------------------------------------------|

|               |        |
|---------------|--------|
| Current per W | 31,77% |
| Power Factor  | 22,31% |
| Voltage       | 21,49% |
| THDU          | 14,18% |
| Delta P%      | 10,25% |

Η εικόνα αυτή είναι απολύτως λογική, αφού διαφορετικοί τύποι driver επηρεάζουν ουσιαστικά τον συντελεστή ισχύος, το ρεύμα εισόδου και την αρμονική συμπεριφορά των προϊόντων. Η σχετική απόκλιση ισχύος  $\Delta P\%$  δεν φαίνεται να αποτελεί το κυρίαρχο χαρακτηριστικό για την αναγνώριση των τεχνικών διαμορφώσεων.

Τα anomaly flags που προέκυψαν από την εφαρμογή του Isolation Forest αξιοποιούνται στο επόμενο υποκεφάλαιο για τη συγκριτική αξιολόγηση με τα alarms των διαγραμμάτων SPC. Με τον τρόπο αυτό, η ανάλυση μεταβαίνει από την ανεξάρτητη αξιολόγηση των ML μοντέλων στη διερεύνηση της συμπληρωματικότητάς τους προς τις παραδοσιακές μεθόδους Στατιστικού Ελέγχου Διεργασιών.

### 4.3 Συγκριτική αξιολόγηση μεταξύ SPC και ML

Μετά την ανεξάρτητη παρουσίαση των αποτελεσμάτων από την εφαρμογή του Στατιστικού Ελέγχου Διεργασιών και των μεθόδων Μηχανικής Μάθησης, στο παρόν υποκεφάλαιο πραγματοποιείται η συγκριτική αξιολόγηση των δύο προσεγγίσεων. Η σύγκριση δεν αντιμετωπίζει το SPC και το ML ως πλήρως ισοδύναμα εργαλεία, καθώς οι δύο μέθοδοι δεν εστιάζουν απαραίτητα στην ίδια διάσταση της ποιότητας. Το SPC αξιολογεί κυρίως τη στατιστική συμπεριφορά μίας διεργασίας ως προς τη φυσική της μεταβλητότητα, ενώ τα ML μοντέλα αξιοποιούν ταυτόχρονα περισσότερα χαρακτηριστικά, με στόχο την αναγνώριση ηλεκτρικών προτύπων ή πολυμεταβλητών ανωμαλιών.

Για τη 2<sup>η</sup> περίπτωση εφαρμογής των μοντέλων ML που αποτελεί και την ουσιαστική σύγκριση και αξιολόγηση μεταξύ SPC και μεθόδων ML, εφαρμόσαμε Isolation Forest για την ανίχνευση πολυμεταβλητών ανωμαλιών τόσο σε συνολικό επίπεδο όσο και ανά configuration. Η ανάλυση πραγματοποιήθηκε με διαφορετικές τιμές contamination (0.01, 0.02,...0.05), ενώ η βασική σύγκριση με τα SPC alarms έγινε μέσω των anomaly flags.

Η σύγκριση με τα SPC I-MR alarms έδειξε περιορισμένη συμφωνία μεταξύ των δύο προσεγγίσεων. Για τον Isolation Forest με global contamination 5%, εντοπίστηκαν 335 anomalies, από τις οποίες μόνο 16 συνέπιπταν με τα alarms του SPC I-MR. Η εφαρμογή

Isolation Forest ανά configuration βελτίωσε σημαντικά τη συμφωνία με τα SPC I-MR alarms. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 17.

**Πίνακας 17 Δείκτες συμφωνίας Isolation Forest με SPC Alarms**

| Reference label       | ML method          | TP | FP  | FN  | Precision | Recall | F1    |
|-----------------------|--------------------|----|-----|-----|-----------|--------|-------|
| <b>SPC I-MR alarm</b> | IF Global 5%       | 16 | 319 | 271 | 4,78%     | 5,57%  | 5,14% |
| <b>SPC I-MR alarm</b> | IF By Config 5%    | 31 | 304 | 256 | 9,25%     | 10,80% | 9,97% |
| <b>SPC I-MR alarm</b> | GMM Low Confidence | 8  | 262 | 279 | 2,96%     | 2,79%  | 2,87% |

Παρατηρούμε ότι ακόμη και με IF ανά configuration, έχουμε πολύ χαμηλή συμφωνία και επικάλυψη μεταξύ των μοντέλων. Η χαμηλή συμφωνία δεν ερμηνεύεται απαραίτητα ως αποτυχία του ML ή του SPC. Αντίθετα, καταδεικνύει ότι το SPC και το ML πιθανόν εντοπίζουν διαφορετικού τύπου αποκλίσεις. Το SPC I-MR βασίζεται κυρίως στη σχετική απόκλιση ισχύος (I) και στη μεταβολή μεταξύ διαδοχικών παρατηρήσεων (MR), ενώ ο Isolation Forest εξετάζει το πολυμεταβλητό ηλεκτρικό προφίλ κάθε τεμαχίου. Συνεπώς, συμπεραίνουμε ότι ένα προϊόν μπορεί να είναι πολυμεταβλητά «ασυνήθιστο» χωρίς να εμφανίζει SPC alarm στη συγκεκριμένη μεταβλητή, ή αντίστροφα να εμφανίζει μονομεταβλητό SPC alarm χωρίς να καταχωρείται ως πολυμεταβλητή ανωμαλία.

Η ανάλυση επικάλυψης μεταξύ SPC I-MR alarms και IF Global 5% επιβεβαιώνει την παραπάνω εικόνα. Από τις 6.692 παρατηρήσεις, 16 χαρακτηρίστηκαν ταυτόχρονα ως SPC alarm και ML anomaly, 271 είχαν μόνο SPC alarm, 319 είχαν μόνο ML anomaly και 6.086 δεν είχαν καμία ένδειξη απόκλισης.

**Πίνακας 18 Αριθμός Alarms SPC και Isolation Forest**

| Κατηγορία         | Πλήθος    |
|-------------------|-----------|
| SPC I-MR only     | 271       |
| IF Global 5% only | 319       |
| <b>Both</b>       | <b>16</b> |
| Neither           | 6.086     |

Η περιορισμένη επικάλυψη των δύο προσεγγίσεων υποστηρίζει τη λογική της συμπληρωματικότητας. Το ML δεν αναπαράγει απλώς τα SPC alarms, αλλά εντοπίζει διαφορετικές περιοχές του πολυδιάστατου χώρου δεδομένων. Αυτό είναι ιδιαίτερα σημαντικό στο πλαίσιο της παρούσας εργασίας, καθώς ζητούμενο δεν είναι η αντικατάσταση του SPC,

αλλά η διερεύνηση του εάν η Μηχανική Μάθηση μπορεί να διανθίσει το SPC και να αξιοποιήσει επιπλέον πληροφορία στην παρακολούθηση της διεργασίας.

Με βάση τα αποτελέσματα διακρίνουμε 3 περιπτώσεις:

- ML & SPC Alarm
- SPC Alarm
- ML Alarm

Οι περιπτώσεις όπου υπάρχει ταυτόχρονα SPC alarm και ML anomaly είναι οι πιο ισχυρές ενδείξεις αποκλίνουσας συμπεριφοράς, καθώς η παρατήρηση εντοπίζεται τόσο από μονομεταβλητό στατιστικό έλεγχο όσο και από πολυμεταβλητό αλγόριθμο ανίχνευσης ανωμαλιών. Οι περιπτώσεις αυτές μπορούν να αντιμετωπιστούν ως υψηλότερης προτεραιότητας για τεχνική διερεύνηση, καθώς συνδυάζουν alarms σε δύο διαφορετικούς μηχανισμούς εντοπισμού αποκλίσεων.

Οι περιπτώσεις όπου εμφανίζεται μόνο SPC alarm υποδηλώνουν ότι η σχετική απόκλιση ισχύος ή η μεταβολή μεταξύ διαδοχικών παρατηρήσεων αποκλίνει από τη φυσική μεταβλητότητα της διεργασίας, χωρίς όμως το συνολικό πολυμεταβλητό προφίλ του τεμαχίου να θεωρείται ασυνήθιστο από το ML μοντέλο. Αυτό μπορεί να σημαίνει ότι η απόκλιση είναι έντονη σε μία συγκεκριμένη μεταβλητή, αλλά δεν συνοδεύεται από αντίστοιχη διαφοροποίηση στις υπόλοιπες ηλεκτρικές παραμέτρους.

Αντίθετα, οι περιπτώσεις όπου εμφανίζεται μόνο ML anomaly flag είναι ιδιαίτερα ενδιαφέρουσες από την πλευρά της πολυμεταβλητής ανάλυσης. Σε αυτές τις περιπτώσεις, το τεμάχιο μπορεί να μην υπερβαίνει τα στατιστικά όρια ελέγχου της μεταβλητής που παρακολουθείται από το SPC, αλλά ο συνδυασμός των ηλεκτρικών χαρακτηριστικών του να αποκλίνει από το σύνηθες πρότυπο και μπορεί να καταδείξει διαφορετικές αστοχίες στην παραγωγική διαδικασία (π.χ. ένα μοντέλο είναι σωστά ρυθμισμένο στην ισχύ, αλλά έχει λάθος driver ή πλακέτα). Το γεγονός αυτό αποτελεί σημείο που το ML μπορεί να προσδώσει προστιθέμενη αξία, καθώς μπορεί να αναδεικνύει ασυνήθιστες σχέσεις μεταξύ μεταβλητών που δεν είναι ορατές σε μονομεταβλητά διαγράμματα.

Τέλος, οι παρατηρήσεις στις οποίες δεν εμφανίζεται ούτε SPC alarm ούτε ML anomaly αποτελούν τις περιπτώσεις όπου και οι δύο προσεγγίσεις συμφωνούν ότι η ηλεκτρική

συμπεριφορά βρίσκεται εντός του αναμενόμενου πλαισίου. Η κατηγορία αυτή είναι η πολυπληθέστερη, γεγονός που είναι αναμενόμενο σε μία παραγωγική διεργασία όπου η συντριπτική πλειοψηφία των προϊόντων αναμένεται να παρουσιάζει κανονική συμπεριφορά.

#### 4.4 Συνολική αξιολόγηση σκοπιμότητας υιοθέτησης ML

Η συνολική αξιολόγηση των αποτελεσμάτων δείχνει ότι οι μέθοδοι Μηχανικής Μάθησης δύνανται να προσφέρουν πρόσθετη πληροφορία στη διαδικασία παρακολούθησης ποιότητας, χωρίς όμως να αντικαθιστούν άμεσα τα παραδοσιακά εργαλεία SPC. Το βασικό πλεονέκτημα του SPC παραμένει η απλότητα, η ευκολία στην ερμηνεία και η σαφής στατιστική λογική των διαγραμμάτων ελέγχου. Τα SPC alarms μπορούν να αξιοποιηθούν άμεσα από ένα σύστημα ποιότητας, καθώς συνδέονται με συγκεκριμένες μεταβλητές, συγκεκριμένα όρια ελέγχου και αναγνωρίσιμα πρότυπα διεργασιακής μεταβλητότητας.

Στον αντίποδα, οι μέθοδοι ML προσφέρουν πλεονεκτήματα κυρίως όπου απαιτείται πολυμεταβλητή προσέγγιση. Τα αποτελέσματα της PCA, του clustering, της supervised ταξινόμησης και του Isolation Forest δείχνουν ότι το ηλεκτρικό αποτύπωμα του τελικού ελέγχου εμπεριέχει πληροφορία που μπορεί να αξιοποιηθεί για την αναγνώριση τεχνικών διαμορφώσεων και την ανίχνευση ασυνήθιστων ηλεκτρικών προτύπων. Η προστιθέμενη αξία του ML εντοπίζεται κυρίως στην ικανότητά του να εξετάζει συνδυασμούς χαρακτηριστικών, οι οποίοι δεν αποτυπώνονται εύκολα σε μονομεταβλητά διαγράμματα ελέγχου.

Ωστόσο, η πρακτική υιοθέτηση ML σε ένα βιομηχανικό σύστημα ποιότητας απαιτεί προσεκτικό σχεδιασμό. Η σημαντική διαφοροποίηση της απόδοσης μεταξύ stratified random split και PO-based holdout που διαπιστώσαμε δείχνει ότι τα μοντέλα μπορούν να επηρεαστούν από production order effects, batch effects και άνιση κάλυψη των τεχνικών διαμορφώσεων ή άλλους παράγοντες οι οποίοι δεν φαίνεται άμεσα να επηρεάζουν τα αποτελέσματα και μπορούν να οδηγήσουν σε λανθασμένη ερμηνεία της μεταβλητότητας. Συνεπώς, πριν από οποιαδήποτε επιχειρησιακή εφαρμογή, απαιτείται επαρκές και αντιπροσωπευτικό ιστορικό δείγμα δεδομένων, ισχυρός έλεγχος διαρροής πληροφορίας, συνεχής παρακολούθηση της απόδοσης των μοντέλων και περιοδική επανεκπαίδευση όπου απαιτείται.

Με βάση τα παραπάνω, η υιοθέτηση ML κρίνεται περισσότερο σκόπιμη ως συμπληρωματικό επίπεδο ανάλυσης πάνω στο υφιστάμενο πλαίσιο SPC και όχι ως πλήρης αντικατάστασή του.

Σε ένα τέτοιο σχήμα, το SPC δύναται να συνεχίσει να λειτουργεί ως πρωτεύων μηχανισμός παρακολούθησης της διεργασίας, ενώ το ML μπορεί να χρησιμοποιηθεί για διερεύνηση πολυμεταβλητών ανωμαλιών, για την αναγνώριση τεχνικών προτύπων και την υποστήριξη πιο στοχευμένης ανάλυσης αποκλίσεων.

## Κεφάλαιο 5 – Συμπεράσματα, Περιορισμοί και Μελλοντική Έρευνα

### 5.1 Συνολική αποτίμηση της εργασίας

Η παρούσα εργασία είχε ως αντικείμενο τη συγκριτική αξιολόγηση παραδοσιακών μεθόδων Στατιστικού Ελέγχου Διεργασιών και μεθόδων Μηχανικής Μάθησης στο πλαίσιο της Διαχείρισης Ποιότητας σε βιομηχανικό περιβάλλον. Η μελέτη βασίστηκε σε πραγματικά δεδομένα τελικού ηλεκτρικού ελέγχου φωτιστικών σωμάτων LED, τα οποία αξιοποιήθηκαν για την εφαρμογή διαγραμμάτων SPC και αλγορίθμων Μηχανικής Μάθησης.

Η βασική λογική της εργασίας είναι η διερεύνηση του κατά πόσο οι ML-SPC προσεγγίσεις μπορούν να λειτουργήσουν συμπληρωματικά προς τα παραδοσιακά εργαλεία παρακολούθησης διεργασιών. Για τον λόγο αυτό, η ανάλυση οργανώθηκε σε δύο επίπεδα: αφενός στην εφαρμογή παραδοσιακών διαγραμμάτων ελέγχου σε μεταβλητές σχετικής απόκλισης και αφετέρου στην αξιοποίηση πολυμεταβλητών αλγορίθμων για την αναγνώριση ηλεκτρικών προτύπων και την ανίχνευση ανωμαλιών.

Τα αποτελέσματα ανέδειξαν ότι π ΣΕΔ εξακολουθεί να αποτελεί ιδιαίτερα χρήσιμο εργαλείο για την παρακολούθηση της στατιστικής σταθερότητας της διεργασίας. Τα I-MR και EWMA διαγράμματα παρείχαν σαφή σήματα σχετικά με τη μεταβλητότητα των επιλεγμένων ομάδων προϊόντων, ενώ οι δείκτες διεργασιακής ικανότητας βοήθησαν στην αποτύπωση της σχέσης μεταξύ φυσικής μεταβλητότητας και ορίων προδιαγραφών.

Από την πλευρά της Μηχανικής Μάθησης, τα αποτελέσματα έδειξαν ότι το ηλεκτρικό αποτύπωμα των προϊόντων περιέχει πληροφορία που μπορεί να αξιοποιηθεί για την αναγνώριση διαφορετικών τεχνικών διαμορφώσεων. Η εφαρμογή PCA, Ομαδοποίησης και Επιβλεπόμενης Ταξινόμησης έδειξε ότι οι συνδυασμοί ηλεκτρικών χαρακτηριστικών μπορούν να αναδείξουν διαφοροποιήσεις μεταξύ τύπων driver και τεχνικών διαμορφώσεων. Ωστόσο, η σημαντική μείωση της απόδοσης στο PO-based holdout ανέδειξε ότι για να επιτευχθεί η γενίκευση των ML μοντέλων σε νέες παραγωγικές συνθήκες απαιτούνται σημαντικά περισσότερα δεδομένα και καλύτερη εκπαίδευση των μοντέλων σε σχέση με την απόδοση που παρατηρείται σε τυχαίο stratified split.

Η σύγκριση μεταξύ SPC alarms και ML anomaly flags έδειξε περιορισμένη άμεση επικάλυψη, γεγονός που δεν ερμηνεύεται ως αποτυχία μίας από τις δύο μεθόδους. Αντίθετα, η περιορισμένη συμφωνία υποδηλώνει ότι το SPC και το ML εντοπίζουν διαφορετικές πτυχές της διεργασιακής συμπεριφοράς. Το SPC αποτυπώνει κυρίως μονομεταβλητές ή σειριακές αποκλίσεις από τη φυσική μεταβλητότητα, ενώ το ML μπορεί να εντοπίζει πολυμεταβλητές ανωμαλίες που προκύπτουν από τον συνδυασμό περισσότερων ηλεκτρικών χαρακτηριστικών. Συνεπώς, η Μηχανική Μάθηση φαίνεται να έχει αξία ως πρόσθετο επίπεδο ανάλυσης, όχι όμως ως άμεση αντικατάσταση των παραδοσιακών διαγραμμάτων ελέγχου.

## 5.2 Κύρια συμπεράσματα

Από την ανάλυση που πραγματοποιήθηκε προκύπτουν τα ακόλουθα βασικά συμπεράσματα, τα οποία παραθέτουμε σε μορφή διακριτών σημείων για διευκόλυνση του αναγνώστη:

- Το SPC παραμένει κατάλληλο και χρήσιμο εργαλείο για την παρακολούθηση της στατιστικής σταθερότητας των διεργασιών, ιδιαίτερα όταν εφαρμόζεται σε ομοιογενείς ομάδες προϊόντων ή σε μεταβλητές σχετικής απόκλισης.
- Τα ML μοντέλα μπορούν να αξιοποιήσουν το πολυμεταβλητό ηλεκτρικό αποτύπωμα των προϊόντων και να αναδείξουν τεχνικές διαφοροποιήσεις που δεν είναι εύκολα εμφανείς μέσω μονομεταβλητών διαγραμμάτων ελέγχου.
- Η υψηλή απόδοση των supervised ML μοντέλων σε stratified random split δείχνει ότι υπάρχει εκμεταλλεύσιμη πληροφορία στα ηλεκτρικά χαρακτηριστικά, όμως η χαμηλότερη απόδοση στο PO-based holdout δείχνει ότι η γενίκευση σε νέες παραγωγικές συνθήκες απαιτεί ιδιαίτερη προσοχή και πιθανότατα περισσότερα δεδομένα.
- Η χαμηλή επικάλυψη μεταξύ SPC alarms και ML anomaly flags υποστηρίζει τη λογική της συμπληρωματικότητας. Οι δύο προσεγγίσεις δεν εντοπίζουν τις ίδιες παρατηρήσεις, καθώς βασίζονται σε διαφορετική λογική.
- Η πρακτική αξιοποίηση της Μηχανικής Μάθησης σε συστήματα Ποιότητας προϋποθέτει σωστή επιλογή χαρακτηριστικών, μεγάλο πλήθος δεδομένων, επαρκή κάλυψη πιθανών διαφορετικών τεχνικών διαμορφώσεων και συνεχή επαναξιολόγηση της απόδοσης των μοντέλων.

### 5.3 Περιορισμοί της έρευνας

Η παρούσα εργασία υπόκειται σε ορισμένους περιορισμούς, οι οποίοι πρέπει να ληφθούν υπόψη κατά την ερμηνεία των αποτελεσμάτων. Πρώτον, τα δεδομένα προέρχονται από τον τελικό ηλεκτρικό έλεγχο προϊόντων και όχι από όλα τα ενδιάμεσα στάδια της παραγωγικής διαδικασίας. Συνεπώς, η ανάλυση δεν περιλαμβάνει άμεσα παραμέτρους που μπορεί να επηρεάζουν τη διεργασία σε προηγούμενα στάδια, όπως συνθήκες συναρμολόγησης, παρτίδες εξαρτημάτων, χειριστές, μηχανές ή περιβαλλοντικές συνθήκες.

Δεύτερον, κάθε εγγραφή αντιστοιχεί σε έναν τελικό έλεγχο ενός προϊόντος και όχι σε επαναλαμβανόμενες μετρήσεις του ίδιου τεμαχίου. Για τον λόγο αυτό, η εφαρμογή SPC βασίστηκε κυρίως σε I-MR charts και στη λογική διαδοχικής παραγωγικής ροής. Η απουσία φυσικών υποομάδων περιορίζει τη δυνατότητα εφαρμογής κλασικών  $\bar{X}$ -R ή  $\bar{X}$ -S διαγραμμάτων, εκτός από περιπτώσεις όπου μπορούν να σχηματιστούν τεχνικά ομοιογενείς υποομάδες.

Τρίτον, η ύπαρξη διαφορετικών οικογενειών προϊόντων, μοντέλων και διαμορφώσεων απαιτεί προσεκτική προεπεξεργασία και κανονικοποίηση των δεδομένων. Μολονότι η χρήση σχετικών αποκλίσεων περιορίζει το πρόβλημα της σύγκρισης μεταξύ ανόμοιων ομάδων, δεν αναιρεί πλήρως την πολυπλοκότητα που προκύπτει από διαφορετικές τεχνικές διαμορφώσεις και διαφορετικές παραγωγικές συνθήκες.

Τέταρτον, η αξιολόγηση των ML μοντέλων επηρεάζεται από την ανισορροπία των κλάσεων και από την άνιση κατανομή των configurations και των Production Orders. Η διαφορά μεταξύ random split και PO-based holdout δείχνει ότι η απόδοση των μοντέλων εξαρτάται σημαντικά από τον τρόπο διαχωρισμού των δεδομένων και από το κατά πόσο το σετ ελέγχου αντιπροσωπεύει πραγματικά νέες, διαφορετικές παραγωγικές συνθήκες.

Πέμπτο, τα SPC alarms δεν πρέπει να ερμηνεύονται ως καθολική αλήθεια για τη συγκριτική αξιολόγηση με τα ML anomaly flags. Τα alarms αποτελούν στατιστικό σημείο αναφοράς και όχι αντικειμενική ετικέτα αστοχίας. Για τον λόγο αυτό, οι δείκτες συμφωνίας μεταξύ SPC και ML πρέπει να ερμηνεύονται ως δείκτες σύγκλισης ή διαφοροποίησης των δύο προσεγγίσεων και όχι ως κλασικοί δείκτες πρόβλεψης πραγματικής αστοχίας.

Τέλος, στην παρούσα φάση δεν πραγματοποιήθηκε πλήρης οικονομική αξιολόγηση μέσω Κόστους Κακής Ποιότητας. Η έλλειψη επαρκών κοστολογικών δεδομένων δεν επέτρεψε τη σύνδεση των σημάτων SPC ή των ML anomalies με συγκεκριμένα κόστη απόρριψης, επανελέγχου, επανακατεργασίας ή καθυστέρησης. Επομένως, η αξιολόγηση της σκοπιμότητας υιοθέτησης ML περιορίστηκε κυρίως σε τεχνικό και λειτουργικό επίπεδο.

## 5.4 Προτάσεις για μελλοντική έρευνα

Με βάση τα αποτελέσματα και τους περιορισμούς της παρούσας εργασίας, προκύπτουν αρκετές δυνατότητες για μελλοντική έρευνα. Αρχικά, θα είχε σημαντικό ενδιαφέρον η επέκταση της ανάλυσης σε μεγαλύτερο αριθμό οικογενειών προϊόντων και configurations, ώστε να αξιολογηθεί εάν τα ευρήματα που προέκυψαν για την επιλεγμένη ομάδα προϊόντων μπορούν να γενικευθούν και σε άλλες κατηγορίες προϊόντων. Επιπλέον, μελλοντική έρευνα θα μπορούσε να ενσωματώσει πρόσθετα δεδομένα από προηγούμενα στάδια της παραγωγικής διαδικασίας. Η σύνδεση των δεδομένων τελικού ελέγχου με πληροφορίες όπως παρτίδες εξαρτημάτων, drivers, LED boards, γραμμές παραγωγής, χειριστές, χρόνος παραγωγής και περιβαλλοντικές συνθήκες θα μπορούσε να βελτιώσει σημαντικά την ερμηνευτική ισχύ τόσο των SPC όσο και των ML προσεγγίσεων.

Ιδιαίτερο ενδιαφέρον παρουσιάζει επίσης η ανάπτυξη πιο ολοκληρωμένων ML-SPC συστημάτων, στα οποία τα ML anomaly scores θα ενσωματώνονται σε λογική παρακολούθηση διεργασιών. Για παράδειγμα, θα μπορούσαν να εξεταστούν υβριδικές προσεγγίσεις όπου οι βαθμολογίες ανωμαλίας από αλγόριθμους όπως Isolation Forest, Autoencoders ή άλλες μέθοδοι ανίχνευσης ανωμαλιών παρακολουθούνται με διαγράμματα ελέγχου ή EWMA-type μηχανισμούς. Μελλοντική εργασία θα μπορούσε επίσης να δώσει μεγαλύτερη έμφαση στη χρονική διάσταση των δεδομένων. Η χρήση αυστηρότερων temporal validation schemes, η αξιολόγηση τάσης στη συμπεριφορά της διεργασίας και η παρακολούθηση της απόδοσης των ML μοντέλων σε μεταγενέστερες παραγωγικές περιόδους θα μπορούσαν να προσφέρουν πιο ρεαλιστική εικόνα για τη βιωσιμότητα τέτοιων συστημάτων σε πραγματικό περιβάλλον παραγωγής.

Επιπλέον, η σύνδεση των τεχνικών αποτελεσμάτων με οικονομικά δεδομένα αποτελεί σημαντική κατεύθυνση για μελλοντική έρευνα. Η ενσωμάτωση στοιχείων κόστους, όπως κόστος απόρριψης, επανελέγχου, επανακατεργασίας, καθυστέρησης ή εγγυήσεων, θα επέτρεπε

την αξιολόγηση της συμβολής SPC και ML όχι μόνο από τεχνική άποψη, αλλά και από την οπτική του Κόστους Κακής Ποιότητας.

Τέλος, θα ήταν χρήσιμη η πιλοτική εφαρμογή ενός συστήματος ML-SPC σε πραγματικό χρόνο ή σχεδόν πραγματικό χρόνο. Ένα τέτοιο σύστημα θα μπορούσε να λειτουργεί συμπληρωματικά προς τα υφιστάμενα διαγράμματα SPC, παρέχοντας ειδοποιήσεις για πολυμεταβλητές ανωμαλίες και υποστηρίζοντας πιο στοχευμένες αποφάσεις από τα στελέχη ποιότητας. Η πρακτική εφαρμογή θα απαιτούσε, ωστόσο, συνεχή έλεγχο της απόδοσης του μοντέλου, περιοδική επανεκπαίδευση και σαφείς κανόνες ερμηνείας των ειδοποιήσεων.

Συμπερασματικά, η εργασία ανέδειξε ότι ο παραδοσιακός Στατιστικός Έλεγχος Διεργασιών και οι μέθοδοι Μηχανικής Μάθησης μπορούν να λειτουργήσουν συμπληρωματικά στο πλαίσιο της βιομηχανικής Διαχείρισης Ποιότητας. Το SPC παρέχει ένα στατιστικά θεμελιωμένο και άμεσα ερμηνεύσιμο πλαίσιο παρακολούθησης της διεργασίας, ενώ η Μηχανική Μάθηση μπορεί να προσθέσει πολυμεταβλητή οπτική και να αναδείξει σύνθετα ηλεκτρικά πρότυπα. Η εφαρμογή ML δεν πρέπει να αντιμετωπίζεται ως αυτόματη αντικατάσταση των παραδοσιακών εργαλείων ποιότητας, αλλά ως τεχνολογική επέκταση που απαιτεί σωστά δομημένα δεδομένα, προσεκτική μεθοδολογία και συνεχή αξιολόγηση. Με τις κατάλληλες προϋποθέσεις, ο συνδυασμός μεθόδων SPC και ML μπορεί να ενισχύσει τη δυνατότητα έγκαιρης αναγνώρισης αποκλίσεων, να βελτιώσει την κατανόηση της διεργασίας και να υποστηρίξει πιο τεκμηριωμένες αποφάσεις στη Διαχείριση Ποιότητας.

## Βιβλιογραφικές Αναφορές

- Bersimis, S., Psarakis, S., & Panaretos, J. (2007). Multivariate statistical process control charts: An overview. *Quality and Reliability Engineering International*, 23(5), 517–543. <https://doi.org/10.1002/qre.829>
- Biegel, T., Jourdan, N., Hernandez, C., Cviko, A., & Metternich, J. (2022). Deep learning for multivariate statistical in-process control in discrete manufacturing: A case study in a sheet metal forming process. *Procedia CIRP*, 107, 422–427. <https://doi.org/10.1016/j.procir.2022.05.002>
- Boaventura, L. L., Ferreira, P. H., & Fiaccone, R. L. (2022). On flexible statistical process control with artificial intelligence: Classification control charts. *Expert Systems with Applications*, 194, Article 116492. <https://doi.org/10.1016/j.eswa.2021.116492>
- Chou, S.-H., Chang, S., Tsai, T.-R., Lin, D. K. J., Xia, Y., & Lin, Y.-S. (2020). Implementation of statistical process control framework with machine learning on waveform profiles with no gold standard reference. *Computers & Industrial Engineering*, 142, Article 106325. <https://doi.org/10.1016/j.cie.2020.106325>
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- Deming, W. E. (1986). *Out of the crisis*. MIT Press.
- Escobar, C. A., Cantoral-Ceballos, J. A., & Morales-Menendez, R. (2025). Quality 4.0: Learning quality control, the evolution of SQC/SPC. *Quality Engineering*, 37(1), 92–117. <https://doi.org/10.1080/08982112.2024.235651>
- Evans, J. R., & Lindsay, W. M. (2020). *Managing for quality and performance excellence* (11th ed.). Cengage Learning.
- Hric, P., & Sabahno, H. (2024). Developing machine learning-based control charts for monitoring different GLM-type profiles with different link functions. *Applied Artificial Intelligence*, 38(1). <https://doi.org/10.1080/08839514.2024.2362511>

Juran, J. M., & Godfrey, A. B. (Eds.). (1999). *Juran's quality handbook* (5th ed.). McGraw-Hill.

Kazmi, M. W., & Noor-ul-Amin, M. (2025). Integrating machine learning based EWMA control charts for multivariate process monitoring. *Computers & Industrial Engineering*, 204, Article 111131. <https://doi.org/10.1016/j.cie.2025.111131>

Montgomery, D. C. (2019). *Introduction to statistical quality control* (8th ed.). Wiley.

Perdikis, T., Celano, G., & Chakraborti, S. (2026). Practical considerations for implementing the Isolation Forest EWMA control chart in Phase II process monitoring. *Statistics and Computing*, 36, Article 9. <https://doi.org/10.1007/s11222-025-10766-9>

Schiffauerova, A., & Thomson, V. (2006). A review of research on cost of quality models and best practices. *International Journal of Quality & Reliability Management*, 23(6), 647–669. <https://doi.org/10.1108/02656710610672470>

Shewhart, W. A. (1931). *Economic control of quality of manufactured products*. D. Van Nostrand Company.

Wang, J., & Liu, L. (2024). A new multivariate control chart based on the isolation forest algorithm. *Quality Engineering*, 36(2), 390 - 406. <https://doi.org/10.1080/08982112.2023.2220773>

Zan, T., Liu, Z., Su, Z., Wang, M., Gao, X., & Chen, D. (2020). Statistical process control with intelligence based on the deep learning model. *Applied Sciences*, 10(1), Article 308. <https://doi.org/10.3390/app10010308>

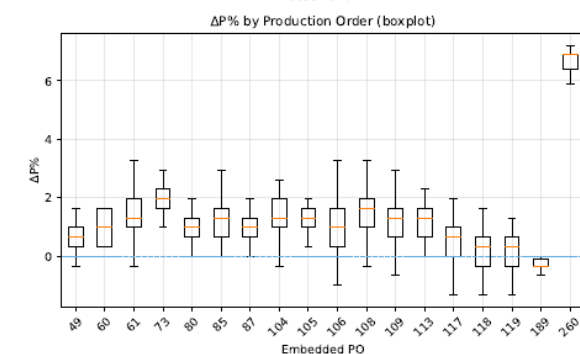
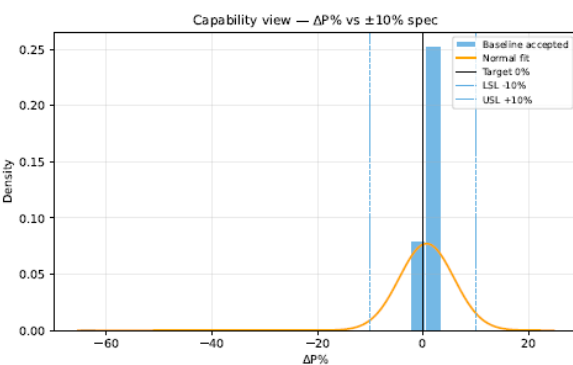
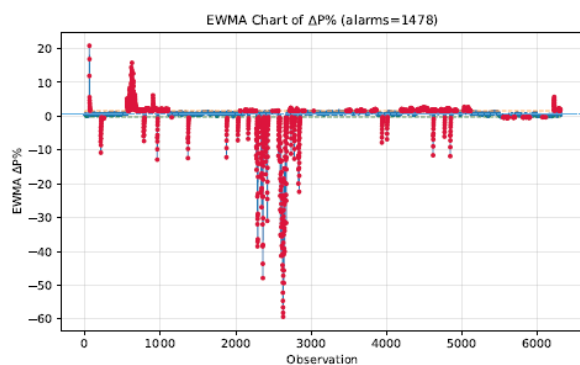
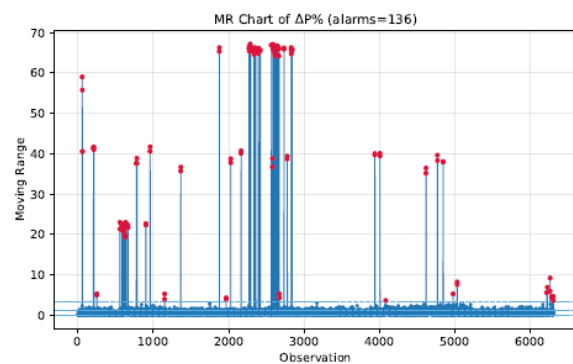
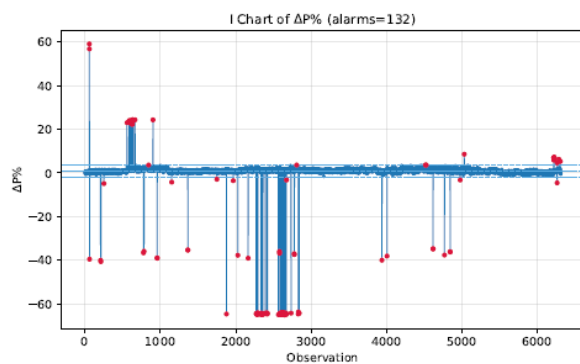
Zhou, Y., Jin, R., & Qiu, P. (2025). Machine learning control charts for monitoring spatio-temporal data streams. *Quality and Reliability Engineering International*, 41, 2373–2384. <https://doi.org/10.1002/qre.3809>

## ΠΑΡΑΡΤΗΜΑΤΑ

### Παράρτημα Α – Διαγράμματα & Σχήματα

A | Target 30.5W  
SPC diagrams including I, MR, EWMA, Capability and ΔP% by PO

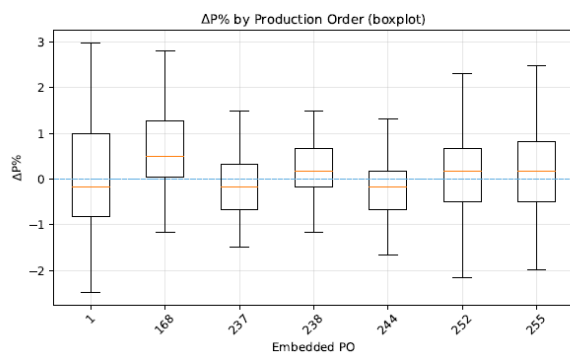
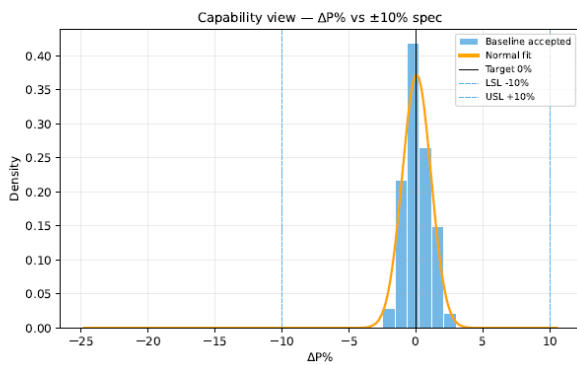
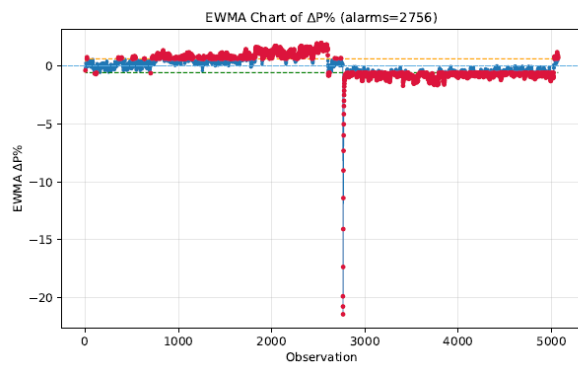
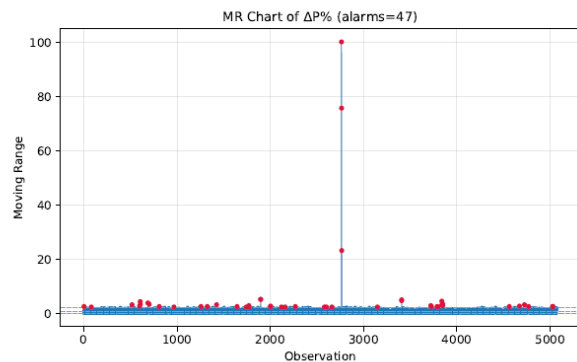
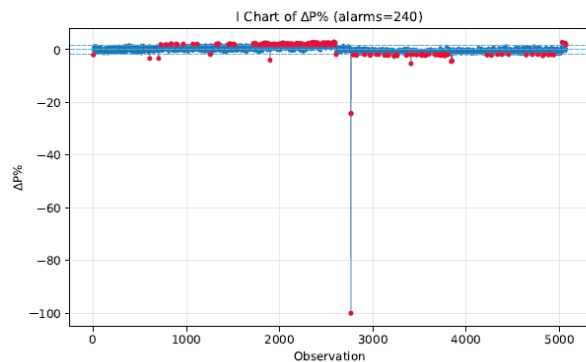
|          |       |            |      |             |               |                     |               |
|----------|-------|------------|------|-------------|---------------|---------------------|---------------|
| n total  | 6310  | n baseline | 6258 | Rejected    | 52            | Spec out $\pm 10\%$ | 109           |
| Mean ΔP% | 0.308 | Std ΔP%    | 7.26 | Cp / Cpk    | 2.941 / 2.850 | Pp / Ppk            | 0.459 / 0.445 |
| I alarms | 132   | MR alarms  | 136  | EWMA alarms | 1478          | PO count            | 18            |



Selection rationale:  
Main high-volume/reject case

A | Target 60.5W  
SPC diagrams including I, MR, EWMA, Capability and ΔP% by PO

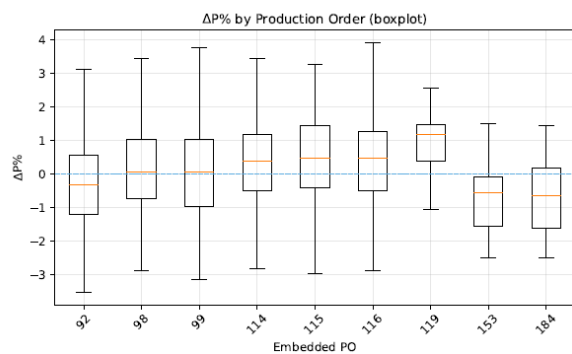
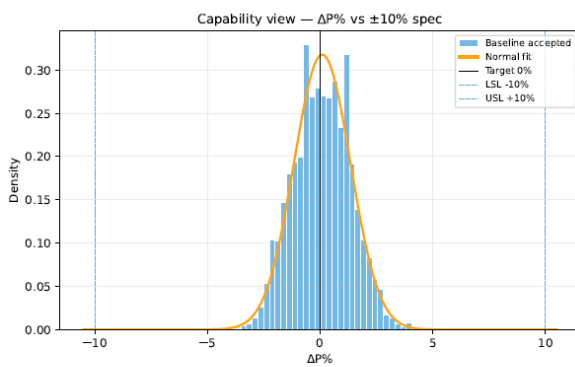
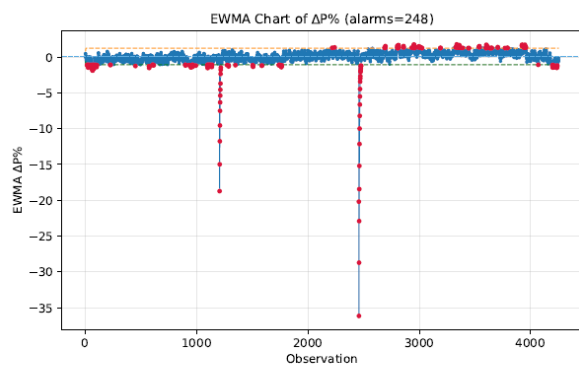
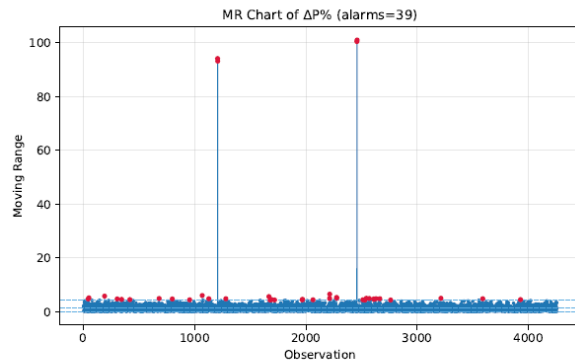
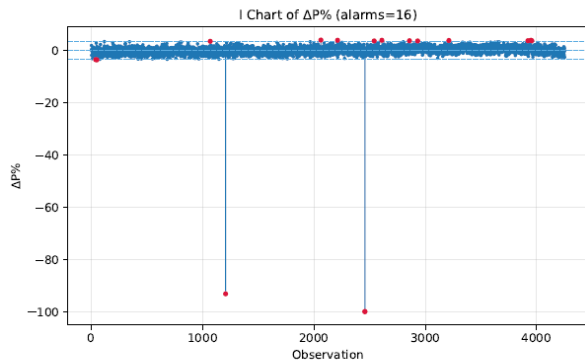
|          |       |            |       |             |               |               |               |
|----------|-------|------------|-------|-------------|---------------|---------------|---------------|
| n total  | 5072  | n baseline | 5071  | Rejected    | 1             | Spec out ±10% | 3             |
| Mean ΔP% | 0.014 | Std ΔP%    | 1.768 | Cp / Cpk    | 5.493 / 5.485 | Pp / Ppk      | 1.886 / 1.883 |
| I alarms | 240   | MR alarms  | 47    | EWMA alarms | 2756          | PO count      | 7             |



Selection rationale:  
Stable high-volume comparison

B | Target 12.5W  
SPC diagrams including I, MR, EWMA, Capability and ΔP% by PO

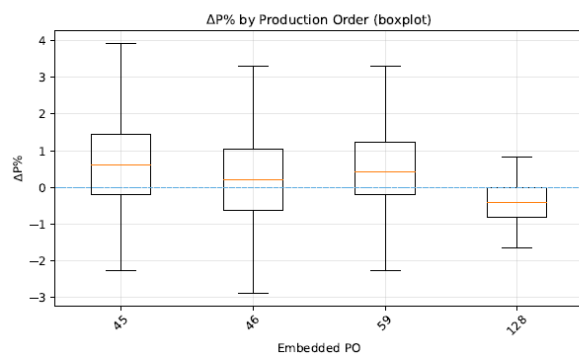
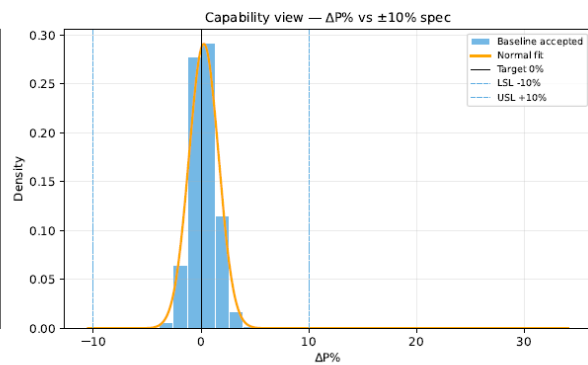
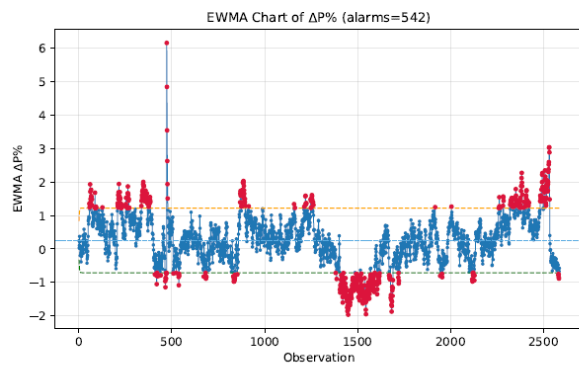
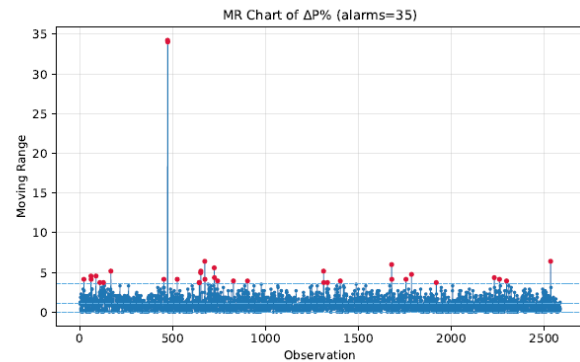
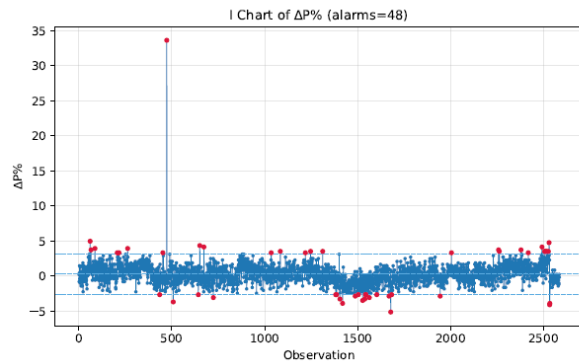
|          |       |            |       |             |               |                     |               |
|----------|-------|------------|-------|-------------|---------------|---------------------|---------------|
| n total  | 4252  | n baseline | 4249  | Rejected    | 3             | Spec out $\pm 10\%$ | 3             |
| Mean ΔP% | 0.022 | Std ΔP%    | 2.886 | Cp / Cpk    | 2.775 / 2.769 | Pp / Ppk            | 1.155 / 1.152 |
| I alarms | 16    | MR alarms  | 39    | EWMA alarms | 248           | PO count            | 9             |



Selection rationale:  
STR family representative

B | Target 48.5W  
SPC diagrams including I, MR, EWMA, Capability and ΔP% by PO

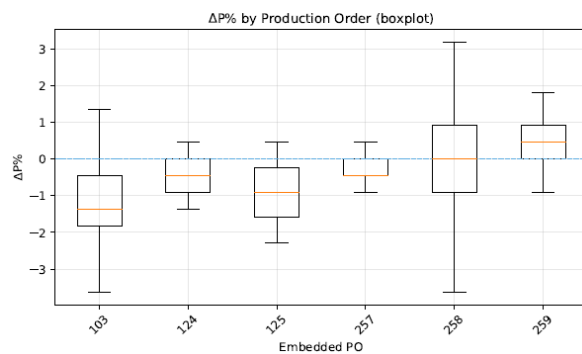
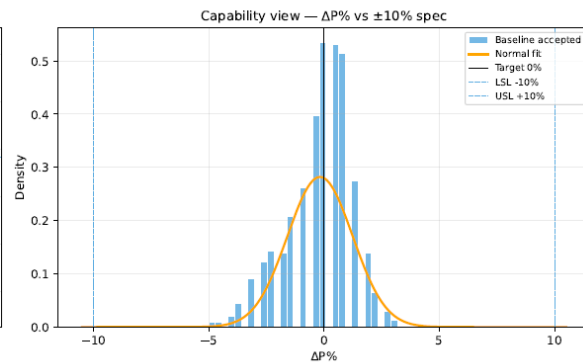
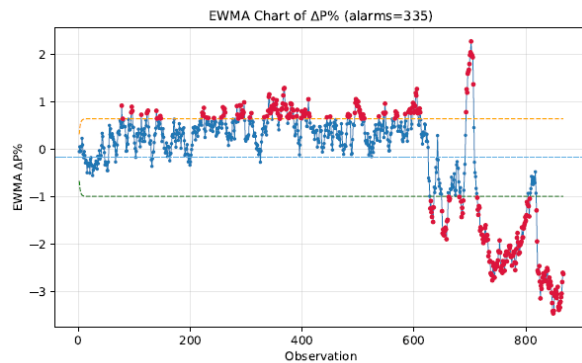
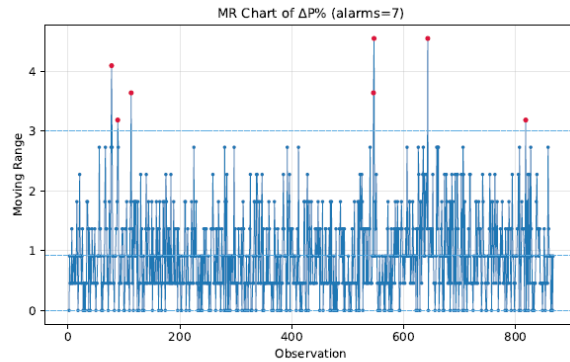
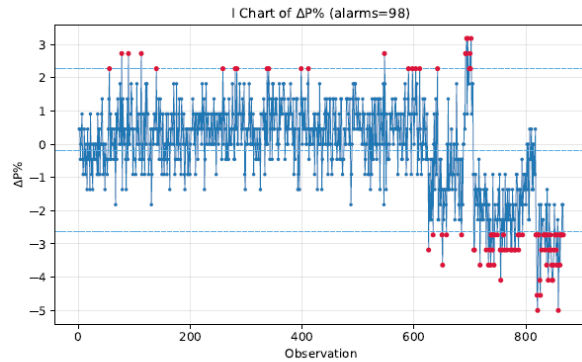
|          |       |            |       |             |               |                     |               |
|----------|-------|------------|-------|-------------|---------------|---------------------|---------------|
| n total  | 2585  | n baseline | 2585  | Rejected    | 0             | Spec out $\pm 10\%$ | 1             |
| Mean ΔP% | 0.257 | Std ΔP%    | 1.372 | Cp / Cpk    | 3.456 / 3.368 | Pp / Ppk            | 2.430 / 2.367 |
| I alarms | 48    | MR alarms  | 35    | EWMA alarms | 542           | PO count            | 4             |



Selection rationale:  
STR stable comparison

C | Target 220.0W  
SPC diagrams including I, MR, EWMA, Capability and ΔP% by PO

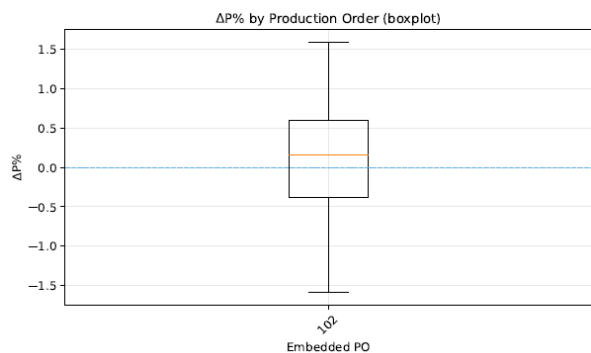
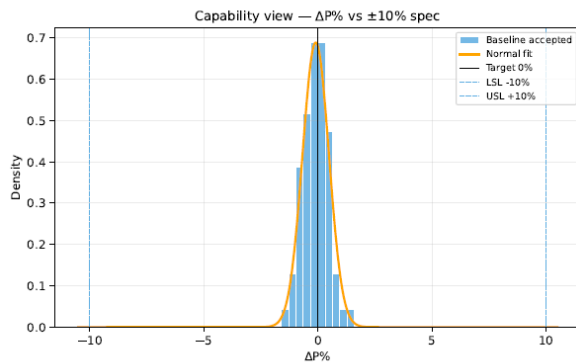
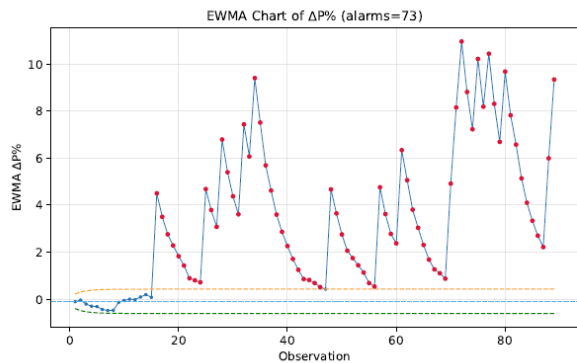
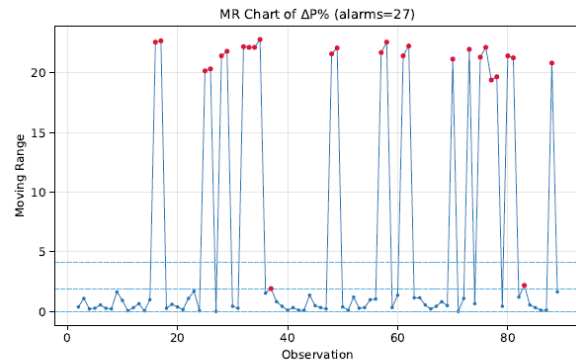
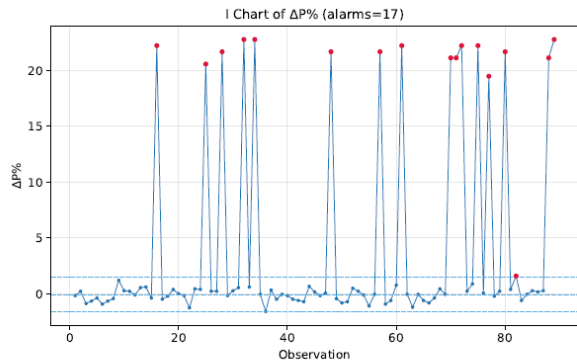
|          |        |            |       |             |               |               |               |
|----------|--------|------------|-------|-------------|---------------|---------------|---------------|
| n total  | 867    | n baseline | 867   | Rejected    | 0             | Spec out ±10% | 0             |
| Mean ΔP% | -0.176 | Std ΔP%    | 1.415 | Cp / Cpk    | 4.086 / 4.014 | Pp / Ppk      | 2.356 / 2.315 |
| I alarms | 98     | MR alarms  | 7     | EWMA alarms | 335           | PO count      | 6             |



Selection rationale:  
PLB family representative

C | Target 182.5W  
SPC diagrams including I, MR, EWMA, Capability and ΔP% by PO

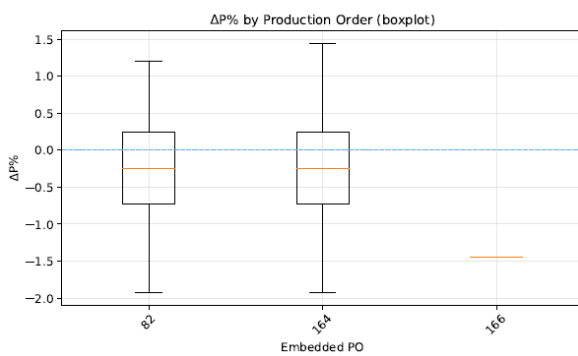
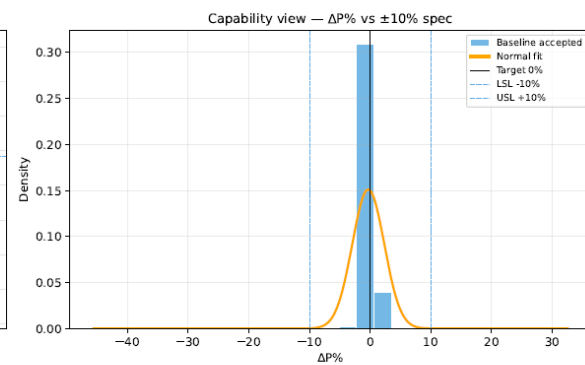
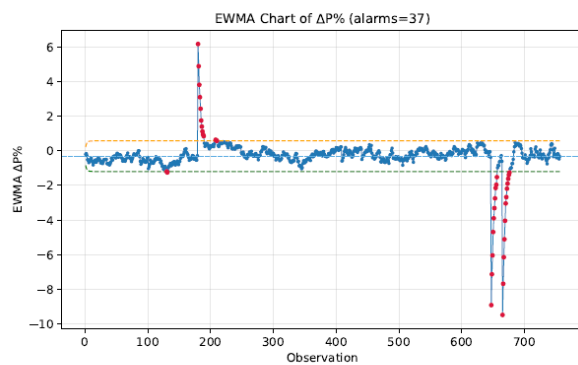
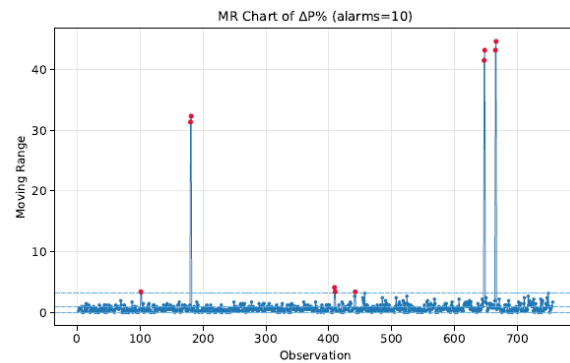
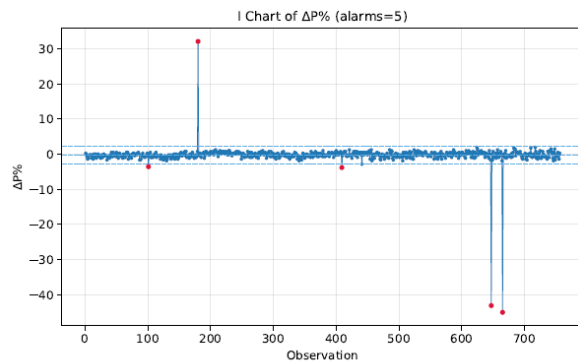
|          |       |            |       |             |               |               |               |
|----------|-------|------------|-------|-------------|---------------|---------------|---------------|
| n total  | 89    | n baseline | 73    | Rejected    | 16            | Spec out ±10% | 16            |
| Mean ΔP% | 3.821 | Std ΔP%    | 8.431 | Cp / Cpk    | 0.907 / 0.560 | Pp / Ppk      | 0.395 / 0.244 |
| I alarms | 17    | MR alarms  | 27    | EWMA alarms | 73            | PO count      | 1             |



Selection rationale:  
High-reject PLB special case

A | Target 41.5W  
SPC diagrams including I, MR, EWMA, Capability and ΔP% by PO

|          |        |            |       |             |               |               |               |
|----------|--------|------------|-------|-------------|---------------|---------------|---------------|
| n total  | 756    | n baseline | 756   | Rejected    | 0             | Spec out ±10% | 3             |
| Mean ΔP% | -0.308 | Std ΔP%    | 2.642 | Cp / Cpk    | 3.759 / 3.643 | Pp / Ppk      | 1.262 / 1.223 |
| I alarms | 5      | MR alarms  | 10    | EWMA alarms | 37            | PO count      | 3             |



Selection rationale:  
Additional PLS comparison case

## Παράρτημα Β – Κώδικας SPC & ML

```
6  # %% [01] Libraries and analysis settings
7  from pathlib import Path
8
9  import numpy as np
10 import pandas as pd
11
12 from sklearn.preprocessing import StandardScaler
13 from sklearn.decomposition import PCA
14 from sklearn.cluster import MiniBatchKMeans
15 from sklearn.mixture import GaussianMixture
16 from sklearn.ensemble import IsolationForest
17 from sklearn.metrics import adjusted_rand_score, silhouette_score
18
19 RANDOM_STATE = 42
20
21 DATA_PATH = Path("ML DATASHEET 60W ONLY.xlsx")
22 OUTPUT_DIR = Path("outputs_pls60w_spc_ml")
23 OUTPUT_DIR.mkdir(exist_ok=True)
24
25 ML_FEATURES = [
26     "Delta_P_pct",
27     "PF_decimal",
28     "THDu_pct",
29     "Voltage_V",
30     "Current_per_W",
31 ]
32
33 CONFIG_COL = "Configuration_ID"
34 RESULT_COL = "Result"
35
36 # %% [02] Data loading and basic filtering
37 df = pd.read_excel(DATA_PATH)
38
39 # The selected dataset already contains the PLS 60W product group
40 # and the technical configuration fields used in the thesis.
41 df = df.copy()
42
43 df["Date"] = pd.to_datetime(df["Date"], errors="coerce")
44 df["Time_str"] = df["Time"].astype(str)
45
46 # Sort records so that SPC and EWMA-based comparisons follow the test order.
47 df = df.sort_values(["Date", "Time_str", "Test_Order", "Source_Row"]).reset_index(drop=True)
48
49 print("Total records:", len(df))
50 print("Configurations:", df[CONFIG_COL].nunique())
51 print("Drivers:", df["Driver Type"].nunique())
52 print("LED boards:", df["LED Board Type"].nunique())
53
54 # %% [03] Feature engineering and leakage control
55 # The raw measured power is not used directly in ML, because nearby target powers
56 # exist within the same PLS 60W group. The normalized deviation ΔP% is used instead.
57 df["Delta_P_pct"] = (
58     (df["Measured_Power_W"] - df["Target_Power_W"])
59     / df["Target_Power_W"]
60     * 100
61 )
62
63 # Current is expressed relative to measured power to reduce direct scale effects.
64 df["Current_per_W"] = df["Input_Current_mA"] / df["Measured_Power_W"]
65
66 # Variables such as Configuration_ID, Driver Type, LED Board Type, P0,
67 # Result and SPC alarms are kept for evaluation only and are not used as ML inputs.
68 context_columns = [
69     "Serial_Number",
70     "Embedded_P0",
71     "Configuration_ID",
72     "Driver Type",
73     "LED Board Type",
74     "Result",
75     "I_MR_I_Alarm",
76     "I_MR_MR_Alarm",
77     "EWMA_Alarm",
78 ]
```

```

80 # %% [04] ML dataset preparation and standardization
81 # Keep only records with complete numerical features.
82 ml_df = df.dropna(subset=ML_FEATURES).copy()
83
84 # Combined SPC alarm used later for comparison with ML anomaly flags.
85 ml_df["SPC_I_MR_Alarm"] = (
86     ml_df["I_MR_I_Alarm"].astype(bool)
87     | ml_df["I_MR_MR_Alarm"].astype(bool)
88 )
89
90 # Baseline used for fitting unsupervised models.
91 # Rejected records and incomplete rows are not used to define the normal baseline.
92 baseline_df = ml_df[ml_df["Baseline_Flag"].astype(bool)].copy()
93
94 scaler = StandardScaler()
95 X_baseline = scaler.fit_transform(baseline_df[ML_FEATURES])
96 X_all = scaler.transform(ml_df[ML_FEATURES])
97
98 print("Rows with complete ML features:", len(ml_df))
99 print("Baseline rows:", len(baseline_df))
100

```

```

101 # %% [05] PCA for electrical signature visualization
102 pca = PCA(n_components=3, random_state=RANDOM_STATE)
103 pca_scores_all = pca.fit_transform(X_all)
104
105 ml_df["PC1"] = pca_scores_all[:, 0]
106 ml_df["PC2"] = pca_scores_all[:, 1]
107 ml_df["PC3"] = pca_scores_all[:, 2]
108
109 pca_explained = pd.DataFrame({
110     "Component": ["PC1", "PC2", "PC3"],
111     "Explained_Variance_Ratio": pca.explained_variance_ratio_,
112     "Cumulative_Explained_Variance": np.cumsum(pca.explained_variance_ratio_),
113 })
114
115 pca_loadings = pd.DataFrame(
116     pca.components_.T,
117     index=ML_FEATURES,
118     columns=["PC1", "PC2", "PC3"],
119 ).reset_index(names="Feature")
120

```

```

121 # %% [06] Clustering for configuration signature assessment
122 n_configs = ml_df[CONFIG_COL].nunique()
123 n_clusters = min(max(n_configs, 2), 8)
124
125 kmeans = MiniBatchKMeans(
126     n_clusters=n_clusters,
127     random_state=RANDOM_STATE,
128     batch_size=1024,
129     n_init=20,
130 )
131 ml_df["KMeans_Cluster"] = kmeans.fit_predict(X_all)
132
133 gmm = GaussianMixture(
134     n_components=n_clusters,
135     covariance_type="full",
136     random_state=RANDOM_STATE,
137 )
138 ml_df["GMM_Cluster"] = gmm.fit_predict(X_all)
139 ml_df["GMM_Max_Probability"] = gmm.predict_proba(X_all).max(axis=1)
140
141 cluster_quality = pd.DataFrame({
142     "Method": ["KMeans", "GMM"],
143     "n_clusters": [n_clusters, n_clusters],
144     "ARI_vs_Configuration": [
145         adjusted_rand_score(ml_df[CONFIG_COL], ml_df["KMeans_Cluster"]),
146         adjusted_rand_score(ml_df[CONFIG_COL], ml_df["GMM_Cluster"]),
147     ],
148     "Silhouette": [
149         silhouette_score(X_all, ml_df["KMeans_Cluster"]),
150         silhouette_score(X_all, ml_df["GMM_Cluster"]),
151     ],
152 })
153

```

```

154 # %% [07] Isolation Forest for multivariate anomaly detection
155 global_iforest = IsolationForest(
156     n_estimators=300,
157     contamination=0.02,
158     random_state=RANDOM_STATE,
159     n_jobs=-1,
160 )
161
162 global_iforest.fit(X_baseline)
163
164 ml_df["IF_Global_Score"] = global_iforest.decision_function(X_all)
165 ml_df["IF_Global_Anomaly"] = global_iforest.predict(X_all) == -1
166
167 # Configuration-level Isolation Forest is closer to Phase I SPC logic:
168 # each configuration is evaluated against its own normal baseline.
169 ml_df["IF_ByConfig_Score"] = np.nan
170 ml_df["IF_ByConfig_Anomaly"] = False
171
172 for config_id, group in ml_df.groupby(CONFIG_COL):
173     base_idx = group[group["Baseline_Flag"].astype(bool)].index
174     if len(base_idx) < 100:
175         continue
176
177     group_idx = group.index
178     cfg_scaler = StandardScaler()
179     X_cfg_base = cfg_scaler.fit_transform(ml_df.loc[base_idx, ML_FEATURES])
180     X_cfg_all = cfg_scaler.transform(ml_df.loc[group_idx, ML_FEATURES])
181
182     cfg_iforest = IsolationForest(
183         n_estimators=300,
184         contamination=0.02,
185         random_state=RANDOM_STATE,
186         n_jobs=-1,
187     )
188     cfg_iforest.fit(X_cfg_base)
189
190     ml_df.loc[group_idx, "IF_ByConfig_Score"] = cfg_iforest.decision_function(X_cfg_all)
191     ml_df.loc[group_idx, "IF_ByConfig_Anomaly"] = cfg_iforest.predict(X_cfg_all) == -1

```

```

193 # %% [08] SPC vs ML agreement tables
194 def agreement_table(reference_col: str, prediction_col: str) -> pd.DataFrame:
195     table = pd.crosstab(
196         ml_df[reference_col].astype(bool),
197         ml_df[prediction_col].astype(bool),
198         rownames=[reference_col],
199         colnames=[prediction_col],
200         dropna=False,
201     )
202     return table.reset_index()
203
204 spc_vs_global_iforest = agreement_table("SPC_I_MR_Alarm", "IF_Global_Anomaly")
205 spc_vs_cfg_iforest = agreement_table("SPC_I_MR_Alarm", "IF_ByConfig_Anomaly")
206
207 kmeans_vs_config = pd.crosstab(
208     ml_df["KMeans_Cluster"],
209     ml_df[CONFIG_COL],
210     rownames=["KMeans_Cluster"],
211     colnames=[CONFIG_COL],
212 )
213
214 gmm_vs_config = pd.crosstab(
215     ml_df["GMM_Cluster"],
216     ml_df[CONFIG_COL],
217     rownames=["GMM_Cluster"],
218     colnames=[CONFIG_COL],
219 )
220

```

```

221 # %% [09] Export results
222 output_file = OUTPUT_DIR / "PLS60W_SPC_ML_appendix_outputs.xlsx"
223
224 with pd.ExcelWriter(output_file, engine="openpyxl") as writer:
225     pca_explained.to_excel(writer, sheet_name="PCA_Explained", index=False)
226     pca_loadings.to_excel(writer, sheet_name="PCA_Loadings", index=False)
227     cluster_quality.to_excel(writer, sheet_name="Cluster_Quality", index=False)
228     kmeans_vs_config.to_excel(writer, sheet_name="KMeans_vs_Config")
229     gmm_vs_config.to_excel(writer, sheet_name="GMM_vs_Config")
230     spc_vs_global_iforest.to_excel(writer, sheet_name="SPC_vs_Global_IF", index=False)
231     spc_vs_cfg_iforest.to_excel(writer, sheet_name="SPC_vs_Config_IF", index=False)
232     ml_df[context_columns + ML_FEATURES + [
233         "PC1", "PC2", "PC3",
234         "KMeans_Cluster", "GMM_Cluster", "GMM_Max_Probability",
235         "SPC_I_MR_Alarm",
236         "IF_Global_Score", "IF_Global_Anomaly",
237         "IF_ByConfig_Score", "IF_ByConfig_Anomaly",
238     ]].to_excel(writer, sheet_name="ML_Assignments", index=False)
239
240 print("Exported:", output_file)

```

## Επιβλεπόμενη Ταξινόμηση

```

6 # %% [01] Libraries and analysis settings
7 from pathlib import Path
8
9 import numpy as np
10 import pandas as pd
11
12 from sklearn.compose import ColumnTransformer
13 from sklearn.pipeline import Pipeline
14 from sklearn.preprocessing import StandardScaler
15 from sklearn.model_selection import train_test_split, GroupShuffleSplit
16 from sklearn.metrics import (
17     accuracy_score,
18     balanced_accuracy_score,
19     f1_score,
20     classification_report,
21     confusion_matrix,
22 )
23 from sklearn.linear_model import LogisticRegression
24 from sklearn.ensemble import RandomForestClassifier, ExtraTreesClassifier
25
26 RANDOM_STATE = 42
27
28 DATA_PATH = Path("ML_DATASHEET_60W_ONLY.xlsx")
29 OUTPUT_DIR = Path("outputs_pls60w_supervised")
30 OUTPUT_DIR.mkdir(exist_ok=True)
31
32 ML_FEATURES = [
33     "Delta_P_pct",
34     "PF_decimal",
35     "THDu_pct",
36     "Voltage_V",
37     "Current_per_W",
38 ]
39
40 TARGETS = ["Driver_Type", "Configuration_ID"]

```

```

42 # %% [02] Data loading and feature preparation
43 df = pd.read_excel(DATA_PATH)
44
45 df["Delta_P_pct"] = (
46     (df["Measured_Power_W"] - df["Target_Power_W"])
47     / df["Target_Power_W"]
48     * 100
49 )
50
51 df["Current_per_W"] = df["Input_Current_mA"] / df["Measured_Power_W"]
52
53 # Keep only complete records for the selected electrical features and labels.
54 model_df = df.dropna(subset=ML_FEATURES + TARGETS + ["Embedded_P0"]).copy()
55
56 print("Rows used for supervised learning:", len(model_df))
57 print("Driver classes:", model_df["Driver_Type"].nunique())
58 print("Configuration classes:", model_df["Configuration_ID"].nunique())

```



```
60 # %% [03] Leakage control and class-size filtering
61 # The model must learn from electrical measurements only.
62 # Target labels such as Driver Type and Configuration_ID are never used as input features.
63 # Target Power, raw Measured Power, P0 and SPC alarms are also excluded from the feature set.
64
65 MIN_CLASS_SIZE = 50
66
67 def filter_small_classes(data: pd.DataFrame, target_col: str) -> pd.DataFrame:
68     counts = data[target_col].value_counts()
69     valid_classes = counts[counts >= MIN_CLASS_SIZE].index
70     return data[data[target_col].isin(valid_classes)].copy()
```

```
72 # %% [04] Model definitions
73 # Balanced class weights are used because the number of records differs by class.
74 models = {
75     "Logistic Regression": LogisticRegression(
76         max_iter=2000,
77         class_weight="balanced",
78         random_state=RANDOM_STATE,
79     ),
80     "Random Forest": RandomForestClassifier(
81         n_estimators=300,
82         class_weight="balanced",
83         random_state=RANDOM_STATE,
84         n_jobs=-1,
85     ),
86     "Extra Trees": ExtraTreesClassifier(
87         n_estimators=300,
88         class_weight="balanced",
89         random_state=RANDOM_STATE,
90         n_jobs=-1,
91     ),
92 }
93
94 def make_pipeline(model):
95     return Pipeline(steps=[
96         ("scaler", StandardScaler()),
97         ("model", model),
98     ])
```

```
100 # %% [05] Evaluation helper
101 def evaluate_classifier(model_name, estimator, X_train, X_test, y_train, y_test, split_name, target_name):
102     estimator.fit(X_train, y_train)
103     y_pred = estimator.predict(X_test)
104
105     metrics = {
106         "Target": target_name,
107         "Split": split_name,
108         "Model": model_name,
109         "Accuracy": accuracy_score(y_test, y_pred),
110         "Balanced_Accuracy": balanced_accuracy_score(y_test, y_pred),
111         "Macro_F1": f1_score(y_test, y_pred, average="macro"),
112         "Weighted_F1": f1_score(y_test, y_pred, average="weighted"),
113         "Train_Size": len(y_train),
114         "Test_Size": len(y_test),
115     }
116
117     report = pd.DataFrame(
118         classification_report(y_test, y_pred, output_dict=True, zero_division=0)
119     ).T.reset_index(names="Class")
120
121     cm = pd.DataFrame(
122         confusion_matrix(y_test, y_pred, labels=sorted(y_test.unique())),
123         index=sorted(y_test.unique()),
124         columns=sorted(y_test.unique()),
125     )
126
127     return metrics, report, cm, estimator
```



```
129 # %% [06] Stratified random split
130 all_metrics = []
131 all_reports = {}
132 all_confusion_matrices = {}
133 fitted_models = {}
134
135 for target_col in TARGETS:
136     data = filter_small_classes(model_df, target_col)
137
138     X = data[ML_FEATURES]
139     y = data[target_col].astype(str)
140
141     X_train, X_test, y_train, y_test = train_test_split(
142         X,
143         y,
144         test_size=0.30,
145         random_state=RANDOM_STATE,
146         stratify=y,
147     )
148
149     for model_name, model in models.items():
150         pipe = make_pipeline(model)
151         metrics, report, cm, fitted = evaluate_classifier(
152             model_name, pipe, X_train, X_test, y_train, y_test,
153             split_name="Stratified random split",
154             target_name=target_col,
155         )
156         all_metrics.append(metrics)
157         all_reports[(target_col, "random", model_name)] = report
158         all_confusion_matrices[(target_col, "random", model_name)] = cm
159         fitted_models[(target_col, "random", model_name)] = fitted
160
161 # %% [07] PO-based holdout split
162 # This split is stricter because complete Production Orders are held out.
163 # It tests whether the electrical signature generalizes beyond neighbouring records
164 # from the same production batch.
165 for target_col in TARGETS:
166     data = filter_small_classes(model_df, target_col)
167
168     X = data[ML_FEATURES]
169     y = data[target_col].astype(str)
170     groups = data["Embedded_PO"].astype(str)
171
172     splitter = GroupShuffleSplit(
173         n_splits=1,
174         test_size=0.30,
175         random_state=RANDOM_STATE,
176     )
177
178     train_idx, test_idx = next(splitter.split(X, y, groups=groups))
179
180     X_train, X_test = X.iloc[train_idx], X.iloc[test_idx]
181     y_train, y_test = y.iloc[train_idx], y.iloc[test_idx]
182
183     for model_name, model in models.items():
184         pipe = make_pipeline(model)
185         metrics, report, cm, fitted = evaluate_classifier(
186             model_name, pipe, X_train, X_test, y_train, y_test,
187             split_name="PO-based holdout",
188             target_name=target_col,
189         )
190         all_metrics.append(metrics)
191         all_reports[(target_col, "po_holdout", model_name)] = report
192         all_confusion_matrices[(target_col, "po_holdout", model_name)] = cm
193         fitted_models[(target_col, "po_holdout", model_name)] = fitted
```



```
195 # %% [08] Feature importance from tree-based models
196 importance_rows = []
197
198 for (target_col, split_name, model_name), fitted in fitted_models.items():
199     final_model = fitted.named_steps["model"]
200
201     if not hasattr(final_model, "feature_importances_"):
202         continue
203
204     for feature, importance in zip(ML_FEATURES, final_model.feature_importances_):
205         importance_rows.append({
206             "Target": target_col,
207             "Split": split_name,
208             "Model": model_name,
209             "Feature": feature,
210             "Importance": importance,
211         })
212
213 feature_importance = pd.DataFrame(importance_rows)
214
215 # %% [09] Export supervised learning results
216 metrics_df = pd.DataFrame(all_metrics)
217 output_file = OUTPUT_DIR / "PLS60W_supervised_appendix_outputs.xlsx"
218
219 with pd.ExcelWriter(output_file, engine="openpyxl") as writer:
220     metrics_df.to_excel(writer, sheet_name="Supervised Metrics", index=False)
221     feature_importance.to_excel(writer, sheet_name="Feature Importance", index=False)
222
223     for key, report in all_reports.items():
224         target_col, split_name, model_name = key
225         sheet = f"Report_{target_col[:6]}_{split_name[:6]}_{model_name[:6]}"
226         report.to_excel(writer, sheet_name=sheet[:31], index=False)
227
228     for key, cm in all_confusion_matrices.items():
229         target_col, split_name, model_name = key
230         sheet = f"CM_{target_col[:6]}_{split_name[:6]}_{model_name[:6]}"
231         cm.to_excel(writer, sheet_name=sheet[:31])
232
233 print("Exported:", output_file)
```

## Παράρτημα Γ – Δείγμα Πρωτογενών Δεδομένων

| Σειριακός Αριθμός | Μέτρηση          | Κατάσταση | Ημέρα    | Ώρα      | P (W) | S (VA) | Q (VAR) | PF (%) | THDu (%) | THDi (mA) | Idiff (mA) | Itau (mA) | U (V) | I (mA) | cosφ (%) |
|-------------------|------------------|-----------|----------|----------|-------|--------|---------|--------|----------|-----------|------------|-----------|-------|--------|----------|
| 20242098166160024 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:37:17  | 40,9  | 43,3   | 14,06   | 0,95   | 3,2      | 16        | 0,05       | 0,004     | 250   | 173    | 0,95     |
| 20242098166160025 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:38:22  | 41,1  | 43,4   | 14,05   | 0,95   | 3,2      | 17        | 0,05       | 0,004     | 251   | 173    | 0,95     |
| 20242098166160022 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:39:21  | 40,7  | 43,1   | 14,09   | 0,94   | 3,1      | 18        | 0,05       | 0,004     | 249   | 173    | 0,95     |
| 20242098166160012 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:40:17  | 40,7  | 43,1   | 14,19   | 0,94   | 3,2      | 18        | 0,06       | 0,004     | 250   | 173    | 0,95     |
| 20242098166160013 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:41:20  | 40,8  | 43,2   | 13,97   | 0,95   | 3,1      | 17        | 0,05       | 0,004     | 250   | 173    | 0,95     |
| 20242098166160020 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:47:18  | 40,8  | 43,1   | 13,86   | 0,95   | 3,4      | 16        | 0,05       | 0,004     | 249   | 173    | 0,95     |
| 20242098166160009 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:48:33  | 42,1  | 44,2   | 13,38   | 0,95   | 3,1      | 16        | 0,06       | 0,004     | 251   | 176    | 0,96     |
| 20242098166160018 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:49:37  | 41    | 43,2   | 13,56   | 0,95   | 2,9      | 16        | 0,05       | 0,004     | 249   | 173    | 0,95     |
| 20242098166160019 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:50:47  | 40,8  | 43,1   | 13,79   | 0,95   | 3,2      | 16        | 0,05       | 0,004     | 249   | 173    | 0,95     |
| 20242098166160008 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:51:40  | 41,1  | 43,4   | 13,76   | 0,95   | 3,2      | 16        | 0,05       | 0,004     | 249   | 174    | 0,95     |
| 20242098166160005 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:57:27  | 41,8  | 43,8   | 13,17   | 0,95   | 2,9      | 18        | 0,05       | 0,004     | 250   | 175    | 0,96     |
| 20242098166160018 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:58:27  | 41    | 43,1   | 13,11   | 0,95   | 3,2      | 18        | 0,05       | 0,004     | 249   | 173    | 0,96     |
| 20242098166160010 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 9:59:22  | 40,6  | 42,9   | 13,61   | 0,95   | 2,9      | 17        | 0,05       | 0,004     | 248   | 173    | 0,95     |
| 20242098166160027 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:00:29 | 41,3  | 43,6   | 14,24   | 0,95   | 2,8      | 16        | 0,05       | 0,004     | 250   | 175    | 0,95     |
| 20242098166160011 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:01:33 | 41,5  | 43,5   | 13,22   | 0,95   | 3,1      | 19        | 0,05       | 0,004     | 250   | 174    | 0,96     |
| 20242098166160003 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:02:47 | 41,4  | 43,2   | 12,66   | 0,96   | 2,8      | 16        | 0,06       | 0,004     | 250   | 173    | 0,96     |
| 20242098166160006 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:04:50 | 41,7  | 43,8   | 13,34   | 0,95   | 3,2      | 20        | 0,05       | 0,004     | 248   | 176    | 0,96     |
| 20242098166160017 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:05:50 | 40,7  | 43     | 13,8    | 0,95   | 3,2      | 17        | 0,05       | 0,004     | 249   | 173    | 0,95     |
| 20242098166160007 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:06:48 | 41,4  | 43,6   | 13,65   | 0,95   | 3,1      | 16        | 0,05       | 0,004     | 250   | 174    | 0,95     |
| 20242098166160026 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:07:44 | 41,1  | 43,5   | 14,03   | 0,95   | 3,1      | 17        | 0,05       | 0,004     | 250   | 174    | 0,95     |
| 20242098166160026 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:08:45 | 41,1  | 43,3   | 13,56   | 0,95   | 3,0      | 16        | 0,06       | 0,004     | 249   | 173    | 0,95     |
| 20242098166160014 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:09:47 | 40,9  | 43,1   | 13,39   | 0,95   | 3,0      | 17        | 0,05       | 0,004     | 249   | 173    | 0,96     |
| 20242098166160004 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:11:13 | 40,8  | 43,1   | 13,77   | 0,95   | 3,3      | 19        | 0,06       | 0,004     | 248   | 174    | 0,95     |
| 20242098166160002 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:12:17 | 41,2  | 43,3   | 13,46   | 0,95   | 4,1      | 21        | 0,05       | 0,004     | 248   | 175    | 0,96     |
| 20242098166160001 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:13:19 | 40,7  | 42,9   | 13,45   | 0,95   | 3,3      | 18        | 0,05       | 0,004     | 248   | 173    | 0,95     |
| 20242098166160021 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:14:30 | 40,6  | 42,9   | 13,91   | 0,95   | 2,8      | 17        | 0,05       | 0,004     | 249   | 173    | 0,95     |
| 20242098166160023 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:15:37 | 41,2  | 43,2   | 12,99   | 0,95   | 2,6      | 16        | 0,05       | 0,004     | 250   | 173    | 0,96     |
| 20242098164160293 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:17:51 | 41,8  | 45,1   | 17,1    | 0,93   | 2,4      | 23        | 0,05       | 0,004     | 250   | 180    | 0,94     |
| 20242098164160295 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:18:55 | 41,8  | 45,1   | 17,02   | 0,93   | 2,3      | 23        | 0,05       | 0,004     | 250   | 180    | 0,94     |
| 20242098164160290 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:20:29 | 41,1  | 44,4   | 16,81   | 0,93   | 2,4      | 24        | 0,05       | 0,004     | 249   | 179    | 0,94     |
| 20242098164160300 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:21:25 | 41,9  | 45,2   | 17,07   | 0,93   | 2,5      | 23        | 0,06       | 0,004     | 250   | 181    | 0,94     |
| 20242098164160283 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:23:29 | 41,8  | 44,8   | 16,3    | 0,93   | 2,6      | 23        | 0,05       | 0,004     | 250   | 179    | 0,94     |
| 20242098164160282 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:24:30 | 42,1  | 45,3   | 16,75   | 0,93   | 2,5      | 23        | 0,06       | 0,004     | 250   | 181    | 0,94     |
| 20242098164160287 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:25:48 | 41,8  | 45,3   | 17,52   | 0,92   | 2,3      | 24        | 0,05       | 0,004     | 252   | 179    | 0,93     |
| 20242098164160289 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:27:03 | 41,6  | 45     | 17,24   | 0,92   | 2,3      | 24        | 0,05       | 0,004     | 253   | 178    | 0,93     |
| 20242098164160296 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:30:00 | 41    | 44,5   | 17,23   | 0,92   | 2,4      | 24        | 0,05       | 0,004     | 252   | 177    | 0,93     |
| 20242098164160297 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:30:54 | 40,7  | 44     | 16,52   | 0,93   | 2,2      | 24        | 0,05       | 0,004     | 251   | 175    | 0,94     |
| 20242098164160298 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:31:51 | 41,7  | 45,1   | 17,22   | 0,92   | 2,6      | 25        | 0,05       | 0,004     | 251   | 179    | 0,94     |
| 20242098164160256 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:32:44 | 42,1  | 45,4   | 17,23   | 0,93   | 2,5      | 24        | 0,06       | 0,004     | 252   | 180    | 0,94     |
| 20242098164160281 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:33:38 | 40,8  | 44,4   | 17,3    | 0,92   | 2,6      | 25        | 0,05       | 0,004     | 250   | 177    | 0,93     |
| 20242098164160280 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:34:33 | 41,2  | 44,6   | 17,13   | 0,92   | 2,8      | 24        | 0,05       | 0,004     | 252   | 177    | 0,93     |
| 20242098164160285 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:35:29 | 41,2  | 44,7   | 17,29   | 0,92   | 2,8      | 23        | 0,05       | 0,004     | 251   | 178    | 0,93     |
| 20242098164160284 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:36:25 | 41,5  | 45     | 17,27   | 0,92   | 2,5      | 23        | 0,05       | 0,004     | 252   | 179    | 0,93     |
| 20242098164160286 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:37:23 | 41,7  | 44,8   | 16,59   | 0,93   | 2,5      | 23        | 0,05       | 0,004     | 252   | 178    | 0,94     |
| 20242098164160288 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:38:33 | 41,3  | 44,7   | 17,09   | 0,92   | 2,5      | 26        | 0,05       | 0,004     | 251   | 178    | 0,94     |
| 20242098164160290 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:39:36 | 40,9  | 44,2   | 16,65   | 0,93   | 2,4      | 25        | 0,05       | 0,004     | 251   | 176    | 0,94     |
| 20242098164160291 | Διαρροές & Ισχύς | Αποδέκτο  | 7/1/2025 | 10:40:32 | 41,6  | 44,9   | 17,05   | 0,93   | 2,8      | 24        | 0,05       | 0,004     | 251   | 179    | 0,94     |
| 20242098167160030 | Διαρροές & Ισχύς | Αποδέκτο  | 8/1/2025 | 11:46:58 | 40,1  | 47,6   | 25,7    | 0,84   | 2,9      | 55        | 0,05       | 0,004     | 251   | 189    | 0,88     |
| 20242098167160032 | Διαρροές & Ισχύς | Αποδέκτο  | 8/1/2025 | 11:48:01 | 40,7  | 48,6   | 26,5    | 0,84   | 3,0      | 54        | 0,06       | 0,004     | 252   | 193    | 0,88     |
| 20242098167160033 | Διαρροές & Ισχύς | Αποδέκτο  | 8/1/2025 | 11:48:56 | 40,2  | 48     | 26,2    | 0,84   | 2,8      | 55        | 0,05       | 0,004     | 253   | 190    | 0,88     |
| 20242098167160034 | Διαρροές & Ισχύς | Αποδέκτο  | 8/1/2025 | 11:49:59 | 40,3  | 47,8   | 25,8    | 0,84   | 2,4      | 54        | 0,05       | 0,004     | 253   | 189    | 0,88     |
| 20242098167160035 | Διαρροές & Ισχύς | Αποδέκτο  | 8/1/2025 | 11:51:06 | 40,2  | 47,6   | 25,5    | 0,84   | 2,8      | 52        | 0,05       | 0,004     | 252   | 189    | 0,88     |

**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.