Hellenic Open University

Postgraduate Course

Postgraduate Dissertation

The analysis and forecasting of energy consumption in greenhouse crops, using statistical and predictive models, with the aim of optimizing energy efficiency

Georgios Voutsas

Supervisor: Stylianos Psarakis

Patras, Greece, May 2026

The analysis and forecasting of energy consumption in greenhouse crops, using statistical and predictive models, with the aim of optimizing energy efficiency

Georgios Voutsas

Supervising Committee

Supervisor:

Stylianos Psarakis

Department of Statistics, Athens University of Economics and Business

Co-Supervisor:

Georgios Androulakis

Department of Business Administration, University of Patras

Patras, Greece, May 2026

“Acknowledgents and / or Dedication”

Abstract

Energy consumption in greenhouse cultivation is analysed and predicted with the daily data for the whole year 2025. Outdoor temperature, greenhouse indoor temperature, relative humidity, solar radiation and energy consumption are included in the data set. The research used descriptive statistics, correlation analysis and interpretive analysis of the major factors of energy demand. The results indicated that the biggest influencing variable is that of the ambient temperature, with the other variables, such as the ambient humidity, solar radiation and indoor temperature, having an auxiliary role in shaping the energy load. The study showed marked seasonality in the consumption, with higher consumption during the late winter and early spring period. The results validate the relevance of predictive analysis towards the optimization of energy efficiency, GHPs operational stability and administrative decisions.

Keywords: greenhouse, energy consumption, energy forecast, microclimate, energy efficiency

Περίληψη

Η εργασία επικεντρώνεται στην ανάλυση και την πρόβλεψη της ενέργειας που χρησιμοποιείται στην παραγωγή θερμοκηπίων με βάση ένα πλήρες έτος ημερήσιων δεδομένων το 2025. Τα δεδομένα αποτελούνται από την εξωτερική θερμοκρασία, την εσωτερική θερμοκρασία του θερμοκηπίου, τη σχετική υγρασία, την ηλιακή ακτινοβολία και την κατανάλωση ενέργειας. Οι μελέτες βασίστηκαν στην περιγραφική στατιστική, τη μελέτη συσχέτισης και την ερμηνευτική ανάλυση των πρωταρχικών παραγόντων που διαμορφώνουν τη ζήτηση ενέργειας. Αυτά τα ευρήματα αποκάλυψαν ότι η εξωτερική θερμοκρασία είναι ο πιο ισχυρός προγνωστικός παράγοντας της κατανάλωσης, ενώ η εσωτερική θερμοκρασία, η υγρασία και η ηλιακή ακτινοβολία είναι συμπληρωματικές στον ορισμό του ενεργειακού φορτίου. Η ανάλυση έδειξε ότι είναι σαφώς εποχιακό, με την κατανάλωση να αυξάνεται στα τέλη του χειμώνα και στις αρχές της άνοιξης. Τα αποτελέσματα αποδεικνύουν τη σημασία της προγνωστικής ανάλυσης για τη βελτιστοποίηση της ενεργειακής απόδοσης, τη σταθερότητα της λειτουργίας του θερμοκηπίου και τη βοήθεια στη λήψη διοικητικών αποφάσεων.

Λέξεις-κλειδιά: θερμοκήπιο, κατανάλωση ενέργειας, πρόβλεψη ενέργειας, μικροκλίμα, ενεργειακή απόδοση

Contents

List of Tables

List of Figures

1 Introduction

1.1 Energy Management in the Agricultural Sector

Energy management is a key strategic point today in modern agriculture in which there is a growing pressure for agricultural production, reducing its impacts on the environment and costs, while increasing productivity. Agriculture consumes a large share of final energy supply both directly in the use of energy and indirectly in the production of fertilizers, machinery, infrastructure etc. (Food and Agriculture Organization [FAO], 2011). The need for systematic planning of energy use on farms has become more urgent in recent years as energy costs have increased and energy markets have fluctuated and as commitments have been made to reduce climate change.

The use of energy in agriculture can be closely associated to mechanization, irrigation, heating, cooling, storage and processing. Energy is one of the most significant costs of operation in intensive farming systems, particularly in horticulture and under cover crops. Energy management, therefore, can be an environmental issue, but it also can play a major role in the profitability and resilience of the farm. The International Energy Agency (IEA, 2017) reports that enhancing energy efficiency of agricultural systems is one of the most cost-effective methods for reducing GHG emissions and ensuring food security.

This data-driven agriculture approach further underlines the need for energy monitoring and optimisation. The application of statistical (and predictive) analyses to energy flows is made possible by the high resolution, time-based data obtained from precision farming technologies, sensor networks and automated control systems. These advances are part of the broader context of sustainability agendas, where resource efficiency, decarbonization and digital transformation of food production systems have their focus (European Commission, 2019).

Energy management in the agricultural sector is shifting from a "fire-fighting" cost control mode to an energy optimization process. It has more and more become a system that mixes performance evaluation, predictive modeling and scenario simulation, thus enabling decision making on the basis of evidence. The use of quantitative methods to help analyse the energy consumption patterns on farms can help farm managers and policy-makers detect inefficiencies, forecast energy demand in a different climate and plan interventions to boost economic and environmental performance.

For this reason, analytical research on energy use in agriculture is not a mere technical activity, but an integral part of the strategy for sustainable rural development.

1.2 Energy demand and operational characteristics of greenhouse crops

Of all agricultural production systems, greenhouse production is one of the most energy demanding. Unlike field production, greenhouse crop production requires the management of the greenhouse environment to provide optimum conditions for crop growth. The total energy use of a greenhouse unit (GHU) is influenced by heating, ventilation, air circulation, cooling, water pumping for irrigation, and supplemental lighting and automation systems (Boulard et al., 2011).

Typically, the energy consumption is dominated by heating, particularly in temperate and cold climate regions. The energy requirements for heating varies with outside temperature, solar radiation, wind velocity, greenhouse design and insulation level, and the heat requirement of the crop. In heated greenhouses, the consumption of energy for greenhouse heating can be as high as 70–90% of the total energy consumed in the greenhouse, depending on the climatic conditions and the technological level of the greenhouse (Pérez-Parra et al., 2004). In Mediterranean countries the need of heating is lower than in northern parts of Europe but as a result of the seasonality and the growing demands for production quality, the energy consumption is still substantial. Energy needs are also dependent on the operational structure of the greenhouses. Optimum limits of temp and humidity vary over different crops. For instance, tomato production typically involves daytime temperature range of 20-25°C with controlled relative humidity in order to prevent disease development, thus requiring continuous monitoring, and climate control interventions (Boulard et al., 2011). In addition, energy requirements change for plants at different growth stages: young plants or flowering and fruiting stages can have specific demand for the microclimate.

Solar radiation has a twofold role. However on the other hand, it's also passive solar, which during the sunny times, reduces the need for heating. However, the opposite can occur when the internal temperature is above the desired set-point, which will require more cooling and increased ventilation. The energy use is complex and non-linear as a result of dynamic interaction between external meteorological conditions and internal control mechanisms. This complexity makes the use of quantitative analysis methods in greenhouses especially appropriate because rule-of-thumb based management cannot account for the relationships between different climatic factors.

Meanwhile, technical advances like automated climate control platforms and built-in monitoring systems enable data – including temperature, humidity, radiation and energy usage – to be gathered continuously. This high-time resolution data created by these systems are used for statistical analysis and to make predictive models. So it is important to understand the energy demand in greenhouses and its operational and climatic aspects, in order to develop reliable models for energy forecasting, which can support energy optimization strategies.

1.3 Problem Statement and Research Gap

However, there is a large research gap regarding the integration of greenhouse climate control and energy saving technologies and the use of real operational greenhouse data with predictive statistics and machine learning models to optimize energy use on the farm scale. A considerable quantity of the current literature is based on the technical development of the design, including improving the use of the insulation material and heating systems, or on general simulated models, which are designed under controlled experimental conditions (Pérez-Parra et al., 2004; Boulard et al., 2011).

While these methods have made an important step in understanding the physical mechanism of energy behavior in greenhouses, many methods lack in examining the empirical validation with historical data from commercial greenhouse production units. In reality all energy use depends not only on physical and technical parameters, but also related to administrative decisions, operational patterns, climatic system regulation strategies and climatic variability at the local level. Thus, the correctness of prediction can be lowered if the model (theoretical/generalized) is implemented without adequate calibration to actual operating conditions.

In addition, the use of machine learning techniques has increased in the agricultural sector, especially in applications like yield forecasting, disease detection, irrigation systems, etc. (Kamilaris & Prenafeta-Boldú, 2018), but is still relatively limited in the application of energy use in greenhouses. No systematic comparative studies exist between the classical statistical models (e.g., multiple linear regression) and the more complex machine learning models, either in terms of their predictive performance or the understanding of the results.

Hence, there exists a research gap in this systematic analysis and prediction of energy use in a greenhouse crop through the combined use of statistical and machine learning analysis with the use of real operation data. Filling this gap is important for the effective design and development of sound decision support systems for greenhouse energy use optimization and cost reduction to cope with various environmental conditions.

1.4 Research Objectives and Research Questions

In this study, the aim is to analyse and predict the energy use in greenhouse crops quantitatively, based on actual data from the greenhouses and advanced analysis methods. To explore the factors affecting energy demand and compare the prediction abilities of these factors, a combination of statistical modeling and machine learning techniques have been used.

The research goals in this study are the following ones:

1. To identify and quantify the most important climatic and greenhouse operational factors influencing energy use.
2. To create and test statistical and machine learning models for short term energy demand forecasting.
3. To check the performance of the models by the well-known performance indices including coefficient of determination, mean absolute error and the root mean square error.
4. To model other operating scenarios, for the purpose of aiding energy optimisation.

Based on the above objectives, the study attempts to answer the following research questions:

1. Which environmental and operational parameters are important for the energy use in greenhouses?
2. How accurate is the multiple linear regression method in comparison with the machine learning models to predict energy demand?

3. What are conclusions based on predictive modelling and scenario analysis regarding energy savings?

The process of setting goals and formulating research questions makes the study methodologically coherent, and provides the opportunity to make the empirical research logically linked to the practical application in the field of greenhouse energy management.

1.5 Research Structure

This work presents six major chapters that are structured in a logical order beginning with the theoretical aspects and then moving to the empirical ones and concluding with the findings.

A literature review on energy use in greenhouses, monitoring and control technologies and statistical analysis methods for energy demand forecasting and application of machine learning in agriculture are provided in Chapter 2.

Chapter 3 presents the research methodology, the process used to gather data, the definition of variables, the operational mapping of variables, the model development process and evaluation process.

Chapter 4 discusses the statistical analysis of the data and the results and highlights the factors influencing energy consumption.

The results of the predictive models are assessed in chapter 5 which also contains simulations for alternative operating scenarios.

Lastly, Chapter 6 concludes the key findings, discusses their implications and makes recommendations for future research.

2 Literature Review

2.1 Energy Consumption Standards in Greenhouse Production Systems

Greenhouse production systems have high integrated energy use, with a high variability because environmental conditions need to be regulated artificially to maintain optimal growing conditions for the crop. Unlike outdoor agriculture where the crops are exposed to the natural climatic conditions, greenhouses are considered as a closed microclimate system. To maintain the internal climate, a continuous energy supply is required for the heating, cooling, ventilation, irrigation and artificial lighting system. The importance of each of these components varies with geographical location, building design, type of crop and the technology level (Boulard et al., 2011).

In temperate and cold zones, the energy process that typically accounts for the most energy is usually heating. The purpose of heating systems is to keep the temperature of the rooms within a certain range, during the period of low temperature in the outside environment, the temperature range is set according to the temperature demand of the crops. The heating demand depends on the heat loss coefficient of the building, the properties of the insulation material, as well as weather conditions outside and desired conditions inside. Conventional systems typically operate with gas or oil boilers and more advanced systems use cogeneration units or renewable energy. It has been reported that up to 70–90% of the direct energy use in greenhouses in northern European countries is available for heating (Pérez-Parra et al., 2004). Due to night temperature drops and seasonal variability winter heating is an important cost component even in Mediterranean countries.

The second primary energy use is for cooling and ventilation, particularly in hot climates or in summer. Overheating in the plant can cause less efficiency in photosynthesis and can negatively influence plant growth. Ventilation can be passive (from roof openings and side windows) or active (by means of fans and exhaust systems). Mechanical systems use more electricity, but provide a more accurate control over the internally controlled environment. Solar radiation intensity, greenhouse orientation, shading systems and crop transpiration rates have a large impact on cooling demand. The combination of heating and cooling load implies complicated consumption patterns on a daily and seasonal scale.

Electricity is consumed during irrigation system implementation, pumping and distribution. Their contribution to the energy use is typically small compared to heating or cooling, but can be important in large systems or hydroponic systems that have continuously running recirculating pumps. Precision farming technologies are designed to maximise the use of water and nutrient resources, minimise the wastage of resources, but can lead to an increase in basic consumption because of the additional control equipment.

In modern high tech greenhouses, artificial light has increasingly become an important energy parameter. To even out production and extend the photoperiod, supplementary lighting is provided. As a higher energy efficient and spectral control alternative to high-pressure sodium systems, LED systems are slowly replacing these types of light systems. However, lighting can be a significant consumer of electricity especially in areas with a low sun angle in winter (International Energy Agency (2017).

The energy mix is reflected in the structure of energy costs as well as in market conditions. Since energy costs are one of the largest variable costs depending on the type of crop and climate (Boulard et al., 2011), there is a significant potential for savings through the implementation of energy efficient technologies. Fuel costs can influence heating bills, and electricity rates are influenced by the tariff and regulatory measures.

Fixed and variable costs must be separated in a cost analysis. Fixed costs are things like investments in heating, lighting and automation systems, and variable costs are related to such expenses as fuel, electricity, and maintenance. To optimize the economy, the working and investments must be optimized. From this overall picture, it is clear that energy use, in greenhouses, is a multi-factorial and climate dependent phenomenon and requires quantitative analytical methods to be understood and predicted.

2.2 Environmental and Climatic Factors Determining Energy Demand

The need for energy in the greenhouse is the basic result of the relationship between the climatic conditions required for best crop growth and the outside climatic environment. As opposed to the conventional buildings, greenhouses are semi transparent constructions with the aim of maximising the uptake of solar radiation, thus at the same time introducing a very high thermal variability. Thus, the use of energy is a consequence of constant adjustments to the internal state relative to changes in environment (Boulard et al., 2011).

Air temperature, solar radiation, wind speed and direction, relative humidity and seasonal variability are the main external climatic variables. Of these the most important factor affecting heating demand is external temperature. Heating system(s) are necessary if ambient temperature goes below desired crop temperature, to make up for heat loss by conduction, convection and radiation from the greenhouse shell. These losses are dependent on how well the structure is insulated and how leaky the structure is, and the type of covering material. In cold climates, more time is spent in the cooling conditions which have a high impact on fuel consumption, particularly at night when there is no sun gain (Pérez-Parra et al., 2004).

The Solar radiation has a double and dynamic role. Incident shortwave radiation during the day will introduce passive heating, which decreases the necessity for artificial heating. However, if the amount of radiation is too great, it can lead to over-heating, particularly in well-insulated or air-tight buildings. In these situations active cooling or ventilation is needed to control the target temperature thresholds. Radiation can have either a negative or positive impact on energy demand as a whole, dependent on the intensity and duration of radiation, and the interaction of radiation with the internal control. Radiation also has a seasonal variation which affects the annual energy profile.

The convective heat losses from the greenhouse surface depend upon wind speed and direction. The more wind the more air will get in, which will increase the heat exchange between the inside and outside and resulted in more heating demand. The sensitivity of the energy use to wind conditions is affected by the structure design, including roof geometry, ventilation opening design, etc. Heating and ventilation requirements are indirectly affected by latent heat fluxes and plant transpiration, which depend on the relative humidity.

Requirements for the indoor climate will depend on crop physiology and production objectives. Optimum temperature, moisture, and light necessary for maximum photosynthesis and yield quality vary from one species to another. A stable temperature regime is important, for example, during the growing of tomato, the humidity levels need to be controlled to prevent fungal infections and to smoothen the fruit development (Boulard et al., 2011). These indoor settings have to be monitored and acted upon by control systems continually. When deviations from the desired limits occur, productivity (yield) may be sacrificed; thus, a trade-off between the energy expended and agronomic yield occurs. Consumption patterns are non-linear and time-dependent because of the interplay between the outdoor and indoor climate. For example, on a cold day in the winter with a sunny day, there is a temporary decrease in the need to heat; on a clear night, heat is lost so quickly that more fuel is needed. Likewise, in warm, but cloudy weather, ventilation requirements may be reduced with high ambient temperatures. The examples show how these interactions can explain the energy demand, which is not the result of any single dimension, but the resulting combination of meteorological influences and control system reactions. It is crucial to understand these environmental factors in order to have an accurate prediction of energy consumption. To account for the variability on daily and seasonal scales, statistical and predictive models should consider the direct effects of climate as well as interactions between the effects. Implementing the high resolutions data from the weather stations and indoor climate sensor helps to improve the forecasting and enables adaptive energy optimization strategies. In this context, the outdoor-to-indoor climate interface represents the primary means of transfer of environmental variability to energy demand in the greenhouse.

2.3 Monitoring and Control Technologies in Greenhouses

The development of greenhouse production towards superior and data-driven production systems is directly related to the development of monitoring and control technologies. In modern greenhouses the environmental conditions are continuously measured, analyzed, and regulated by automatic control systems that can be considered as cyber-physical systems. These technologies form the essential part of the energy optimization strategies and by using them, the energy of the microclimate can be managed immediately and predictively (Van Straten et al., 2010; Shamshiri et al., 2018).

The basic level of monitoring systems are the sensor networks. Real-time measurements are taken at the various sensors for key environmental conditions including indoor and outdoor temperature, relative humidity, solar radiation, CO_2 concentration, soil moisture and substrate electrical conductivity. Further, sensors can measure wind speed, precipitation and energy consumption via smart meters. The development of wireless sensor networks has increased the flexibility and scalability of systems, lowered the installation costs, and enabled detailed spatial characterization of conditions inside of greenhouse compartments (Li et al., 2010; Ruiz-Garcia et al., 2009). Data collected at high time resolution (e.g. on an hourly or minute basis) are used for the statistical analysis and the development of prediction models (Morais et al., 2008).

Sensors are crucial for accurate energy management, and for proper calibration. Inaccuracies in measurement or calibration drift can result in incorrect control commands, which can cause unneeded heating, cooling or ventilation energy consumption. Data validation and filtering procedures, and anomaly detection techniques are therefore taken on board in the management systems (Wolfert et al., 2017). Modern platforms offer services like data storage in Cloud infrastructures and remote access, providing managers with the ability to intervene whenever needed and monitor the performance (Tzounis et al., 2017).

The second technology is climate control. They translate the sensor data to operational instructions for actuators like boilers, heat exchangers and ventilation windows, fans, shading system, irrigation pumps and artificial light. Simple threshold rules are used to define control logic, or advanced methods like model-based predictive control can be used which aim to optimise multiple variables concurrently (Van Straten et al., 2010; Ferentinos, 2018). For instance, if the temperature is below a set level, the heating is switched on, and if the temperature and/or humidity is higher, the ventilation will be increased accordingly.

One of those most important considerations has been model based predictive control which can consider system dynamics and minimise energy use while still observing agronomic constraints. With the inclusion of meteorological forecasts and crop growth models these can be adjusted in advance and prevent for example energy intensive corrective measures (Oldewurtel et al., 2012), which would be needed if heating or ventilation strategies would not be adjusted.

Monitoring and control functions are embedded in commercial greenhouse management platforms to be used in integrated decision support systems. One of the most common ones, for instance, is the Priva climate control system, prevalent across Europe. This system comprises sensor networks, central data processing and automated actuator control and is used to monitor and regulate temperature, humidity, carbon dioxide, irrigation and energy flows (Priva, 2022). It offers realtime dashboards and historical data analysis, as well as customisable control strategies based on the crop and greenhouse set-up. Energy meters are able to be interfaced, providing monitoring of fuel and electricity use to enable benchmarking and cost control.

Other companies, like Hoogendoorn and Ridder, propose similar solutions, increasingly using machine learning and artificial intelligence algorithms to aid decisions (Shamshiri et al., 2018; Wolfert et al., 2017). These adaptive algorithms can use past data to continuously adjust setpoints, based on crop response and system efficiency. Incorporation of monitoring/control technologies results in greater transparency in energy utilization. With continual monitoring, inefficient usage, abnormal consumption patterns and technical issues can be identified. In parallel, digital platforms offer orderly data sets, which are used in creating statistical and predictive models, for developing an efficient and sustainable greenhouse energy management system.

2.4 Statistical models for energy forecasting

Despite the emergence of more complex algorithmic approaches, statistical models continue to be a major benchmark for energy demand forecasting due to their simplicity, relative transparency in method and lower data demands. But for greenhouses, where energy use is related to climatic and operational parameters that can be measured, statistical methods can predict energy use, and also identify the key parameters. This is crucial for administrative decisions as Montgomery et al. (2012); Hyndman & Athanasopoulos (2021) states that a model that is well predicting and there is difficulty in understanding its meaning can be translated to practical energy management policy.

Typically, the correlation analysis is done first. Coefficients like Pearson capture the strength and direction of the relationship between energy consumption and the outside temperature, the solar radiation, the relative humidity or the wind speed. But its usefulness is as a diagnostic tool, and it is not enough to predict. It cannot be concluded from correlation that one is the cause of the other or that there are not other causal interactions or time lags. Correlation coefficients in greenhouse data may be used to oversimplify the meaning because multiple environmental factors are correlated in many greenhouses.

The linear and particularly multiple linear regression are the simplest statistical forecasting techniques, since they enable the joint evaluation of the influence of a number of independent variables on energy demand. One advantage of using them is that they have interpretable coefficients, statistical tests of significance and measures of fit including the R2. For energy demand forecasting, regression is proven to be useful if relationships are not very nonlinear and are relatively stable (Maaouane et al., 2021; Montgomery et al., 2012). It is not useful, however, in situations where multicollinearity, heteroscedasticity, or strong interaction effects between the variables are present, which is often the case in a controlled microclimate situation.

A Shift in focus from relationships to time dependencies (time series) e.g. ARIMA models, etc. They have the main advantage of using autocorrelation, seasonality, and trend, which are significant in greenhouse energy use because of the diurnal and seasonal variations. Time series can be especially useful for short-term forecasting when the behavior of the past is highly repetitive, as noted by Hyndman and Athanasopoulos (2021). Their weakness however is that they often do worse when there is a structural change, for instance change of the control strategy, equipment, or when it is very dry or very wet.

It thus deems that statistical models are not "outdated" with regard to machine learning. Instead, they provide an essential analytical foundation and a benchmark. They are mainly used for energy greenhouse systems due to their interpretability, lower computational burden and applicability on modest sample sizes in spite of their limitations in terms of nonlinearity and dynamic complexity of energy systems (Maaouane et al., 2021; Hyndman & Athanasopoulos, 2021).

2.5 Machine Learning Applications in Smart Agriculture

In the context of smart agriculture, machine learning has been playing a pivotal role in the process of revealing patterns from the vast, heterogeneous and high-frequency data sets collected from different sources such as sensors, weather stations, cameras, and automated control systems. In contrast to classical statistical models, the models of machine learning do not always have to be linearly or normally distributed and this is one of the reasons why it is so attractive in an agricultural system where many of the relationships between input and output variables are probably not linear and dynamic. But they're not always the best. Their utility relies on the amount of data available, how they are trained and, importantly, whether the increase in accuracy is worth the loss of interpretability (Kamilaris & Prenafeta-Boldú, 2018; Wolfert et al., 2017).

Random Forest is among the most popular methods used in smart agriculture as it is relatively accurate, noise robust, and has little risk of overfitting. It is a Tree Aggregation technique which is suitable for taking into account nonlinearities and interactions between the variables without the need of a priori theoretically accurate specification of their shape. It has also exhibited good predictive performance in energy-related applications in smart farms, particularly when all the three factors (environmental, operational and temporal) are supplied as inputs (Venkatesan et al., 2022). But it has a limit to its relative power. It gives some idea of the variance importance within the variables, but it does not give the pure parametric interpretation of regression; it also can be affected when a very fine temporal prediction is needed with strong serial dependence.

SVM and in particular SVR has been successfully used for the problem of stable prediction on complex but not necessarily large structures. The major benefit of their use is that they use kernels in their method, and thus can handle nonlinear relationships, and still have good generalization properties even on moderate sized samples. SVM has been utilized in classification problems in rural areas, such as predicting microclimate and analyzing sensory data, and proved to be very satisfactory in many cases (Kamilaris & Prenafeta-Boldú, 2018). However, they are not quite as easy to use as the end user, as the hyper-parameters must be carefully set, and sometimes are not the most suitable if the data are very large and/or multimodal. The most powerful but at the same time most demanding category is neural networks. Their usefulness is in learning complex multi-level representations, which is essential for problems like predicting temperature, humidity or energy loads in a dynamic greenhouse environment. In recent years, the research of greenhouse microclimate prediction has shown that architectures such as LSTM, Attention-LSTM and Autoformer can outperform the linear prediction models, mainly because they can model temporal dependence and nonlinear variation, which is difficult for simpler models (Li et al., 2024; Ahn et al., 2024). The problem is that as they become more predictive, they need more data, higher quality data, are more compute intensive and are less explainable.

It can be seen that there is no best model for all smart agriculture applications through comparative studies. However, the performance of machine learning is strongly problem dependent, as well as data dependent, and it is influenced by the operational context, as observed by Kamilaris and Prenafeta-Boldú (2018) among others. As such, in a greenhouse use, deep learning methods may outperform other types of models at making complex temporal predictions however this does not imply that they are necessarily the best model to apply for decision support. Interpretability and stability and operational integration are as important as error reduction, in administrative and energy contexts.

So, the conclusion is complex, a result of critical reading of the literature. Machine learning isn't the end of classical approaches, it's in addition to them. Random forest algorithms are frequently effective and applicable, SVM in medium scale and controlled problems, and neural network is more efficient in the presence of rich temporal data and high accuracy demand. The most developed approach in greenhouse energy use is not to take the most complicated model simply because of the lack of a critical evaluation of the various options in terms of accuracy, interpretability and operational utility (Venkatesan et al., 2022; Ahn et al., 2024).

2.6 Energy efficiency, operational optimization and saving strategies in greenhouses

Energy in the greenhouse cannot be just described in terms of consumption. What is important is the question of how this consumption can be reduced without compromising the yield of foods, product quality and stability of microclimate. Thus, energy efficiency in the greenhouse is not synonymous with the use of the minimum amount of energy, but with using the optimum balance of energy input, environmental control and production. International literature indicates that the effectiveness of these passive measures, more efficient equipment and advanced control strategy highly depends on climate, crop and energy cost structure (Ahamed et al., 2019; Trépanier et al., 2025).

First, to talk about "energy efficiency" in the greenhouse, it must be defined in an operational term and not just Accounting. While a greenhouse can have lower consumption, it could be less efficient if this is done by damaging the microclimate, by diminishing growth or by increasing the risk of diseases. Therefore, energy efficiency is better evaluated with respect to production, quality and maintaining desired limits of the environment. This logic is reflected both in the reviews of saving techniques and in the latest comparative study, which demonstrated that "the best" is not the one that would simply save the most kWh, but rather the one that would improve the overall energy and economic performance of the system (Ahamed et al., 2019 and Trépanier et al., 2025).

On the passive, the literature agrees, that the energy demand reduction should begin with the greenhouse as an edifice. This includes the use of thermal curtains, improved covering material, double shells, improved sealing, shading, correct orientation of installation and thermal mass. The rationale of such practices is easy to see: they help minimize heat losses during the winter or unwanted heat gains during the summer, which decreases the amount that mechanical systems need to address. Thermal screens have been proven to be one of the most effective interventions under review, and optimization studies, although dated, have demonstrated that interventions like double thermal curtains or double glazing can significantly impact the energy demand up to 60% on average, per measure under very favorable conditions (Vadiee & Martin, 2014; Trépanier et al., 2025). Numerous comparative studies have been conducted to assess the consumption savings that can be achieved by fitting and/or replacing thermal curtains, including the most recent ones by Trépanier et al. (2025), which indicate that consumption savings of around 17.7% to 26.5% are available, depending on the case and climate.

But there are also some drawbacks to passive solutions. Raising the thermal resistance of the shell can decrease losses, but can also decrease light transmission or, because of increased heat removal problems, can be a burden to summer operation. Similarly, shading will be beneficial under high-irradiance and/or hot weather but will become detrimental under high incidence by restricting the amount of photosynthetically active radiation (PAR) and have a negative impact on the crop. Thus, passive strategies are only useful when they are part of a logic of climate, and crop adaptation and not as general “recipes” for saving (Allali et al., 2025; Ahamed et al., 2019).

Active strategies level is the focus on equipment and operation not shell. This incorporates more effective heating systems, heat pumps, better ventilation and cooling systems, dynamic setpoint managing, more effective LED lighting and more precise irrigation. The significance of these measures is that they are acting directly on how the greenhouse responds to changes by external and internal conditions. The conversion from HP sodium to LED is one of the most significant technological shifts that has happened in recent years as it not only represents a decrease in energy use but also increased spectral control. The studies conducted by Katzin et al. (2021) and by Trépanier et al. (2025) demonstrate that the shift to LED can result in significant energy savings, and that if only a single measure is possible, LED lighting is often the most effective measure to reduce energy costs per unit of production.

But even more important is the use of advanced control strategies. Optimization in the greenhouse requires more than just the efficiency of each machine, it also requires the time machine is used, the method of use, and the extent to which machine is used. It can be seen from the literature that model-based predictive control of heating, ventilation, dehumidification and lighting can be achieved using the dynamic model, meteorological forecasting and optimization algorithms, which results in better synchronisation. This helps to minimize unnecessary, costly corrective actions and eliminate unnecessary overconsumption resulting from simple on/off logic or static setpoints. For operational optimization, recent studies demonstrate the ability of the data-driven and stochastic approaches based on model predictive control to enhance energy performance while enforcing agronomic constraints (Mahmood et al., 2023, Kim et al., 2025, Benzzine et al., 2025).

But the savings is not without its drawbacks and are not boundless, as outlined above. Possibly the most important aspect of serious literature reviewing is this one. To reduce the control temperature may reduce the amount of heating consumed, but may increase the humidity in the protected space, encourage fungus attack or reduce crop growth rates. Thus, high-intensity physical and/or mechanical cooling may decrease thermal stress, but also may add an electrical load or may cause other microclimate parameters to become unstable. Contrary to this is highlighted in recent literature about cooling strategies, the need for balance between plant thermal comfort, energy consumption and environmental performance, in particular in warm climates and in a context of climate uncertainty (Allali et al., 2025; Chen et al., 2025).

Thus, energy use in greenhouses should be considered as a multi-criteria and dynamic issue. The passive strategies act as a means for cutting down on the basic energy consumption, the active strategies act to optimize how energy is used in operation while the advanced control systems enable more accurate adaptation to changing conditions. But no one technology is enough alone. The combination of technological interventions, real-world operational data and research of the impacts not only on consumption, but also on production and quality and economic performance, leads to a very significant level of optimization. This is why the step from the description of greenhouses' energy demands to the research of intervention strategies is needed for any research related to the real energy optimization of greenhouses (Ahamed et al., 2019; Trépanier et al., 2025; Mahmood et al., 2023).

2.7 Economic dimension of energy consumption and support for administrative decisions

Energy consumption in greenhouse is not only an ecological issue but also a technical issue, and one of the most significant economic issues, which has a strong impact on competitiveness and sustainability of the activity. Energy is one of the key operating expenses of greenhouses, as mentioned in international literature, particularly in high-tech greenhouses and/or in cold climates where heating, cooling and artificial lighting are extensively utilized. This implies that tiny fluctuations in energy costs or in system efficiency can have a huge effect on profit, production stability and the producer's competitiveness on the market. The need for considering the energy strategy in relation to production is also emphasized in the latest literature as an energy-saving measure may not be the most economic if there are no energy savings and the performance of the process is decreased and/or high investment costs are incurred (Trépanier et al., 2025; Dimitropoulou et al., 2025). Hence, the economic performance of greenhouses of different technological levels has been studied by reference studies, not only related to a single technical parameter, but to the combination of energy use, water consumption and emission, and production structure (Hopwood et al., 2024).

On this perspective, the forecasting of energy consumption is materialized from manager's point of view as long as it is applied to energy planning and decision making processes realized on different scales. Firstly, it will allow the heating, ventilation and lighting systems to be better planned for operation, especially if their use is affected by factors such as solar radiation and external temperatures. Second, it can be used as a starting point for energy planning, to control the peak load requirement and energy unproductive costs. Thirdly, it can help maintain equipment because the potential of forecasting deviations from a normal pattern of energy use can give early warning signs of equipment wear or malfunctioning. Finally, it is a risk management instrument because it allows the analysis of a high demand situation ex ante and can be used to adjust the operational strategy before the problem occurs. In this literature evidence is provided that uncertainty in meteorological forecasts can cause problems, not only in the accuracy of the energy demand forecast, but also in the commercial and/or operational application of the forecast, for example in energy market management and load timing, in the greenhouse sector. Therefore, forecasting models are calculation models, as well as inculcators of operational preparation and decision support models (Payne et al., 2022; Golzar et al., 2021).

The importance of energy-related performance measurement and benchmarking is also particularly relevant as energy-related data can only be relevant to management if it can be compared and evaluated systematically. Some indication of the total annual or daily use of energy is helpful but is not sufficient to compare different size greenhouses or those with a different technology or even production orientation. Therefore, indicators like consumption per meter, per kg of product, per growing season or per degree day are used in the literature to generate a comparable evaluation framework. These indicators can allow the manager to detect any deviations, determine if the consumption is "normal" in comparison with other similar facilities and monitor how technical interventions or changes in strategies affect consumption over time. The significance of benchmarking has been emphasised in classic greenhouse energy efficiency studies and in more recent studies with techno-economic reference points for various levels of technology. Omid et al. (2011) showed that it is possible to compare energy inputs and performance in order to determine technical efficiency, and Hopwood et al. (2024) offer detailed benchmarks for energy, water, emissions and financial performance for low, medium and high technology greenhouses. Therefore, performance indicators are not a supporting supplement, but are a tool that is necessary for evaluating the performance of the farm and positioning it strategically (Omid et al., 2011; Hopwood et al., 2024).

Lastly, uncertainty needs to be recognized to support management decisions. Forecasting models are not intended to give absolute certainty, but are to give a probability, and there is always a range of variability when it comes to the weather, the energy market, the performance of the equipment, and the biology of the crop. However, the use of models must be prudent on the part of the management. The best forecast is one that is well used in the decision-making process in conditions of uncertainty; not necessarily the one with the smallest error. There is growing literature on energy systems that recognizes that more consideration for uncertainty in decision making improves economic resilience and leads to better outcomes than a single ‘certain’ forecast. This means, for the greenhouse context, the need for scenarios, confidence intervals, alternative operating strategies, and an interpretation of the results to complement the judgment, not replace it. The true value of energy models therefore lies not just in their computing capabilities, but in how they are incorporated within a management framework that marries data, technical knowledge, risk assessment and more (Payne et al., 2022; McDonald et al., 2024). In this aspect the economic aspects of energy use are not an add-on to the technical aspects of green house energy forecasting but rather an integral part of energy forecasting and a foundation of energy use in greenhouses.

2.8 Sustainability, environmental impacts and decarbonization of greenhouse production

Energy use in greenhouses is not only a quantity that affects the functioning of the greenhouse, but also has an impact on the environmental footprint of greenhouse production and the sustainability of the agri-food system. Because there is intensive use of energy for heating, cooling, ventilation and lighting, the environmental impact of a greenhouse is not only affected by the amount of energy that is used, but also by the quality of energy that is used. In the current literature, energy efficiency is more and more perceived as emissions; system's resilience and adaptation to decarbonisation policies. This dimension is necessary to ensure that the evaluation of energy performance is technically valuable but also environmentally complete (Soussi et al., 2025; Campana et al., 2025).

Firstly, the connection between energy and carbon footprint in greenhouses is very important. The environmental impact varies among the different generation technologies, here used in the same form (1kWh) which includes natural gas, LPG, biomass, heat pump generation or PVs (photovoltaic generation). Life cycle literature demonstrates that highly automated greenhouses have a high production capacity and stability, but also requires a higher use of energy, and thus a higher amount of CO_2 produced if the greenhouse energy is derived from fossil fuels. A recent LCA performed in Japan for a tomato greenhouse was able to estimate the emissions at 3.67kg CO_2 per 1kg tomato, of which 55.6% was from the production and combustion of LPG for heating. The same study indicated that the change from LPG to wood chips could achieve up to 49.1% less emissions, highlighting the importance of the energy mix in addition to the overall amount consumed (Taguchi et al., 2024). In addition, a recent study in Switzerland is in line with this and shows that fossil fuels used for heating have a high carbon footprint and that non-fossil fuel solutions like heat pumps, biogas, waste heat recovery, and biomass burners are important emission reduction measures (Solgi et al., 2025).

This is even more evident if an entire system (as opposed to any individual machine) is considered when evaluating energy efficiency. A Green House could be environment harmful if the energy saved or continued to be used is from high carbon emission, but it could be useful to reduce use of energy in total. Therefore, the idea of sustainable greenhouse production cannot be considered as synonymous with the logic of technical saving, but with changeover to a lower carbon energy system. It is the financial side that is most highlighted in the recent literature on "nearly zero-energy" greenhouses: improvement of the environment must be combined with the use of renewable energy and, importantly, a smart control and redesign of the greenhouse so that environmental improvement is not only logistical (Soussi et al., 2025).

Second, greenhouses must be part of the wider framework of the climate policy. The European Green Deal is a plan to drive the transition of the European economy to become modern, competitive and resource efficient, with the target of cutting emissions by at least 55% by 2030 and legally binding climate neutrality by 2050. Meanwhile, the European Commission emphasizes that more than three quarters of the European Union's greenhouse gas emissions are generated by energy production and consumption (European Commission, n.d.-a, n.d.-b) and thus the decarbonisation of the energy system is a prerequisite for reaching climate targets. For greenhouses, that doesn't just mean that energy management is a cost factor, it is also a question of transition towards sustainable agriculture, resilient supply chains and less carbon intensive production systems. This is especially significant, since the greenhouse is at the nexus of food, energy, water and climate adaptation.

Thirdly, an increasing number of greenhouse crops are decarbonized by using renewable energy sources and hybrid energy systems. PV, geothermal, biomass, heat pumps, waste heat recovery and thermal energy storage are primary choices offered by the literature. These solutions are not necessarily mutually exclusive. Instead, hybrid schemes are becoming the norm, where technologies are used in complementation to each other to provide thermal and electrical greenhouse requirements in a better and more environmentally friendly manner. Analyses conducted over the past two years highlight that the selection of the suitable renewable technology does not correspond to a one size fits all solution, but is dependent on climatic conditions, resource and type and form of energy load, and crop. Solar energy, in particular, is also very flexible and ubiquitous, and heat recovery and biomass energy can play a key role in regions with a significant thermal load instead of the electric one (Campana et al., 2025; Soussi et al., 2025).

Meanwhile, the environmental discussion shouldn't be restricted only to heating emissions. In more recent comparative studies the influence of the greenhouse and that of electricity, fuels and fertilizers as well as the interactions between these factors and efficiency also determine the overall environmental performance. A study on the evaluation of different greenhouse types for vegetable production revealed that there are significant differences not only in terms of energy consumption but also in the overall environmental impact indicating the need for a holistic approach to sustainability assessment, and not relying on a single indicator (Saadi et al., 2025).

But this is where the “green” technology stops. Technology is not a sufficient prerequisite for sustainability. Even if a greenhouse implements PV, heat pumps or biomass and still does not run in a truly sustainable way, if the investment is financially not feasible, the local climate is not suitable for the particular system, management is lacking or the cultivation practice is still too intensive in terms of energy consumption. From the literature, it is now obvious that consideration should be given to environmental improvement at the same time as economic viability and feasibility, as well as knowing what is available and what is compatible with producers' actual operating patterns. However, renewable sources do not necessarily remove all environmental impacts or be the most viable option in all spatial and production scenarios when emissions are significantly reduced (Taguchi et al., 2024; Campana et al., 2025).

Overall, the nexus of energy, sustainability and decarbonisation is now an integral part of the greenhouse literature. Energy efficiency is not only measured in terms of costs or technical functioning, but also emissions and climate policy and resilience of agri-food system. Thus, a study on energy consumption in greenhouses only has sense if considered as part of this wider environmental and policy context, in which consumption reduction, as well as energy mix change and realistic technological transition are considered as interdependent strategies (European Commission, n.d.-a, n.d.-b; Soussi et al., 2025).

2.9 Gaps in the literature and theoretical foundation of the present study

An extensive and relatively well-established literature exists in the field of greenhouse energy, as well as in respect of energy consumption patterns, climatic factors that drive energy demand, monitoring and control technologies, and capabilities of statistical and computational energy models. But this growing body of knowledge does not point towards the complete methodological and theoretical maturity of the research field. Conversely, a survey of the recent literature reveals that there is a lack of information on types of data applied to models, selection of variables, comparison of models, and, importantly, conversion of research results into practical energy management. This is where the theoretical underpinning for the current study becomes vital – simply creating the next forecasting app is not enough; a targeted intervention is needed that fills key research gaps in existing literature, as identified by both Soussi et al. (2025) and Payne et al. (2022).

The first, and critical, gap is methodological. Many of the sources of greenhouse information are derived from experimental or modelled or idealised scenarios which are not necessarily all that representative of the actual complexity of a greenhouse production system. This does not mean that these studies are not useful but it does constrain the potential for generalizing the results of these studies into actual decision making environments. In the most developed control and optimization tasks, the focus is more on the control algorithm or controller performance rather than on the validity of the input data per se. The recent literature on energy-efficient greenhouse control, though very sophisticated, still uses modeling frameworks with a high level of verifiable information, and the forecasting literature indicates that even the errors in forecasts of the weather can significantly impact the energy demand forecast and the resultant commercial or operational uncertainty (Kim & You, 2025; Payne et al., 2022). Hence, there is still gap for studies that would use daily data sets full, continuous and operationally organized which would provide realistic variability over the entire annual cycle.

A second area of gap is the variables in the models themselves. Literature is biased towards the use of temperature as the primary forecastable variable (or as the primary variable controlled). While the temperature is indeed an important factor in determining energy requirements, a one-sided focus on temperature can result in a simplified microclimatic and operational system. Greenhouses are not just thermal systems but multifactorial systems in which the relative humidities, the solar radiation and the internal temperature interact with the ventilation dynamics and the operational control logic. As stated in recent reviews on smart agriculture in greenhouses and new models for environment control (Chen et al., 2025; Li et al., 2025), such multivariate approaches combining both external and internal parameters are exactly required and missing in current models. It was therefore found that there is a gap in research for developing models using a coherent combination of the microclimatic and energy variables, instead of using limited single-factor schemes.

The third gap identified is in the comparative assessment of models. Much of the research that is relevant is presented with the implementation of some model as evidence of methodological advances, without the development of a suitable comparative logic. That is, one must not only provide a demonstration that a model “works”. It needs to be systematically compared with alternative approaches, not only in terms of accuracy, but also in terms of interpretability, stability and usefulness for the end user. It has been found in the machine learning literature that a machine learning algorithm's performance is problem-dependent, and data structure, as well as the forecasting time horizon, is also a crucial factor, in addition, more recent comparative studies have shown that different machine learning architectures are more suitable for different microclimatic parameters or time horizons (Kamilaris & Prenafeta-Boldú, 2018; Ahn et al., 2024). The actual value of a study, in this case, does not so much consist of merely using the “most complex” model but rather the systematic comparison of statistical and machine learning approaches, with a clear selection process based not only on the error but also the operational utilization of the results.

The fourth and most important gap is related to how the forecast is used in administration. Although the technical literature has made great advances in the prediction of loads, microclimatic parameters and energy behavior, very often it remains at the level of the technical performance of the model and does not translate its results at the level of greenhouse management. That is, the information on how the forecasts can be practically applied to energy planning, setpoint optimization, risk management, preventive maintenance and cost and sustainability decisions in energy operation is not yet clear. The literature about the errors in weather forecast and the uncertainty in energy forecast already demonstrated that it is not only a computational problem, but also an administrative problem, as the deviations in energy forecast will impact the operation and management strategy of the energy system (Payne et al., 2022). Thus there is still a gap between the interfacing of technical forecasting and actual decision making, and this is crucial in a study which is not purely academic in nature but is intended to help in administrative decision making.

In this background the present study is seen as a focused answer to the above gaps. First, it relies on daily observations, throughout one year (365 consecutive days) in order to account for seasonality, transient periods and the daily change in energy within one operational framework. Second, it does not involve only one variable or only technical ones but uses several external variables (temperature, solar radiation) and internal ones (greenhouse temperature, relative humidity) related to the daily energy requirements. Thirdly, it combines both the interpretative and predictive logics by comparing statistical and machine learning methods and aiming at accuracy but also interpretability. Fourthly, it connects modeling with administration use, with emphasis laid on how well the modeling results can enable energy planning and operational optimisation.

The theoretical value of the research is not the application of known techniques to a new data set but in an attempt to tie technical forecasting, multivariate interpretation and administrative utility together in a coherent greenhouse.

3 Methodology

3.1 Research approach and study design

This research was methodologically quantitative, empirical and non-experimental, with the objective of analyzing and predicting daily energy consumption in the greenhouse cultivation using real data obtained in the operation. Since the subject of the study is about measurable variables, certain observation units and the assessment of relationships between the dependent and independent variables, the choice of quantitative approach was considered appropriate. At the same time, the study has an applied approach, because it strives not only to explain how the greenhouse is behaving energetically, but to create a model that can be used to predict the behavior of the greenhouse and make management decisions. As stated in the literature review on smart agriculture, the added value of the sensor data and analytical models lies in connecting them to operational decisions, and to actual improvements in the efficiency of farming systems (Wolfert et al., 2017).

The research design is a retrospective and observational research because no experimental change in operating conditions of the greenhouse was done and the data is already recorded for one full year (1 calendar year). Since there is no experimental manipulation, the study is not intended to produce a cause and effect conclusion, but statistical investigation of relationships, the assessment of correlation, statistical modelling and prediction. This kind of design is said to be especially suitable when the objective is to study the behaviour of an actual production system rather than under laboratory conditions (Hyndman & Athanasopoulos, 2021).

The stages of the research were planned in sequence methodologically. The data set collection, control and organisation were initially done. This was followed by descriptive and exploratory analysis with the aim of understanding the dispersion, distribution and seasonality of the variables. Next, correlation analysis and regression models were used to interpret the relationship between energy consumption and environmental factors. Finally, predictive models based on a chronological forecasting logic were created, and a comparison was made regarding the predictive power of the various methodological approaches. This is in line with what is suggested by the literature on forecasting (e.g., Hyndman & Athanasopoulos, 2021; Montgomery et al., 2012) where interpretive and predictive phases should be performed together and explanation and prediction should not be confused.

3.2 Description of the data set and time coverage

This study data spans 2025 calendar year (i.e., January 1, 2025 to December 31, 2025). Therefore, the analysis used 365 daily observation – all days of the year – with no sampling or sub-periods and no selectivity. A full annual cycle is methodologically essential, as it permits a description of the seasonal trend as well as warm/cold season influences and also the transitional seasons which include changes in solar radiation and energy demand. For time series problems whose seasonal variation is an important mechanism, it is believed that full-year temporal coverage is required for reliable representation of the time series (Hyndman & Athanasopoulos, 2021).

The unit of analysis was daily. There are two reasons for this option. First, it allows temporal analysis in sufficient detail to reveal important shifts in energy use and to link its use to day-to-day environmental variability. Second, it reduces the excessive short-term (minute or hourly) noise which can impact the minute or hourly data but not necessarily have significant management implications. The daily unit is sufficient for research and operational activities when the target is to predict demand and provide decision support for energy planning, performance evaluation or deviation detection (Wolfert et al., 2017).

The data are collected during the actual operation of the greenhouse system, and include materials that are directly associated with the formation of energy consumption, such as important environmental and operating variables. It is appreciated that use of real field data will be very important in the study of greenhouse systems, since commercial production situations are more complex than ideal or experimental situations. Literature on greenhouse control and optimization indicates that management needs to be based on actual data obtained from operation of the greenhouse, as the influence of equipment, climate change and control setup occur simultaneously and interact with each other (van Straten et al., 2010).

3.3 Study variables

The outcome variable in this study is the daily energy use in "kw.Hour." This variable represents the daily energy demand of the greenhouse and is taken as an important variable for interpretation and prediction. It was selected as it is a direct indicator of the operational efficiency, economical cost and energy behaviour of the system. Typically, the dependent variable in the literature on energy consumption prediction is a continuous numerical value, so that regression, time series and machine learning algorithms could be applied (Hyndman & Athanasopoulos, 2021).

The independent variable are the external temperature, the internal temperature of greenhouse, the relative humidity and the solar radiation. The external temperate is indicative of the climatic situation of the outside environment and is most likely to impact on heating or cooling requirements. The internal temperature is a representation of the real condition of the greenhouse and is related to the culture and the interventions made. Relative humidity is an important microclimate factor because it influences ventilation, condensation and the conditions for plant growth. In the day time, solar radiation has a passive heating effect but can in some instances be correlated to increased ventilation or cooling requirement. The relevance of these variables for operation in greenhouses has been well studied in the case of sensor networks and in the case of microclimate control systems (Li et al., 2010; Ferentinos et al., 2017).

This was done at the daily level; that is, each row in the dataset represents a calendar day in 2025, and holds one value for each variable. This format enables the analysis of a cross-sectional time series with exogenous factors, that is, a chronological series affected by other variables in the same time period, as well as by the time structure of that series. Exogenous variables are important to add to a time series model as long as it is known that the behavior of the dependent variable relates to the observable external mechanism (as is the case with energy consumption related to climatic conditions) as it is in the theory of forecasting.

4 Research Analysis

This dataset describes a synthetic, but internally consistent, daily greenhouse operation for the entire year 2025. It contains 365 observations, one per day of the year, and 5 variables: outdoor temperature, indoor temperature in the greenhouses, relative humidity, solar radiation and energy consumption. However, an advantage of this data set is that it has a full annual cycle, so it is possible to assess the seasonality, operating responses, and energy load changes, without a time gap. Concurrently, its synthetic nature means that the results of the analysis can be understood as an analysis of a realistic operating scenario and should not be generalised to any actual greenhouse.

4.1 Initial Assessment of the Greenhouse

The first data gathering provides a good and valuable picture of the overall greenhouse operation for 2025. The descriptive statistics shows that the mean for the outdoor temperature is 18.8°C and the mean for the indoor temperature is 23.1°C. This is not a numerical discovery but is significant. Conversely, it reveals that greenhouse is not a passive building that responds to the surrounding climate but is an actively-control microclimate system. This shows that thermal control interventions are applied (most often during the colder seasons) with the goal of maintaining conditions suitable for the growth of the crop, as the average temperature is always higher inside the structure as compared to the outside temperature. Meanwhile, relative humidity is at 78.8°, which is fully compatible with the greenhouse environment. Greenhouses have several features unique to other building or industrial environments: Their humidity levels are generally higher than those found in other types of structures because of the transpiration of the plants; They have irrigation systems and generally provide a microclimate just as the crop needs to function normally. Thus this value cannot be used as an indicator of malfunction, but rather as an element that verifies that the dataset represents the realistic greenhouse context. Solar energy contribution is measured by the average solar radiation which is 289.0 W/m2, at the same time, sufficient for year-round use, but not avoiding the need for energy topping. The average daily energy consumption is 20.2 kWh, further emphasizing the fact that energy is being used systematically to regulate and stabilize the inside environment.

What is more telling are the values at each of the extremes. The external temperature shows a range from 33.2°C to 4.1°C which indicates that the greenhouse climatic environment has a significant seasonal variation and is suited to the Mediterranean climatic conditions. These different outside temperatures are significant in particular because they present a challenging condition on which the internal control of the greenhouse has to act. The internal temperature on the other hand (from 18.0°C to 27.3°C, see above) is in much narrower limits. This is probably one of the most significant features of the whole descriptive analysis in that it demonstrates that what happens inside the greenhouse is a significant smoothing of the thermal variations outside the greenhouse. The internal microclimate therefore, is maintained within functionally and biologically acceptable limits, while the environment can vary greatly. This constitutes a clear indication of effective control, but at the same time it also indicates a significant energy burden to maintain this stability.

The energy usage is also fairly diverse, between 12.2 and 38.3 kWh per day. The range is broad enough to illustrate that energy is not just for a 'base load' function of the system, but is applied as necessary to meet seasonal and microclimatic needs. This means that the GH is run differently based on outside conditions; more energy is used when it's not a good outside environment, and less when it is a good outside environment.

Table 1 Descriptive statistics of key greenhouse variables

Variable	**Mean Value**	**Minimum Value**	**Maximum Value**	**Observation**
External temperature (°C)	18.8	4.1	33.2	Marked seasonal variation
Internal temperature (°C)	23.1	18.0	27.3	Clearly more stable than the external temperature
Relative humidity (%)	78.8	—	—	High, consistent with a greenhouse environment
Solar radiation (W/m²)	289.0	—	—	Strong seasonality
Energy consumption (kWh)	20.2	12.2	38.3	Variable load, with winter and transitional periods predominating

The table indicates that the greenhouse is a controlled microclimate system. The fact that the external temperature does not vary so much as the internal temperature suggests that there is thermal regulation.

4.2 Seasonality and annual cycle

The dataset reveals a clear and strong seasonal structure, thus signifying that the energy behaviour in the greenhouse can be adequately related to the energy changes in the external environment which in turn can be related to its yearly cycle. The annual change in external temperature is normal, lower in the winter and early spring, higher in the summer and early autumn. This image suggests a Mediterranean type climate pattern with relatively mild, but cold enough to impact upon heating needs, winters and warm, high solar loading summers. Of particular analytical interest is that the maximum external temperature does not happen, as it should have in a completely typical seasonal cycle, in July or August, but at the end of September. This is not an inherent reason to consider this a "magic number" violation. Rather, it reveals that the synthetic dataset is enriched with realistic diurnal variability, enabling warm episodes to be included that are not necessarily part of the tight definition of summer. This is important because it allows this data set to become more “real” in that the data is not necessarily symmetrically distributed through the seasons in “ideal” fashion with extremes at high and low temperatures. Seasonality is even more evident when considering energy use on a monthly average basis. The peak values are reported in March (28.8 kWh), April (27.7 kWh) and February (26.1 kWh). On the other hand, July and August are the rainiest months of the year with a minimum of 14.8 and 15.0 kWh. This distribution is especially telling: It indicates that the 'greenhouse energy load' does not always occur in the coldest calendar week of winter, but moves to the late winter/early spring seasons. This is of fundamental importance because it indicates that the energy performance of the greenhouse is not only a function of absolute temperature but also of the dynamic relationship between temperature, radiation, humidity and microclimate control parameters needed.

The following picture has more than one interpretation. To explain the first point, the need for heating is still very great during February - April as the temperature outside is still relatively low with high demands during night and early morning hours. Meanwhile, the amount of solar radiation slowly starts to rise relative to deep winter, leading to higher diurnal instabilities. It is not enough that the greenhouse is constantly heated, but must adjust to varying conditions that occur in the same day. This means that energy use will not only vary according to the lows in the environment, but also with the frequency and intensity which the microclimate control needs to be activated to maintain the production limits.

Second, there is a greater thermal instability in spring than in winter. The greenhouse must moderate the fact that nights are cold, and must control the more rapid fluctuation in night lows and daytime sun and higher temperatures. This change imposes greater load on the control system, the microclimate should not fluctuate a great deal, but should stay as close to the acceptable range as possible. That is, the energy required in the spring is not just for heating but also for compensation. This is why the maximum average consumptions are not only in the center of winter.

Third, the synthetic model appears to have been assembled with the operational realism operation in mind, that is, the demand for microclimate control is higher during periods of increased demand. Transitional seasons in practice, are not as straightforward as they sound, since the system is not of a clean nature (heating only, or ventilation only) but is more a complex one with more frequent changes needed. So the higher use during the spring can be seen as a sign of not shortage but rather of enhanced use through an intensive operational activity of maintaining a stable and productive microclimate.

Conversely, the incoming energy use in the summer months is significantly reduced although solar radiation and outside temperatures are high. This indicates that the energy load of the summer season in this dataset is lower than that of the winter and transitional seasons, implying that the cooling, ventilation and air circulation load is not higher than the heating and microclimate stabilization load for the colder seasons. This observation is relevant because it is consistent with this greenhouse being operating primarily as a strong winter/transient energy dependent system. On the whole, the seasonality of the data set plays not only a descriptive role but also an essential part in the energy functioning of the greenhouse. The consumption also varies with time but does not follow the calendar but rather it is the effect of the combination between temperature, radiation and control dynamics as seen in the annual cycle. The latter is crucial for this observation, as well for the later phases of the analysis, to underline the importance of using approaches that consider the seasonality, chronological sequence and functional complexity of the system.

Table 2 Indicative monthly energy consumption averages

Month	**Average Energy Consumption (kWh)**	**Interpretation**
February	26.1	High load due to the continued need for heating
March	28.8	Highest monthly consumption, indicating a demanding transitional period
April	27.7	Sustained high load due to daily thermal instability
July	14.8	Very low consumption, reflecting thermally favourable outdoor conditions
August	15.0	Similarly low energy demand to July

The energy load is not maximum in winter, but predominantly in the late winter and early spring. This demonstrates the seasonality of this region to be highly challenging for microclimate control.

4.3 Relationship between outdoor temperature and energy consumption

A most significant relationship that can be drawn from the data set is the strong negative correlation between the outdoor temperature and daily energy consumption. The correlation co-efficient, ~ -0.86 suggests that the variation of the outdoor temperature with respect to the base temperature is closely related to the variation of greenhouse energy load, and is also highly negative in value. The energy used decreases as the outdoor temperature rises. This relationship isn't a secondary statistical finding, it's the central operating mechanism of this system. In fact it is an expression of the cost of heating and regulating a constant internal microclimate for most of the energy used for most of the coldest months of the year of a greenhouse. The significance of the discovery is straightforward and profound. If the external temperature is low, it uses more energy to reduce energy losses from the greenhouse and keep the internal temperature at an optimum level for growing the crops. Conversely, the external environment is warmed up and closer to the desired thermal operating limits, the cooling requirements lessen dramatically. Therefore, the outside temperature is the principal exogenous factor of energy demand. In other words, the energy load in the greenhouse does not simply occur at random, but rather in a predictable fashion as a function of the magnitude of the difference between the natural external conditions and the desired microclimate.

This relationship is also manifested by the very extreme daily observations. On April 10th, the external temperature is just 4.5°C and the peak energy consumption is set at 38.3 kWh: very high energy consumption values are found on March 13th, March 27th and February 21st, which are days of very low external temperature. The above special cases clearly indicate that as the outside environment becomes colder, the greenhouse must expend a lot more energy to maintain the internal environment. Conversely, the lowest consumption rate, at 12.2 kWh, is on July 6 with an ambient temp of 24.5°C. It is a very illustrative day, because it illustrates that at the climate conditions of the outside air very close to the thermal requirements of the crop, the greenhouse is at its lowest energy level, covering primarily basic or auxiliary loads and not intense heating loads.

This relationship has also important managerial implications. Outdoor temperature is the most important variable from an operational point of view to predict the daily energy consumption. A structureless predictive model that would try to predict consumption without a component reflecting outdoor temperature would be incomplete both in interpretation and in structure. Even if it included additional factors, such as solar radiation, humidity or indoor temperature, it would be lacking the one factor which appears to have a major influence on the load variation. This implies, that in practice, daily weather forecasts, and especially forecasts for outdoor temperature can be regarded as an important instrument for energy planning, cost forecasting and management of heating equipment. In general, the most evident and relevant outcome of the analysis is the strong negative correlation between the outdoor temperature and energy use. It demonstrates that the greenhouse is energy sensitive to changes in external climate, and that a relatively good consistency of description and prediction of the greenhouse behavior is possible if the heat load of the outdoor environment is considered. This discovery is the foundation to every later analysis, be it for statistical modelling or for building more comprehensive prediction instruments.

4.4 Internal temperature and operational stability

The correlation coefficient of the internal temperature of the greenhouse with the energy used in the day is also about -0.76, which is a negative correlation. That might not at first glance be as self-evident as the connection between those outside temperatures and consumption, since one might think that the higher the temperature inside, the more energy would be needed to get it there. The meaning of internal temperature in this data set is more complicated, however. This variable is not only the ultimate greenhouse control goal, but it is also an indirect measure of the overall seasonal/operational condition of the greenhouse. In the warmer external conditions and external environment, the higher temperature can be kept in the greenhouse, and the energy burden is obviously reduced. Hence, the rise of internal temperature does not always mean higher consumption, but instead, a time that is more advantageous to the system from the external thermal environment.

This will mean that the internal temperature should NOT be looked at by itself. Most factors that influence the final value of the internal temperature in a fully controlled greenhouse are external weather, solar radiation, thermal behavior of shell, ventilation strategy and, naturally, the application of energy for heating or cooling. Thus, although it has a negative correlation with consumption, this does not render energy unimportant, but does indicate that the internal temperature is a reflection of better environmental conditions. For instance, during the summer months the greenhouse air temperature level may be set higher than in the winter months, with expending less energy for air temperature maintenance, given the climatic condition by itself that enables the desired microclimate to be maintained. In a completely opposite way, in the colder months, despite a considerable increase in consumption, the internal temperature can still be relatively moderate, since the system is not able to compensate the thermal losses.

Perhaps even more significant is that the internal temperature ranges in a much smaller range than the external range. In detail, the temperature range is 18.0°C - 27.3°C with an external range of 4.1°C - 33.2°C, that is, the external temperature range exceeds 29 degrees. This disparity is highly indicative of the functioning of the greenhouse. It demonstrates that the system cannot simply let the internal environment fluctuate along with the strong external fluctuation, but needs systematic intervention to control the thermal uncertainty and keep the microclimate within the operation limits. This smoothing ability is actually one of the most significant functions that the greenhouse serves as a controlled agricultural system.

This particular result may be expressed in a more substantial manner: Energy use "purchases" thermal stability. That is, one does not expend energy just to raise the temperature, but rather to keep the internal environment from fluctuating as a result of the extremes of the outside climate. More important for the greenhouse production, as the stability of the microclimate is directly connected with the photosynthetic activity and with the normal growth of plants, disease prevention, maintaining the quality of the end product. That is why, energy use should be measured, not only quantitatively, in a quantifiable number of kilowatt hours, but also functionally, as a means of securing a stable, productively suitable environment.

The internal temperature is thus a very convenient variable from an analytical and management point of view. On the other hand, it is an indicator for evaluating the effectiveness of the greenhouse control. Conversely, it supports the interpretation of the energy load as it indicates the degree of conversion of energy consumption to thermal stability. Based on this, a greenhouse can have a fairly stable internal temperature despite outside temperature fluctuations and can still be considered to be operationally efficient, even if it consumes more during a particular period. The analysis of the internal temperature has shown that energy management in the greenhouse is not only to minimize the load, but also to be able to keep the thermal balance condition necessary for the sustainable operation of the greenhouse.

Table 3 Basic correlations of energy consumption with environmental variables

Relationship	Correlation Coefficient (r)	Direction of Relationship	Interpretation
External temperature – Energy consumption	-0.86	Strong negative	As the external temperature increases, energy consumption decreases
Internal temperature – Energy consumption	-0.76	Strong negative	Higher internal temperature often coincides with more favourable external conditions
Relative humidity – Energy consumption	0.18	Weak positive	Possible indirect association through microclimate management
Solar radiation – Energy consumption	-0.10	Weak negative	Limited direct linear relationship, although an indirect effect may be possible

The most important variable used in predicting energy consumption is the outside temperature. Solar radiation and humidity seem to act more as complementary or modifying factors.

4.5 Humidity and microclimate control

There is a weaker, but not insignificant, positive correlation between energy consumption and relative humidity of about 0.18. This is because slightly more consumption is seen during wetter days, although this correlation is not as strong as temperature. Caution should be used in interpreting this. Humidity isn't something that works on its own, but rather as part of a microclimate. More humid air is related to less solar radiation, cooler outside air temperature, demand for increased air ventilation and dehumidification, and overall more challenging conditions for control.

The importance here isn't just the actual strength of the correlation per se, but that humidity provides new information. Energy management is not just about temperature in a greenhouse. Even if humidity is a minor component of all the variation in consumption, if the system must ventilate or control the humidity to protect the crop, then it is operationally important.

4.6 Solar radiation and passive energy support

There is a negative relationship between consumption and solar radiation, but this is weaker than that between consumption and temperature at around -0.10. At the first sight this indicates that the greater the radiation the lower the consumption, which is not unreasonable, given that solar energy is passive thermal support. The simple correlation, though, is not very strong, indicating that radiation does not contribute to the energy load in a simple linear manner.

The justification is that the role of the solar radiation is two fold. During the winter it can help lower the demand for heating. However, it's not the case in summer when high radiation levels are associated with elevated ventilation or cooling requirements. Therefore, its impact is not clear. It is this duality that appears to be embodied in the dataset. For instance, in summer, on one day, there can be very high radiation but very low consumption due to the temperature regime that the outdoor temperatures provide. In spring or fall, however, radiation may be a significant compensating factor. Radiation is therefore likely to gain significance in multivariate analysis as compared to simple correlation.

4.7 Multifactorial interpretation of consumption

Daily energy consumption as the dependent variable, and the four environmental variables as independent variables, the multiple linear regression model accounts for about 75% of the total variation of energy consumption. This is a very high proportion for daily energy data and indicates the dataset is well structured and consistent. It is the outdoor temperature that turns out to be the most important explanatory variable. Statistically significant additional information is provided by solar radiation. Humidity seems to play a role which is lower but still significant, while the indoor temperature becomes less statistically significant when added with the others.

There is by no means a lack of significance for indoor temperature. It implies to share information with outdoor temperature and with the overall seasonal dynamics. That is, the information is duplicated. It is methodologically significant since it illustrates that in a predictive mode multicollinearity has to be controlled and the importance of individual factors has to be avoided. The central idea for the administration is that consumption is primarily dependent on the external thermal load as well as the secondary impacts of microclimatic load and the presence of the sun.

Table 4 Extreme daily energy consumption observations

Date	External Temperature (°C)	Internal Temperature (°C)	Solar Radiation (W/m^2)	Energy Consumption (kWh)	Comment
10/04/2025	4.5	20.5	402.1	38.3	Maximum consumption, with very low external temperature
13/03/2025	4.1	22.3	252.5	35.3	Very high load on a cold day
27/03/2025	6.3	20.8	334.0	34.7	Increased consumption during a transitional period
21/02/2025	7.2	20.2	280.8	34.4	High consumption, reflecting a substantial

					heating requirement
06/07/2025	24.5	25.2	472.3	12.2	Minimum consumption, under thermally favourable conditions

The extremes values indicate that the greenhouse is being overloaded with energy when the outside temperature is very different from the desired inside temperatures. On the other hand, if the temperature of the external environment is close to the set thermal requirement of the crop, the consumption is almost restricted to the base load.

4.8 Greenhouse operational reading

These data show the greenhouse to be a three-energy regime system. The first is the cold regime (January-April) when consumption is high and the primary reason is heating and internal heat control. The other is the summer regime, which extends from June through September, in which consumption drops near to the base load and energy is used more in the process of ventilation, circulation, maybe a little cooling and general infrastructure. The third is the transitional autumn regime, in which consumption increases somewhat, but not yet to spring/winter levels.

This means there are real-world implications. The most promising period for energy savings involves the transition seasons, as well as the cold half of the year, when usage is higher. Here, it should be logical to test other scenarios, as for example better insulation, dynamic setpoint control, thermal curtains, using a heat retention strategy at night, or more accurate daily load forecasting using weather forecast information.

4.9 Application of multiple linear regression and predictive evaluation

To integrate theoretical considerations on statistical models with the practical aspects of the work, a multiple linear regression model was used with the daily energy consumption of the greenhouse as the dependent variable. The outside temperature, the inside temperature, the relative humidity and the solar radiation were taken as independent variables. The variables were selected based on the bibliographic basis and the results obtained from the descriptive and correlative analysis, in which the external temperature is the most basic variable that influences the energy load, and the humidity and solar radiation work as complementary.

The model had the following general form:

$$Energy_i = \beta_0 + \beta_1 OutsideTemp_i + \beta_2 InsideTemp_i + \beta_3 Humidity_i + \beta_4 SolarRadiation_i + \varepsilon_i$$

Where $Energy_i$ is the daily energy consumption in kWh,

$OutsideTemp_i$ is the external temperature,

$InsideTemp_i$ is the internal greenhouse temperature,

$Humidity_i$ is the relative humidity,

$SolarRadiation_i$ is the solar radiation, and

ε_i is the error term.

$$Energy_i = 45.2456 - 0.6472(OutsideTemp_i) - 0.1923(InsideTemp_i) - 0.0886(Humidity_i) - 0.0052(SolarRadiation_i) + \varepsilon_i$$

The equation indicates that, for a given set of variables, if the outside temperature rises 1°C, the daily energy use decreases about 0.65kWh. The most stable and statistically significant factor affecting energy consumption is the external temperature. Solar radiation also negatively and statistically significantly affects, which shows its passive support as a solar energy source. The relative humidity also exhibits a statistically significant, but less pronounced, effect. In contrast, internal temperature does not play a significant role in the multivariable model, likely because it is co-linear with external temperature and the overall seasonal operation of the GH.

Table 5 Multiple Linear Regression Results for Daily Energy Consumption

Variable	Coefficient B	Standard Error	t	p-value	Interpretation
Constant	45.2456	4.8148	9.3973	0.0000	Baseline level of the model when all independent variables are equal to zero.
External temperature (°C)	-0.6472	0.0417	-15.5237	0.0000	An increase in external temperature significantly reduces energy consumption.
Internal temperature (°C)	-0.1923	0.1808	-1.0638	0.2881	The effect is not statistically significant in the multivariable model.
Relative humidity (%)	-0.0886	0.0372	-2.3794	0.0179	An increase in humidity is associated with a small but statistically significant reduction in energy consumption in the model.
Solar radiation (W/m²)	-0.0052	0.0011	-4.7692	0.0000	Increased solar radiation slightly reduces energy demand.

The data has showed that the model has high explanatory power, the value of R^2 is 0.7501. This indicates that four environmental data considered in the model account for about 75% of the variation in the daily energy consumption. This is considered a very good result for daily energy-consumption information, because, in addition to measurable climatic parameters, demand is also affected by operational decisions, technical features of the equipment, and possible random day-to-day fluctuations.

Table 6 Model Fit Summary

Indicator	Value
Number of observations	365
R^2	0.7501
Adjusted R^2	0.7473
Mean Absolute Error (MAE)	2.4058
Root Mean Squared Error (RMSE)	2.9450
Mean actual energy consumption (kWh)	20.1723
Minimum energy consumption (kWh)	12.2000
Maximum energy consumption (kWh)	38.3000

The adjusted R^2 value of 0.7473 remains very close to the original R^2, indicating that the independent variables provide substantial explanatory information rather than being included merely to artificially improve model fit. In addition, the MAE indicates that the average absolute deviation between predicted and actual values is approximately 2.41 kWh, while the RMSE is approximately 2.95 kWh. These values indicate satisfactory predictive performance, considering that daily energy consumption ranges from 12.2 to 38.3 kWh.

4.10 Checking regression assumptions

Basic diagnostics were done to determine if the basic assumption of multiple linear regression was satisfied before the final interpretation of the model. In particular, linearity of the relationships, independence of the errors, normality of the residuals, multicollinearity and the presence of extreme or highly influential observations were investigated.

Table 7 Tests of Key Assumptions of the Regression Model

Assumption	Test	Value / Finding	Criterion	Conclusion
Linearity	Scatterplots and residual plots	Visual inspection	Approximately linear relationship and random distribution of residuals	Acceptable, with reservations
Independence of errors	Durbin–Watson	1.321	Value close to 2	Partial positive autocorrelation
Normality of residuals	Shapiro–Wilk	p = 0.0260	p > 0.05	Minor deviation from normality
Normality of residuals	Jarque–Bera	p = 0.0427	p > 0.05	Minor deviation from normality
Multicollinearity	VIF	Maximum VIF = 4.091	VIF < 5	Acceptable
Multicollinearity	Tolerance	Minimum tolerance = 0.244	Tolerance > 0.2	Acceptable
Outlying observations	Standardized residuals	Maximum |standardized residual| = 3.112	Ideally < 3	One borderline observation

Influential observations	Cook's Distance	Maximum = 0.032	< 1	No seriously influential observations
Homoscedasticity	Breusch–Pagan	p = 0.2164	p > 0.05	No strong indication of heteroscedasticity

Analysis of the negatives of the scatterplot and the residual plot revealed that there is a definite negative and approximately linear relationship between the external temperature and the energy used. The rest of the variables showed less correlation as yield, humidity and solar radiation are more supplementary or modifying factors. There was no evidence of systematic bias in the residual plot but given that the data showed seasonality the full temporal structure of the energy consumption data has not been fully captured by the linear model.

The Durbin–Watson statistic was used to test the independence of the errors. The value of 1.321 is less than the desirable 2, which means that there is positive correlation among the residuals. This is not a fault of the model but it does indicate that since the data are a daily time series, there are likely to be some similarities in the energy consumption patterns from day-to-day. Thus, the multiple linear regression model is still useful as a first approximation to an explanatory and predictive tool in the analysis, but in the future the use of time-series models or the introduction of temporal variables as lags in the models could be further explored.

The normality of residuals has been analyzed by a histogram, QQ plot and Shapiro wilk and Jarque bera statistical tests. The p-values for both tests were < 0.05 which means that the residuals do not exactly fit a normal distribution. The deviation, however, does not seem to be significant since there is not too much skewness. Thus, the normality assumption can be considered as partially met and the results should be interpreted carefully.

The multicollinearity was tested using VIF and tolerance statistics. All VIF values were less than 5 and all tolerance values were greater than 0.2, thus showing that there was no serious multicollinearity issue. The highest VIF values were found for external and internal temperature and was expected since the two variables are functionally and seasonally related. However, the values are still within acceptable ranges.

Table 8 Multicollinearity Assessment Using VIF and Tolerance

Variable	VIF	Tolerance	Conclusion
External temperature (°C)	4.062	0.246	Acceptable
Internal temperature (°C)	4.091	0.244	Acceptable
Relative humidity (%)	1.240	0.806	Acceptable
Solar radiation (W/m²)	1.226	0.816	Acceptable

Finally, standardized residuals and Cook's Distance were used to look for the occurrence of outlying and highly influential observations. There was one observation with a standardized residual slightly above the limits of ±3, suggestive of a possible extreme daily deviation. The maximum Cook's distance, however, was 0.032 which is far less than the critical value of 1. Thus, no observation had a significant impact on the overall estimation of the model. This meant that no observations were deemed necessary to exclude from the data set.

4.11 Monthly and Seasonal Analysis

Table 9 Monthly Mean Daily Energy Consumption

Month	**Mean Energy Consumption (kWh)**
January	24.04
February	26.11
March	28.81
April	27.69
May	23.15
June	17.04
July	14.77
August	15.03
September	15.88
October	16.01
November	16.35
December	17.66

There is definite seasonality in the energy use that appears in the table. The maximum mean was recorded during the late winter and early spring (March and April). The lowest values on the other hand were recorded in July and August. This finding supports the idea that, on the whole, there are more energy requirements in the transitional and colder seasons, when better microclimatic control is needed in the greenhouse.

Table 10 Seasonal Distribution of Daily Energy Consumption

Season	Mean (kWh)	Median (kWh)	Minimum	Maximum	Standard Deviation
Winter	22.48	20.40	14.50	34.40	5.07
Spring	26.54	27.35	15.20	38.30	4.80
Summer	15.60	15.10	12.20	22.90	2.11
Autumn	16.08	16.00	12.60	19.60	1.53

Results from the seasonal analysis showed that spring and winter were the two seasons with the greatest mean energy use. This indicates that there are greater needs for thermal stabilisation in transitional spring conditions. Lower and more stable values were obtained in the summer and autumn, which reflected less energy burden.

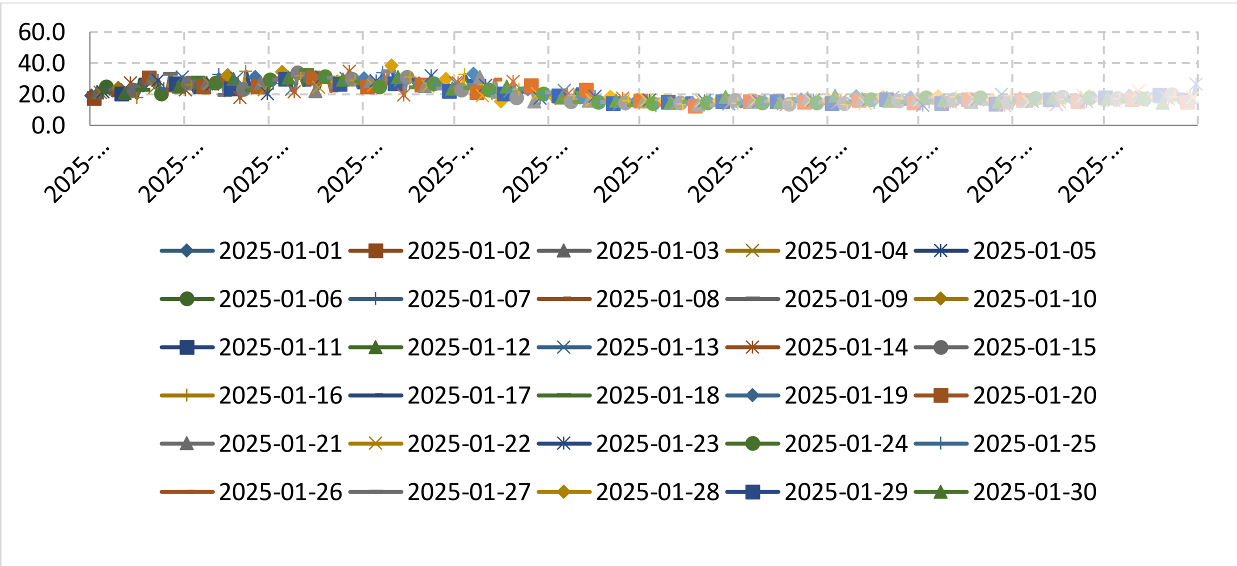
Table 11 Correlations with Energy Consumption

Variable	Correlation with Energy Consumption
External temperature	-0.855
Internal temperature	-0.762
Relative humidity	0.178
Solar radiation	-0.103

The energy consumption was most negatively correlated to external temperature. This translates to less demand for energy support with a rise in outside temperature. Other strong negative relationships were with internal temperature, but interpretation is more complex as it is dependent on the operation of the greenhouse control system. Humidity and solar radiation had lower correlations indicating that these are more secondary.

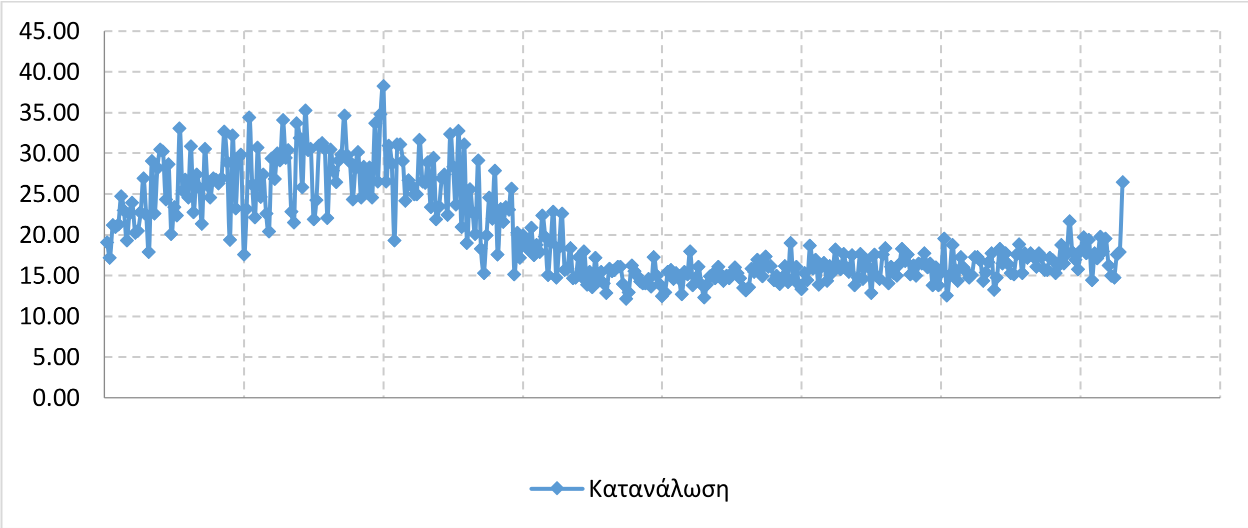
The next figures show the time trends for energy demand, as well as the relationship between the variables and the fit of the regression model. These figures supplement and augment the tables, and can be used to obtain direct visual perception of seasonality, the negative correlation between external temperature and energy consumption, and the overall accuracy of the model.

Figure 1 Daily variation in greenhouse energy consumption during 2025.



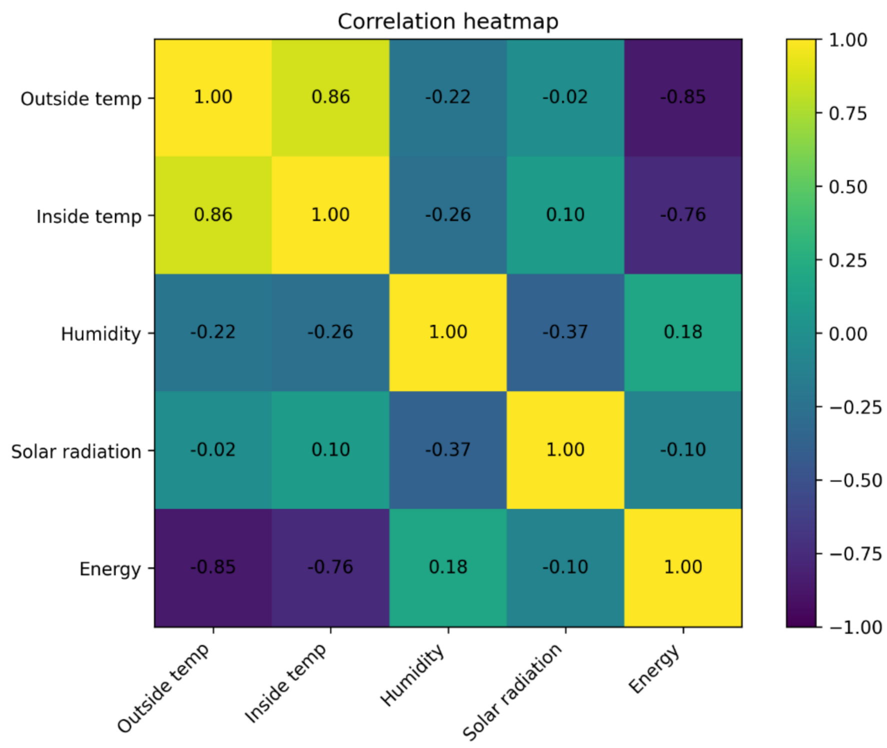
The figure represents the amount of energy used in kWh throughout the calendar year. There is a clear seasonality with higher values during late winter and spring period and lower values during summer.

Figure 2 Relationship between outdoor temperature and daily energy consumption.



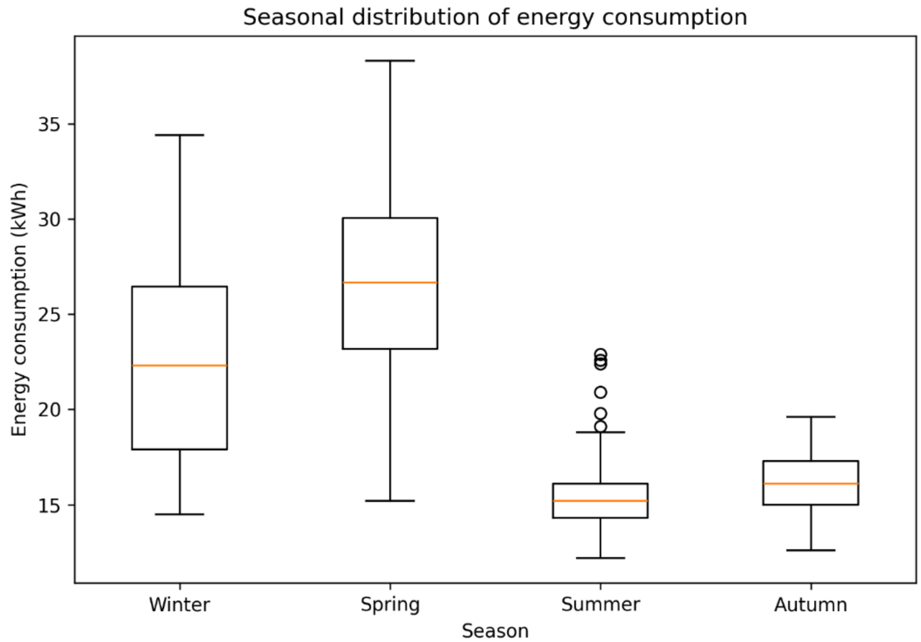
The scatter graph indicates a high negative correlation between outside temperature and energy use. The daily consumption shows a reduction with increasing outdoor temperature, which validates the thermal load dominant.

Figure 3 Seasonal variation of daily energy consumption.



It can be seen from the box plot that the consumption is greater and more variable in spring and winter, and less variable and smaller in summer and autumn.

Figure 4 Actual and predicted energy consumption values.



The chart shows the actual consumption and the consumptions predicted by the regression. Two lines are relatively close to each other, showing that the model has a good predictive performance.

Figure 5 Average monthly daily energy consumption of the greenhouse during 2025

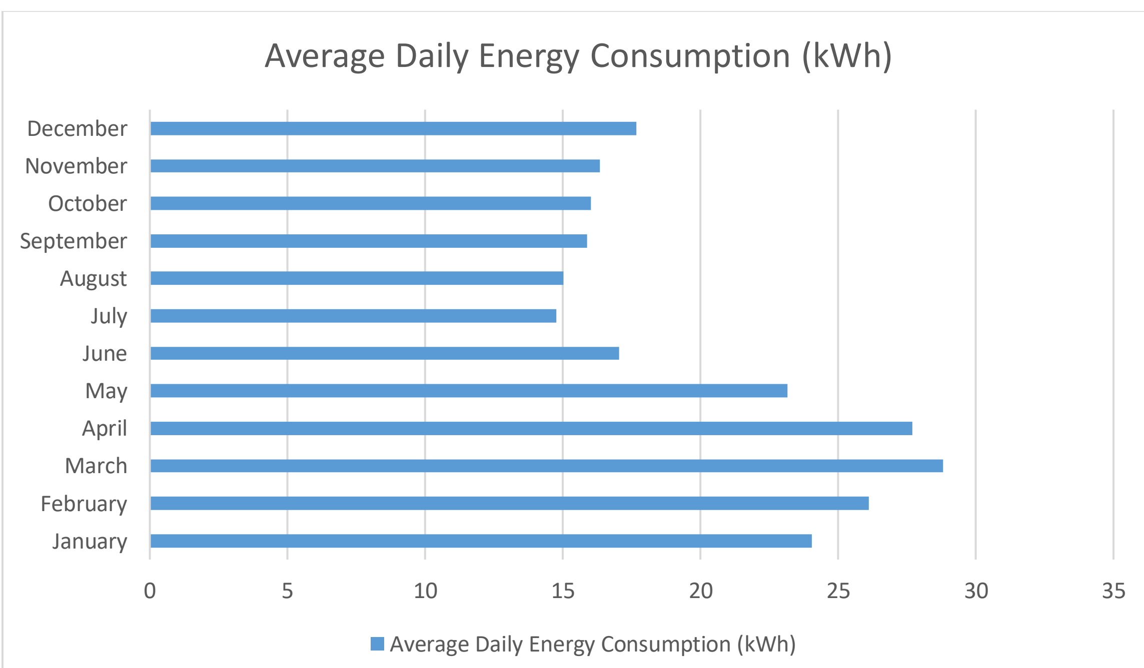


Figure 6 Actual and Predicted Daily Energy Consumption of the Greenhouse during 2025

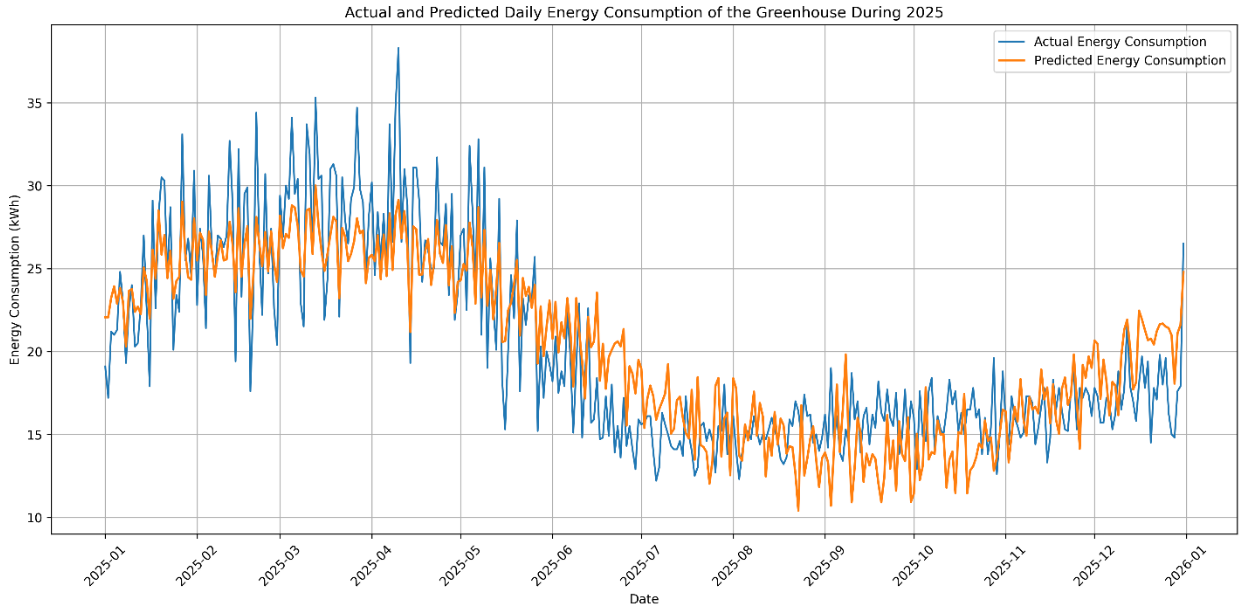


Figure 7 Residuals versus predicted energy consumption.

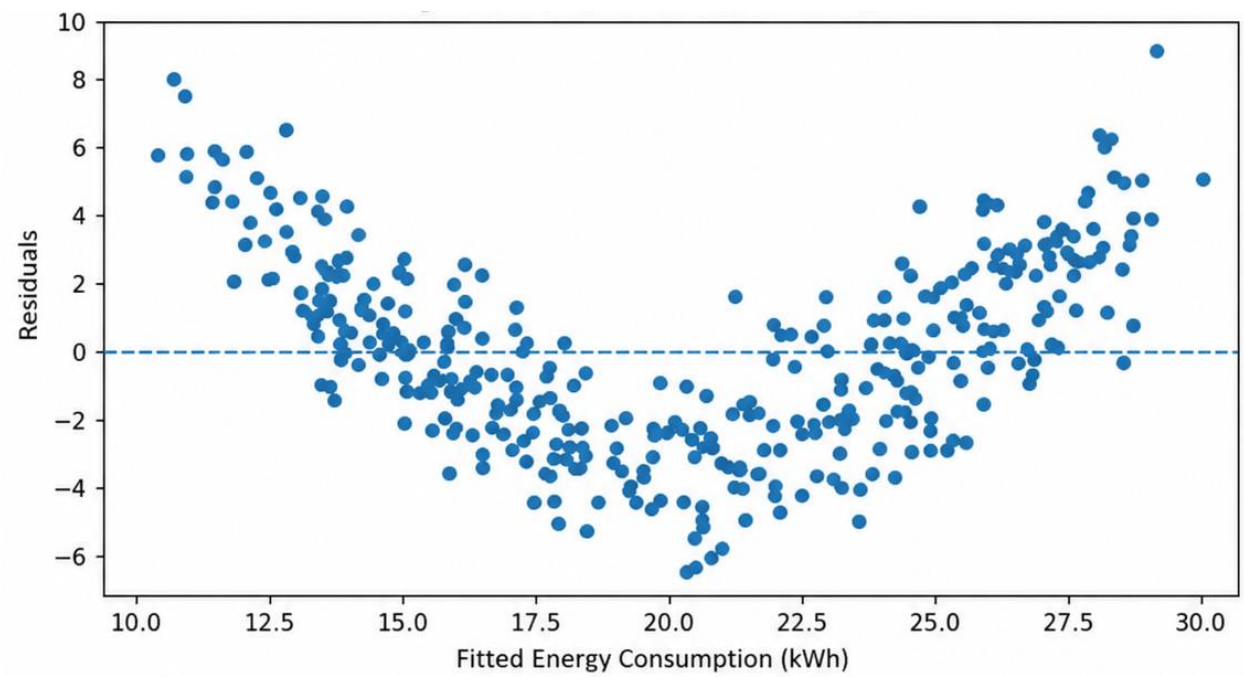


Figure 8 Histogram of regression residuals.

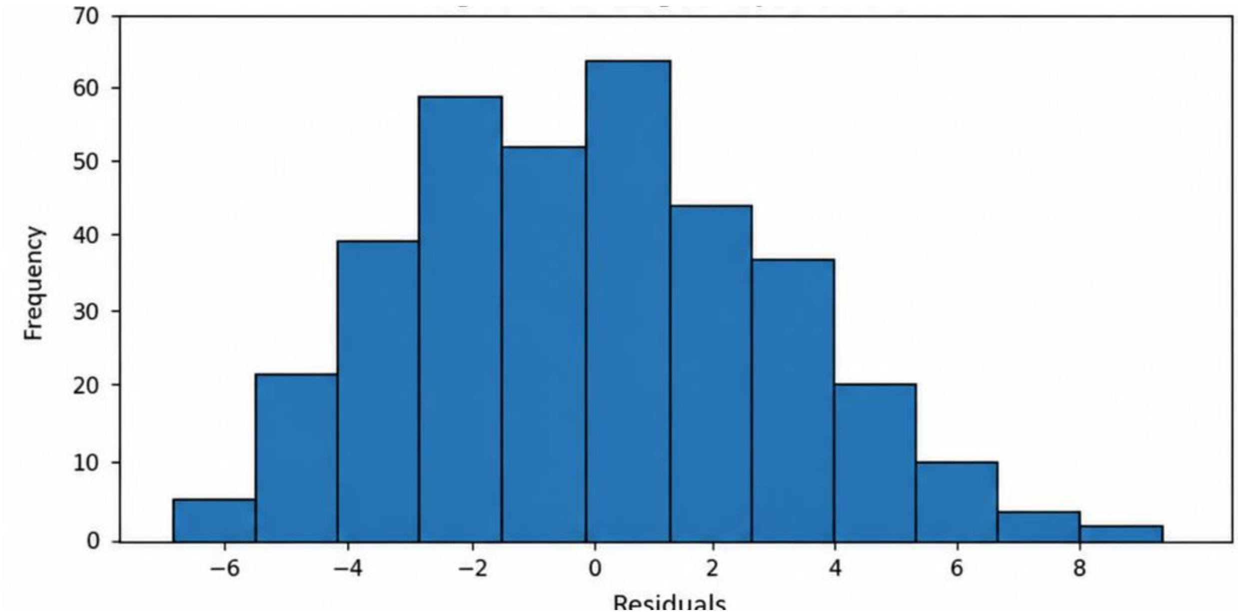


Figure 9 Q–Q plot of standardized residuals from the multiple linear regression model.

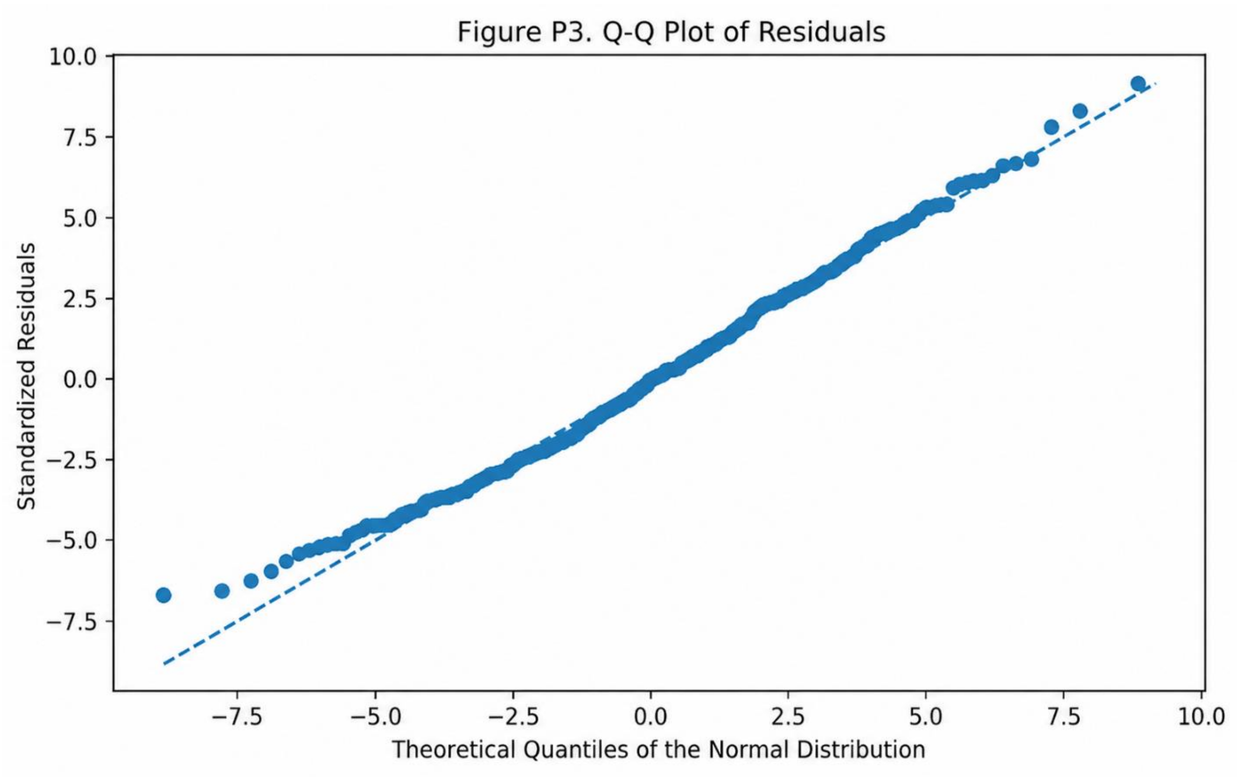
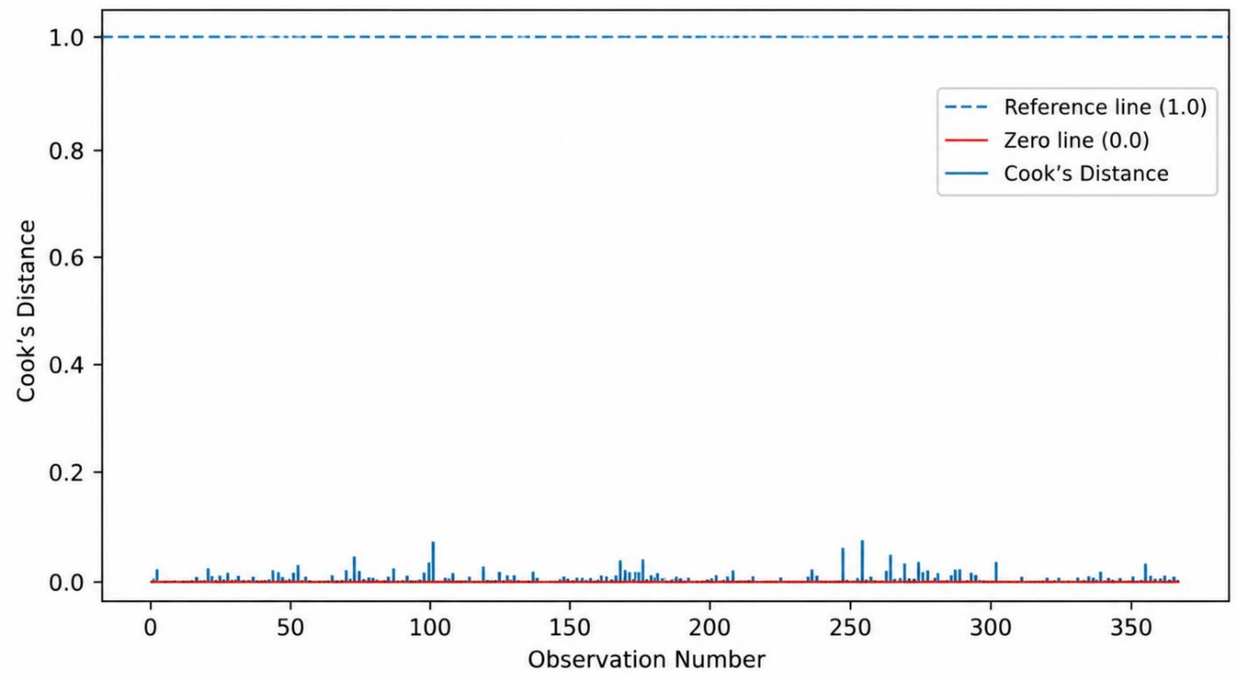


Figure 10 Cook's Distance by Observation



The regression diagnostic graphs indicate the general validity of the model for interpretation and prediction of daily energy use, although some limitations with the data are anticipated given the 'time series' nature of the data. The plot of residuals vs. the predicted values has no evident strong systematic form; i.e., there is no obvious violation of the assumption of linearity, nor does it suggest a serious heteroscedasticity problem. The residuals histogram and Q-Q plot indicate that the errors are reasonably normally distributed with some minor deviation at the extreme ends which is to be expected in daily energy data. Lastly, the Cook's Distance plot indicates that here no observation has an undue influence on the model as the largest value is very small. So diagnostic plots also increase the reliability of the regression and in the future models of the time series can also be taken into account to better reflect the time dependence.

Control of the presence of extreme and influential observations was performed by the analysis of the standardized residuals and the Cook's Distance index. The results indicated one marginal observation with absolute value of the standardized residual just above the traditional criterion of ±3. However, the maximum value of Cook's Distance was 0.032, i.e. significantly lower than the reference limit of 1. This implies that no observation has an excessive influence on the estimation of the model coefficients. As a result, no extreme or influential observations were identified and methodologically there was no reason to discard any observations.

5 Conclusions and Suggestions for Future Research

In this study, the energy use analysis and prediction were conducted based on the daily data for whole year 2025 in the greenhouse cultivation. The method of the research was descriptive statistics, investigation of relationship between the variables and predictive logic; and also to capture the consumption patterns and to draw conclusions useful for energy management and for administrative decisions. The central thesis of the work was that energy in the greenhouse is not a case of merely operating costs, but rather an amount of energy which is bound with microclimate, production stability, economic viability and environment impact. This perspective is validated with the analysis of the results.

The results of this study yield a first central conclusion: Greenhouse energy use is highly seasonal and strongly influenced by external climatic factors. The 2025 data indicated distinctly different energy consumption during colder and warmer months, evidencing a rise in energy demand during colder outdoor temperatures, and a definite reduction in energy use during warmer months. This is consistent with published studies, which found that heating is still the most important factor in energy consumption, particularly when indoor and outdoor climate conditions are quite different from the set points of the greenhouse microclimate (Ahmed et al., 2019; Boulard et al., 2011). The present study provides an empirical verification of this: the difference in temperature between the inside and outside of the greenhouse indicates that energy was utilized not only as a passive response to the weather, but to stabilize the microclimate, with the inside temperature range being noticeably smaller than the outside.

The other major result relates to the relative influence of environmental factors. The most critical factor in the daily energy use was outside temperature, and solar radiation and humidity were rather determinants of the outside temperature than independent dominant factors in energy use. This is in line with literature, where the most important parameter to take into account to drive the heating load is the outdoor temperature, and the solar radiation has a dual role as it can reduce the heating load or increase the cooling and ventilation load depending on the season, structure and greenhouse control strategy (Pérez-Parra et al., 2004; Chen et al., 2025). The present study confirms this stance, in that it is neither "positive" nor "negative" in terms of consumption, but a factor that interacts with temperature and the internal control needs. In this sense, relative humidity was not insignificant, however its influence was more indirect and this fact helps to substantiate that micro climate management, in the case of the greenhouse is multi factorial and not purely related to the temperature.

This was further supported by application of multiple linear regression. The model was able to account for about 75% of the variation in the daily energy consumption, suggesting that the rudimentary set of environment-related variables is adequate to a large extent in the sense of the model for understanding and predicting the energy load. The temperature outdoors turned out to be the most important negative consumption factor, and humidity and solar radiation played a complementary role. This result indicates that regression can be used not only as a theoretical method, but also as a practical method for energy planning in the greenhouse and for load forecasting and administrative decision making.

The third important conclusion is that the energy behavior of the greenhouse can't be explained with univariate logic. While the outdoor temperature was a primary factor, the overall consumption was determined by the effect of outdoor temperature on indoor temperature, humidity and radiation coupled with the operational management. This is especially relevant, as a key theoretical position of contemporary literature was confirmed, that greenhouses are dynamic cyber-physical systems (Li et al., 2025; Shamshiri et al., 2018), in which the use of energy is related to a combination of events in both the environment and the technology, and also events in administration. The present study supports this aspect and expresses the view that development of greenhouse prediction tools cannot be restricted to the simplified schemes and needs a multivariate basis. This conclusion has two implications. Researchers should note from a research perspective, the need for comprehensive models. It demonstrates, from the manager's perspective, that no single indicator like outside temperature or historical load will suffice as a measure for good greenhouse energy policy; many variables need to be monitored and correlated continually.

The methodological contribution of the study was to emphasize the importance of the full year of daily data. The use of days for all years in 2025 enabled the seasonal variation to be well captured, as well as the transition season stages between winter, spring, summer and autumn operation. This is very important given that a lot of literature is written on the basis of shorter monitoring period or idealised data that tend to miss some of the greenhouse system dynamics over the year (Soussi et al., 2025; Payne et al., 2022). Based on the synthetic but realistically structured data set employed in the present study, it is evident that complete temporal coverage is of a great analytical advantage. It can help to identify stable patterns, to interpret extremes, to show periods of increased risk and to link forecast and operational planning.

Another important aspect of the discussion will be the interpretability/accuracy relationship. The current study was designed not so that just a “good” forecast outcome was produced, but a study of the energy behaviour that was both interpretable and usable for decision making purposes. This is especially relevant to the context of statistical models and machine learning. It has been found from the relevant literature that at the level of forecasting, more sophisticated models like the neural networks, LSTM or stochastic control models perform better when compared to simple models in the presence of nonlinearities or high time dependency (Ahn et al., 2024; Kim & You, 2025; Li et al., 2024). Yet, the same literature emphasizes that the operational value of a model is not intricately dependent on the reduction of the error only, but also on the ability of the model to be easily understood and incorporated into the decision making and translated into practical energy management policy (Kamilaris & Prenafeta-Boldú, 2018; Wolfert et al., 2017). This is corroborated by the present study. If a higher accuracy model of a greenhouse cannot be made transparent for operational interpretation, a model which makes it clear that consumption increases due to a combination of the considered factors of low outdoor temperature, low radiation and a requirement for maintaining a greenhouse at constant indoor environment, may be more useful for greenhouse management.

The current study also reiterated the literature on the relationship between energy use and the administrative/economical aspect of greenhouse production. The results indicated that there are several levels where energy demand forecasting may be useful practically: at the energy planning level, energy cost control, to avoid peak demands and energy equipment monitoring. This dimension acknowledges the fact that energy is not just an input in the operation but also a factor of competitiveness, as suggested by the literature. Cost of energy, as well as its volatility, plays a crucial role, particularly in the economic viability of the unit, as demonstrated in studies examining the profitability of greenhouses in Greece and the techno-economic performance of various levels of technology. (Dimitropoulou et al., 2025; Hopwood et al., 2024). This paper indirectly validates this stilted attitude, since the energy variability observed in the data is not only a technical reality, but a possible operational risk that can impact planning and profit levels.

At the same time, the results of the study are consistent with the current literature on sustainability and decarbonization. Since most consumption is in times of higher heating requirements, that is where the focus of any emission reduction strategy should be. The literature on environmental assessment of greenhouse gas emissions revealed that most of the emissions are linked with the heating units based on fossil fuels, and decarbonization efforts could involve reducing their consumption but also changing the mix of energy sources, such as using biomass, heat pumps or geothermal energy, etc. (Taguchi et al., 2024; Solgi et al., 2025; Campana et al., 2025). In this regard, the present work does not provide new LCA but gives the operational context for LCA: it indicates when and under which conditions the energy consumption increases and thus where the environmental improvement intervention would be important.

Theoretically, the research highlights the fact that GH energy management is a systemic issue. It's not just a matter of selecting a superior fuel, and not even only a matter of a more precise algorithm. It's all about the synergy of environment and crop needs, and control technology and economic sense and sustainability objectives. The importance of considering the greenhouse as a system and not as a collection of different subsystems, which are evaluated separately, has already been highlighted in the literature on multi-criteria climate control and on models at the intersection of food, water and energy (Benzzine et al., 2025; Golzar et al., 2021). That this is so is confirmed by the results from the current research. The result of the analysis indicated that consumption factors are not independent from each other and that the saving strategies are not neutral. A decrease in consumption is possible by altering the control temperature or another ventilation concept, but also influences the quality of the microclimate or the biological stability of the crop. Optimization is thus regarded as a balance and not as a one-dimensional minimization.

Some conclusions can be drawn from the above. First, the energy consumption of greenhouses is strongly seasonal and is mainly determined by the external thermal conditions, but is not limited to them. Second, to have a stable internal microclimate is an important source of energy burden and at the same time a condition necessary for the stability of production. Third, multivariate and chronologically coherent data is an enormous advantage for the interpretation and forecasting of energy demand. Fourth, predictive models are only useful when they can't be just technical performance – when they can be operational and administrative utility. Fifth, energy efficiency should be assessed in terms of economy and environmental sustainability and not as a sole technical parameter.

There are indeed limitations of the study. One was that the data set was synthetic, although it was based on realistic operating logic. This indicates that the conclusions are powerful in the methodological and theoretical aspects and cannot be copied blindly to all real greenhouse units. The second constraint is that the number of variables needed to explain consumption was limited and there were no variables for crop types, growth stage, equipment types or real energy market prices. Thirdly, the analysis has been conducted daily but not hourly. This helps in operational understanding, whilst at the same time eliminating intraday peaks, which may be relevant if a more detailed control strategy or load management scenario is desired.

Such restrictions result in recommendations for further research. An initial and evident way forward is to repeat the analysis with actual operational data from a commercial greenhouse, to check how well the trends found in this analysis apply in a greenhouse environment with true administrative, technical and economic uncertainties. An alternate approach is to cover a time scale much longer than a year. Multi-year datasets would allow for better detection of temporal changes, assessment of extreme climate events and more robust assessment of the stability of the models under different climatic conditions. Another way, is to introduce other variables, for instance the speed of the wind, the stage of the crop, CO2, type of heating system, lighting programme and real energy cost data. A more enriched framework would help better predict but also more fully interpret the energy behavior.

A fourth view relates to the systematic comparison of various forecasting models having an equal basis for evaluation. As seen from literature, there is no universally best model, and the performance is problem dependent, the data and the time horizon (Ahn et al., 2024 and Kamilaris & Prenafeta-Boldú, 2018). Hence, it is important to compare different models (multiple linear regression, ARIMA, Random Forest, Support vector Regression, LSTM, newer transformer based models) not only on basis of forecast error but also on basis of Interpretability, Computational requirement and Administrative utility for future studies. A fifth view is to integrate energy forecasting with actual control systems such as model predictive control or data-driven optimization solutions. The literature revealed that the optimum use of forecasting is when the forecast is directly embedded in the system rather than an analytical treatment (Kim & You, 2025; Mahmood et al., 2023). Last, and most important, is the direction of energy analysis to include the economic and environmental assessments. Future work may integrate consumption forecast with cost calculation, energy tariffs, sensitivity analysis, carbon footprint and analysis of investment scenarios of technologies (LEDs, heatpumps, thermal curtains, geothermal, biomass). Only in this way the discussion on the energy efficiency of greenhouses can be brought to a level of strategic sustainability from technical description.

To summarize, this work has demonstrated that energy use in greenhouses can be predicted to a wide extent if it is considered a function of microclimatic and/or seasonal relationships rather than an isolated technical variable. The results agree with the literature available internationally, which suggests that external temperature is the most dominant factor, that microclimatic control has a multifactorial character, that forecasting has a great impact on operational optimization and that there is a need to couple energy management with economic and environmental sustainability. The study's major novel aspect is that it provides a framework for the interpretation of an energy analysis not only as a map of energy use, but for prediction and decision making. In this sense, the study proves that the greenhouse energy study is not just a quantitative study on the energy use in greenhouses but also about the potential use of knowledge on energy use for more efficient, more resilient and more sustainable greenhouse production.

Reference List

Ahamed, M. S., Guo, H., & Tanino, K. (2019). Energy saving techniques for reducing the heating cost of conventional greenhouses. *Biosystems Engineering, 178*, 9–33. https://doi.org/10.1016/j.biosystemseng.2018.10.017

Ahn, J. Y., Cho, S., Kim, S., Kim, Y. J., Lim, J., Choi, Y., & Cho, Y. Y. (2024). Evaluating time-series prediction of temperature, relative humidity, and CO2 concentration in a greenhouse using linear, recurrent, and transformer-based models. *Agronomy, 14*(3), 417. https://doi.org/10.3390/agronomy14030417

Allali, F. E., Fatnassi, H., Demrati, H., Errais, R., Bekkali, M. E., Bensheikh, N., Abdenouri, N., Zine, N., El Khaldi, M., & Aharoune, A. (2025). Greenhouse cooling systems: A systematic review of research trends, challenges, and recommendations for improving sustainability. *Cleaner Engineering and Technology, 26*, Article 100973. https://doi.org/10.1016/j.clet.2025.100973

Benzzine, H., Maestre, J. M., Rodríguez, F., & Ocampo-Martinez, C. (2025). Greenhouse climate control at the food–water–energy nexus: A knowledge-driven multi-objective MPC approach. *Energies, 18*(23), Article 6219. https://doi.org/10.3390/en18236219

Boulard, T., Raeppel, C., Brun, R., Lecompte, F., Hayer, F., Carmassi, G., & Gaillard, G. (2011). Environmental impact of greenhouse tomato production in France. *Agronomy for Sustainable Development, 31*(4), 757–777. https://doi.org/10.1007/s13593-011-0031-3

Campana, E., Ciriello, M., Rouphael, Y., & De Pascale, S. (2025). A new paradigm in greenhouse horticulture: Integrating resource efficiency for sustainable intensification. *Environment, Development and Sustainability*. Advance online publication. https://doi.org/10.1007/s10668-025-06861-6

Chen, S., Liu, A., Tang, F., Hou, P., Lu, Y., & Yuan, P. (2025). A review of environmental control strategies and models for modern agricultural greenhouses. *Sensors, 25*(5), Article 1388. https://doi.org/10.3390/s25051388

Dimitropoulou, A.-M. N., Giannini, E. N., & Maroulis, Z. B. (2025). A greenhouse profitability model: The effect of the energy system. *Energies, 18*(17), Article 4748. https://doi.org/10.3390/en18174748

European Commission. (2019). *The European Green Deal*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0640

European Commission. (n.d.). *The European Green Deal*. European Union. Retrieved April 15, 2026, from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture, 145*, 311–318. https://doi.org/10.1016/j.compag.2018.01.009

Ferentinos, K. P., Katsoulas, N., Tzounis, A., Bartzanas, T., & Kittas, C. (2017). Wireless sensor networks for greenhouse climate and plant condition assessment. *Biosystems Engineering, 153*, 70–81.

Food and Agriculture Organization of the United Nations. (2011). *Energy-smart food for people and climate*. FAO. https://www.fao.org/3/i2454e/i2454e.pdf

Golzar, F., Heeren, N., Hellweg, S., & Roshandel, R. (2021). Optimisation of energy-efficient greenhouses based on an integrated energy demand-yield production model. *Biosystems Engineering, 202*, 1–15. https://doi.org/10.1016/j.biosystemseng.2020.11.012

Hopwood, W., Lopez-Reyes, Z., Bantan, A., Vietti, C., Al-Shahrani, D., Al-Harbi, A., Qaryouti, M., Davies, P., Tester, M., Wing, R., & Waller, R. (2024). Benchmarking techno-economic performance of greenhouses with different technology levels in a hot humid climate. *Biosystems Engineering, 244*, 177–199. https://doi.org/10.1016/j.biosystemseng.2024.06.005

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. https://otexts.com/fpp3/

International Energy Agency. (2017). *Energy access outlook 2017: From poverty to prosperity*. IEA.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Katzin, D., Marcelis, L. F. M., & van Mourik, S. (2021). Energy savings in greenhouses by transition from high-pressure sodium to LED lighting. *Applied Energy, 281*, Article 116019. https://doi.org/10.1016/j.apenergy.2020.116019

Kim, J., & You, F. (2025). Energy-efficient greenhouse climate control using Gaussian process-based stochastic model predictive control. *Applied Energy, 391*, Article 125841. https://doi.org/10.1016/j.apenergy.2025.125841

Li, K., Shi, J., Hu, C., & Xue, W. (2025). The intelligentization process of agricultural greenhouse: A review of control strategies and modeling techniques. *Agriculture, 15*(20), Article 2135. https://doi.org/10.3390/agriculture15202135

Li, M., Liu, H., & Wang, X. (2010). Design of a greenhouse monitoring and control system based on wireless sensor network. *Computers and Electronics in Agriculture, 70*(1), 201–208. https://doi.org/10.1016/j.compag.2009.09.010

Li, X., Zhang, Y., Li, D., Zhang, L., Wu, T., & Zhang, T. (2024). Forecasting greenhouse air and soil temperatures: A multi-step time series method based on deep learning and environmental data fusion. *Computers and Electronics in Agriculture, 218*, 108631. https://doi.org/10.1016/j.compag.2023.108631

Maaouane, M., El Mghouchi, Y., El Bouardi, A., Choulli, Z., & Amrany, Y. (2021). Modelling industry energy demand using multiple linear regression analysis based on consumed quantities of goods. *Energy, 225*, 120270. https://doi.org/10.1016/j.energy.2021.120270

Mahmood, F., Govindan, R., Bermak, A., Yang, D., & Al-Ansari, T. (2023). Data-driven robust model predictive control for greenhouse temperature control and energy utilisation assessment. *Applied Energy, 343*, Article 121190. https://doi.org/10.1016/j.apenergy.2023.121190

McDonald, M. A., & Maravelias, C. T. (2024). Sequential decision-making under uncertainty for long-term energy transition planning. *iScience, 27*(12), Article 111288. https://doi.org/10.1016/j.isci.2024.111288

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). Wiley.

Morais, R., Valente, A., Serôdio, C., Couto, C., & Correia, J. H. (2008). A wireless sensor network for smart irrigation and environmental monitoring. *Computers and Electronics in Agriculture, 62*(2), 120–132. https://doi.org/10.1016/j.compag.2008.01.003

Oldewurtel, F., Parisio, A., Jones, C. N., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B., & Morari, M. (2012). Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings, 45*, 15–27. https://doi.org/10.1016/j.enbuild.2011.09.022

Omid, M., Ghojabeige, F., Delshad, M., & Ahmadi, H. (2011). Energy use pattern and benchmarking of selected greenhouses in Iran using data envelopment analysis. *Energy Conversion and Management, 52*(1), 153–162. https://doi.org/10.1016/j.enconman.2010.06.058

Payne, H. J., Hemming, S., van Rens, B. A. P., van Henten, E. J., & van Mourik, S. (2022). Quantifying the role of weather forecast error on the uncertainty of greenhouse energy prediction and power market trading. *Biosystems Engineering, 224*, 1–15. https://doi.org/10.1016/j.biosystemseng.2022.09.009

Pérez-Parra, J., Baeza, E., Montero, J. I., & Bailey, B. J. (2004). Natural ventilation of parral greenhouses. *Biosystems Engineering, 87*(3), 355–366. https://doi.org/10.1016/j.biosystemseng.2003.12.006

Priva. (2022). *Priva climate control systems*. Priva B.V. https://www.priva.com

Ruiz-Garcia, L., Lunadei, L., Barreiro, P., & Robla, J. I. (2009). A review of wireless sensor technologies and applications in agriculture and food industry: State of the art and current trends. *Computers and Electronics in Agriculture, 68*(1), 1–15. https://doi.org/10.1016/j.compag.2009.04.008

Saadi, H., Behnia, M., Taki, M., & Kaab, A. (2025). A comparative study on energy use and environmental impacts in various greenhouse models for vegetable cultivation. *Environmental and Sustainability Indicators, 25*, Article 100553. https://doi.org/10.1016/j.indic.2024.100553

Shamshiri, R. R., Kalantari, F., Ting, K. C., Thorp, K. R., Hameed, I. A., Weltzien, C., Ahmad, D., & Shad, Z. M. (2018). Advances in greenhouse automation and controlled environment agriculture: A transition to plant factories and urban agriculture. *International Journal of Agricultural and Biological Engineering, 11*(1), 1–22. https://doi.org/10.25165/j.ijabe.20181101.3210

Solgi, H., Rezaei, F., Burg, V., & Roshandel, R. (2025). Carbon footprint analysis of non-fossil heating strategies for agricultural greenhouses in Switzerland. *Energy and Climate Management, 4*, Article 101304. https://doi.org/10.1016/j.ecmx.2025.101304

Soussi, A., Zero, E., Ouammi, A., Zejli, D., Zahmoun, S., & Sacile, R. (2025). Smart greenhouse farming: A review towards near zero energy consumption. *Discover Cities, 2*, Article 55. https://doi.org/10.1007/s44327-025-00096-w

Taguchi, J., Hondo, H., & Moriizumi, Y. (2024). Life cycle CO2 emissions analysis of a high-tech greenhouse horticulture utilizing wood chips for heating in Japan. *Sustainability, 16*(9), Article 3692. https://doi.org/10.3390/su16093692

Trépanier, M.-P., Gosselin, L., & Jørgensen, B. N. (2025). Best combinations of energy-efficiency measures in greenhouses considering energy consumption, yield, and costs: Comparison between two cold climate cities. *Applied Energy, 382*, Article 125163. https://doi.org/10.1016/j.apenergy.2024.125163

Vadiee, A., & Martin, V. (2014). Energy management strategies for commercial greenhouses. *Applied Energy, 114*, 880–888. https://doi.org/10.1016/j.apenergy.2013.08.089

Van Straten, G., van Willigenburg, L. G., van Henten, E. J., & van Ooteghem, R. J. C. (2010). *Optimal control of greenhouse cultivation*. CRC Press. https://doi.org/10.1201/9781420059613

Venkatesan, S. K., Lim, J., Ko, H., & Cho, Y. Y. (2022). A machine learning based model for energy usage peak prediction in smart farms. *Electronics, 11*(2), 218. https://doi.org/10.3390/electronics11020218

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming: A review. *Agricultural Systems, 153*, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023

Appendix I - Dataset

Greenhouse Synthetic Dataset 2025					
date	outside_temp_C	inside_temp_C	humidity_%	solar_radiation_W_m2	energy_consumption_kWh
1/1/2025	17.0	21.8	81.8	142.5	19.1
2/1/2025	17.6	22.0	79.4	101.7	17.2
3/1/2025	16.0	21.4	82.9	40.0	21.2
4/1/2025	14.3	22.7	82.0	83.6	21.0
5/1/2025	15.2	21.3	88.7	110.1	21.3
6/1/2025	13.7	21.1	88.0	133.3	24.8
7/1/2025	16.2	21.0	82.3	90.2	23.0
8/1/2025	19.2	21.3	89.5	93.6	19.3
9/1/2025	14.5	22.5	81.1	133.1	22.6
10/1/2025	14.0	24.0	83.5	80.0	24.0
11/1/2025	16.6	21.3	88.2	40.0	20.3
12/1/2025	16.1	22.8	85.0	40.0	20.5
13/1/2025	15.3	22.4	90.7	142.3	22.8
14/1/2025	12.6	21.0	83.7	111.2	27.0
15/1/2025	14.7	21.7	79.2	53.3	22.5
16/1/2025	16.3	22.1	85.5	171.9	17.9
17/1/2025	11.1	21.3	79.9	144.3	29.1
18/1/2025	13.1	23.6	78.1	177.7	22.6
19/1/2025	9.4	19.1	71.2	130.9	28.2
20/1/2025	10.8	22.0	86.5	99.9	30.5
21/1/2025	9.2	20.9	85.3	134.5	30.3
22/1/2025	13.1	21.5	85.1	127.5	24.4
23/1/2025	10.4	21.9	85.8	122.3	28.7
24/1/2025	14.1	22.8	86.8	166.7	20.1
25/1/2025	13.7	22.1	81.3	130.5	23.4
26/1/2025	12.7	22.3	84.9	139.0	22.4
27/1/2025	6.7	21.2	77.6	177.1	33.1
28/1/2025	11.5	20.3	80.1	164.6	25.5
29/1/2025	12.6	23.1	85.9	108.3	26.8
30/1/2025	12.9	23.8	81.7	145.3	24.6
31/1/2025	8.6	21.4	75.6	162.1	30.9
1/2/2025	11.1	22.8	80.4	205.2	22.8
2/2/2025	9.7	19.5	81.0	169.7	27.4
3/2/2025	10.0	21.8	75.9	214.8	25.8
4/2/2025	14.6	21.6	81.8	188.3	21.4
5/2/2025	9.8	20.1	81.4	115.5	30.6
6/2/2025	11.6	22.7	74.1	158.1	26.1
7/2/2025	13.8	21.5	75.7	186.7	24.6

8/2/2025	10.0	21.8	85.9	226.7	27.0
9/2/2025	11.1	20.3	76.3	134.7	26.8
10/2/2025	11.5	21.8	81.2	178.0	26.3
11/2/2025	11.3	21.9	83.7	144.0	27.0
12/2/2025	8.0	20.1	83.2	194.5	32.7
13/2/2025	11.1	21.3	77.0	133.1	28.9
14/2/2025	14.2	23.7	76.9	213.5	19.4
15/2/2025	6.9	19.8	80.2	236.2	32.2
16/2/2025	12.8	21.2	83.4	189.6	23.3
17/2/2025	10.8	21.1	79.8	121.3	29.5
18/2/2025	8.9	21.7	74.3	224.5	29.9
19/2/2025	15.4	24.1	80.8	290.5	17.6
20/2/2025	12.2	22.3	83.6	205.2	23.2
21/2/2025	7.2	20.2	80.5	280.8	34.4
22/2/2025	10.3	21.6	77.1	206.0	26.2
23/2/2025	11.5	22.9	79.7	231.6	22.2
24/2/2025	9.5	19.9	83.0	137.3	30.7
25/2/2025	11.6	23.2	80.8	245.9	24.7
26/2/2025	9.7	21.8	76.2	148.3	27.4
27/2/2025	11.4	22.2	78.4	269.0	22.6
28/2/2025	13.3	22.7	77.4	235.1	20.4
1/3/2025	7.9	21.3	73.9	249.5	29.4
2/3/2025	10.1	23.1	80.7	170.9	26.9
3/3/2025	8.4	20.0	79.2	356.1	30.0
4/3/2025	9.8	21.6	78.8	175.8	29.2
5/3/2025	6.4	20.4	81.2	226.3	34.1
6/3/2025	7.9	20.1	75.9	166.3	29.5
7/3/2025	8.8	20.7	81.1	160.9	30.4
8/3/2025	11.5	22.0	79.7	311.7	22.9
9/3/2025	12.1	21.5	79.4	331.8	21.5
10/3/2025	5.9	19.6	83.6	334.4	33.7
11/3/2025	7.2	20.0	74.3	297.2	31.9
12/3/2025	10.8	23.3	71.3	304.5	25.9
13/3/2025	4.1	22.3	78.8	252.5	35.3
14/3/2025	7.9	21.3	80.2	250.0	30.4
15/3/2025	8.8	21.8	86.6	311.5	30.6
16/3/2025	12.2	21.6	81.9	208.1	21.9
17/3/2025	10.8	22.0	74.5	299.2	24.3
18/3/2025	8.2	21.0	81.7	317.7	31.0
19/3/2025	8.1	19.9	75.8	257.3	31.3
20/3/2025	8.4	20.0	73.1	315.2	30.6
21/3/2025	12.8	22.4	88.2	312.2	22.1
22/3/2025	7.9	20.6	80.1	313.3	30.5
23/3/2025	8.2	20.0	82.0	367.6	27.9
24/3/2025	9.9	21.2	80.7	415.2	26.5

25/3/2025	8.7	20.7	86.0	408.6	29.2
26/3/2025	8.5	21.7	79.3	366.6	29.9
27/3/2025	6.3	20.8	83.5	334.0	34.7
28/3/2025	9.0	20.6	74.7	329.6	29.8
29/3/2025	8.0	21.0	73.0	435.8	29.0
30/3/2025	12.0	22.1	83.6	327.2	24.4
31/3/2025	10.8	21.3	87.3	150.8	28.2
1/4/2025	9.1	20.6	80.8	462.7	30.2
2/4/2025	10.9	22.7	78.5	274.4	24.6
3/4/2025	8.4	21.3	79.0	327.8	28.4
4/4/2025	11.9	23.4	74.5	400.2	25.4
5/4/2025	9.3	19.2	90.4	91.4	28.3
6/4/2025	10.8	22.1	80.6	445.8	24.6
7/4/2025	6.2	20.5	80.8	347.7	33.7
8/4/2025	10.3	22.5	79.2	446.2	26.6
9/4/2025	5.3	20.6	83.0	439.5	34.8
10/4/2025	4.5	20.5	80.8	402.1	38.3
11/4/2025	8.8	21.4	81.3	272.0	26.6
12/4/2025	7.4	20.9	73.9	273.0	31.0
13/4/2025	10.1	21.0	80.7	244.7	28.8
14/4/2025	15.4	23.7	85.8	373.9	19.3
15/4/2025	7.8	21.6	76.1	338.5	31.1
16/4/2025	8.4	20.3	78.4	306.3	31.1
17/4/2025	10.5	23.0	81.0	428.2	29.1
18/4/2025	11.3	21.7	82.0	355.1	24.2
19/4/2025	9.7	21.7	80.7	306.8	26.7
20/4/2025	9.7	20.7	72.3	346.1	26.1
21/4/2025	12.1	21.8	80.1	406.5	25.0
22/4/2025	11.7	21.9	73.3	318.8	25.0
23/4/2025	7.9	20.2	75.7	309.0	31.7
24/4/2025	10.4	22.1	80.5	236.5	26.6
25/4/2025	10.8	22.3	76.6	353.9	26.4
26/4/2025	8.1	20.7	83.8	188.8	28.9
27/4/2025	11.5	21.7	90.7	308.1	23.4
28/4/2025	8.8	18.0	83.3	454.0	29.5
29/4/2025	13.5	22.8	80.5	511.9	21.9
30/4/2025	11.7	21.1	74.9	536.9	23.5
1/5/2025	11.5	22.7	75.1	478.0	27.0
2/5/2025	9.9	21.9	73.8	541.3	27.4
3/5/2025	11.2	22.5	75.6	402.9	22.5
4/5/2025	6.7	20.7	74.4	492.8	32.4
5/5/2025	8.9	20.8	76.0	471.0	28.4
6/5/2025	12.8	22.6	76.8	562.0	23.7
7/5/2025	6.7	20.4	74.9	318.1	32.8
8/5/2025	14.2	22.1	72.3	409.4	21.0

9/5/2025	7.9	20.2	73.5	463.2	31.1
10/5/2025	14.3	23.4	79.1	332.5	19.0
11/5/2025	10.4	22.3	78.0	459.2	25.6
12/5/2025	14.6	23.4	81.3	414.4	22.8
13/5/2025	13.1	22.1	82.1	280.5	20.1
14/5/2025	9.1	20.9	77.2	375.4	29.2
15/5/2025	16.2	22.8	80.2	518.4	18.3
16/5/2025	16.8	23.4	71.5	560.4	15.3
17/5/2025	13.2	23.1	86.5	406.8	20.0
18/5/2025	12.8	22.7	82.4	455.8	24.6
19/5/2025	13.2	21.8	77.3	310.2	22.0
20/5/2025	11.3	21.6	73.6	334.4	27.9
21/5/2025	16.7	24.0	78.4	368.0	17.6
22/5/2025	12.7	22.2	70.4	401.8	23.2
23/5/2025	14.1	22.2	74.8	358.5	21.6
24/5/2025	12.4	23.5	80.2	327.3	23.4
25/5/2025	13.0	23.6	77.4	537.9	23.1
26/5/2025	11.5	23.0	75.5	513.2	25.7
27/5/2025	18.0	22.8	81.4	524.3	15.2
28/5/2025	14.6	24.4	72.5	376.3	20.3
29/5/2025	17.6	23.4	78.4	517.4	17.2
30/5/2025	15.3	23.3	75.3	517.9	20.0
31/5/2025	13.7	22.5	74.3	460.5	19.2
1/6/2025	14.8	22.0	79.9	688.6	18.2
2/6/2025	14.4	21.5	74.3	425.9	20.9
3/6/2025	16.0	23.3	84.9	564.9	17.5
4/6/2025	15.2	23.1	75.1	490.8	18.8
5/6/2025	15.5	22.6	85.6	477.2	17.9
6/6/2025	13.0	21.8	73.7	555.9	22.4
7/6/2025	14.6	21.8	80.5	595.9	19.8
8/6/2025	20.9	23.1	72.6	566.9	15.1
9/6/2025	15.3	21.3	73.0	300.7	19.1
10/6/2025	14.5	23.7	77.7	617.9	22.9
11/6/2025	18.1	23.9	79.6	482.4	14.8
12/6/2025	21.0	25.8	70.8	624.9	18.5
13/6/2025	14.0	23.3	82.0	451.6	22.6
14/6/2025	17.3	22.6	74.5	546.3	15.7
15/6/2025	16.4	22.3	78.4	536.6	15.9
16/6/2025	13.7	19.7	80.8	361.6	18.4
17/6/2025	20.1	22.7	80.0	489.6	14.7
18/6/2025	18.4	22.8	70.4	432.6	14.8
19/6/2025	18.8	24.1	83.3	635.9	17.3
20/6/2025	16.9	23.6	79.4	589.3	14.9
21/6/2025	20.1	23.0	79.2	135.9	18.0
22/6/2025	17.8	23.1	73.3	444.3	13.9

23/6/2025	19.0	23.3	70.5	309.3	15.5
24/6/2025	16.7	24.3	71.3	602.9	13.6
25/6/2025	16.6	21.8	80.3	354.1	17.2
26/6/2025	23.2	24.4	75.3	640.2	14.3
27/6/2025	18.8	22.2	73.4	614.1	15.5
28/6/2025	20.9	24.3	74.6	340.9	14.1
29/6/2025	20.3	24.3	83.0	501.2	12.9
30/6/2025	19.4	23.5	75.8	375.1	15.9
1/7/2025	19.4	22.3	75.5	530.7	15.6
2/7/2025	22.5	24.8	81.8	629.4	15.7
3/7/2025	20.3	23.3	79.1	668.6	16.1
4/7/2025	20.8	23.3	77.5	479.4	16.1
5/7/2025	21.4	23.3	74.9	568.1	14.0
6/7/2025	24.5	25.2	70.0	472.3	12.2
7/7/2025	23.4	24.3	72.0	487.7	13.0
8/7/2025	22.8	25.2	74.3	402.7	16.3
9/7/2025	20.6	23.6	76.2	608.2	15.6
10/7/2025	18.8	22.8	79.0	472.3	15.0
11/7/2025	24.7	24.7	78.7	477.7	14.3
12/7/2025	25.0	24.4	77.6	415.7	14.1
13/7/2025	22.3	25.1	77.2	400.6	14.1
14/7/2025	24.2	24.3	74.1	205.9	14.6
15/7/2025	25.0	25.0	67.7	434.6	13.7
16/7/2025	25.2	24.6	74.5	483.9	17.3
17/7/2025	25.6	23.8	73.2	549.8	15.0
18/7/2025	22.3	25.1	73.7	340.9	14.0
19/7/2025	27.4	24.3	75.6	512.5	12.5
20/7/2025	20.7	23.7	72.1	472.1	13.0
21/7/2025	26.1	25.4	71.7	528.9	15.5
22/7/2025	25.3	23.7	76.7	631.5	15.7
23/7/2025	26.4	24.3	77.2	522.7	14.6
24/7/2025	29.1	25.0	73.7	586.0	15.3
25/7/2025	28.2	24.5	78.6	362.9	14.6
26/7/2025	21.8	22.2	77.0	407.5	12.7
27/7/2025	20.6	23.8	72.9	477.7	15.5
28/7/2025	27.0	24.9	76.6	488.7	15.2
29/7/2025	22.5	24.9	84.0	479.7	18.0
30/7/2025	25.2	24.3	74.0	269.2	13.8
31/7/2025	27.5	25.4	80.8	550.8	14.8
1/8/2025	21.4	22.3	83.4	257.4	16.1
2/8/2025	20.3	23.2	83.4	477.5	14.0
3/8/2025	26.4	25.6	72.7	589.4	12.3
4/8/2025	26.0	25.0	84.4	521.8	13.9
5/8/2025	25.4	24.4	77.2	322.6	14.9
6/8/2025	26.2	24.1	75.8	441.7	15.2

7/8/2025	24.1	24.8	75.5	440.3	14.7
8/8/2025	22.6	24.6	79.0	252.5	16.1
9/8/2025	26.1	24.6	76.8	364.4	15.2
10/8/2025	24.2	23.9	76.2	258.8	14.4
11/8/2025	22.6	25.0	81.5	480.5	15.0
12/8/2025	28.1	27.3	72.2	565.7	14.7
13/8/2025	26.8	24.7	72.3	371.6	15.3
14/8/2025	28.0	25.3	79.4	288.8	16.0
15/8/2025	24.6	24.8	80.6	203.8	15.3
16/8/2025	25.6	22.9	81.9	505.6	14.7
17/8/2025	24.8	24.8	72.2	399.3	13.5
18/8/2025	25.2	24.1	81.2	298.1	13.2
19/8/2025	28.0	26.1	77.7	265.4	13.6
20/8/2025	25.6	24.7	83.6	430.9	15.9
21/8/2025	28.6	24.4	71.7	281.2	15.5
22/8/2025	28.6	25.1	77.2	469.6	17.0
23/8/2025	32.9	25.8	74.8	378.7	16.4
24/8/2025	24.4	24.2	70.4	347.9	14.9
25/8/2025	30.2	25.3	73.3	350.1	17.4
26/8/2025	27.8	24.0	81.4	364.3	16.1
27/8/2025	28.1	24.7	78.5	121.2	16.2
28/8/2025	24.6	25.4	73.2	474.8	14.5
29/8/2025	27.1	26.6	78.9	412.5	15.0
30/8/2025	30.2	25.9	79.8	350.0	14.0
31/8/2025	28.2	25.2	80.2	283.4	14.8
1/9/2025	28.7	26.0	72.6	256.3	16.2
2/9/2025	27.8	25.9	75.6	431.0	14.2
3/9/2025	31.4	26.9	79.5	385.7	19.0
4/9/2025	28.6	25.9	73.8	230.3	14.6
5/9/2025	23.2	24.2	77.0	144.6	16.1
6/9/2025	27.0	25.2	79.6	372.2	13.9
7/9/2025	23.8	24.6	83.3	234.1	13.4
8/9/2025	20.6	23.0	73.8	216.9	15.3
9/9/2025	27.5	25.2	65.8	422.7	14.5
10/9/2025	32.2	25.1	70.5	466.3	18.7
11/9/2025	29.0	23.6	73.1	477.7	15.9
12/9/2025	26.0	24.1	79.4	145.1	17.0
13/9/2025	26.6	23.2	79.8	270.0	13.9
14/9/2025	31.8	25.3	79.6	117.0	16.1
15/9/2025	29.4	25.0	76.3	156.1	16.6
16/9/2025	29.1	23.9	77.6	348.8	14.4
17/9/2025	28.9	25.6	72.1	274.3	16.2
18/9/2025	29.1	25.8	76.8	215.3	15.4
19/9/2025	31.0	25.5	73.5	329.0	18.2
20/9/2025	30.4	26.7	86.1	365.2	16.3

21/9/2025	29.5	26.7	78.7	314.7	15.8
22/9/2025	26.4	23.7	78.5	92.7	17.7
23/9/2025	30.3	26.1	71.8	253.8	16.0
24/9/2025	27.3	25.3	74.4	286.6	15.5
25/9/2025	31.7	25.9	69.3	386.5	17.5
26/9/2025	25.8	25.5	73.2	262.8	13.8
27/9/2025	28.6	24.5	73.7	323.4	14.8
28/9/2025	28.9	24.6	80.1	253.9	17.7
29/9/2025	25.6	24.0	75.2	262.6	14.6
30/9/2025	33.2	25.0	73.2	293.6	17.0
1/10/2025	32.5	25.5	70.7	311.3	16.0
2/10/2025	27.6	23.8	72.3	261.9	12.9
3/10/2025	30.7	24.5	76.3	318.4	17.6
4/10/2025	29.7	25.6	75.1	259.7	14.9
5/10/2025	22.1	24.2	82.0	227.5	14.6
6/10/2025	29.3	25.4	73.6	268.0	17.6
7/10/2025	28.4	23.5	82.5	212.6	18.4
8/10/2025	28.7	24.4	78.1	237.2	14.1
9/10/2025	25.8	24.2	77.2	234.7	16.1
10/10/2025	27.8	25.0	68.3	270.9	15.3
11/10/2025	27.9	24.4	73.3	191.4	15.0
12/10/2025	31.3	26.1	79.8	214.2	16.4
13/10/2025	29.1	24.9	86.9	85.1	18.3
14/10/2025	28.2	25.0	77.2	266.4	16.8
15/10/2025	31.9	25.4	80.0	223.3	17.6
16/10/2025	26.6	25.4	74.5	271.9	15.2
17/10/2025	27.0	24.5	75.1	261.7	16.3
18/10/2025	23.3	25.3	73.5	259.9	15.0
19/10/2025	31.7	25.8	80.5	228.6	16.5
20/10/2025	30.1	25.3	78.7	211.7	16.5
21/10/2025	29.9	25.2	75.3	249.9	17.8
22/10/2025	29.2	25.5	75.7	218.0	16.0
23/10/2025	27.7	25.6	81.6	137.9	16.5
24/10/2025	27.9	25.7	78.1	221.1	13.8
25/10/2025	26.6	25.0	77.7	95.9	16.0
26/10/2025	26.7	25.4	85.1	180.2	13.8
27/10/2025	27.2	24.9	78.9	201.1	15.4
28/10/2025	30.7	26.1	81.1	70.9	19.6
29/10/2025	28.2	25.0	82.8	233.8	12.6
30/10/2025	26.6	23.1	81.6	246.7	15.0
31/10/2025	25.2	25.5	81.2	67.4	18.8
1/11/2025	24.9	24.6	81.8	148.2	15.8
2/11/2025	30.4	24.3	74.0	197.5	14.4
3/11/2025	27.6	24.7	76.4	169.1	17.3
4/11/2025	26.3	23.9	72.0	112.9	16.0

5/11/2025	25.2	26.2	84.3	125.2	15.5
6/11/2025	23.2	24.7	71.9	150.3	14.8
7/11/2025	25.6	23.4	81.1	210.8	15.1
8/11/2025	27.9	25.3	77.1	110.6	17.3
9/11/2025	24.6	24.1	76.7	122.9	17.3
10/11/2025	24.9	26.0	78.7	128.5	16.9
11/11/2025	24.7	23.7	80.4	175.4	14.4
12/11/2025	25.4	22.8	82.8	157.7	15.4
13/11/2025	21.0	22.6	89.5	90.3	16.7
14/11/2025	24.3	24.1	83.7	69.7	17.8
15/11/2025	22.6	23.9	83.3	155.5	13.3
16/11/2025	26.8	23.1	76.9	186.8	14.8
17/11/2025	22.5	24.5	84.1	94.5	18.3
18/11/2025	25.7	25.9	80.1	129.6	16.5
19/11/2025	28.0	23.3	79.9	102.2	17.8
20/11/2025	23.2	25.2	80.7	89.3	16.4
21/11/2025	22.4	23.7	81.0	112.0	15.3
22/11/2025	24.2	26.3	80.3	121.7	15.2
23/11/2025	23.6	24.7	82.2	116.1	17.6
24/11/2025	20.9	22.4	80.7	85.4	18.9
25/11/2025	24.4	25.3	75.7	165.8	15.3
26/11/2025	28.1	25.3	82.1	151.4	17.8
27/11/2025	22.3	23.6	75.5	76.7	17.2
28/11/2025	22.3	24.8	76.7	160.2	17.8
29/11/2025	20.0	22.1	89.0	90.1	17.4
30/11/2025	23.3	22.8	72.0	75.9	16.1
1/12/2025	19.2	24.1	78.0	118.5	17.8
2/12/2025	19.4	22.1	86.3	64.6	17.3
3/12/2025	25.2	24.0	74.3	115.3	15.7
4/12/2025	19.5	22.7	92.1	113.6	15.7
5/12/2025	24.3	23.1	71.2	109.1	17.2
6/12/2025	25.3	23.4	88.0	81.9	16.9
7/12/2025	22.0	22.4	86.6	165.6	15.3
8/12/2025	22.5	23.8	84.1	136.7	16.2
9/12/2025	25.8	24.6	79.8	112.9	18.8
10/12/2025	20.3	23.4	82.6	118.1	16.5
11/12/2025	19.1	21.9	80.8	40.0	17.7
12/12/2025	17.1	22.5	87.2	40.0	21.7
13/12/2025	20.4	23.0	77.3	105.1	17.9
14/12/2025	23.8	23.4	80.9	89.8	17.0
15/12/2025	22.3	24.3	85.9	78.0	15.8
16/12/2025	17.4	21.5	79.2	70.7	18.1
17/12/2025	17.5	23.1	81.6	61.4	19.7
18/12/2025	18.2	22.7	83.4	81.2	17.8
19/12/2025	20.0	22.5	77.3	85.9	19.4

20/12/2025	18.6	21.3	86.7	126.5	14.5
21/12/2025	19.5	22.5	84.3	78.0	17.8
22/12/2025	19.5	21.8	77.5	67.3	17.1
23/12/2025	17.8	22.8	82.6	74.0	19.8
24/12/2025	18.3	22.2	80.5	61.3	18.0
25/12/2025	18.7	21.6	81.6	50.2	19.6
26/12/2025	17.8	23.1	81.4	127.1	16.3
27/12/2025	19.1	23.0	79.6	83.3	15.0
28/12/2025	22.4	23.7	84.8	119.7	14.8
29/12/2025	19.0	23.2	80.6	48.0	17.6
30/12/2025	17.5	22.0	87.7	59.6	17.9
31/12/2025	13.0	21.7	82.1	114.5	26.5