



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ  
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

## Μοντέλα Πρόβλεψης Ισχύος σε Έξυπνα Δίκτυα Χαμηλής Τάσης με Χρήση Τεχνικών Μηχανικής και Βαθιάς Μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αφροδίτη Μπλίκια

**Επιβλέπων:** Δημήτριος Ασκούνης  
Καθηγητής ΕΜΠ

Αθήνα, Ιούλιος 2023





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

• ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ  
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**Μοντέλα Πρόβλεψης Ισχύος σε Έξυπνα Δίκτυα Χαμηλής  
Τάσης με Χρήση Τεχνικών Μηχανικής και Βαθιάς Μάθησης**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αφροδίτη Μπλίκια

**Επιβλέπων:** Δημήτριος Ασκούνης,  
Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 12 Ιουλίου 2023.

.....

Δημήτριος Ασκούνης

Καθηγητής ΕΜΠ

.....

Ιωάννης Ψαρράς

Καθηγητής ΕΜΠ

.....

Χρυσόστομος Δούκας

Καθηγητής ΕΜΠ

Αθήνα, Ιούλιος 2023

.....  
Αφροδίτη Μπλίκια

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Αφροδίτη Μπλίκια, 2023

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.



## Περίληψη

Η βραχυπρόθεσμη πρόβλεψη ισχύος (STPF) και φορτίου (STLF) σε ηλεκτρικά δίκτυα και ιδιαίτερα σε έξυπνα ηλεκτρικά δίκτυα είναι κρίσιμη, καθώς συνεισφέρει στην βελτιστοποίηση της λειτουργίας τους τόσο σε επίπεδο αξιοπιστίας και απόδοσης όσο και σε επίπεδο κόστους. Πρόκειται για ένα ζήτημα που παρουσιάζει πολλές προκλήσεις εξ' αιτίας των πολλαπλών ηλεκτρικών φορτίων που είναι συνδεδεμένα στο δίκτυο και την εξάρτηση της ισχύος τους, ενεργού και αέργου, από πληθώρα εξωγενών μεταβλητών. Στην παρούσα εργασία γίνεται χρήση μοντέλων μηχανικής μάθησης (ML) και βαθιάς μάθησης (DL) –συγκεκριμένα των μοντέλων Light Gradient Boosting Machine (LightGBM), Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS) και Temporal Convolutional Networks (TCN). Επιπλέον δοκιμάζεται και μια μέθοδος ensemble στο μοντέλο N-BEATS με χρήση πολλαπλών instances του μοντέλου με διαφορετική αρχικοποίηση βαρών. Τα μοντέλα εξετάζονται σε σχέση με την προβλεπτική τους ικανότητα σε πολλαπλά σύνολα δεδομένων που αφορούν χρονοσειρές ηλεκτρικής ενέργειας χαμηλής τάσης οι οποίες έχουν συλλεχθεί από έξυπνους μετρητές (smart meters) στο Terni της Ιταλίας. Ερευνάται, επιπλέον, η επίδραση εξωγενών μεταβλητών, όπως τα καιρικά δεδομένα και οι χρονικές μεταβλητές, στην προβλεπτική ικανότητα των μοντέλων. Στο πλαίσιο της ανακάλυψης πρώιμων μοτίβων και συσχετίσεων μεταξύ χρονοσειρών αλλά και για την ενδεχόμενη βελτίωση των μεταγενέστερων προβλέψεων διεξάγεται μία εκτενής διερευνητική ανάλυση δεδομένων (EDA). Κατά τη διάρκεια της πειραματικής διαδικασίας γίνεται προσπάθεια πρόβλεψης χρονοσειρών που αφορούν σε σύνολα δεδομένων ενεργού και αέργου ισχύος, ωριαίας ανάλυσης, που μετρήθηκαν από smart meters τοποθετημένα σε κομβικές θέσεις ενός έξυπνου δικτύου. Ο χρονικός ορίζοντας των προβλέψεων εκτείνεται στη μία ημέρα. Από τα αποτελέσματα των πειραμάτων συμπεραίνεται ότι το πιο αποδοτικό μοντέλο κατά μέσο όρο αναδεικνύεται μέσα από τη μέθοδο μηχανικής μάθησης LightGBM με καλύτερες επιδόσεις από τις αρχιτεκτονικές βαθιάς μάθησης στις περισσότερες περιπτώσεις, αλλά και σχετικά χαμηλές υπολογιστικές απαιτήσεις. Παρατηρήθηκε επίσης ότι η τεχνική ensemble στο N-BEATS οδηγεί το μοντέλο σε πολύ καλύτερα αποτελέσματα, με βελτίωση της απόδοσής του στην πρόβλεψη της ενεργούς ισχύος έως 16% με βάση το MASE. Ένα σημαντικό αποτέλεσμα επίσης έγκειται στο ότι –σε αντίθεση με το N-BEATS– για την εξασφάλιση της βέλτιστης απόδοσης του μοντέλου LightGBM είναι αναγκαία η χρήση εξωγενών μεταβλητών. Τα αποτελέσματα της εργασίας μπορούν να αξιοποιηθούν από τον διαχειριστή του έξυπνου συστήματος ηλεκτρικής ενέργειας ως i) ένα εργαλείο πρόβλεψης μηχανικής μάθησης για την αποδοτικότερη λειτουργία του δικτύου και ii) ως έναν οδηγό επιλογής κατάλληλων μεθόδων πρόβλεψης στο πλαίσιο του έξυπνου δικτύου, ο οποίος λαμβάνει υπόψη τόσο την ακρίβεια όσο και τη βιωσιμότητα τους.

**Λέξεις-Κλειδιά:** Βαθιά Μάθηση, Βραχυπρόθεσμη Πρόβλεψη Χρονοσειρών, Έξυπνα Δίκτυα, Έξυπνοι μετρητές, Μηχανική Μάθηση, Νευρωνικά Δίκτυα, Πρόβλεψη Ισχύος Βραχυπρόθεσμου Χρόνου, Χαμηλή Τάση, Χρονικά Συνελκτικά Δίκτυα (TCN), Ensemble, Gradient boosting, N-BEATS

## Abstract

Short-term power (STPF) and short term load (STLF) forecasting in power grids and especially in smart grids is critical, as it contributes to the optimization of their operation both in terms of reliability and performance as well as in terms of costs. STPF and STLF are challenging tasks due to the multiple electrical loads connected to the grid and the dependence of their powers, active and reactive, on a multitude of exogenous variables. In this paper, a comparative study of machine learning (ML) and deep learning (DL) models is conducted – namely Light Gradient Boosting Machine (LightGBM), Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS) and Temporal Convolutional Networks (TCN) models. Additionally, an ensemble model with N-BEATS is evaluated using individual models with different weight initializations, models with different random states. The models are evaluated with respect to their forecasting performance, on day-ahead forecasts, in multiple data sets concerning electricity time series, collected by smart meters placed at crucial nodes of a smart grid. From the experimental results it is concluded that the most efficient model on average emerges through the LightGBM machine learning method with better performance than deep learning architectures in most cases and relatively low computational requirements. It is also observed that the N-BEATS ensemble leads the model to much better results, with an active power forecast accuracy improvement of about 16%. Note here that, contrary to N-BEATS, the use of exogenous variables is necessary to ensure optimal performance of the LightGBM. The results of this work can be leveraged by smart power system operators as i) an effective machine learning forecasting tool for more efficient grid operation and ii) a guide for selecting appropriate forecasting methods in the smart grid context, which takes into account both their accuracy and sustainability.

**Keywords:** Deep Learning (DL), Ensemble, LightGBM, Low Voltage, Machine Learning (ML), N-BEATS, Neural Networks, Smart Grids, Smart Meters, Short-Term Power Forecasting, Temporal Convolutional Networks (TCN), Timeseries Forecasting

## Περιεχόμενα

Περίληψη .....	6
Abstract .....	7
<b>1.Εισαγωγή .....</b>	<b>13</b>
1.1. Πλαίσιο .....	13
1.2 Σκοπός .....	14
1.3 Διάρθρωση εργασίας .....	14
1.4. Σχετική Εργασία .....	14
<b>2.Βιβλιογραφική Ανασκόπηση.....</b>	<b>16</b>
2.1. Συστήματα Ηλεκτρικής Ενέργειας .....	16
2.1.1. Εισαγωγή .....	16
2.1.2. Σύστημα Ηλεκτρικής Ενέργειας.....	16
2.1.3. Ανάγκη πρόβλεψης φορτίου.....	18
2.1.4 Αδυναμίες του Παραδοσιακού Συστήματος Ηλεκτρικής Ενέργειας .....	18
2.1.5. Το έξυπνο δίκτυο .....	19
2.2 Χρονοσειρές και Τεχνικές Προβλέψεων .....	22
2.2.1. Εισαγωγή στις Χρονοσειρές.....	22
2.2.2. Ποιοτικά Χαρακτηριστικά Χρονοσειρών .....	22
2.2.3 Αποσύνθεση χρονοσειρών .....	24
2.2.4. Βασική Στατιστική Ανάλυση .....	27
2.2.5. Στατιστική Ανάλυση Ακριβείας Προβλέψεων .....	30
2.3. Μηχανική Μάθηση.....	33
2.3.1 Εισαγωγή.....	33
2.3.2 Αλγόριθμοι Μηχανικής Μάθησης.....	34
2.3.2. Μοντέλα Μηχανικής και Βαθιάς Μάθησης .....	34
2.3.3. Μέθοδος Ensemble.....	38
<b>3. Μεθοδολογία .....</b>	<b>39</b>
3.1. Σύνολο Δεδομένων .....	39
3.2. Ζητούμενο.....	39
3.3 Διαχείριση Δεδομένων.....	40
3.3.1. Εργαλεία .....	40
3.3.2. Προεπεξεργασία Δεδομένων .....	40
3.3.3. Ανάλυση Δεδομένων .....	45
3.4. Διαδικασία Πρόβλεψης .....	56
3.4.1. Επιλογή Συμμεταβλητών.....	56
3.4.2. Επιλογή και Εκπαίδευση Μοντέλων .....	57
3.4.3. Αξιολόγηση Μοντέλων .....	60
<b>4. Αποτελέσματα .....</b>	<b>62</b>
4.1. Αποτελέσματα Μοντέλων .....	62
4.1.1. Αποτελέσματα Μοντέλου LightGBM.....	62
4.1.2. Αποτελέσματα Μοντέλου N-BEATS.....	64
4.1.3. Αποτελέσματα Μοντέλου TCN.....	66
4.2. Οπτικοποίηση Αποτελεσμάτων – Συγκριτική Ανάλυση .....	67
4.2.1. Ενεργός Ισχύς W4 .....	67
4.2.2. Άεργος Ισχύς W4.....	71



4.2.3. Ενεργός Ισχύς W6 .....	73
4.2.4. Άεργος Ισχύς W6.....	75
4.2.5. Βέλτιστα Μοντέλα.....	76
<b>5. Συμπεράσματα .....</b>	<b>78</b>
<b>6. Μελλοντική Εργασία – Προεκτάσεις .....</b>	<b>81</b>
<b>7. Κατάλογος Βιβλιογραφίας .....</b>	<b>83</b>

Εικόνα 1 - Βασικές δομές ενός ΣΗΕ .....	17
Εικόνα 2 - Παραδείγματα Χρονοσειρών με έντονα ποιοτικά χαρακτηριστικά .....	24
Εικόνα 3 -Αποσύνθεση χρονοσειράς στα βασικά της χαρακτηριστικά [15] .....	27
Εικόνα 4 -Η βασική αρχιτεκτονική του N-BEATS.....	36
Εικόνα 5 - Η δομή μιας στοίβας ενός TCN σε univariate σύστημα (Πηγή: [40]) .....	38
Εικόνα 6 - Τοπολογία δικτύου .....	39
Εικόνα 7 - Ποσοστό κενών τιμών για τον W4, ανά χρόνο, μήνα, μέρα της εβδομάδας και ώρα της ημέρας .....	41
Εικόνα 8 - Η σχέση του βάρους $w$ και της απόστασης $d[i]$ .....	43
Εικόνα 9 - Ποσοστό κενών τιμών στα δεδομένα ισχύος, ανά μετρητή, μετά την εφαρμογή των διαδικασιών συμπλήρωσης .....	43
Εικόνα 10 - Εντοπισμός outliers στα δεδομένα ισχύος.....	44
Εικόνα 11 - Στατιστική Περιγραφή Δεδομένων μετρητών W3 και W5 .....	45
Εικόνα 12 - Στατιστική Περιγραφή Δεδομένων μετρητών W2, W4, W6.....	45
Εικόνα 13 - Πλήθος αρνητικών τιμών στην ενεργό ισχύ του W4 ανά την ώρα της ημέρας που παρουσιάζονται .....	47
Εικόνα 14 - Πλήθος θετικών και αρνητικών τιμών της ενεργούς ισχύος του W4, σε λογαριθμική κλίμακα.....	47
Εικόνα 15 - Πλήθος θετικών τιμών στην ενεργό ισχύ του W4 ανά την ώρα της ημέρας που παρουσιάζονται την ενεργό ισχύ του φωτοβολταϊκού. ....	48
Εικόνα 16 - Γραφική αναπαράσταση δεδομένων ισχύος τριών ημερών .....	48
Εικόνα 17 - Γραφική αναπαράσταση δεδομένων ρεύματος τριών ημερών .....	49
Εικόνα 18 - Γραφική αναπαράσταση δεδομένων τάσης τριών ημερών .....	49
Εικόνα 19 - Ωριαία διαγράμματα μέσης τιμής ισχύος για τον W4 ανά χρόνο.....	50
Εικόνα 20 - Ωριαία διαγράμματα μέσης τιμής ισχύος για τον W6, ανά χρόνο.....	50
Εικόνα 21 - Γραφική αναπαράσταση συντελεστή ισχύος του W4 σε διάστημα τριών ημερών. ....	51
Εικόνα 22 - Πλήθος αρνητικών τιμών του συντελεστή ισχύος του W4 ανά ημέρα της εβδομάδας και ώρα της ημέρας .....	51
Εικόνα 23 - Γραφική αναπαράσταση συντελεστή ισχύος του W6 σε διάστημα τριών ημερών. ....	52
Εικόνα 24 - Πλήθος αρνητικών τιμών συντελεστή ισχύος W6 ανά ώρα της ημέρας.....	53
Εικόνα 25 - Πίνακας συσχέτισης δεδομένων ισχύος .....	54
Εικόνα 26 - Scatterplots μετεωρολογικών δεδομένων και δεδομένων ισχύος του W6 .....	55
Εικόνα 27 - Scatterplots μετεωρολογικών δεδομένων και δεδομένων ισχύος του W4 .....	56
Εικόνα 28- Γραφική εξήγηση μεθόδου Early Stopping .....	59
Εικόνα 29 – Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της ενεργούς ισχύος του W4 .....	68
Εικόνα 30 – Στιγμιότυπο πρόβλεψης της ενεργούς ισχύος του W4 .....	69
Εικόνα 31 – Σύγκριση του MAE των μοντέλων για την πρόβλεψη της ενεργούς ισχύος του W4 ανά χρονικό διάστημα της ημέρας .....	69

Εικόνα 32 – Σύγκριση του MAE των μοντέλων για την πρόβλεψη της ενεργούς ισχύος του W4 ανά εργάσιμη ή μη ημέρα.....	70
Εικόνα 33 – Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της αέργου ισχύος του W4 .....	71
Εικόνα 34 - Στιγμιότυπο πρόβλεψης της αέργου ισχύος του W4 .....	72
Εικόνα 35 - Σύγκριση του MAE των μοντέλων για την πρόβλεψη της αέργου ισχύος του W4 ανά χρονικό διάστημα της ημέρας .....	73
Εικόνα 36 - Σύγκριση του MAE των μοντέλων για την πρόβλεψη της ενεργού ισχύος του W6 ανά χρονικό διάστημα της ημέρας .....	73
Εικόνα 37 -Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της ενεργού ισχύος του W6 .....	74
Εικόνα 38 - Στιγμιότυπο πρόβλεψης της ενεργού ισχύος του W6.....	75
Εικόνα 39 - Στιγμιότυπο πρόβλεψης της ενεργού ισχύος του W6.....	75
Εικόνα 40 - Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της αέργου ισχύος του W6 .....	76

---

*Κατάλογος Πινάκων*

---

Πίνακας 1 - Διαφορές Παραδοσιακού και Έξυπνου Δικτύου.....	22
Πίνακας 2 - Τιμές υπερπαραμέτρων μοντέλου N-BEATS .....	59
Πίνακας 3 -Τιμές παραμέτρων μοντέλου TCN.....	60
Πίνακας 4 -Δείκτες ακριβείας πρόβλεψης της ενεργούς ισχύος του φορτίου του παρόχου με το μοντέλο LightGBM, ανά περίπτωση.....	62
Πίνακας 5 - Δείκτες ακριβείας πρόβλεψης της αέργου ισχύος του φορτίου του παρόχου με το μοντέλο LightGBM, ανά περίπτωση.....	62
Πίνακας 6 - Δείκτες ακριβείας πρόβλεψης της ενεργούς ισχύος του φωτοβολταϊκού με το μοντέλο LightGBM ανα περίπτωση.....	63
Πίνακας 7 - Δείκτες ακριβείας πρόβλεψης της αέργου ισχύος του φωτοβολταϊκού με το μοντέλο LightGBM ανα περίπτωση.....	63
Πίνακας 8 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W4 με το μοντέλο N-BEATS ανα περίπτωση.....	64
Πίνακας 9 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W6 με το μοντέλο N-BEATS ανά περίπτωση.....	65
Πίνακας 10 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W4 με το μοντέλο TCN ανά περίπτωση.....	66
Πίνακας 11 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W4 με το μοντέλο TCN ανά περίπτωση.....	66
Πίνακας 12 - Βέλτιστα μοντέλα για κάθε χρονοσειρά.....	76

# 1.Εισαγωγή

## 1.1. Πλαίσιο

Η ολοένα αυξανόμενη ζήτηση ενέργειας σε παγκόσμιο επίπεδο καθώς και η ενσωμάτωση περισσότερων ανανεώσιμων πηγών ενέργειας στα ηλεκτρικά δίκτυα, καθιστούν επιτακτική την ανάγκη για αύξηση της απόδοσης, της αξιοπιστίας αλλά και του ελέγχου των ενεργειακών δικτύων. Η παραγωγή ενέργειας είναι ένα πολυδιάστατο ζήτημα και συνδέεται με ποικίλες παραμέτρους οικονομικής, περιβαλλοντικής αλλά και κοινωνικής φύσεως. Με την εισαγωγή των ηλεκτρικών οχημάτων στην καθημερινότητα εκατομμυρίων ανθρώπων, τη συνεχή χρήση κλιματιστικών και συσκευών θέρμανσης λόγω της κλιματικής αλλαγής αλλά και την ευρύτερη ποικιλία ηλεκτρικών συσκευών που χρησιμοποιούνται, η ζήτηση ενέργειας έχει φτάσει σε πολύ υψηλά επίπεδα, γεγονός που οδηγεί σε υψηλότερες αιχμές κατανάλωσης. Η κάλυψη αυτών των αιχμών αποτελεί πρόκληση για τις εταιρείες παραγωγής ηλεκτρικής ενέργειας. Επιπρόσθετα, η παραγωγή ενέργειας από ανανεώσιμες πηγές, και ιδιαίτερα από διακοπτόμενες πηγές, όπως η αιολική και η ηλιακή ενέργεια, χαρακτηρίζεται από ευμεταβλητότητα και αστάθεια. Το σύστημα παραγωγής ηλεκτρικής ενέργειας καθίσταται, έτσι, λιγότερο αξιόπιστο καθώς αυξάνεται η αβεβαιότητα. Κλειδί για την αντιμετώπιση αυτών των προβλημάτων και τη βελτίωση της απόδοσης και της αξιοπιστίας του ηλεκτρικού δικτύου, είναι η σταδιακή μετατροπή του σε ένα έξυπνο δίκτυο, με την ενσωμάτωση ψηφιακών τεχνολογιών.

Τα έξυπνα δίκτυα είναι τα ηλεκτρικά δίκτυα που χρησιμοποιούν αναπτυγμένες τεχνολογίες για τη βελτιστοποίηση της παραγωγής, της διανομής και της κατανάλωσης της ηλεκτρικής ενέργειας. Δομικό στοιχείο των έξυπνων δικτύων αποτελεί η προηγμένη μετρητική υποδομή (Advanced Metering Infrastructure). Πρόκειται για μια συστάδα τεχνολογιών, με βασική αυτή των έξυπνων μετρητών (smart meters). Οι μετρητές αυτοί, είναι συσκευές που συλλέγουν δεδομένα και πραγματοποιούν μετρήσεις σε πραγματικό χρόνο, με την επιθυμητή συχνότητα. Τα βασικά πλεονεκτήματα τους έναντι των συμβατικών μετρητών είναι ότι παρέχουν αμφίδρομη ροή δεδομένων τόσο προς τον καταναλωτή όσο και προς τον παραγωγό. Δίνεται, έτσι, η δυνατότητα αφενός στους καταναλωτές να ελέγχουν την κατανάλωσή τους, αφού βλέπουν σε πραγματικό χρόνο το κόστος λειτουργίας κάθε συσκευής, και αφετέρου στους παρόχους ενέργειας να υπολογίζουν καλύτερα το ισοζύγιο ισχύος, μέσω της αλληλεπίδρασης ζήτησης και παραγωγής, να ανιχνεύουν και να διορθώνουν ταχύτερα τυχόν βλάβες, προσδίδοντας έτσι, τελικά, μεγαλύτερη αξιοπιστία στο ηλεκτρικό δίκτυο και μειώνοντας το κόστος λειτουργίας. Επιπλέον, στοιχείο των έξυπνων δικτύων αποτελεί η ενσωμάτωση ανανεώσιμων πηγών ενέργειας και συστημάτων αποθήκευσης ενέργειας, ώστε το δίκτυο να μη βασίζεται αποκλειστικά σε ορυκτά καύσιμα και να είναι λιγότερο ευάλωτο σε διακοπές.

Μαζί με τα έξυπνα δίκτυα, για την καλύτερη απόδοση και αξιοπιστία του δικτύου, κρίσιμη είναι και η ανάπτυξη προγνωστικών μοντέλων για την πρόβλεψη της ενέργειας. Αξιοποιώντας τα δεδομένα που παρέχονται από τα έξυπνα δίκτυα για την κατανάλωση της ενέργειας, τη

συμπεριφορά των καταναλωτών αλλά και δεδομένα όπως τα μετεωρολογικά, που είτε προέρχονται από το ίδιο το δίκτυο, μέσω αισθητήρων, είτε από εξωτερικές πηγές, κ.α., είναι δυνατόν να εντοπιστούν μοτίβα και να αναπτυχθούν μοντέλα πρόβλεψης, που δίνουν την δυνατότητα ακριβέστερου προγραμματισμού και σχεδίασης των συστημάτων.

Συνοψίζοντας, σε μια εποχή που οι τιμές και η ζήτηση της ενέργειας βρίσκονται σε πρωτοφανή υψηλά επίπεδα και οι ανανεώσιμες πηγές ενέργειας διεισδύουν όλο και περισσότερο στον ενεργειακό τομέα, κρίνονται αναγκαίες τόσο η ανάλυση όσο και η πρόβλεψη του φορτίου ώστε να εξασφαλιστεί η εύρυθμη λειτουργία αλλά και η εξέλιξη του κλάδου της ενέργειας.

## 1.2 Σκοπός

Σκοπός αυτής της εργασίας είναι αφενός η βραχυπρόθεσμη πρόβλεψη πολλαπλών συνόλων δεδομένων που αφορούν χρονοσειρές ηλεκτρικής ενέργειας και συλλέγονται από smart meters, αξιοποιώντας την τεχνητή νοημοσύνη και αφετέρου η ανάλυση, ερμηνεία και βιωσιμότητα των μεθόδων.

Τα βασικά ερωτήματα που τίθενται προς έρευνα είναι ποια μοντέλα μηχανικής ή βαθιάς μάθησης είναι πιο αποδοτικά στην πρόβλεψη αυτών των δεδομένων, ποιες είναι οι παράμετροι που βελτιστοποιούν την απόδοσή τους αλλά και ποιες είναι οι συμμεταβλητές που επιδρούν θετικά στην επίδοσή τους.

Η απάντηση των παραπάνω ερωτημάτων προκύπτει τόσο μέσα από τη βιβλιογραφική έρευνα όσο και από το πρακτικό κομμάτι της εργασίας. Τα συμπεράσματα της βιβλιογραφικής έρευνας χρησιμοποιούνται σε συνδυασμό με τεχνικές ανάλυσης δεδομένων για τον προσδιορισμό των μοντέλων που παρουσιάζουν το μικρότερο σφάλμα στην πρόβλεψη των χρονοσειρών.

## 1.3 Διάρθρωση εργασίας

Στο Κεφάλαιο 2 της εργασίας παρουσιάζεται η βιβλιογραφική έρευνα, περιγράφοντας τρία βασικά κομμάτια που αφορούν την εργασία, τα συστήματα ηλεκτρικής ενέργειας, τις χρονοσειρές και τις τεχνικές προβλέψεων, καθώς και μια βασική εισαγωγή στη μηχανική μάθηση και τα μοντέλα που θα χρησιμοποιηθούν.

Στο Κεφάλαιο 3 παρουσιάζεται η μεθοδολογία που ακολουθήθηκε ενώ στο Κεφάλαιο 4 τα αποτελέσματα που προέκυψαν από τις διάφορες μεθόδους.

Στο Κεφάλαιο 5 παρουσιάζονται τα συμπεράσματα όπως αυτά εξήχθησαν από τα αποτελέσματα του Κεφαλαίου 4 ενώ στο Κεφάλαιο 6 παρουσιάζονται οι μελλοντικές προεκτάσεις της εργασίας.

## 1.4. Σχετική Εργασία

Η πρόβλεψη ηλεκτρικής ενέργειας αποτελεί ένα θέμα που ερευνάται πολύ από την επιστημονική κοινότητα. Στην βιβλιογραφία οι έννοιες της πρόβλεψης φορτίου και της πρόβλεψης ενέργειας και ισχύος πολλές φορές συγχέονται καθώς είναι άμεσα συνδεδεμένες. Στις σχετικές εργασίες που ακολουθούν αναφέρονται εργασίες που εξερευνούν όλες αυτές τις έννοιες. Τα μοντέλα που

έχουν χρησιμοποιηθεί για την πρόβλεψη ηλεκτρικής ισχύος και φορτίου μπορούν να χωριστούν σε δύο κατηγορίες [1]: Η πρώτη κατηγορία αφορά συμβατικά στατιστικά μοντέλα και βασικά μοντέλα μηχανικής μάθησης (ML), ενώ η δεύτερη αφορά μοντέλα βαθιάς μάθησης (DL), state of the art τεχνικές μηχανικής (ML) και βαθιάς (DL) μάθησης καθώς και υβριδικά μοντέλα.

Σχετικά με την πρώτη κατηγορία, πολλά στατιστικά μοντέλα έχουν δοκιμαστεί για την πρόβλεψη ισχύος και φορτίου. Συγκεκριμένα, βασικό μοντέλο αποτελούσε, αρχικά, το ARIMA αφού χρησιμοποιήθηκε ευρύτατα τόσο στην πρόβλεψη ισχύος [2] όσο και στην βραχυπρόθεσμη πρόβλεψη φορτίου [3]. Ωστόσο, το μοντέλο αυτό παρουσιάζει αδυναμία στην ενσωμάτωση εξωγενών μεταβλητών και σύντομα οι εργασίες επεκτάθηκαν στην αξιολόγηση των μοντέλων ARMAX [4], [5] και ARIMAX [6], ώστε να καλυφθεί αυτή η αδυναμία, αφού η ενσωμάτωση εξωγενών μεταβλητών, όπως είναι τα καιρικά δεδομένα και οι χρονικές μεταβλητές είναι συχνά απαραίτητη για την αποδοτική πρόβλεψη ισχύος και φορτίου. Συγκεκριμένα στην εργασία [4] εξετάστηκε η απόδοση του μοντέλου ARMAX στην βραχυπρόθεσμη (1-day ahead) πρόβλεψη της ισχύος ενός φωτοβολταϊκού μέγιστης ισχύος 2.1kW, συνδεδεμένου σε δίκτυο, χρησιμοποιώντας ως εξωγενείς μεταβλητές την θερμοκρασία, το ποσοστό βροχόπτωσης, την ηλιοφάνεια και την υγρασία. Η εργασία καταλήγει με το συμπέρασμα ότι το μοντέλο ARMAX παρουσιάζει σαφώς καλύτερη απόδοση από το ARIMA μοντέλο. Επιπλέον, στην πρόβλεψη ισχύος και φορτίου έχει χρησιμοποιηθεί και η μέθοδος της εκθετικής εξομάλυνσης [7],[8].

Στα βασικά μοντέλα μηχανικής μάθησης που έχουν χρησιμοποιηθεί για πρόβλεψη ισχύος και φορτίου ανήκουν κυρίως οι μηχανές υποστήριξης διανυσμάτων (support vector machines, SVM) [9], [10] καθώς και τεχνητά νευρωνικά δίκτυα (ANN) [11]–[13]. Ακόμα, για την επίλυση του προβλήματος της παρούσας εργασίας έχει γίνει ευρεία χρήση μοντέλων της οικογένειας δέντρων αποφάσεων και συγκεκριμένα, των random forest [14] και των gradient boosting regression trees [15].

Σχετικά με τα μοντέλα της δεύτερης κατηγορίας, για την βραχυπρόθεσμη πρόβλεψη ισχύος και φορτίου έχουν εκτενώς χρησιμοποιηθεί τόσο μοντέλα βαθιάς μάθησης και state-of-the-art μέθοδοι όσο και υβριδικά μοντέλα. Συχνά χρησιμοποιούνται υβριδικά μοντέλα με Temporal Convolutional Networks (TCN) ή Recurrent Neural Network (RNN) [16], [17], [18] για να μπορέσουν να εντοπίσουν αποτελεσματικότερα συσχετισμούς μεταξύ των δεδομένων ισχύος και των μετεωρολογικών δεδομένων. Τέλος άλλη μια οικογένεια μοντέλων που έχουν χρησιμοποιηθεί, πρόσφατα, για το πρόβλημα πρόβλεψης φορτίου βραχυπρόθεσμου χρόνου είναι τα νευρωνικά δίκτυα Γενετικών Αλγορίθμων (GA-NN) [29], [30], καθώς οι γενετικοί αλγόριθμοι μειώνουν σημαντικά τον χρόνο εκπαίδευσης των μοντέλων [19] και επιλέγουν τις παραμέτρους εκπαίδευσης των υβριδικών μοντέλων, νευρωνικών δικτύων Γενετικών Αλγορίθμων [20].

## 2.Βιβλιογραφική Ανασκόπηση

### 2.1. Συστήματα Ηλεκτρικής Ενέργειας

#### 2.1.1. Εισαγωγή

Η εμφάνιση της ηλεκτροδότησης έχει, δικαίως, χαρακτηριστεί ως το πιο σημαντικό επίτευγμα της Μηχανικής του 20ου αιώνα. Η σημασία της ηλεκτρικής ενέργειας είναι ανεκτίμητη, καθώς αποτελεί δομικό στοιχείο της σύγχρονης κοινωνίας, διευκολύνοντας την οικονομική ανάπτυξη, την τεχνολογική καινοτομία και βελτιώνοντας την ποιότητα ζωής των ανθρώπων σε όλο τον κόσμο. Η παραγωγή της ηλεκτρικής ενέργειας ξεκίνησε στα τέλη του 19ου αιώνα, με την πρώτη μονάδα παραγωγής ηλεκτρικού ρεύματος να δημιουργείται στο Λονδίνο, το 1882. Στην Ελλάδα ωστόσο έφτασε επτά χρόνια αργότερα, το 1889, οπότε και φωτίστηκε για πρώτη φορά το κέντρο της Αθήνας. Η εξέλιξη της παραγωγής ήταν σταδιακή και το 1950, με την ίδρυση της ΔΕΗ, οι δραστηριότητες παραγωγής, μεταφοράς και διανομής συγκεντρώθηκαν σε έναν δημόσιο φορέα. Τότε ξεκίνησαν οι διαδικασίες για την αξιοποίηση των εγχώριων πηγών ενέργειας και η ενοποίηση των δικτύων μεταφοράς ηλεκτρικής ενέργειας σε ένα εθνικό διασυνδεδεμένο Σύστημα. Το σύστημα αυτό διατηρεί μέχρι σήμερα τα βασικά του στοιχεία, παρά τις τεχνολογικές εξελίξεις των τελευταίων χρόνων. [21]

#### 2.1.2. Σύστημα Ηλεκτρικής Ενέργειας

Ως Σύστημα Ηλεκτρικής Ενέργειας ορίζεται ένα δίκτυο από εγκαταστάσεις παραγωγής, μεταφοράς και διανομής ηλεκτρικής ενέργειας, που εξασφαλίζει την παροχή ενέργειας προς τους καταναλωτές, σε όλη την περιοχή που εξυπηρετεί το δίκτυο [21].

Οι βασικές λειτουργίες ενός τέτοιου δικτύου είναι οι ακόλουθες:

- Παραγωγή Ηλεκτρικής Ενέργειας

Στο σύστημα παραγωγής ηλεκτρικής ενέργειας ανήκουν όλες οι μονάδες ηλεκτροπαραγωγής που ηλεκτροδοτούν το σύστημα. Το σύστημα παραγωγής αποτελείται κυρίως από γεννήτριες και από μετασχηματιστές ισχύος. Οι γεννήτριες χρησιμοποιώντας διάφορες πηγές όπως ορυκτά καύσιμα, φυσικό αέριο, ανανεώσιμες πηγές ενέργειας κ.α., μετατρέπουν την εκάστοτε μορφή ενέργειας - ανάλογα με την πηγή που αξιοποιούν - σε ηλεκτρική και οι μετασχηματιστές χρησιμοποιούνται για την μεταφορά του ρεύματος, ανυψώνοντας την τάση σε υψηλή, ώστε να ελαχιστοποιούνται οι απώλειες.

- Μεταφορά Ηλεκτρικής Ενέργειας

Το σύστημα μεταφοράς ηλεκτρικής ενέργειας είναι υπεύθυνο για την σύνδεση των



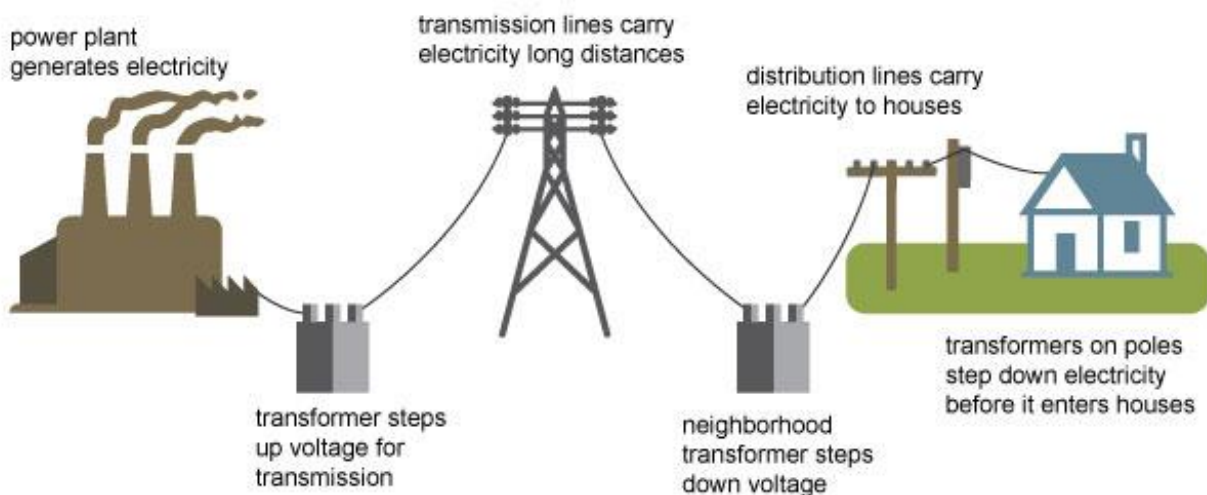
σταθμών παραγωγής και την μεταφορά της ηλεκτρικής ενέργειας από τα σημεία παραγωγής στα σημεία κατανάλωσης. Αποτελείται από τις γραμμές μεταφοράς υψηλής και υπερυψηλής τάσης, που είναι στην πλειονότητά τους εναέριες και σπανιότερα υπόγειες, με τυπικά επίπεδα μεταφοράς 220kV έως τα 750kV, τους υποσταθμούς ζεύξης των δικτύων αυτών και τους υποσταθμούς μετασχηματισμού μεταξύ των διαφόρων επιπέδων τάσεων που χρησιμοποιούνται στο δίκτυο μεταφοράς. Στο σύστημα μεταφοράς θα συμπεριλάβουμε και τις γραμμές υπομεταφοράς, που μεταφέρουν υποβιβασμένη τάση στους υποσταθμούς διανομής.

- Διανομή Ηλεκτρικής Ενέργειας

Το σύστημα διανομής είναι υπεύθυνο για την παραλαβή της ηλεκτρικής ενέργειας από τις γραμμές μεταφοράς και τη διανομή της στους καταναλωτές. Χωρίζεται στην πρωτογενή διανομή - ή διανομή μέσης τάσης - και στην δευτερογενή διανομή - ή διανομή χαμηλής τάσης. Περιλαμβάνει τους σταθμούς διανομής στους οποίους γίνεται υποβιβασμός της τάσης, είτε από υψηλή σε μέση είτε από μέση σε χαμηλή. Στη Μέση Τάση συνδέονται παροχές ισχύος άνω των 100kW, όπως βιοτεχνίες και βιομηχανίες (τριφασικές παροχές), ενώ η Χαμηλή Τάση τροφοδοτεί καταναλωτές, κυρίως οικιακούς, και μεγάλο μέρος των πελατών εμπορικής χρήσης (μονοφασικές παροχές). Στο ελληνικό ΣΗΕ, η Μέση Τάση βρίσκεται στα 20kV(παλιότερα στα 15kV), και η Χαμηλή Τάση στα 190V – 380V.

Στην Εικόνα 1 φαίνεται η σύνδεση των τριών βασικών δομών του ΣΗΕ.

## Electricity generation, transmission, and distribution



Source: Adapted from National Energy Education Development Project (public domain)

Εικόνα 1 - Βασικές δομές ενός ΣΗΕ

### 2.1.3. Ανάγκη πρόβλεψης φορτίου

Όπως έχει αναφερθεί ήδη, σκοπός ενός συστήματος ηλεκτρικής ενέργειας είναι η εξασφάλιση της ποιοτικής και αδιάλειπτης παροχής ηλεκτρικής ισχύος στους καταναλωτές. Για να συμβεί αυτό, είναι κρίσιμο κάθε στιγμή η παραγωγή ηλεκτρικής ενέργειας να είναι ίση με το φορτίο που καταναλώνεται, αλλά και να υπάρχει η αναγκαία εφεδρεία για να αντισταθμιστούν τυχόν διαλείψεις.

Το φορτίο ενός ΣΗΕ χωρίζεται σε τρεις κατηγορίες: το φορτίο αιχμής, το κυμαινόμενο και το φορτίο βάσης. Το φορτίο αιχμής αντιστοιχεί στο φορτίο που καταναλώνεται στην πολύ υψηλή ζήτηση και διαρκεί ένα πολύ σύντομο διάστημα της ημέρας. Καλύπτεται από μονάδες παραγωγής με γρήγορη αντίδραση στην ανάληψη φορτίου, όπως για παράδειγμα υδροηλεκτρικές μονάδες που εκμεταλλεύονται τη φυσική διαδικασία του κύκλου του νερού. Το φορτίο βάσης αντιστοιχεί στο ελάχιστο φορτίο που καταναλώνεται καθημερινά και δεν είναι παράμετρος της εποχής, της ώρας της μέρας κ.λ.π. αλλά αποτελεί σταθερή συνιστώσα. Καλύπτεται συνήθως από μεγάλες μονάδες παραγωγής, γνωστές ως και μονάδες βάσης (π.χ. λιγνιτικές), αλλά μπορεί να καλυφθεί και από συνδυασμό μονάδων σταθερής και μεταβλητής παραγωγής, αρκεί να εξασφαλίζεται ότι ο συνδυασμός τους παράγει αξιόπιστα συνεχή ηλεκτρική ισχύ. [22] Τέλος, το κυμαινόμενο φορτίο είναι το φορτίο που μεταβάλλεται συναρτήσει του χρονικού διαστήματος της ημέρας, της εποχής, των καιρικών συνθηκών αλλά και από άλλους εξωτερικούς παράγοντες.

Οι δυσκολίες που αντιμετωπίζουν οι εταιρείες διαχείρισης του φορτίου έγκεινται στο γεγονός ότι το συνολικό φορτίο των καταναλωτών είναι συνεχώς μεταβαλλόμενο. Συνήθως, για τον προσδιορισμό του αναμενόμενου φορτίου οι εταιρείες διαχείρισης λαμβάνουν υπόψη την ιστορική ζήτηση ενέργειας, τις τάσεις ανάπτυξης καθώς και την πρόβλεψη της μελλοντικής ζήτησης. Προκύπτει λοιπόν ότι η πρόβλεψη φορτίου είναι ζωτικής σημασίας στάδιο για την εξασφάλιση τόσο της εύρυθμης λειτουργίας όσο και της αποδοτικότητας του συστήματος της ηλεκτρικής ενέργειας.

### 2.1.4 Αδυναμίες του Παραδοσιακού Συστήματος Ηλεκτρικής Ενέργειας

Το παραδοσιακό σύστημα ηλεκτρικής ενέργειας στην Ελλάδα, αλλά και στις περισσότερες αναπτυγμένες χώρες, έχει αρκετές αδυναμίες, που με την πάροδο του χρόνου και την αλλαγή των απαιτήσεων και των αναγκών γίνονται ολοένα και περισσότερο αντιληπτές.

Το παραδοσιακό ηλεκτρικό δίκτυο βασίζεται στην κεντροποιημένη παραγωγή (centralized power generation), δηλαδή σε μεγάλους σταθμούς παραγωγής ηλεκτρικής ενέργειας, που βρίσκονται συνήθως μακριά από τα αστικά κέντρα, για περιβαλλοντικούς λόγους. Προκύπτει έτσι η ανάγκη μεταφοράς της ηλεκτρικής ενέργειας σε μεγάλες αποστάσεις, γεγονός που συντελεί σε καταγεγραμμένες απώλειες ενέργειας λόγω των γραμμών μεταφοράς και των διαφόρων μετασχηματιστών που χρησιμοποιούνται για την ανύψωση της τάσης. Οι απώλειες αυτές έχουν οικονομικό αλλά και περιβαλλοντικό αντίκτυπο καθώς γίνεται καύση ορυκτών

καυσίμων για την παραγωγή ενέργειας που τελικά ένα ποσοστό της χάνεται και δεν χρησιμοποιείται από τους καταναλωτές.

Ένα ακόμη σημαντικό μειονέκτημα του παραδοσιακού ΣΗΕ είναι η έλλειψη ευελιξίας και ελαστικότητας, καθώς η ροή της επικοινωνίας είναι μονόδρομη, από τον σταθμό παραγωγής προς το δίκτυο διανομής και στη συνέχεια στον καταναλωτή. Η έλλειψη αμφίδρομης επικοινωνίας οδηγεί σε μια σειρά προβλημάτων. Από τη μία μεριά δεν δίνεται η δυνατότητα στον καταναλωτή να ελέγχει σε μεγάλο βαθμό την κατανάλωσή του σε πραγματικό χρόνο και να επιστρέφει το περίσσειμα ενέργειας πίσω στο δίκτυο και από την άλλη οι εταιρείες διαχείρισης ενέργειας δεν μπορούν να παρακολουθούν την κατάσταση του δικτύου και τις ανάγκες των καταναλωτών σε πραγματικό χρόνο. Έτσι δεν μπορούν να προβλέπουν αποδοτικά τη ζήτηση της ενέργειας και να εντοπίζουν άμεσα τα προβλήματα που εμφανίζονται στην παροχή ενέργειας.

Οι παραπάνω αδυναμίες του παραδοσιακού ΣΗΕ οδήγησαν στην ανάγκη ψηφιοποίησης του δικτύου και στην γέννηση της ιδέας του έξυπνου δικτύου (smart grid).

#### 2.1.5. Το έξυπνο δίκτυο

Το δίκτυο της επόμενης γενιάς, ή αλλιώς έξυπνο δίκτυο (smart grid), προορίζεται να αντιμετωπίσει τις προαναφερθείσες αδυναμίες του παραδοσιακού συστήματος ηλεκτρικής ενέργειας.

Σύμφωνα με την Ευρωπαϊκή επιτροπή “European Technology Platform”, ως έξυπνο δίκτυο ορίζεται το ηλεκτρικό δίκτυο το οποίο μπορεί έξυπνα να ενσωματώνει τις ενέργειες όλων των συνδεδεμένων χρηστών που μπορεί είτε να παράγουν/ καταναλώνουν ηλεκτρική ισχύ είτε και τα δύο, με σκοπό την αποδοτική, οικονομική και ασφαλή παροχή ηλεκτρικών προμηθειών. Ένα έξυπνο δίκτυο χρησιμοποιεί καινοτόμα προϊόντα και υπηρεσίες σε συνδυασμό με έξυπνη καταγραφή, έλεγχο, επικοινωνία και τεχνολογίες αυτοϊασης (self-healing) ώστε να: [23]

- Διευκολύνει τη σύνδεση και την λειτουργία των μονάδων παραγωγής όλων των μεγεθών και τεχνολογιών
- Επιτρέπει στους καταναλωτές να συμμετέχουν στη βελτιστοποίηση της λειτουργίας του συστήματος
- Παρέχει στους καταναλωτές περισσότερη πληροφορία και την επιλογή ελέγχου της κατανάλωσής τους
- Μειώνει σημαντικά το περιβαλλοντικό αντίκτυπο ολόκληρου του συστήματος παροχής ηλεκτρικής ενέργειας
- Παρέχει ενισχυμένα επίπεδα αξιοπιστίας και ασφάλειας της παροχής ισχύος

Τα βασικά χαρακτηριστικά που περιγράφουν ένα έξυπνο δίκτυο είναι τα ακόλουθα:

- Βελτίωση Αξιοπιστίας

Η αξιοπιστία ενός δικτύου έγκειται στην ικανοποίηση των αναγκών των καταναλωτών. Σε ένα έξυπνο δίκτυο είναι δυνατός ο γρήγορος εντοπισμός βλαβών και η αυτοϊαση του

συστήματος. Με τον όρο αυτοΐαση εννοούμε την δυνατότητα του ΕΔ να εντοπίζει γρήγορα και αυτόματα, να απομονώνει το προβληματικό τμήμα του και να αντιμετωπίζει τις βλάβες και τις διαταραχές χωρίς την ανθρώπινη παρέμβαση. Αυτό κατορθώνεται με τον συνδυασμό προηγμένων αισθητήρων, δικτύων επικοινωνίας και συστημάτων ελέγχου που παρακολουθούν συνεχώς το δίκτυο και καταγράφουν ασυνήθιστες συμπεριφορές. Όταν μια βλάβη εντοπίζεται, αυτοματοποιημένοι αλγόριθμοι χρησιμοποιούνται για να αναλύσουν τα δεδομένα, να απομονώσουν το πρόβλημα και να αποφασίσουν τις ενέργειες στις οποίες πρέπει να προχωρήσουν προκειμένου να επαναφερθεί το σύστημα [24]–[26].

- Διαχείριση Ζήτησης

Το ΕΔ προσφέρει τη δυνατότητα στην πλευρά της ζήτησης, δηλαδή του καταναλωτή, να αλληλοεπιδρά με το σύστημα, χρησιμοποιώντας αμφίδρομη ροή πληροφορίας /δεδομένων. Έτσι, ο καταναλωτής μπορεί να ελέγχει την κατανάλωση του, να επιλέγει οικονομικότερους τρόπους λειτουργίας, μειώνοντας, για παράδειγμα, την κατανάλωσή του τις ώρες αιχμής. Τέτοιο παράδειγμα αποτελούν τα προγράμματα Real Time Pricing. Στα προγράμματα αυτά η τιμή του ρεύματος αλλάζει μέσα στη μέρα ανάλογα με την χονδρική τιμή του ρεύματος αλλά και τις συνθήκες που επικρατούν στο δίκτυο. Κινητοποιούνται, έτσι, οι καταναλωτές να μειώνουν την κατανάλωσή τους τις ώρες που η παραγωγή είναι ακριβή και η ζήτηση υψηλή [27].

Επιπλέον, ο καταναλωτής έχει την δυνατότητα να προσφέρει πίσω στο δίκτυο ενέργεια, που παράγει για παράδειγμα ο ίδιος στο σπίτι του με φωτοβολταϊκά πάνελ, βοηθώντας έτσι στην εξισορρόπηση του φορτίου με εξομάλυνση των αιχμών (στέλνοντας ισχύ πίσω στο δίκτυο τις ώρες υψηλής ζήτησης). Παράλληλα, μέσω της αμφίδρομης ροής δίνεται η δυνατότητα στην παραγωγή να διαχειρίζεται αποδοτικότερα το δίκτυο, γνωρίζοντας πλέον άμεσα τις ανάγκες των καταναλωτών [28].

- Εισαγωγή προηγμένης μετρητικής υποδομής (Advanced Metering Infrastructure) στο δίκτυο

Ο πιο σημαντικός μηχανισμός του έξυπνου δικτύου είναι αυτός της προηγμένης μετρητικής υποδομής. Όπως έχει αναφερθεί στο 1.1, πρόκειται για μια συστάδα τεχνολογιών, που αποτελείται από έξυπνους μετρητές, δίκτυα επικοινωνίας, συστήματα διαχείρισης των δεδομένων των μετρητών (MDMS), καθώς και μέσα για την ενσωμάτωση των συλλεγόμενων δεδομένων σε πλατφόρμες επικοινωνίας. Είναι δηλαδή ο μηχανισμός που παρέχει στο ΕΔ την δυνατότητα της αμφίδρομης ροής πληροφοριών [29]. Ο βασικός τρόπος λειτουργίας της είναι ο ακόλουθος:

Προηγμένοι μετρητές χρησιμοποιούνται για την καταγραφή των δεδομένων με την επιθυμητή συχνότητα. Στη συνέχεια μεταδίδουν τα δεδομένα που έχουν συλλέξει μέσω

συνηθισμένων δικτύων, όπως μέσω PLC<sup>1</sup>, στην υποδοχή της υποδομής. Εκεί, τα συστήματα διαχείρισης δεδομένων, προχωρούν στην αποθήκευση και ανάλυση των συλλεχθέντων δεδομένων ώστε να παραληφθούν σε χρηστική μορφή από τον διαχειριστή του δικτύου [30].

- Χρήση έξυπνων μετρητών

Οι μετρητές είναι ηλεκτρονικές συσκευές ικανές να συλλέγουν και να μετράνε δεδομένα σε χρονικά διαστήματα με την συχνότητα που ορίζει ο διαχειριστής του συστήματος, να επικοινωνούν μέσω απομακρυσμένων κέντρων πληροφοριών με τον διαχειριστή και να μεταδίδουν ποικίλες πληροφορίες σε επιθυμητά χρονικά διαστήματα. Έχουν αμφίδρομη ροή επικοινωνίας, που σημαίνει ότι μπορούν να δεχτούν εντολές και να ενεργήσουν αναλόγως. Από την πλευρά του καταναλωτή, οι έξυπνοι μετρητές με την μετάδοση των δεδομένων στις κατάλληλες πλατφόρμες επικοινωνίας, όπως π.χ. οθόνη συστήματος οικιακής χρήσης ενέργειας, παρουσιάζουν την κατανάλωση σε ένα διαδραστικό περιβάλλον, επιτρέποντας την παρακολούθηση και τον έλεγχο της [30]. Επιπλέον, δίνεται έτσι η δυνατότητα για ακριβή τιμολόγηση, καταργώντας την κατ' εκτίμηση χρέωση.

- Πλήρης ενσωμάτωση ανανεώσιμων πηγών ενέργειας

Σημαντικό πλεονέκτημα του ΕΗΔ έναντι του παραδοσιακού είναι η ευκολότερη ενσωμάτωση ΑΠΕ στο δίκτυο. Γενικά η ενσωμάτωση των ΑΠΕ καθίσταται δύσκολη λόγω της μεταβλητότητας και της ασυνέχειας της ισχύος που παράγουν. Η παραγωγή των ηλιακών, για παράδειγμα, πηγών εξαρτάται αποκλειστικά από την εποχή, την ώρα της ημέρας και τα καιρικά φαινόμενα. Το ΕΔ, μέσω των τεχνολογιών που αναφέρθηκαν προηγουμένως, έχει την δυνατότητα να ελέγξει και να υπολογίσει καλύτερα το ισοζύγιο ισχύος, μέσω αναλύσεων των δεδομένων και πρόβλεψης της παραγόμενης ισχύος [31]. Στον ακόλουθο πίνακα παρατίθενται συγκεντρωτικά οι διαφορές του παραδοσιακού με το έξυπνο δίκτυο, που έχουν αναλυθεί στις παραπάνω παραγράφους.

---

<sup>1</sup> Το σύστημα επικοινωνίας Powerline Carrier είναι ένα σύστημα που χρησιμοποιεί ως μέσο μετάδοσης τις καλωδιώσεις εναλλασσόμενου ηλεκτρικού ρεύματος [79].

	<i>Παραδοσιακό Δίκτυο</i>	<i>Έξυπνο Δίκτυο</i>
<b>Ροή Ισχύος</b>	Μονόδρομη ροή ισχύος από την παραγωγή στην κατανάλωση	Αμφίδρομη ροή
<b>Μετρητές</b>	Παραδοσιακοί Μετρητές	Έξυπνοι μετρητές, με αυτόματη καταγραφή σε πραγματικό χρόνο
<b>Παραγωγή</b>	Συγκεντρωτική παραγωγή	Διεσπαρμένη (distributed) παραγωγή
<b>Πρόβλεψη φορτίου</b>	Μη ακριβής πρόβλεψη, βασισζόμενη σε ιστορικά δεδομένα	Προηγμένα μοντέλα πρόβλεψης, με χρήση δεδομένων πραγματικού χρόνου
<b>Τιμολόγηση</b>	Κατά προσέγγιση	Ακριβής
<b>Ενσωμάτωση ΑΠΕ</b>	Περιορισμένη δυνατότητα για ενσωμάτωση	Πλήρης ενσωμάτωση ΑΠΕ
<b>Ροή Πληροφορίας</b>	Μονόδρομη ροή	Αμφίδρομη ροή
<b>Αποκατάσταση βλαβών</b>	Χειροκίνητη	Αυτόματη
	Φαινόμενα υπερφόρτωσης	Βιωσιμότητα και ασφάλεια

Πίνακας 1 - Διαφορές Παραδοσιακού και Έξυπνου Δικτύου

## 2.2 Χρονοσειρές και Τεχνικές Προβλέψεων

### 2.2.1. Εισαγωγή στις Χρονοσειρές

Χρονοσειρά (timeseries) είναι το σύνολο της καταγραφής παρατηρήσεων της τιμής ενός φυσικού ή άλλου μεγέθους, ανά ίσα διαδοχικά χρονικά διαστήματα. Σημαντικό χαρακτηριστικό των χρονοσειρών είναι ότι οι διαδοχικές αυτές παρατηρήσεις δεν είναι ανεξάρτητες μεταξύ τους. Αυτό το χαρακτηριστικό δίνει τη δυνατότητα της πρόβλεψης των μελλοντικών τιμών από τις προηγούμενες. Παραδείγματα χρονοσειρών εντοπίζονται σε πολλούς επιστημονικούς κλάδους, όπως για παράδειγμα στο χώρο της οικονομίας το ετήσιο ΑΕΠ μιας χώρας, το οποίο, καθώς υπολογίζεται κάθε τρεις μήνες, ανήκει στις τριμηνιαίες χρονοσειρές [32].

### 2.2.2. Ποιοτικά Χαρακτηριστικά Χρονοσειρών

Μέσω της ανάλυσης των μοτίβων και της συμπεριφοράς που παρουσιάζουν οι χρονοσειρές μπορούν να εξαχθούν πολύτιμες πληροφορίες που βοηθούν στη διαδικασία πρόβλεψης. Βασικός, λοιπόν, όρος για την κατανόηση των δεδομένων των χρονοσειρών είναι η εξερεύνηση των

ποιοτικών τους χαρακτηριστικών. Τα βασικά χαρακτηριστικά που προκύπτουν από την επισκόπηση του γραφήματος της χρονοσειράς στο πεδίο του χρόνου είναι τα ακόλουθα:

- **Τάση**

Τάση μιας χρονοσειράς ορίζουμε την μεταβολή του μέσου επιπέδου των τιμών της, σε μακροπρόθεσμο επίπεδο. Μια χρονοσειρά μπορεί να είναι ανοδική, καθοδική ή ακόμη και σταθερή. Η τάση της χρονοσειράς μπορεί να είναι είτε γραμμική είτε μη γραμμική (εκθετική, λογαριθμική κ.λ.π.). Η δυσκολία στον εντοπισμό της τάσης έγκειται στο γεγονός ότι χρειάζεται να υπάρχει επαρκής αριθμός δεδομένων για ένα ικανοποιητικό χρονικό διάστημα ώστε να μπορεί να χαρακτηριστεί ως μακροπρόθεσμο [33].

- **Εποχιακότητα**

Με τον όρο εποχιακότητα μιας χρονοσειράς αναφερόμαστε στην επανάληψη ενός μοτίβου με μια σταθερή συχνότητα, μικρότερη του έτους. Γίνεται λοιπόν αντιληπτό ότι μπορεί να παρατηρηθεί μόνο σε χρονοσειρές με συχνότητα παρατήρησης όχι μεγαλύτερη του ενός έτους. Ένα παράδειγμα εποχιακότητας αποτελεί η παραγωγή ισχύος από τα φωτοβολταϊκά πάνελ, τα οποία έχουν έντονη άνοδο τις πρωινές ώρες της ημέρας και απόλυτη πτώση τις υπόλοιπες. Η εποχιακότητα σε μια χρονοσειρά συνήθως είναι εύκολα κατανοητή και προβλέψιμη, γεγονός που καθιστά τη μέτρησή της σχετικά εύκολη, με αποτέλεσμα να γίνεται δυνατή η απομόνωσή της από την χρονοσειρά, ώστε να προκύψει τελικά η αποεποχικοποιημένη χρονοσειρά [32]. Οφείλεται σε παράγοντες όπως η εποχή, η ώρα της ημέρας και τα καιρικά μοτίβα.

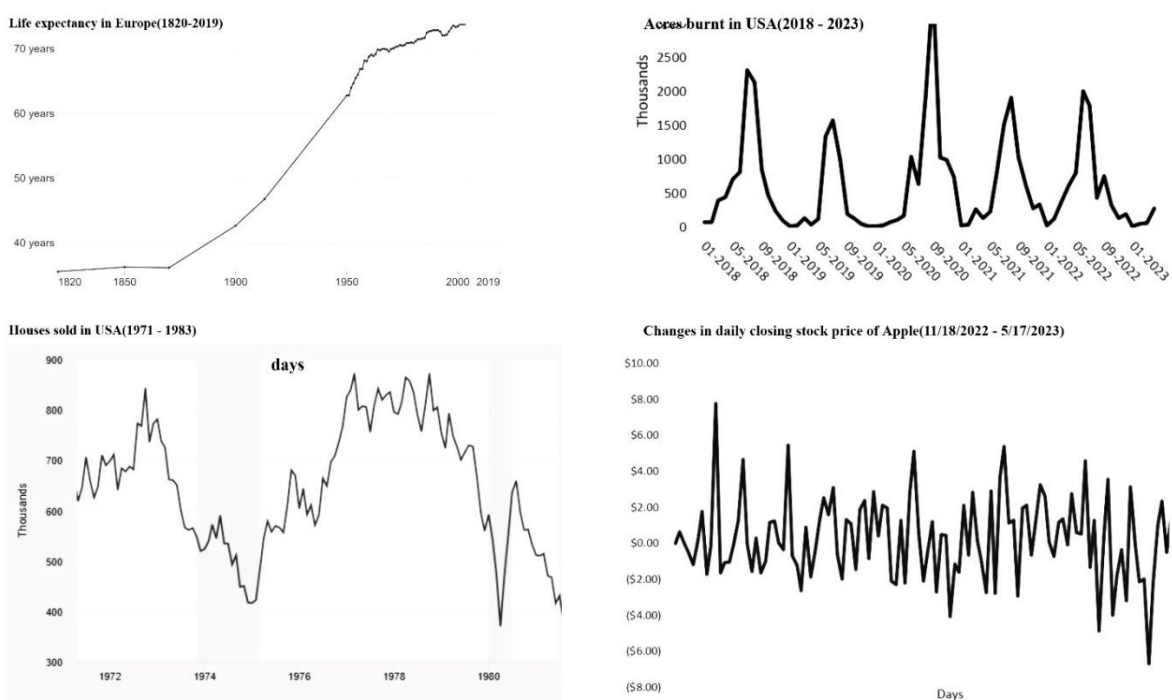
- **Κυκλικότητα**

Ένα ακόμη σημαντικό ποιοτικό χαρακτηριστικό των χρονοσειρών αποτελεί η κυκλικότητα. Πρόκειται για τη διακύμανση που εμφανίζεται στις τιμές μιας χρονοσειράς, καταγράφεται κατά περιόδους, με μήκος περιόδου κατά κανόνα μεγαλύτερο του έτους και οφείλεται σε ειδικούς εξωγενείς παράγοντες, όπως για παράδειγμα οι οικονομικές συνθήκες. Οι περίοδοι, σε αντίθεση με τις περιόδους της εποχιακότητας, δεν είναι απαραίτητα σταθερές. Για παράδειγμα, οι χρονοσειρές των περισσότερων οικονομικών μεγεθών εμφανίζουν κυκλικότητα [33].

- **Τυχαιότητα**

Η τυχαιότητα στη χρονοσειρά, ή αλλιώς οι μη κανονικές διακυμάνσεις, αφορούν στις διακυμάνσεις που απομένουν όταν από τη χρονοσειρά αφαιρεθούν τα ποιοτικά χαρακτηριστικά της τάσης, εποχιακότητας και κυκλικότητας. Μπορεί να οφείλονται είτε σε κάποια εντελώς τυχαία μεταβλητή είτε σε outliers, ή special events, δηλαδή σε απομονωμένες παρατηρήσεις που εμφανίζονται ως απότομες αλλαγές στο πρότυπο συμπεριφοράς της χρονοσειράς, με παροδικό χαρακτήρα. Τέλος, πιθανή είναι και η εμφάνισή της, εξ ' αιτίας μόνιμων ασυνεχειών, ή αλλιώς αλλαγών επιπέδου (level shifts) που επηρεάζουν το μέσο επίπεδο των τιμών της χρονοσειράς με απότομο τρόπο.

Στην Εικόνα 2 παρουσιάζονται κάποιες χρονοσειρές με τα ποιοτικά χαρακτηριστικά που αναλύθηκαν. Όπως φαίνεται, η χρονοσειρά που βρίσκεται πάνω αριστερά αντιστοιχεί στο προσδόκιμο ζωής στην Ευρώπη και η ανοδική τάση της είναι ξεκάθαρη. Η χρονοσειρά που βρίσκεται πάνω δεξιά έχει έντονη εποχιακότητα, με περίοδο 12 μηνών, γεγονός που δικαιολογείται, αφού αντιπροσωπεύει τις καμένες περιοχές των Η.Π.Α οι οποίες εμφανίζουν αύξηση κατά τους καλοκαιρινούς μήνες. Τέλος η σειρά κάτω αριστερά βλέπουμε ότι παρουσιάζει κυκλικότητα, με περίοδο περίπου τα έξι χρόνια, ενώ στην κάτω δεξιά εικόνα, η χρονοσειρά που αντιπροσωπεύει τις αλλαγές στην τιμή κλεισίματος μετοχής της Apple διακρίνεται από έντονη τυχαιότητα.



Εικόνα 2 - Παραδείγματα Χρονοσειρών με έντονα ποιοτικά χαρακτηριστικά

### 2.2.3 Αποσύνθεση χρονοσειρών

Όπως αναφέρθηκε στην προηγούμενη υποενότητα, οι χρονοσειρές έχουν κάποια βασικά ποιοτικά χαρακτηριστικά. Η απομόνωση αυτών των χαρακτηριστικών είναι πολύ βοηθητική τόσο στην κατανόηση όσο και στην ανάλυση των διαφόρων μοτίβων που υπάρχουν στα δεδομένα. Η διαδικασία αυτή, της απομόνωσης, ονομάζεται αποσύνθεση της χρονοσειράς και είναι ιδιαίτερα σημαντική στην πρόβλεψη των δεδομένων. Μαθηματικά η χρονοσειρά μπορεί να εκφραστεί ως συνάρτηση των ποιοτικών της χαρακτηριστικών :



$$Y(t) = F(Tt, Ct, St, Rt)$$

όπου

$Yt$  : η παρατήρηση την χρονική στιγμή  $t$

$Tt$  : η συνιστώσα της τάσης την χρονική στιγμή  $t$

$Ct$  : η συνιστώσα της κυκλικότητας την χρονική στιγμή  $t$

$St$ : η συνιστώσα της εποχιακότητας την χρονική στιγμή  $t$

$Rt$ : η συνιστώσα της τυχαιότητας την χρονική στιγμή  $t$

Οι πιο απλές μορφές της συνάρτησης  $F$  είναι αυτή του προσθετικού και του πολλαπλασιαστικού μοντέλου, οι οποίες διατυπώνονται, αντίστοιχα:

*Προσθετική*

$$Y(t) = T(t) + C(t) + S(t) + R(t)$$

*Πολλαπλασιαστική*

$$Y(t) = T(t) * C(t) * S(t) * R(t)$$

Η προσθετική αποσύνθεση είναι πιο αποτελεσματική σε χρονοσειρές όπου οι μέγιστες τιμές των εποχιακών δεδομένων δεν έχουν μεγάλη απόκλιση, ενώ η πολλαπλασιαστική σε χρονοσειρές που τα εποχιακά δεδομένα έχουν μεγαλύτερες διαφορές στους εποχιακούς κύκλους [34].

Για διευκόλυνση, στην παρακάτω θεωρητική ανάλυση θα θεωρήσουμε χρονοσειρά δεδομένης πολλαπλασιαστικής σχέσης των συνιστωσών.

Η διαδικασία της αποσύνθεσης αποτελείται από επτά βασικά βήματα [32]:

1. Υπολογισμός ενός κινητού μέσου όρου, με βάση το μήκος της εποχιακότητας.

$$KMO(n)t = \frac{1}{n} \sum_{i=-(n \bmod 2)}^{(n \bmod 2)} Y_{t+i} \quad , \quad \text{όπου } n \text{ το μήκος του κινητού μέσου όρου}$$

Λαμβάνουμε έτσι μια καλή εκτίμηση της σειράς τάσης-κύκλου αφού επιλέγοντας το μήκος του ΚΜΟ βάσει του μήκους της εποχιακότητας, απομακρύνεται η εποχιακότητα από τον υπολογιζόμενο κινητό μέσο όρο. Παράλληλα, καθώς η τυχαιότητα αντιπροσωπεύεται από τυχαίες διακυμάνσεις γύρω από τον μέσο όρο των παρατηρήσεων, με τη χρήση του ΚΜΟ η συνιστώσα της τυχαιότητας σχεδόν μηδενίζεται. Η εξάλειψη των συνιστωσών της τυχαιότητας και της εποχιακότητας οδηγεί στη σειρά τάσης κύκλου, οπότε μπορούμε να πούμε ότι:

$$KMO(n) = T * C$$

όπου

$KMO(n)$ : ο κινητός μέσος όρος μήκους  $n$

$T$ : η συνιστώσα της τυχειότητας

$C$ : η συνιστώσα της κυκλικότητας

Αν ωστόσο το μήκος της εποχιακότητας είναι άρτιος αριθμός, είναι προτιμότερο να γίνεται χρήση του κεντρικού κινητού μέσου όρου.

2. Προσδιορισμός των λόγων εποχιακότητας, με τη διαίρεση των πραγματικών παρατηρήσεων με τον ΚΜΟ, που όμως εμπεριέχουν και την τυχειότητα, καθώς είναι:

$$\frac{Y}{KMO(n)} = \frac{T * C * S * R}{T * C} = S * R$$

όπου  $T$ ,  $C$ ,  $S$ ,  $R$  οι συνιστώσες της τάσης, κυκλικότητας, εποχιακότητας και τυχειότητας, αντίστοιχα.

3. Απαλοιφή της τυχειότητας από τους λόγους εποχιακότητας, με τη χρήση του μέσου όρου των δεικτών εποχιακότητας που αντιστοιχούν στις ίδιες περιόδους του εποχιακού κύκλου. Αν για παράδειγμα πρόκειται για χρονοσειρά ημερήσιων δεδομένων, υπολογίζεται ο μέσος όρος των λόγων εποχιακότητας που αντιστοιχούν στην Δευτέρα, στη συνέχεια ο μέσος όρων των λόγων εποχιακότητας που αντιστοιχούν στην Τρίτη κ.ο.κ.. Αυτοί οι μέσοι όροι αναφέρονται συνήθως ως «δείκτες εποχιακότητας».
4. Διαίρεση των πραγματικών δεδομένων με τους αντίστοιχους δείκτες εποχιακότητας για την εύρεση της αποεποχικοποιημένης σειράς.
5. Απαλοιφή της τυχειότητας από την αποεποχικοποιημένη σειρά με τη χρήση ενός κινητού μέσου όρου, συνήθως διπλού κινητού μέσου όρου 3 x 3, για βέλτιστη εξομάλυνση και απαλοιφή της τυχειότητας.
6. Υπολογισμός της τάσης της χρονοσειράς. Εδώ ανάλογα με τα δεδομένα της χρονοσειράς, θα πρέπει να επιλεγθεί το μοντέλο τάσης που την αντιπροσωπεύει περισσότερο. Για παράδειγμα, αν η τάση είναι γραμμική, τότε εφαρμόζεται απλή παλινδρομική παλινδρόμηση.
7. Τέλος με τον προσδιορισμό της τάσης, μπορούμε εύκολα να υπολογίσουμε την κυκλικότητα, αφού έχοντας τη σειρά που περιέχει την τάση και την κυκλικότητα, καθώς

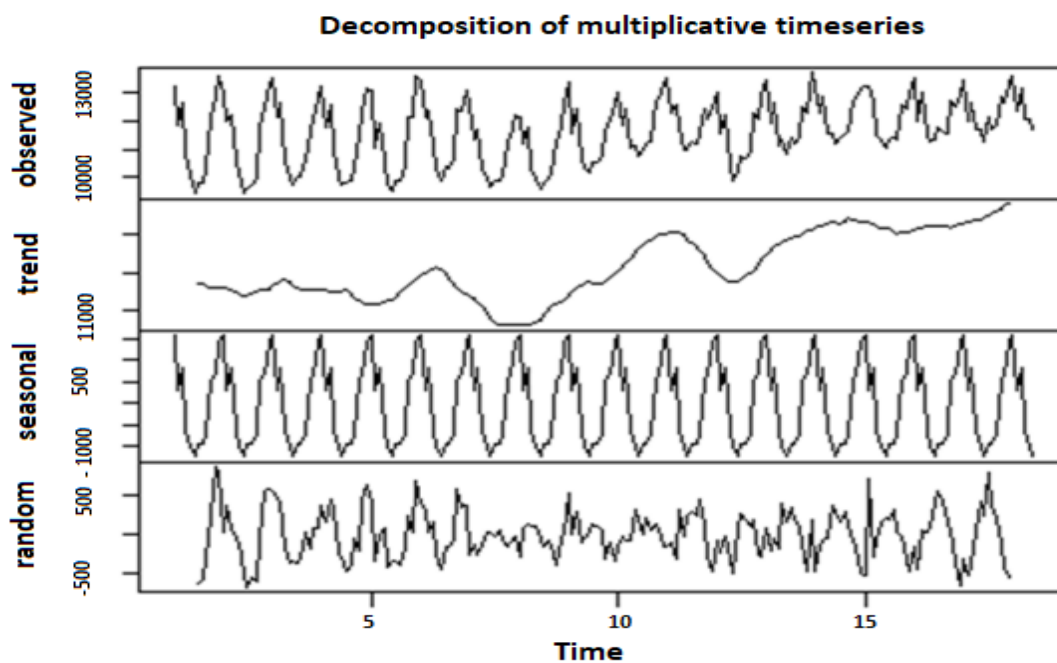
και την συνιστώσα της τάσης, με απλή διαίρεση παίρνουμε την συνιστώσα της κυκλικότητας.

Οι επιμέρους συνιστώσες της χρονοσειράς μετά την απομόνωσή τους από αυτήν, μπορούν να χρησιμοποιηθούν για την πρόβλεψη των μελλοντικών τιμών της:

$$F_{n+i} = T_t * C_t * S_t$$

Όπου  $F$  η πρόβλεψη της τιμής της χρονοσειράς την χρονική στιγμή  $n+i$ .

Η συνιστώσα της τυχαιότητας απουσιάζει από την πρόβλεψη αφού είναι ένας παράγοντας που δεν μπορεί να προβλεφθεί.



Εικόνα 3 -Αποσύνθεση χρονοσειράς στα βασικά της χαρακτηριστικά [15]

#### 2.2.4. Βασική Στατιστική Ανάλυση

Βασική προϋπόθεση για την βαθύτερη κατανόηση των χαρακτηριστικών των δεδομένων μιας χρονοσειράς αλλά και για την πιο σωστή πρόβλεψη των μελλοντικών τιμών της είναι ο υπολογισμός κάποιων βασικών στατιστικών δεικτών. Αυτή η υποενότητα περιλαμβάνει μια εισαγωγή στους βασικούς στατιστικούς δείκτες μιας χρονοσειράς  $Y$ , μεγέθους  $n$  παρατηρήσεων [32].

## Μέση Τιμή (Average)

Πρόκειται για τον απλό γραμμικό μέσο όρο των παρατηρήσεων και μας δείχνει το επίπεδο γύρω από το οποίο αυτές κινούνται.

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

## Μέγιστη και Ελάχιστη Τιμή (Maximum and Minimum)

Οι ακραίες τιμές της χρονοσειράς μας δίνουν μια πρώτη εκτίμηση της διακύμανσης των παρατηρήσεων αλλά και της τυχαιότητας που περιέχουν.

## Τυπική απόκλιση (Standard Deviation)

Εκφράζει το βαθμό που απέχουν οι παρατηρήσεις από την μέση τιμή, το πόσο διεσπαρμένες είναι δηλαδή γύρω από αυτήν.

$$\sigma_{\pi} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}, \text{ όπου } \pi \text{ ο πληθυσμός των παρατηρήσεων}$$

Σε περίπτωση που δεν είναι πρακτικό να υπολογιστεί η τυπική απόκλιση στον πληθυσμό  $\pi$  των παρατηρήσεων, πιθανόν λόγω του μεγάλου μεγέθους του πληθυσμού ή της μη διαθεσιμότητας όλων των παρατηρήσεων, μπορεί να υπολογιστεί η τυπική απόκλιση σε δείγμα των παρατηρήσεων, προσαρμόζοντας τον παρονομαστή της ως εξής:

$$\sigma_{\delta} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1}}, \text{ όπου } \delta \text{ το δείγμα των παρατηρήσεων}$$

Η προσαρμογή αυτή του παρονομαστή είναι γνωστή ως διόρθωση του Bessel (Bessel's correction) και γίνεται για την αμεροληψία του δείκτη[35]

## Διακύμανση (Variance)

Ορίζεται ως το τετράγωνο της τυπικής απόκλισης και εκφράζει την μέση τετραγωνική απόκλιση από την μέση τιμή. Καθώς μετριέται σε τετραγωνικές μονάδες και όχι στην ίδια μονάδα με τις πραγματικές παρατηρήσεις είναι πιο δύσκολα ερμηνεύσιμη από την τυπική απόκλιση, γι' αυτό και συνήθως προτιμάται η τελευταία.

## Συνδιακύμανση (Covariance)

Έστω δύο τυχαίες μεταβλητές  $X, Y$  με μέσες τιμές  $\bar{X}, \bar{Y}$  αντίστοιχα. Ένα μέτρο του τρόπου με τον οποίο οι δύο μεταβλητές μεταβάλλονται από κοινού. Μπορεί να μεταβάλλονται ανάλογα, αντιστρόφως ανάλογα ή να είναι ασυσχέτιστες μεταξύ τους, οπότε μιλάμε για θετική, αρνητική και μηδενική συνδιακύμανση, αντίστοιχα. Μαθηματικά διατυπώνεται ως εξής:

$$\text{cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n [(X_i - \bar{X}) * (Y_i - \bar{Y})]$$

Το σημαντικό μειονέκτημα της συνδιακύμανσης είναι ότι δε δείχνει πόσο ισχυρή είναι η συσχέτιση, αλλά μόνο την κατεύθυνσή της, καθώς εξαρτάται σε μεγάλο βαθμό από την κλίμακα των μεταβλητών [36].

### Συντελεστής Γραμμικής Συσχέτισης (Linear Correlation Coefficient)

Ένας δείκτης που εκφράζει τόσο την κατεύθυνση όσο και τον βαθμό της συσχέτισης μεταξύ δύο τυχαίων μεταβλητών  $X, Y$  είναι αυτός του συντελεστή γραμμικής συσχέτισης, ή αλλιώς του συντελεστή Pearson. Πρόκειται ωστόσο για έναν δείκτη που έχει σχεδιασθεί για να προσδιορίζει την γραμμική συσχέτιση των μεταβλητών και παίρνει τιμές στο διάστημα  $[-1, 1]$ .

$$r_{X,Y} = \frac{\sum_{i=1}^n [(X_i - \bar{X}) * (Y_i - \bar{Y})]}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} * \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Ανάλογα με την τιμή του, διακρίνονται οι εξής περιπτώσεις :

- Αν  $r = \pm 1$ , υπάρχει τέλεια γραμμική συσχέτιση
- Αν  $-1 < r \leq -0.8$  ή  $0.8 \leq r < 1$ , υπάρχει πολύ ισχυρή γραμμική συσχέτιση
- Αν  $-0.8 < r \leq -0.7$  ή  $0.7 \leq r < 0.8$ , υπάρχει ισχυρή γραμμική συσχέτιση
- Αν  $-0.7 < r \leq -0.5$  ή  $0.5 \leq r < 0.7$ , υπάρχει μέση ισχυρή γραμμική συσχέτιση
- Αν  $-0.5 < r \leq -0.3$  ή  $0.3 \leq r < 0.5$ , υπάρχει ασθενής γραμμική συσχέτιση
- Αν  $-0.3 < r < 0.3$ , δεν υπάρχει γραμμική συσχέτιση

### Συντελεστής αυτοσυσχέτισης (Autocorrelation Coefficient, ACF)

Ο συντελεστής αυτοσυσχέτισης δείχνει την συσχέτιση, τόσο τον βαθμό όσο και την κατεύθυνση, μεταξύ παρατηρήσεων της ίδιας μεταβλητής, που υστερούν χρονικά  $k$  περιόδους και παίρνει τιμές στο διάστημα  $[0, 1]$ . Αν η τιμή του είναι κοντά στην μονάδα, δείχνει μεγάλη συσχέτιση ενώ αν είναι κοντά στο μηδέν, δείχνει μηδενική συσχέτιση των παρατηρήσεων χρονικής υστέρησης  $k$ . Ο συντελεστής αυτός είναι σημαντικός για τον εντοπισμό της εποχιακότητας της χρονοσειράς, αφού οι παρατηρήσεις που αντιστοιχούν σε ίδια σημεία του εποχιακού κύκλου θα παρουσιάζουν ισχυρή αυτοσυσχέτιση.

### 2.2.5. Στατιστική Ανάλυση Ακριβείας Προβλέψεων

Έστω ότι το διάνυσμα  $Y$  αποτελεί τα πραγματικά δεδομένα και το διάνυσμα  $F$  τις προβλέψεις πλήθους  $n + h$ , όπου  $n$  το πλήθος των παρατηρήσεων και  $h$  ο ορίζοντας πρόβλεψης<sup>2</sup>. Οι στατιστικοί δείκτες που θα αναλυθούν παρακάτω, αναφέρονται στην ακρίβεια των προβλέψεων πέραν της περιόδου  $n$ , με δεδομένη την ύπαρξη νέων πραγματικών δεδομένων [32]:

- **Στατιστικοί δείκτες εξαρτώμενοι από την κλίμακα (scale-dependent metrics)**

Πρόκειται για στατιστικούς δείκτες που η κλίμακά τους εξαρτάται από την κλίμακα των δεδομένων. Είναι χρήσιμο να υπολογίζονται για την σύγκριση της απόδοσης διαφορετικών μοντέλων σε ίδια δεδομένα, αλλά δεν είναι αξιόπιστος δείκτης για την σύγκριση μοντέλων που προβλέπουν δεδομένα σε διαφορετικές κλίμακες. Σε αυτή την κατηγορία ανήκουν οι παρακάτω δείκτες:

#### **Μέσο Απόλυτο Σφάλμα (MAE, Mean Absolute Error)**

Δηλώνει ένα μέσο μέτρο της αστοχίας της πρόβλεψης, χωρίς να δίνεται ωστόσο ένδειξη για το αν το σφάλμα είναι θετικό ή αρνητικό. Επιπλέον, πρόκειται για ένα δείκτη ευαίσθητο στις ακραίες τιμές (outliers), αφού η ύπαρξη έστω και μίας μπορεί να έχει μεγάλη επίδραση στο μέγεθός του.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - F_i|$$

#### **Μέσο Τετραγωνικό Σφάλμα (MSE, Mean Squared Error)**

Όπως και στην περίπτωση του μέσου απόλυτου σφάλματος, ο δείκτης MSE επηρεάζεται πολύ από ακραίες τιμές, αλλά δίνει ακόμη μεγαλύτερο βάρος στα μεγάλα, αυτά, σφάλματα καθώς τετραγωνίζονται. Χρησιμοποιείται κυρίως για τον υπολογισμό των βέλτιστων παραμέτρων εξομάλυνσης.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - F_i)^2$$

Ο δείκτης αυτός, αν και χρήσιμος, πολλές φορές μπορεί να γίνει ευαίσθητος σε δεδομένα που εμπεριέχουν συνυπάρχοντα μοτίβα και θόρυβο [37].

#### **Ρίζα Μέσου Τετραγωνικού Σφάλματος (RMSE, Root Mean Squared Error)**

Αυτός ο δείκτης συχνά προτιμάται από το μέσο τετραγωνικό σφάλμα, καθώς είναι εκφρασμένος στις ίδιες μονάδες με τα δεδομένα της χρονοσειράς, οπότε είναι πιο εύκολη

---

<sup>2</sup> Ορίζοντας πρόβλεψης είναι το πλήθος των μελλοντικών χρονικών στιγμών που επιδιώκεται να προβλεφθεί.

η ερμηνεία του. Ωστόσο είναι κι αυτός περισσότερο ευαίσθητος στα μεγάλα σφάλματα σε σχέση με το μέσο απόλυτο σφάλμα.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - F_i)^2}$$

▪ **Στατιστικοί δείκτες βασισμένοι σε ποσοστιαία σφάλματα (percentage errors)**

Αυτοί οι δείκτες εκφράζονται σε ποσοστό επί τις εκατό (%) και έχουν το πλεονέκτημα ότι είναι ανεξάρτητοι της κλίμακας των δεδομένων. Λόγω αυτού μπορούν να χρησιμοποιηθούν αξιόπιστα για τη σύγκριση μοντέλων που προβλέπουν διαφορετικής κλίμακας δεδομένα. Επιπλέον, είναι χρήσιμοι όταν οι πραγματικές τιμές  $Y$  είναι ιδιαίτερα υψηλές. Στην κατηγορία αυτοί ανήκουν οι παρακάτω στατιστικοί δείκτες:

**Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (MAPE, Mean Absolute Percentage Error)**

Παίρνει τιμές στο διάστημα  $[0,1]$ , με τις χαμηλότερες τιμές να υποδεικνύουν καλύτερη απόδοση της μεθόδου πρόβλεψης.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - F_i}{Y_i} \right| * 100 \%$$

Ο δείκτης αυτός παρουσιάζει κάποια σημαντικά μειονεκτήματα, κυρίως λόγω του παρονομαστή του. Αρχικά αν έστω και μια χρονική στιγμή η τιμή της χρονοσειράς είναι μηδενική, τότε η τιμή του MAPE είναι ακαθόριστη. Επιπλέον δεν είναι αξιόπιστος δείκτης αν οι πραγματικές τιμές  $Y$  είναι κοντά στο μηδέν, καθώς έστω και μια μικρή διαφορά ανάμεσα στην πραγματική και την προβλεπόμενη τιμή, οδηγεί σε μεγάλα σφάλματα. Επιπλέον, δεν είναι συμμετρικός, καθώς λόγω του τρόπου υπολογισμού του δίνει μεγαλύτερη βαρύτητα στα θετικά σφάλματα.

**Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (SMAPE, Symmetric Mean Absolute Percentage Error)**

Ο δείκτης αυτός είναι μια παραλλαγή του MAPE, σχεδιασμένος έτσι ώστε να καλύπτει την μεροληψία τού MAPE προς τα θετικά σφάλματα, προσδίδοντάς του συμμετρικότητα. Αυτό πραγματοποιήθηκε με την απλή αντικατάσταση του παρονομαστή τού MAPE, με τη μέση τιμή των πραγματικών τιμών  $Y$  και των μελλοντικών  $F$ . [38] Υπολογίζεται, λοιπόν, ως εξής:

$$SMAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - F_i}{\frac{Y_i + F_i}{2}} \right| * 100 (\%) = \sum_{i=1}^n \left| \frac{2 * (Y_i - F_i)}{Y_i + F_i} \right| * 100 (\%)$$

- **Στατιστικοί δείκτες βασισμένοι σε σχετικά σφάλματα (relative errors)**

Σε αυτή την κατηγορία ανήκουν στατιστικοί δείκτες που βασίζονται σε σχετικά σφάλματα που προκύπτουν από τη διαίρεση των σφαλμάτων της μεθόδου πρόβλεψης που θέλουμε να αξιολογήσουμε, με τα σφάλματα μιας άλλης μεθόδου που θεωρούμε σημείο αναφοράς. Συνήθως χρησιμοποιείται η Αφελής (Naïve) μέθοδος, η οποία δίνει ως πρόβλεψη για την επόμενη περίοδο την πραγματική τιμή της τρέχουσας περιόδου.

**Μέσο Απόλυτο Κανονικοποιημένο Σφάλμα (MASE, Mean Absolute Scaled Error)**

Ο δείκτης MASE δίνει την ίδια βαρύτητα στα μικρά και στα μεγάλα σφάλματα, σε αντίθεση με τους δείκτες RMSE και MSE και συγχρόνως είναι απαλλαγμένος από τις περιπτώσεις απροσδιοριστίας των δεικτών MAPE και SMAPE [39]. Υπολογίζεται ως εξής:

$$MASE = \frac{\frac{1}{n} \sum_{i=1}^n |Y_i - F_i|}{\frac{1}{n-1} \sum_{i=2}^n |Y_i - Y_{i-1}|}$$

Όπου ο παρονομαστής αντιπροσωπεύει το μέσο απόλυτο σφάλμα της μεθόδου Naïve. Έτσι, όταν ο δείκτης MASE υπολογισθεί μεγαλύτερος της μονάδας, και συνεπώς ο παρονομαστής είναι μικρότερος του αριθμητή, υποδεικνύεται ότι η μέθοδος Naïve έχει καλύτερα αποτελέσματα ενώ όταν αυτός υπολογισθεί μικρότερος της μονάδας, τότε το μοντέλο πρόβλεψης που εξετάζουμε έχει καλύτερη απόδοση από την μέθοδο Naïve.

- **Άλλοι, τροποποιημένοι, δείκτες**

Στην κατηγορία αυτή ανήκουν οι στατιστικοί δείκτες που έχουν εισαχθεί σε διάφορες ερευνητικές εργασίες και χρησιμοποιούνται για να λύσουν ένα μειονέκτημα του αντίστοιχου δείκτη τον οποίο χρησιμοποιούν ως βάση.

**Κανονικοποιημένη Ρίζα Μέσου Τετραγωνικού Σφάλματος (nRMSE, Normalized Root Mean Square Error)**

Πρόκειται για μια παραλλαγή του κλασσικού στατιστικού δείκτη RMSE. Το βασικό μειονέκτημα που καλείται να αντιμετωπίσει είναι αυτό της εξάρτησης από την κλίμακα των δεδομένων. Υπολογίζεται ως εξής:

$$nRMSE = \frac{1}{\bar{y}} * RMSE = \frac{1}{\bar{y}} * \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - F_i)^2} , \text{ όπου}$$

$\frac{1}{\bar{y}}$  : ο παράγοντας κανονικοποίησης



Συνήθως, ως παράγοντας κανονικοποίησης επιλέγεται η μέση τιμή των πραγματικών δεδομένων ή η μέγιστη μείον την ελάχιστη τιμή ή, τέλος, η μέγιστη τιμή. Επιπλέον, η τιμή αυτή μπορεί να υπολογισθεί είτε στο σύνολο των δεδομένων είτε σε ένα μικρότερο διάστημα [40].

## 2.3. Μηχανική Μάθηση

### 2.3.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα γίνει βιβλιογραφική ανάλυση του τομέα της Μηχανικής Μάθησης (Machine Learning), ενός τομέα που, μαζί με την στατιστική, αποτελεί σημαντική πηγή μεθοδολογιών, οι οποίες βρίσκουν σημαντική εφαρμογή στον τομέα των προβλέψεων. Η Μηχανική Μάθηση είναι μια κατηγορία τεχνητής νοημοσύνης που επιτρέπει στους υπολογιστές να σκέφτονται και να μαθαίνουν χωρίς ανθρώπινη παρέμβαση.

Έχουν δοθεί πολλοί ορισμοί για την Μηχανική Μάθηση. Ο Arthur Samuel, ένας πρωτοπόρος ερευνητής στον τομέα τεχνητής νοημοσύνης, ήταν ο πρώτος που εισήγαγε τον όρο της μηχανικής μάθησης, το 1959. Συνέλαβε την ιδέα της ανάπτυξης ενός πρωτότυπου προγράμματος παιχνιδιού σκακιού που θα μπορούσε να κερδίζει τον εκάστοτε παγκόσμιο πρωταθλητή, για να καταδείξει την δύναμη των ηλεκτρονικών υπολογιστών. Αυτό το πρόγραμμα υπήρξε το πρώτο πρόγραμμα αυτομάθησης και αποτέλεσε το πρώιμο στάδιο της μηχανικής μάθησης [41], [42]. Πιο πρόσφατα, το 1997, ο Tom M. Mitchell έδωσε τον παρακάτω ορισμό [43]:

*Ορισμός : Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από εμπειρία  $E$  ως προς μια κλάση εργασιών  $T$  και ένα μέτρο επίδοσης  $P$ , αν η επίδοσή του σε εργασίες της κλάσης  $T$ , όπως αποτιμάται από το μέτρο  $P$ , βελτιώνεται με την εμπειρία  $E$ .*

Στην μηχανική μάθηση, αντί να τροφοδοτείται ο υπολογιστής με δεδομένα και κανόνες, να τα επεξεργάζεται και σύμφωνα με τους κανόνες να εξάγει αποτελέσματα, τροφοδοτείται με δεδομένα αλλά και με συμπεράσματα και, μέσω αυτών, εξάγει τους κανόνες [44].

Η μηχανική μάθηση βρίσκει εφαρμογή σε μια πληθώρα πεδίων. Στον τομέα της υγείας μια αξιοσημείωτη εφαρμογή είναι, για παράδειγμα, η δημιουργία μοντέλων για την ταξινόμηση των δερματικών βλαβών με την χρήση εικόνων [45]. Δύο ακόμη παραδείγματα εφαρμογής της μηχανικής μάθησης συναντάμε στον τομέα της όρασης των υπολογιστών, με την εισαγωγή της μηχανικής μάθησης για την αναγνώριση προσώπου [46] και στον τομέα της επεξεργασίας της φυσικής γλώσσας με σημαντικές εφαρμογές στην αυτόματη μετάφραση, την εξαγωγή πληροφοριών από κείμενο, την αυτόματη δημιουργία περιλήψεων κ.α. [47].

Τα παραπάνω είναι λίγα μόνο παραδείγματα από τους τομείς που η μηχανική μάθηση βρίσκει εφαρμογή. Γίνεται, λοιπόν, αντιληπτή η δυνατότητα της μηχανικής μάθησης να ενθαρρύνει την καινοτομία, να επαναπροσδιορίσει τον τρόπο που προσεγγίζουμε πολύπλοκα προβλήματα αλλά και να αντιμετωπίσει προκλήσεις που θεωρούνταν αδύνατο να επιλυθούν στο παρελθόν.

### 2.3.2 Αλγόριθμοι Μηχανικής Μάθησης

Οι κλασικές στατιστικές μέθοδοι πρόβλεψης, όπως η ARIMA ή η εκθετική εξομάλυνση, χρησιμοποιούνται πολλές δεκαετίες για την ανάλυση χρονοσειρών, με την πρώτη αναφορά σε αυτές να χρονολογείται το 1970 [48] και το 1956 [49], αντίστοιχα. Ωστόσο, τα κλασικά στατιστικά μοντέλα πρόβλεψης αντιμετωπίζουν κάποιους περιορισμούς, αφού συχνά μπορεί να δυσκολεύονται να συλλάβουν πολύπλοκα μοτίβα, να χειριστούν μη γραμμικές σχέσεις ή να προσαρμοστούν στη μεταβαλλόμενη δυναμική των δεδομένων. Επιπλέον, ο μεγάλος όγκος δεδομένων που τίθεται προς διαχείριση και η ανάγκη αξιοποίησης των big data καθιστά κρίσιμη την αξιοποίηση της υπολογιστικής ισχύος για την δημιουργία καλύτερων και πιο ακριβών προβλέψεων. Εισήχθησαν έτσι τα μοντέλα μηχανικής μάθησης στον χώρο της πρόβλεψης χρονοσειρών. Μια σύγκριση της ακρίβειας που παρουσιάζουν στις προβλέψεις του τα στατιστικά μοντέλα και τα μοντέλα μηχανικής μάθησης έγινε στον M4 διαγωνισμό που πραγματοποιήθηκε το 2018. Από τον διαγωνισμό αυτό, προέκυψε ότι η χρήση υβριδικών μοντέλων, με συνδυασμό στατιστικών μοντέλων και μοντέλων στατιστικής πρόβλεψης [50] ενώ στη συνέχεια, στον επόμενο διαγωνισμό, M5, που διεξήχθη το 2020 όλα τα μοντέλα που είχαν τις κορυφαίες επιδόσεις χρησιμοποίησαν αποκλειστικά μηχανική μάθηση, ξεπερνώντας, έτσι, τις παραδοσιακές στατιστικές μεθόδους [51].

Οι αλγόριθμοι της μηχανικής μάθησης διακρίνονται σε τρεις βασικές κατηγορίες ανάλογα με το επιθυμητό αποτέλεσμα και το είδος του προβλήματος που καλούνται να αντιμετωπίσουν: την επιβλεπόμενη μάθηση (supervised learning), την μη-επιβλεπόμενη μάθηση (unsupervised learning) και την ενισχυτική μάθηση (reinforcement learning) [52].

### 2.3.2. Μοντέλα Μηχανικής και Βαθιάς Μάθησης

Σε αυτήν την υποενότητα θα γίνει αναφορά στα μοντέλα μηχανικής μάθησης, δίνοντας έμφαση σε αυτά που θα χρησιμοποιηθούν στο πειραματικό μέρος της εργασίας.

- **Δένδρα Αποφάσεων**

Τα δένδρα αποφάσεων αποτελούν δενδρικές δομές οι οποίες ταξινομούν το δείγμα με βάση τα χαρακτηριστικά του. Κάθε εσωτερικός κόμβος αναπαριστά ένα από τα χαρακτηριστικά του δείγματος ενώ κάθε φύλλο αναπαριστά την προβλεπόμενη έξοδο [53].

- **Light Gradient Boosting Machine**

Σε αυτή την τεχνική η εκπαίδευση βασίζεται σε δένδρα αποφάσεων. Το Light Gradient Boosting Machine μοντέλο είναι μια επέκταση της τεχνικής Gradient Boosting. Βασική ιδέα της αποτελεί η διαδοχική εκπαίδευση νέων μοντέλων ώστε να διορθώνουν τα λάθη

των προηγούμενων. Πιο συγκεκριμένα, κάθε μοντέλο (που συνήθως βασίζεται σε δένδρα αποφάσεων) εκπαιδεύεται σε μια συνάρτηση της αρνητική παραγώγου της συνάρτησης κόστους (loss function) του προηγούμενου. Συνήθεις συναρτήσεις κόστους που επιλέγονται σε προβλήματα παλινδρόμησης είναι το μέσο τετραγωνικό σφάλμα (MSE) και το μέσο απόλυτο σφάλμα (MAE) [54].

Η διαφορά του Light Gradient Boosting με το Gradient Boosting έγκειται στο ότι στην πρώτη εφαρμόζονται βελτιστοποιήσεις με σκοπό την πιο γρήγορη και αποδοτική λειτουργία της.

Η κυριότερη βελτιστοποίηση είναι αυτή της τεχνικής Gradient-based One-Side Sampling (GOSS), η οποία προέκυψε από την παρατήρηση ότι χαρακτηριστικά με διαφορετικές παραγώγους του σφάλματος παίζουν διαφορετικό ρόλο στη διαδικασία εκμάθησης. Συγκεκριμένα, το κέρδος πληροφοριών (information gain) από χαρακτηριστικά με μεγάλη παράγωγο σφάλματος είναι πολύ μεγαλύτερο. Πιο απλά, πρόκειται για τα χαρακτηριστικά για τα οποία το δένδρο έχει μάθει λιγότερο και άρα η απόδοση του μοντέλου χρειάζεται βελτίωση σε αυτά. Έτσι, επιλέγεται ο χωρισμός των χαρακτηριστικών σε δύο σύνολα: το πρώτο περιέχει χαρακτηριστικά με μεγάλη παράγωγο σφάλματος και το δεύτερο χαρακτηριστικά με μικρή παράγωγο σφάλματος. Στη συνέχεια, κάθε δένδρο που προστίθεται δεν εκπαιδεύεται σε όλα τα χαρακτηριστικά του συνόλου δεδομένων αλλά σε ολόκληρο το πρώτο σύνολο και σε ένα μόνο τυχαίο δείγμα του δεύτερου. Μειώνεται έτσι, πολύ, το σύνολο των χαρακτηριστικών στο οποίο πρέπει κάθε δένδρο να εκπαιδεύεται, οδηγώντας σε πιο γρήγορες προβλέψεις.

Το δεύτερο σημαντικό χαρακτηριστικό του LightGBM είναι η ομαδοποίηση αμοιβαίως αποκλειόμενων χαρακτηριστικών (Exclusive Feature Bundling technique (EFB)). Με την τεχνική αυτή χαρακτηριστικά τα οποία δεν μπορούν να πάρουν μη μηδενική τιμή ταυτόχρονα, ομαδοποιούνται σε ένα κοινό χαρακτηριστικό, μειώνοντας έτσι το σύνολο των χαρακτηριστικών στο οποίο γίνεται η εκπαίδευση και, επομένως, μειώνοντας τον χρόνο εκπαίδευσης.

Τέλος, αντί το δένδρο να δοκιμάζει όλες τις τιμές ενός χαρακτηριστικού για να βρει το καλύτερο σημείο διαχωρισμού, ομαδοποιεί συνεχόμενες τιμές του χαρακτηριστικού σε bins τα οποία χρησιμοποιεί για να φτιάξει ιστογράμματα. Με αυτή την τεχνική μειώνεται η κατανάλωση μνήμης και αυξάνεται η ταχύτητα εκπαίδευσης [55].

- **Neural Basis Expansion Analysis for Time Series forecasting (N -Beats)**

Πρόκειται για ένα μοντέλο βαθιάς μάθησης που βασίζεται στο πολυστρωματικό δίκτυο perceptron ( multi-layer perceptron) καθώς και στα υπολειμματικά δίκτυα (residual networks). Το βασικό δομικό στοιχείο του είναι το Block, το οποίο αποτελείται από 4 επίπεδα πλήρως συνδεδεμένων (fully connected – FC) δικτύων. Πολλά βασικά δομικά στοιχεία – blocks – αποτελούν μια στοίβα (stack), σε διάταξη υπολειμματικής στοίβας. Κάθε block, για παράδειγμα το 1 block, δέχεται μια είσοδο  $x_l$  και επιστρέφει δύο διανύσματα  $\hat{x}_l$  και  $\hat{y}_l$ . Για το πρώτο block του μοντέλου, το  $x_l$  αντιστοιχεί στην συνολική

είσοδο του μοντέλου – ίση με ένα παράθυρο προηγούμενης γνώσης, συγκεκριμένου μήκους, που τελειώνει στην τελευταία καταγεγραμμένη παρατήρηση. Το μήκος αυτού του παραθύρου προτείνεται να τίθεται ως πολλαπλάσιο του ορίζοντα πρόβλεψης. Για κάθε επόμενο block, το  $x_l$  ισούται με την υπολειμματική έξοδο (residual output) του προηγούμενου. Δηλαδή:

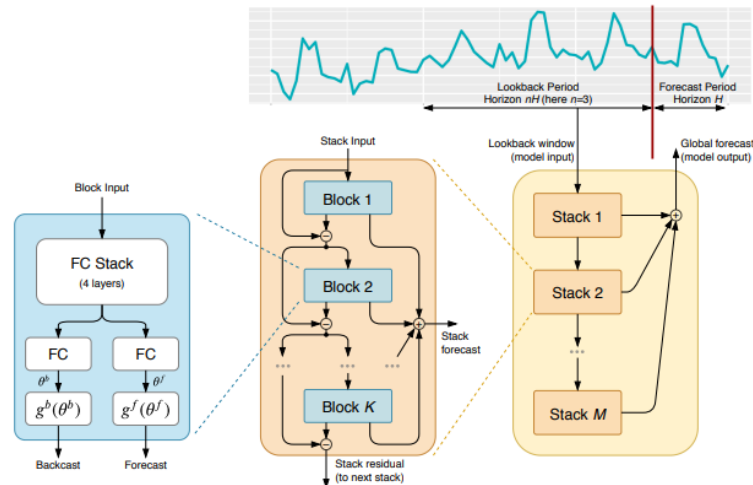
$$x_l = x_{l-1} - \widehat{x}_{l-1}$$

Η πρώτη έξοδος  $\widehat{x}_l$  του block αντιστοιχεί στην εκτίμηση των δεδομένων εισόδου, μήκος ίσου με το παράθυρο προηγούμενης γνώσης ενώ η δεύτερη έξοδος  $\widehat{y}_l$  του block αντιστοιχεί στην πρόβλεψη των δεδομένων μετά την είσοδο, μήκους ίσου με τον ορίζοντα πρόβλεψης (H). Η πρόβλεψη αυτή προστίθεται κάθε φορά στην επόμενη, αρχικά σε επίπεδο στοιβάς και έπειτα στο συνολικό επίπεδο του δικτύου. Έτσι, τελικά, η συνολική πρόβλεψη  $\widehat{y}$  ισούται με το άθροισμα όλων των επιμέρους προβλέψεων:

$$\widehat{y} = \sum_l \widehat{y}_l$$

Η τεχνική αυτή ονομάζεται διπλή υπολειμματική στοιβά (Double Residual Stacking). Οι δύο έξοδοι τού block,  $\widehat{y}_l$  και  $\widehat{x}_l$ , αναφέρονται και ως forward forecast και backtest/backcast, αντίστοιχα [56].

Στην Εικόνα 4 παρουσιάζεται η αρχιτεκτονική του N-BEATS που περιγράφηκε.



Εικόνα 4 -Η βασική αρχιτεκτονική του N-BEATS

### ▪ Χρονικά Συνελκτικά Δίκτυα (*Temporal Convolutional Networks ή TCNs*)

Τα Χρονικά Συνελκτικά Δίκτυα ή TCNs ανήκουν στην οικογένεια των Συνελκτικών Νευρωνικών Δικτύων (CNNs) [57]. Είναι ένας τύπος αρχιτεκτονικής νευρωνικών δικτύων που έχει σχεδιαστεί για την επίλυση προβλημάτων μοντελοποίησης ακολουθιών. Αξιοποιούν τη δύναμη των συνελκτικών λειτουργιών για την καταγραφή χρονικών εξαρτήσεων σε διαδοχικά δεδομένα. Τα TCN έχουν κερδίσει δημοτικότητα λόγω της ικανότητάς τους να μοντελοποιούν μακροπρόθεσμες εξαρτήσεις διατηρώντας παράλληλα την υπολογιστική απόδοση, χρησιμοποιώντας αιτώδεις, διεσταλμένες συνελίξεις για να επεκτείνουν το δεκτικό πεδίο (receptive field) και να συλλαμβάνουν πληροφορίες από διαφορετικά χρονικά βήματα παράλληλα.

Η αιτώδης συνέλιξη διασφαλίζει ότι η έξοδος σε κάθε χρονικό βήμα εξαρτάται μόνο από τα προηγούμενα χρονικά βήματα και όχι από τα μελλοντικά. Αυτή η ιδιότητα είναι ιδιαίτερα σημαντική σε εργασίες μοντελοποίησης ακολουθίας όπου η σειρά των δεδομένων είναι σημαντική, όπως η μετάφραση γλώσσας ή η πρόβλεψη χρονοσειρών. Επιβάλλοντας την αιτιότητα, τα TCN μπορούν να κάνουν προβλέψεις βασισμένες σε προηγούμενες πληροφορίες χωρίς πρόσβαση σε μελλοντικά δεδομένα, κάτι που συμβαίνει συχνά σε σενάρια πραγματικού κόσμου.

Για να επιτευχθεί η αιτιότητα, τα TCN χρησιμοποιούν τεχνικές padding για να προσαρμόσουν το μέγεθος της ακολουθίας εισόδου και να ελέγξουν το δεκτικό πεδίο των συνελκτικών φίλτρων. Συγκεκριμένα, χρησιμοποιούν την τεχνική “zero padding” συμπληρώνοντας την ακολουθία εισόδου με μηδενικά με τέτοιο τρόπο ώστε η έξοδος σε κάθε χρονικό βήμα να επηρεάζεται μόνο από τα προηγούμενα χρονικά βήματα [58].

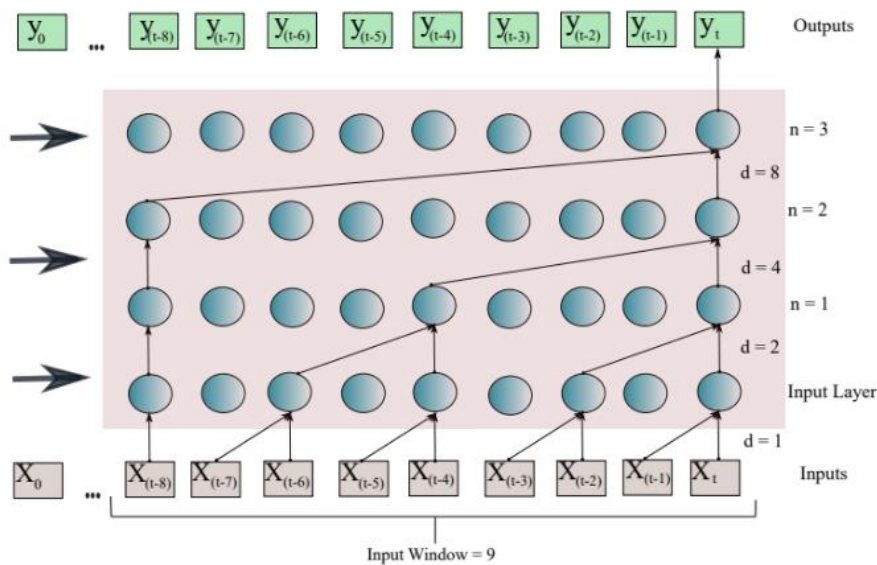
Το πρόβλημα είναι ότι ένα δίκτυο που αποτελείται από αιτώδεις συνελίξεις μπορεί να κοιτάζει πίσω σε ένα ιστορικό μήκος που κλιμακώνεται γραμμικά με το βάθος τού δικτύου[59]. Για την αντιμετώπιση αυτού του προβλήματος, στο TCN χρησιμοποιούνται διεσταλμένες αιτιατές συνελίξεις οι οποίες ενσωματώνουν έναν παράγοντα διαστολής  $d$  (dilation factor), η τιμή του οποίου αυξάνεται εκθετικά με το βάθος τού δικτύου[58]. Συγκεκριμένα, ο παράγοντας διαστολής σε κάθε επίπεδο υπολογίζεται ως εξής:

$$d_l = 2^l$$

Όπου  $l$  το επίπεδο του δικτύου

Αυτό επιτρέπει στα TCN να συλλαμβάνουν πληροφορίες από διαφορετικές χρονικές κλίμακες.

Η παραπάνω αρχιτεκτονική μπορεί να εξελιχθεί, χρησιμοποιώντας στοίβες (Stacks) από επίπεδα (layers), σε βαθιά αρχιτεκτονική. Η δομή μιας τέτοιας στοίβας απεικονίζεται στην



Εικόνα 5 - Η δομή μιας στοίβας ενός TCN σε univariate σύστημα (Πηγή: [40])

### 2.3.3. Μέθοδος Ensemble

Η μέθοδος Ensemble είναι μια τεχνική που συνδυάζει τις προβλέψεις πολλαπλών μεμονωμένων μοντέλων πρόβλεψης για να παράγει πιο ακριβείς προβλέψεις. Με τον συνδυασμό των επιμέρους προβλέψεων μπορεί να απομακρυνθεί η τυχαιότητα των αποτελεσμάτων και να μειωθούν τα σφάλματα. Τα επιμέρους μοντέλα από τα οποία αποτελείται η μέθοδος Ensemble καλούνται μοντέλα βάσης. Αν το είδος των μοντέλων βάσης είναι κοινό, τότε έχουμε ομογενή μοντέλα βάσης ενώ αν τα μοντέλα βάσης είναι διαφορετικά, τότε έχουμε ανομοιογενή μοντέλα βάσης [60].

Ιδιαίτερα χρήσιμη είναι η μέθοδος Ensemble όσον αφορά τα νευρωνικά δίκτυα βαθιάς μάθησης (Deep Neural Networks, καθώς η εκπαίδευση και οι επιδόσεις τους χαρακτηρίζονται από μεγάλη στοχαστικότητα αφού σε κάθε διαδικασία εκπαίδευσης, τα τελικά βάρη που θα προκύψουν μπορεί να είναι διαφορετικά από μια άλλη με ίδιες τιμές υπερπαραμέτρων. Με την χρήση πολλών ίδιων μοντέλων, με διαφορετική αρχικοποίηση των βαρών των νευρωνικών, και με τη χρήση του μέσου όρου των προβλέψεων τους ως την τελική πρόβλεψη, η στοχαστικότητα αυτή μπορεί να αντιμετωπιστεί, οδηγώντας σε ακριβέστερες προβλέψεις [60].

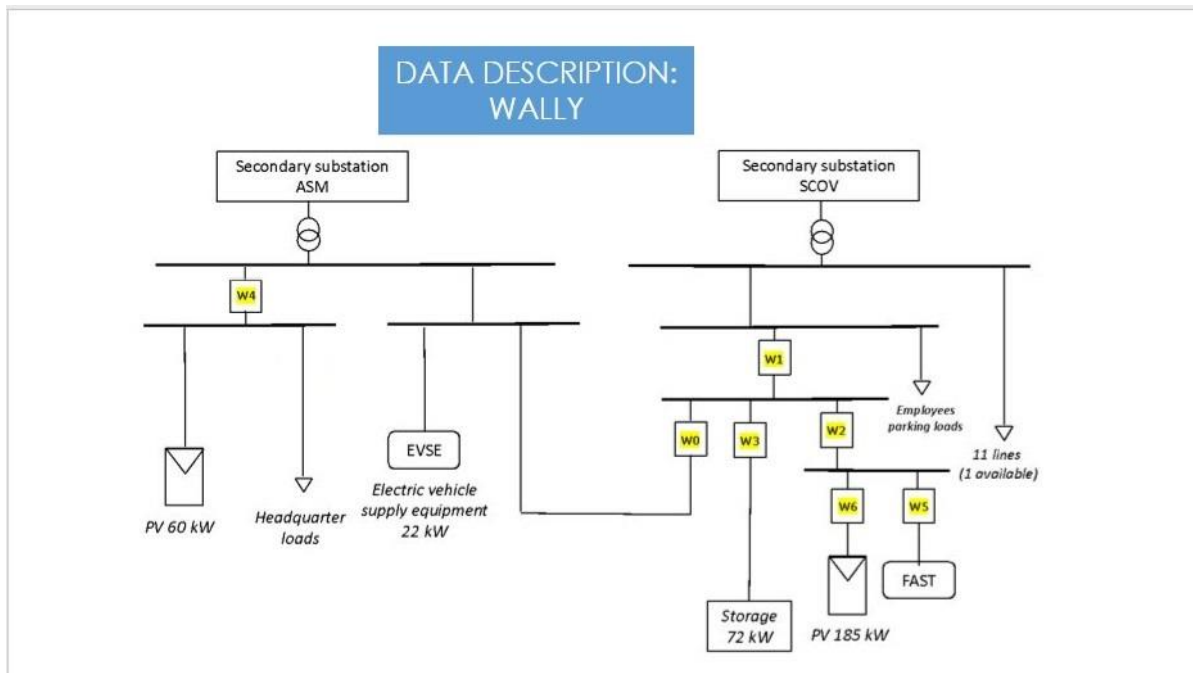
### 3. Μεθοδολογία

#### 3.1. Σύνολο Δεδομένων

Τα δεδομένα που έχουν χρησιμοποιηθεί στο πειραματικό μέρος έχουν διατεθεί από την εταιρεία ASM. Πρόκειται για μια δημόσια επιχείρηση που κατέχει και διαχειρίζεται το δίκτυο διανομής ηλεκτρικής ενέργειας σε μια κομμόπολη της κεντρικής Ιταλίας, το Terni. Τα δεδομένα προέρχονται από πέντε έξυπνους μετρητές που η ASM έχει τοποθετήσει σε κομβικά σημεία ενός έξυπνου δικτύου, προκειμένου να βελτιώσει την απόδοσή του. Οι μετρητές καταγράφουν διαδοχικές μετρήσεις, της ενεργούς (P), αέργου (Q) και φαινόμενης ισχύος (S) καθώς και των τριών φάσεων του ρεύματος (I1-I3) και της τάσης (V1-V3). Συγκεκριμένα οι μετρήσεις γίνονται σε σχεδόν πραγματικό χρόνο (2-10 sec) και στη συνέχεια γίνεται επαναδειγματοληψία (resampling) σε διαστήματα πέντε λεπτών.

Επιπλέον έχουμε στη διάθεση μας μετεωρολογικά δεδομένα για την περιοχή Terni της Ιταλίας και πιο συγκεκριμένα, δεδομένα της ηλιακής ακτινοβολίας, με μονάδα μέτρησης  $J * m^{-2}$ , της νεφοκάλυψης, σε ποσοστό %, και της θερμοκρασίας, μετρημένη σε K, όπως αυτά μας διατέθηκαν από το ευρωπαϊκό πρόγραμμα Copernicus [61]. Οι μετρήσεις αυτές είναι ωριαίες και καταγράφηκαν για τα έτη 2021 και 2022.

#### 3.2. Ζητούμενο



Εικόνα 6 - Τοπολογία δικτύου

Στην Εικόνα 6 φαίνεται η τοπολογία του δικτύου που χρησιμοποιήθηκε, όπου:

W1 – W6: οι έξυπνοι μετρητές

PV 60 kW και PV 185 kW: φωτοβολταϊκό μέγιστης ισχύος 60kW και 185kW, αντίστοιχα

FAST: γρήγορος σταθμός φόρτισης ηλεκτρικών οχημάτων

Storage 72 kW: μπαταρία χωρητικότητας μέγιστης ισχύος 72 kW

Headquarter Loads: το φορτίο του παρόχου

Όπως φαίνεται, ο μετρητής W2 είναι τοποθετημένος στο άκρο του φωτοβολταϊκού 185 kW ενώ ο W5 στο άκρο του γρήγορου σταθμού φόρτισης. Επιπλέον ο μετρητής W2 είναι τοποθετημένος ανάμεσα στον ζυγό στον οποίο συνδέονται το φωτοβολταϊκό μέγιστης ισχύος 185 kW και ο γρήγορος σταθμός φόρτισης, καθώς και στον ζυγό στον οποίο συνδέεται η μπαταρία. Ο μετρητής W3 είναι τοποθετημένος στο άκρο της μπαταρίας ενώ ο μετρητής W4 στον ζυγό στον οποίο συνδέονται τα headquarter loads και το φωτοβολταϊκό μέγιστης ισχύος 60 kW. Σκοπός είναι να προβλεφθεί τόσο η ενεργός όσο και η άεργος ισχύς των μετρητών W2-W6, που όπως θα δούμε στην συνέχεια, ο σκοπός τελικά περιορίζεται στην πρόβλεψη μόνο της ισχύς των μετρητών W6 και W4, ενεργούς και αέργου.

### 3.3 Διαχείριση Δεδομένων

#### 3.3.1. Εργαλεία

Στο στάδιο της προ- επεξεργασίας χρησιμοποιήθηκε κυρίως η βιβλιοθήκη Pandas [62] της Python, για τη διαχείριση, την επεξεργασία και την αποθήκευση των datasets. Επιπλέον χρησιμοποιήθηκαν οι βιβλιοθήκες Matplotlib και Seaborn [63] της Python για την οπτικοποίηση των δεδομένων και η βιβλιοθήκη Darts [64] για τη δημιουργία και επεξεργασία χρονοσειρών καθώς και για την εφαρμογή αλγορίθμων πρόβλεψης.

#### 3.3.2. Προεπεξεργασία Δεδομένων

Σε αυτό το στάδιο έγινε επεξεργασία των δεδομένων, όπως αυτά λήφθηκαν από τη βάση δεδομένων MongoDB, ώστε να αποκτήσουν μια πιο εύκολα διαχειρίσιμη και επεξεργάσιμη μορφή.

Τα δεδομένα αποθηκεύτηκαν σε ένα dataframe, το οποίο τέθηκε υπό επεξεργασία ώστε κάθε στήλη του να αντιπροσωπεύει κάθε, υπό μέτρηση, μέγεθος του κάθε μετρητή και κάθε γραμμή του τις διαδοχικές μετρήσεις. Έγινε καθαρισμός των δεδομένων, με διαγραφή των στηλών που δεν περιείχαν χρήσιμη πληροφορία αλλά και με τον εντοπισμό και την διαχείριση, με κατάλληλες συναρτήσεις, των outliers και των κενών τιμών (NaNs) που υπήρχαν στο σύνολο των δεδομένων.

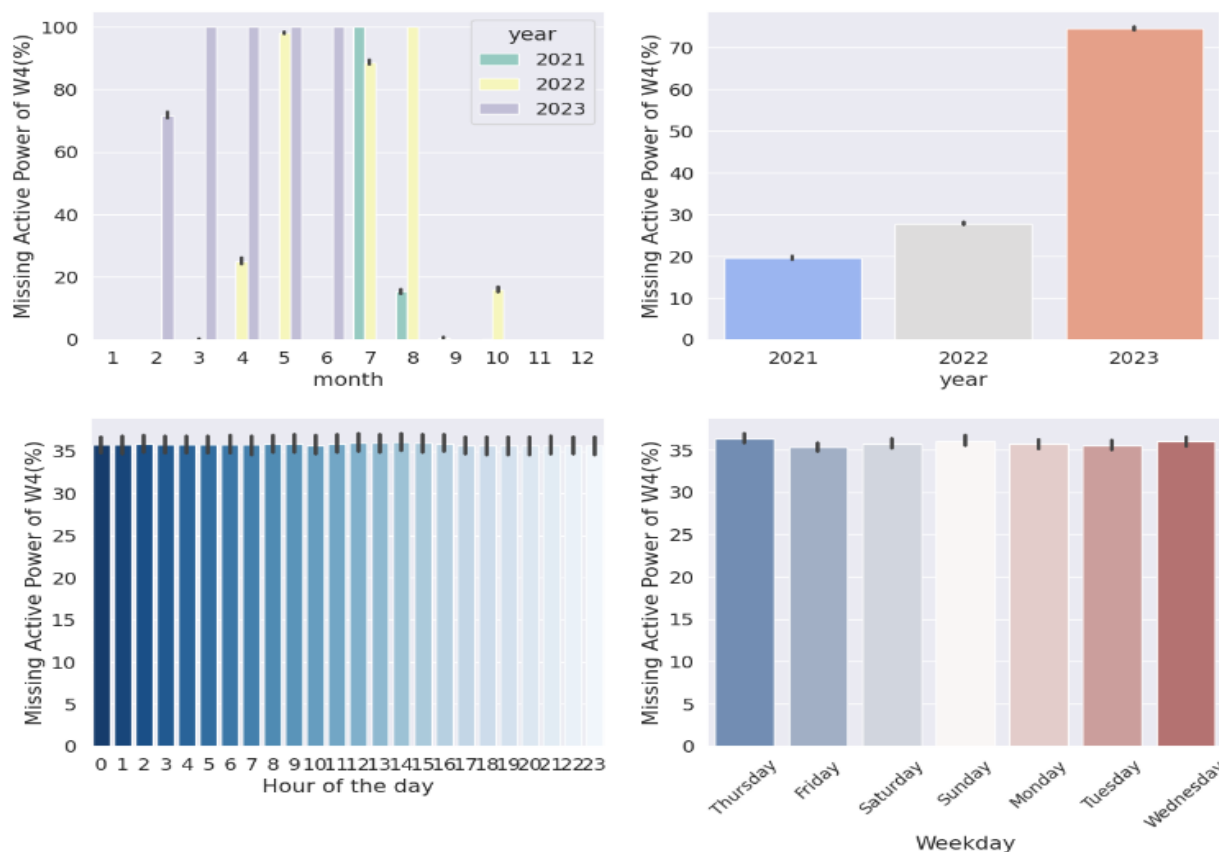
Τέλος, μετά τη διαδικασία ανάλυσης των δεδομένων που περιγράφεται στην επόμενη ενότητα 3.3.3. Ανάλυση Δεδομένων”, πραγματοποιήθηκε επαναδειγματοληψία με συχνότητα μίας ώρας,



ώστε τα δεδομένα των smart meter και τα μετεωρολογικά δεδομένα να είναι σε συμφωνία. Η επαναδειγματοληψία πραγματοποιήθηκε χρησιμοποιώντας τη μέση τιμή των δεδομένων σε κάθε ώρα.

### 3.3.2.1. Διαχείριση Κενών Τιμών

Διαγράμματα όπως αυτό της Εικόνας 7 βοήθησαν στην κατανόηση της κατανομής των κενών τιμών, τόσο ανά μήνα και χρόνο, όσο και ανά ημέρα της εβδομάδας και ώρα της ημέρας.



Εικόνα 7 - Ποσοστό κενών τιμών για τον W4, ανά χρόνο, μήνα, μέρα της εβδομάδας και ώρα της ημέρας

Από τα διαγράμματα αυτά συμπεράναμε ότι υπήρχαν σοβαρές ελλείψεις στα δεδομένα μας, ειδικά στα δεδομένα της ισχύος, καθώς υπήρχαν ολόκληροι διαδοχικοί μήνες με κενές τιμές. Πιο συγκεκριμένα, γνωρίζαμε ότι έχουμε δεδομένα από τον Αύγουστο του 2021 έως τον Φεβρουάριο του 2023. Όπως συμπεράναμε, οι μετρήσεις της ισχύος για τον Αύγουστο του 2022 έλειπαν από όλους τους μετρητές, όπως και το 90% των μετρήσεων του Ιουλίου 2022. Επιπλέον κενά υπήρχαν τόσο στον Αύγουστο του 2021 όσο και στον Οκτώβριο του 2022, όπου λείπει περίπου το 20% των μετρήσεων.

Ειδικότερα για τον μετρητή W4 υπάρχουν επιπλέον ελλείψεις τον Απρίλιο και τον Μάιο του 2022, με περίπου το 22% και το 99%, αντίστοιχα, των μετρήσεων να λείπει. Τέλος ο μετρητής

W2 φαίνεται να έχει σταματήσει να μετρά την ενεργό ισχύ από τα τέλη του Ιουλίου 2022. Παρατηρούμε επιπλέον ότι οι κενές τιμές είναι ομοιόμορφα κατανομημένες τόσο στη διάρκεια της εβδομάδας όσο και στη διάρκεια της ημέρας

Στα δεδομένα ρεύματος και τάσης των μετρητών έχουμε λιγότερες κενές τιμές, με τις περισσότερες να εντοπίζονται το 2021 και ειδικά τις Κυριακές. Ωστόσο αξίζει να σημειωθεί ότι κι εδώ, όπως και στα δεδομένα της ισχύος, εμφανίζεται μεγάλο ποσοστό κενών τιμών τον Απρίλιο και τον Μάιο του 2022, στα δεδομένα που μετρήθηκαν από τον W4. Τέλος βλέπουμε ότι η κατανομή των κενών τιμών είναι ομοιόμορφη κατά τη διάρκεια της ημέρας.

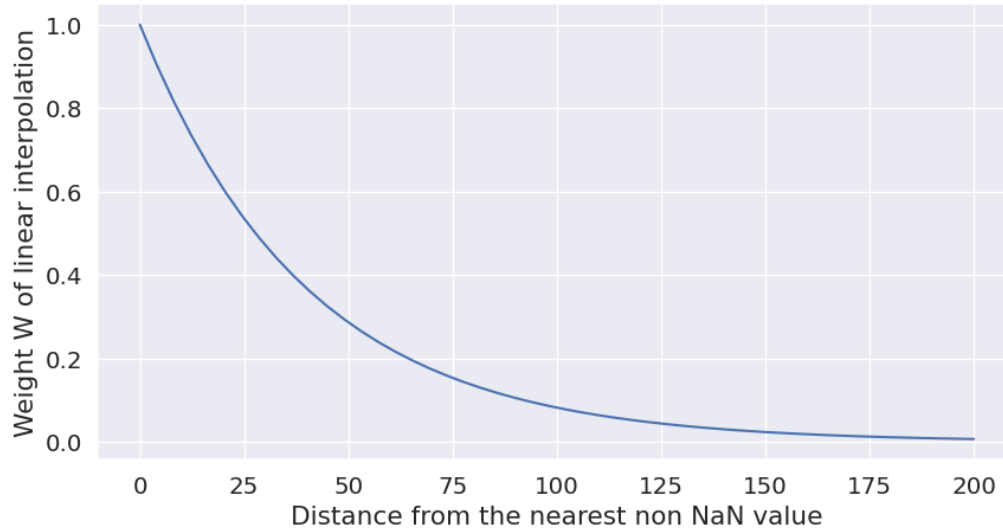
Η συμπλήρωση των μεγάλων κενών των διαδοχικών μηνών αποφεύχθηκε, αφού, όπως γίνεται αντιληπτό, είναι πολύ δύσκολο έως αδύνατο να πραγματοποιηθεί με αξιόπιστο τρόπο. Για τη συμπλήρωση των υπόλοιπων κενών τιμών επιλέχθηκε η εφαρμογή ενός συνδυασμού τεχνικών imputation. Συγκεκριμένα επιλέχθηκε ένα διάστημα 12 διαδοχικών ωρών ως το μέγιστο διάστημα μέσα στο οποίο μπορούμε να συμπληρώσουμε τιμές. Δηλαδή περισσότερες από  $12 * 12 = 144$  διαδοχικές κενές τιμές (διαδοχικά πεντάλεπτα) δε συμπληρώθηκαν και επέστρεψαν ως NaN. Οι υπόλοιπες κενές τιμές συμπληρώθηκαν με τη χρήση ενός σταθμισμένου μέσου όρου των ιστορικών παρατηρήσεων και της γραμμικής παρεμβολής (linear interpolation) [65]. Το βάρος της κάθε μεθόδου εξαρτάται εκθετικά από την απόσταση της τρέχουσας κενής τιμής από την κοντινότερη μη κενή τιμή. Πιο συγκεκριμένα ορίζουμε το βάρος ( $w$ ) της γραμμικής παρεμβολής ως:

$$w = e^{-a*d[i]}$$

όπου  $a = 0.025$

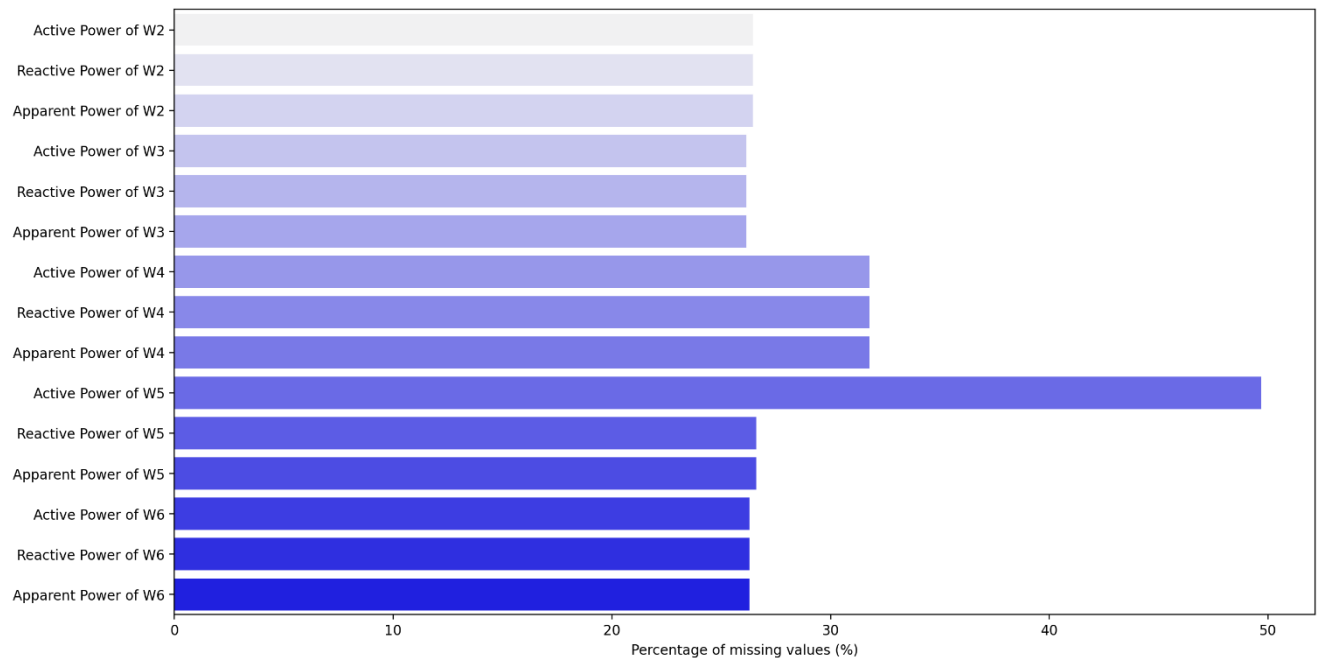
Το βάρος για τις ιστορικές παρατηρήσεις ορίζουμε ως  $1-w$ .

Όσο μεγαλύτερη λοιπόν είναι η απόσταση της εκάστοτε κενής τιμής από την κοντινότερη μη κενή τιμή τόσο ελαττώνεται το βάρος της γραμμικής παρεμβολής και αυξάνεται το βάρος των ιστορικών παρατηρήσεων. Γραφικά η σχέση αυτή αναπαρίσταται στην Εικόνα 8.



Εικόνα 8 - Η σχέση του βάρους  $w$  και της απόστασης  $d[i]$

Μετά την εφαρμογή των παραπάνω διαδικασιών και την κάλυψη 20,000 μετρήσεων, δηλαδή 20,000 πεντάλεπτων που ισοδυναμούν σε περίπου 70 ημέρες, μόνο στα δεδομένα της ισχύος, προκύπτουν τα ποσοστά κενών τιμών ανά μετρητή, όπως αυτά εμφανίζονται στην Εικόνα 9.



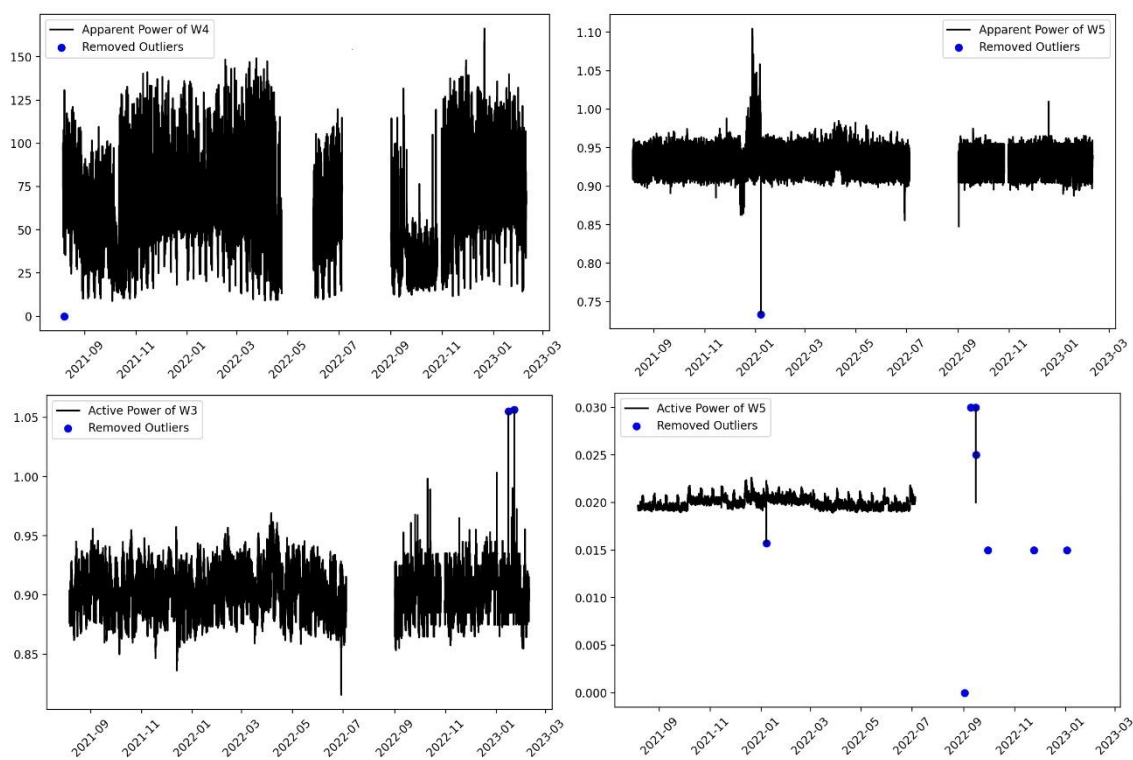
Εικόνα 9 - Ποσοστό κενών τιμών στα δεδομένα ισχύος, ανά μετρητή, μετά την εφαρμογή των διαδικασιών συμπλήρωσης

### 3.3.2.2. Διαχείριση Outliers

Η διαχείριση των outliers (τιμές πολύ μεγαλύτερες ή πολύ μικρότερες από τις γειτονικές τους) έγινε αρχικά με την αντικατάστασή τους από NaNs. Για την επιλογή αυτών των outliers, υπολογίστηκε ο μέσος όρος και η τυπική απόκλιση για κάθε μήνα στο dataset και αφαιρέθηκαν όσες τιμές απέχουν σημαντικά από τον μέσο όρο. Ως προκαθορισμένη τιμή επιλέχθηκε να αφαιρούνται όσες τιμές απέχουν από τον μέσο όρο του μήνα περισσότερο από το δεκαπλάσιο της τυπικής απόκλισης, ώστε να επιλεχθούν οι τιμές εκείνες που είναι πράγματι outliers. Επιπλέον, αφαιρέθηκαν τα μηδενικά, όπου αυτά αποτελούσαν την ελάχιστη τιμή του συνόλου των δεδομένων.

Όπως παρατηρήθηκε, ωστόσο, τα δεδομένα μας είχαν ελάχιστα έως και μηδενικά, σε πολλές περιπτώσεις, outliers. Συγκεκριμένα, μερικά outliers βρέθηκαν μόνο στην φαινόμενη ισχύ του W4 (ένα μηδενικό) και του W5, καθώς και στην ενεργό ισχύ του W3 και του W5. Επιπλέον, αναφορικά με τα δεδομένα του ρεύματος και της τάσης, λίγα outliers βρέθηκαν στα φασικά ρεύματα του W4 και στις φασικές τάσεις του W4 και W5.

Γραφικά, ο εντοπισμός των outliers φαίνεται στην Εικόνα 10.



Εικόνα 10 - Εντοπισμός outliers στα δεδομένα ισχύος

### 3.3.3. Ανάλυση Δεδομένων

#### 3.3.3.1. Στατιστική Περιγραφή Δεδομένων

Για να αποκτήσουμε μια καλύτερη εικόνα των δεδομένων προχωρούμε στην περιγραφή των δεδομένων χρησιμοποιώντας χρήσιμους στατιστικούς δείκτες και συγκεκριμένα τη μέση τιμή, την τυπική απόκλιση καθώς και τη μέγιστη και ελάχιστη τιμή κάθε μεγέθους.

Τα αποτελέσματα παρουσιάζονται στις Εικόνες 11 και 12.

	Active Power of W3	Reactive Power of W3	Apparent Power of W3	Active Power of W5	Reactive Power of W5	Apparent Power of W5
<b>count</b>	140580.000000	140580.000000	140580.000000	95780.000000	139721.000000	139723.000000
<b>mean</b>	0.902387	4.010730	4.144675	0.019967	-0.916296	0.932934
<b>std</b>	0.016463	0.075356	0.075058	0.000580	0.011781	0.013909
<b>min</b>	0.815568	3.572744	3.712278	0.000000	-0.967817	0.733695
<b>25%</b>	0.890496	3.959247	4.093210	0.019521	-0.924894	0.924565
<b>50%</b>	0.903137	4.010615	4.144447	0.019903	-0.915417	0.932990
<b>75%</b>	0.914865	4.062486	4.196164	0.020303	-0.907723	0.941351
<b>max</b>	1.056667	4.329491	4.463091	0.030000	-0.806604	1.104646

Εικόνα 11 - Στατιστική Περιγραφή Δεδομένων μετρητών W3 και W5

	Active Power of W4	Reactive Power of W4	Apparent Power of W4	Active Power of W6	Reactive Power of W6	Apparent Power of W6	Active Power of W2	Reactive Power of W2	Apparent Power of W2
<b>count</b>	129706.000000	129705.000000	129706.000000	140040.000000	138714.000000	140212.000000	137059.000000	137837.000000	139331.000000
<b>mean</b>	56.935606	-2.346907	58.987475	-21.364824	14.691302	27.987838	-21.895014	12.864545	28.316928
<b>std</b>	22.228859	10.860479	21.380883	37.339100	18.315300	40.738522	38.522041	17.017614	40.347044
<b>min</b>	-24.164224	-17.673406	0.000000	-192.649814	-0.092388	0.547251	-195.630865	-1.363166	1.061034
<b>25%</b>	43.368109	-9.739494	45.070357	-28.827266	-0.048556	0.592489	-31.143411	-1.125121	1.535630
<b>50%</b>	53.800019	-6.579050	55.148649	0.042595	0.004958	0.610076	0.856187	-1.065000	2.255000
<b>75%</b>	68.863822	0.564067	70.075270	0.045000	25.240951	39.437437	0.865454	23.119137	39.079225
<b>max</b>	161.608690	48.147457	166.341750	4.656833	83.177921	210.213884	14.330000	75.789100	210.148292

Εικόνα 12 - Στατιστική Περιγραφή Δεδομένων μετρητών W2, W4, W6

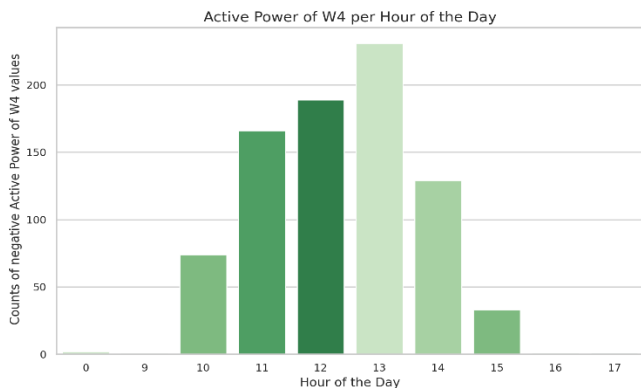
Από την Εικόνα 11, παρατηρείται ότι η ισχύς που μετρήθηκε από τους μετρητές W3 και W5, που βρίσκονται τοποθετημένοι στη μπαταρία και στον γρήγορο σταθμό φόρτισης αντίστοιχα, είναι σχεδόν μηδενική, αφού οι τιμές της ενεργούς ισχύος τους κυμαίνονται στο διάστημα [0.82 kW – 1.1 kW] ενώ της αέργου στο [3.57 VA – 4.46 VA]. Συμπεραίνουμε λοιπόν ότι αυτά τα δύο στοιχεία του έξυπνου δικτύου δεν λειτουργούσαν, γεγονός που επαληθεύτηκε από τον πάροχο των δεδομένων. Όσον αφορά την ενεργό ισχύ που μετρήθηκε από τον W4, βλέπουμε από την Εικόνα 12 ότι το 50% των τιμών της ανήκουν στο διάστημα [43 kW – 68 kW] ενώ το υπόλοιπο 50% μοιράζεται ισόποσα στα διαστήματα [-24 kW – 43 kW], [ 68 kW – 161 kW].

Σχετικά με την ισχύ που μετρήθηκε από τους W2 και W6 παρατηρείται ότι το 50% των μετρήσεων ανήκει σε ένα πολύ περιορισμένο διάστημα γύρω από το μηδέν, γεγονός που δικαιολογείται από το ότι οι ώρες που λειτουργεί το φωτοβολταϊκό και παράγει ισχύ -δηλαδή κυρίως οι ώρες 08:00 -16:00 - είναι πολύ λιγότερες από τις ώρες που δε λειτουργεί. Το αρνητικό πρόσημο ερμηνεύεται ως ισχύς που παράγεται και ρέει πίσω προς το δίκτυο.

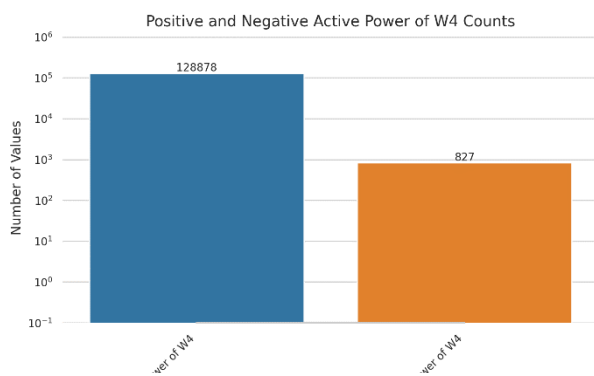
### 3.3.3.2. Ανάλυση ασυνήθιστων τιμών

- *Για την ισχύ που μετρήθηκε από τον W4*

Καθώς οι αρνητικές τιμές της ενεργού ισχύος που μετρήθηκε από τον W4 παρουσιάζουν ενδιαφέρον, προχωρήσαμε σε περαιτέρω διερεύνησή τους. Παρατηρήθηκε ότι οι τιμές αυτές εμφανίζονται κυρίως το διάστημα 10:00 – 15:00 με peak τις 12:00 (Εικόνα 13). Αυτές οι αρνητικές τιμές θα μπορούσαν να ερμηνευτούν ως εξής: Όπως έχουμε ήδη δει, ο μετρητής W4 είναι τοποθετημένος στον ζυγό που είναι συνδεδεμένο το φορτίο του παρόχου καθώς και το φωτοβολταϊκό 60 kW. Η πλειονότητα των τιμών τού W4 είναι θετική που σημαίνει ότι το φορτίο καταναλώνει περισσότερη ισχύ από όση παράγει το φωτοβολταϊκό. Οι λίγες περιπτώσεις που η ισχύς τού W4 είναι αρνητική, καταγράφονται σε ώρες που το φωτοβολταϊκό λειτουργεί στο μέγιστο, οπότε ίσως αυτές οι αρνητικές τιμές να αντιστοιχούν στην περίπτωση που το φωτοβολταϊκό παράγει περισσότερη ισχύ από αυτή που καταναλώνει το φορτίο, με αποτέλεσμα ένα μέρος της ισχύος να επιστρέφει πίσω στο δίκτυο. Επιπλέον, παρατηρήθηκε ότι αυτές οι τιμές αντιστοιχούν σε λιγότερο από το 0.65% του συνόλου των τιμών της (Εικόνα 14).



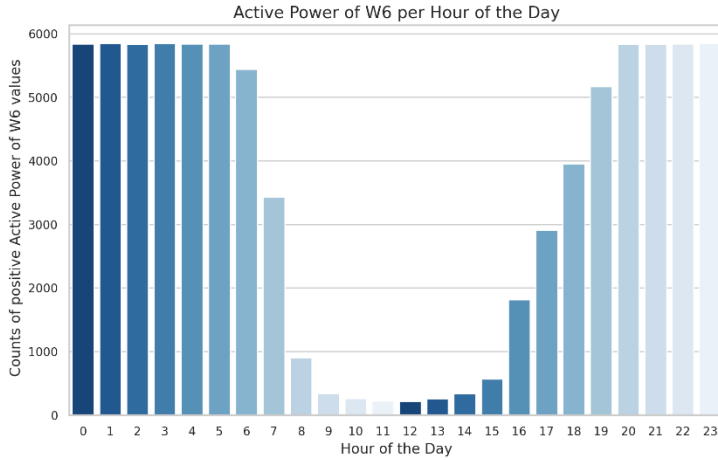
Εικόνα 13 - Πλήθος αρνητικών τιμών στην ενεργό ισχύ του W4 ανά την ώρα της ημέρας που παρουσιάζονται



Εικόνα 14 - Πλήθος θετικών και αρνητικών τιμών της ενεργούς ισχύος του W4, σε λογαριθμική κλίμακα

- Για την ισχύ που μετρήθηκε από τον W6

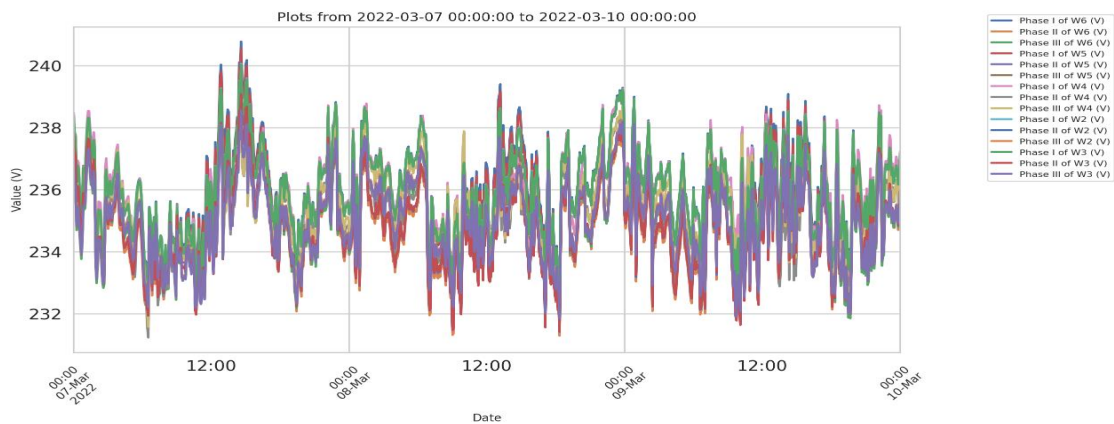
Σχετικά με την ενεργό ισχύ του φωτοβολταϊκού, που μετρήθηκε από τον W6, παρατηρήθηκε ότι παρουσιάζει κάποιες θετικές τιμές κοντά στο μηδέν. Μετά από μελέτη συμπεράναμε ότι οι τιμές αυτές εμφανίζονται τις ώρες της ημέρας που δεν υπάρχει έντονη ηλιακή ακτινοβολία και συγκεκριμένα από τις 17:00 έως τις 7:00, όπως φαίνεται και στην Εικόνα 15. Επιπλέον, οι τιμές αυτές είναι πολύ μικρές και πολύ κοντά στο μηδέν. Μια πιθανή εξήγηση γι' αυτό, είναι ότι το φωτοβολταϊκό και συγκεκριμένα τα εξαρτήματα από τα οποία αποτελείται, όπως για παράδειγμα οι αντιστροφείς του, χρειάζονται κάποια ενεργό ισχύ για να λειτουργήσουν. Έτσι, τις ώρες αυτές, το φωτοβολταϊκό καταναλώνει ένα μικρό ποσό ενεργούς ισχύος.



Εικόνα 15 - Πλήθος θετικών τιμών στην ενεργό ισχύ του W4 ανά την ώρα της ημέρας που παρουσιάζονται την ενεργό ισχύ του φωτοβολταϊκού.

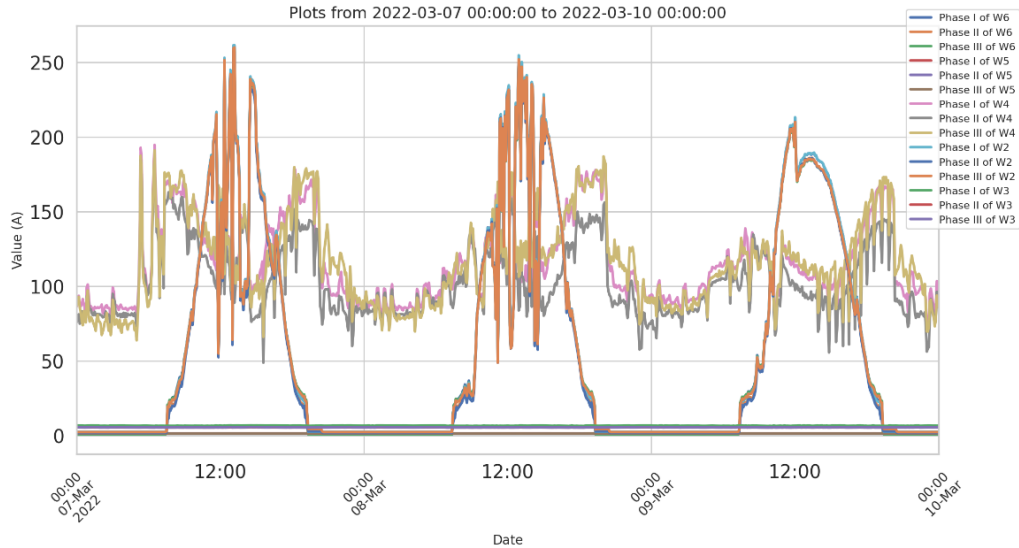
### 3.3.3.3. Γραφική αναπαράσταση δεδομένων

Λόγω του μεγάλου όγκου των δεδομένων, επιλέξαμε, για να αποκτήσουμε μια καλύτερη εικόνα της κατανομής των δεδομένων, να δούμε τα δεδομένα μας γραφικά σε ένα διάστημα τριών ημερών. Οι γραφικές αυτές απεικονίζονται στις Εικόνες 16, 17 και 18.

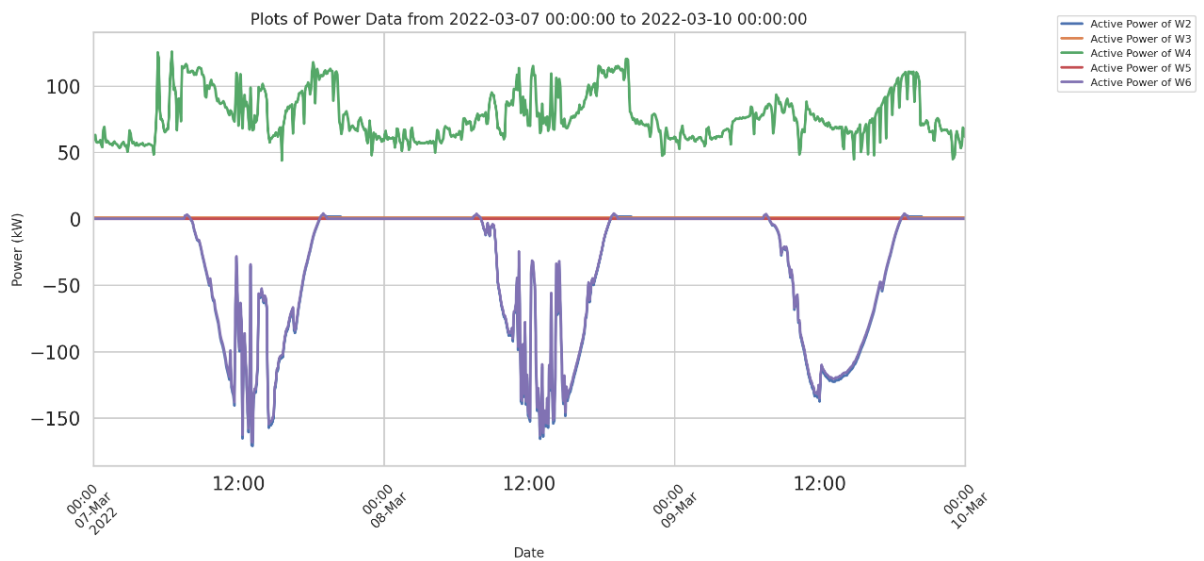


Εικόνα 16 - Γραφική αναπαράσταση δεδομένων ισχύος τριών ημερών





Εικόνα 17 - Γραφική αναπαράσταση δεδομένων ρεύματος τριών ημερών

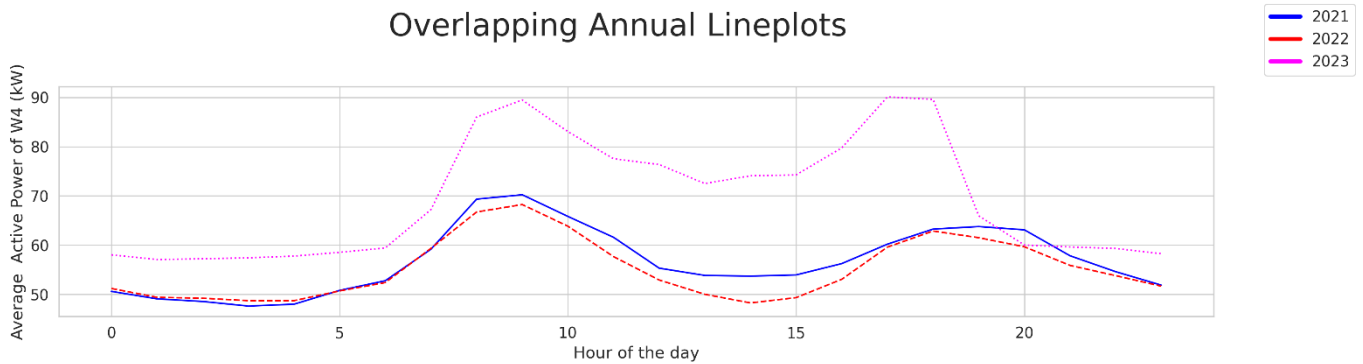


Εικόνα 18 - Γραφική αναπαράσταση δεδομένων τάσης τριών ημερών

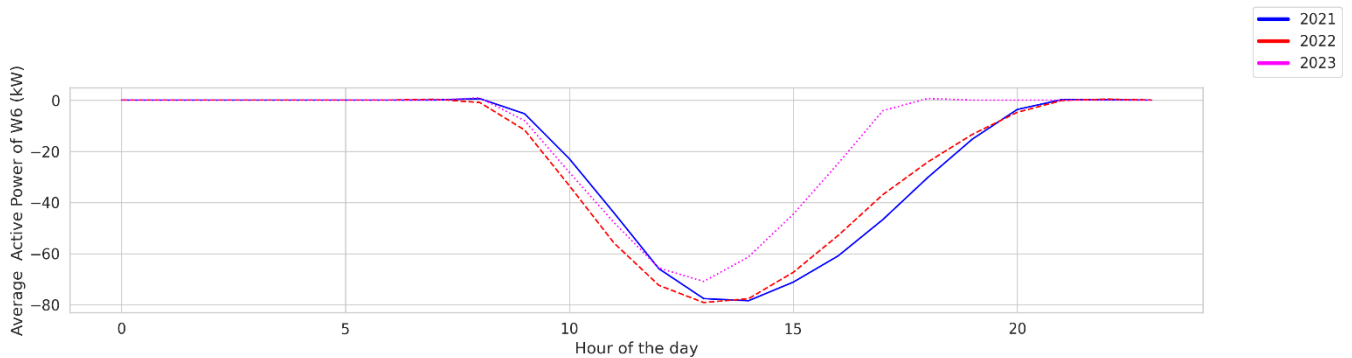
Από το διάγραμμα της γραφικής αναπαράστασης της τάσης όλων των μετρητών παρατηρούμε ότι οι τριφασικές τάσεις ανήκουν στο διάστημα [230 V -240 V]. Βλέπουμε, επιπλέον, ότι τόσο η τάση όσο και το ρεύμα τού μετρητή τού φωτοβολταϊκού (W6) παρουσιάζει μια ελάχιστη, μόνο, διαφοροποίηση από τα αντίστοιχα μεγέθη που μετρήθηκαν από τον μετρητή W2. Το γεγονός αυτό δικαιολογείται εύκολα αν αναλογιστεί κανείς την μη λειτουργία τού γρήγορου σταθμού φόρτισης που αντιστοιχεί στον μετρητή W5. Έστω  $W2_P$ ,  $W5_P$ ,  $W6_P$  η ισχύς που μετρήθηκε από τον μετρητή W2, W5 και W6 αντίστοιχα. Προκύπτει τότε από την τοπολογία του δικτύου ότι:

$$W2_P = W5_P + W6_P \rightarrow W2_P = W6_P$$

Επιπλέον, δημιουργήσαμε διαγράμματα για να παρακολουθήσουμε τον τρόπο μεταβολής κατά τη διάρκεια της ημέρας για τα έτη 2021, 2022 και 2023 (Εικόνα 19, Εικόνα 20). Από αυτά τα διαγράμματα παρατηρούμε αύξηση της κατανάλωσης στα headquarter load και μείωση της παραγωγής τις ώρες αιχμής για το φωτοβολταϊκό κατά το έτος 2023. Οι συμπεριφορές είναι λογικές και αναμενόμενες αν αναλογιστούμε ότι γι' αυτό το έτος διαθέτουμε δεδομένα μόνο για τους μήνες Ιανουάριο και Φεβρουάριο όπου οι χαμηλές θερμοκρασίες αυξάνουν σημαντικά τη ζήτηση και η ηλιακή ακτινοβολία είναι πιο μειωμένη σε σχέση με τους υπόλοιπους, μη χειμερινούς, μήνες του χρόνου.



Εικόνα 19 - Ωριαία διαγράμματα μέσης τιμής ισχύος για τον W4 ανά χρόνο



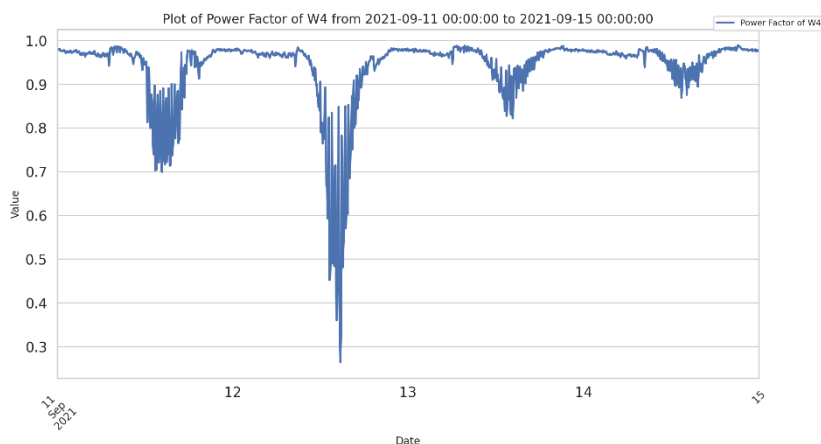
Εικόνα 20 - Ωριαία διαγράμματα μέσης τιμής ισχύος για τον W6, ανά χρόνο

### 3.3.3.4. Υπολογισμός συντελεστή ισχύος

Στο επόμενο στάδιο της ανάλυσης επιλέξαμε να υπολογίσουμε τον συντελεστή ισχύος για κάθε στοιχείο του ηλεκτρικού δικτύου. Έμφαση ωστόσο δόθηκε στο φωτοβολταϊκό και στο φορτίο του κεντρικού παρόχου καθώς, όπως είδαμε, η μπαταρία και ο σταθμός φόρτισης δεν

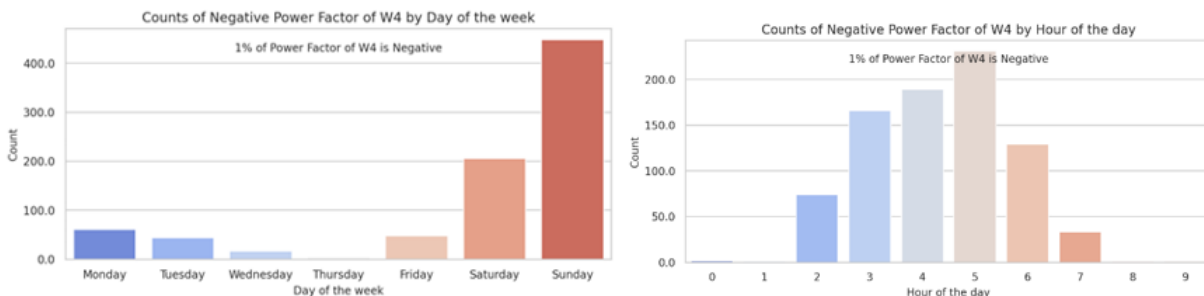
προσφέρουν ισχύ στο ηλεκτρικό δίκτυο, ενώ οι μετρήσεις του W2 είναι σχεδόν οι ίδιες με αυτές του φωτοβολταϊκού. Από την ανάλυση αυτή διεξήχθησαν τα παρακάτω αποτελέσματα.

Για τον συντελεστή ισχύος στο headquarter load αναμένουμε να έχει τιμή κοντά στο 1, αφού θέλουμε μέγιστη απόδοση και όσο το δυνατόν λιγότερες απώλειες στην ισχύ. Όπως φαίνεται στην Εικόνα 21, όπου παρουσιάζεται γραφικά η τιμή του συντελεστή ισχύος στο headquarter load για ένα διάστημα τριών ημερών, σε μεγάλο βαθμό ικανοποιείται η παραπάνω υπόθεση, αφού τις περισσότερες ώρες της ημέρας η τιμή του είναι πολύ κοντά στο ένα.



Εικόνα 21 - Γραφική αναπαράσταση συντελεστή ισχύος του W4 σε διάστημα τριών ημερών

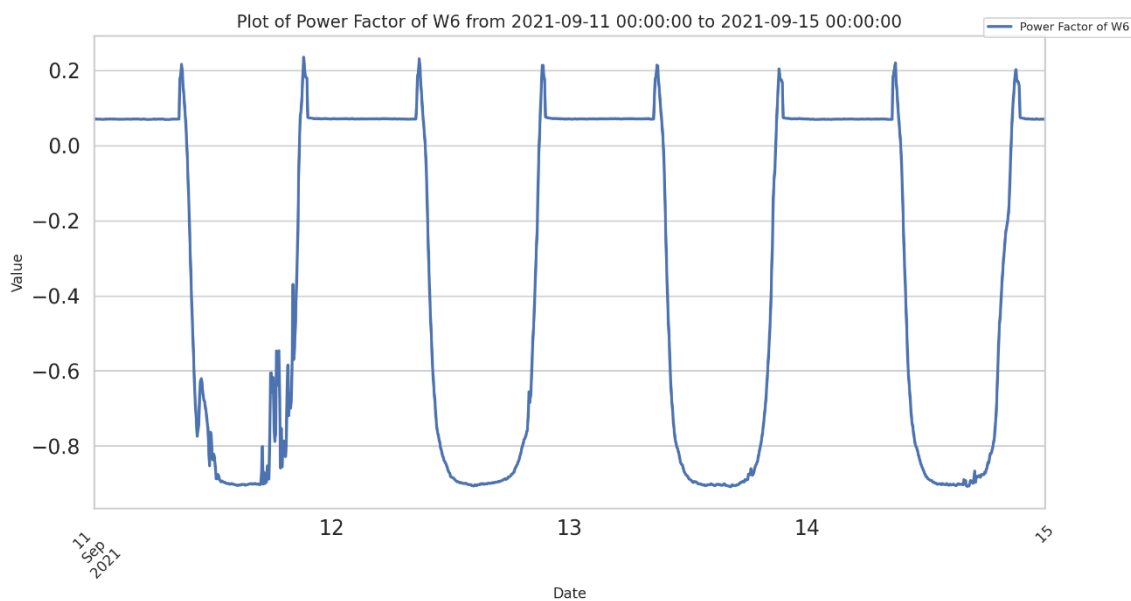
Μέσα από την περιγραφή των στατιστικών του δεικτών είδαμε ότι το 75% των τιμών του ανήκουν στο διάστημα  $[0.957, 0.984]$  ενώ εμφανίζει ελάχιστο στο  $-0.866$ . Σχετικά με τις αρνητικές τιμές, παρατηρήσαμε, έπειτα από ανάλυση, ότι αυτές εμφανίζονται κυρίως τα Σαββατοκύριακα μεταξύ των ωρών 02:00 – 07:00 και αντιστοιχούν σε λιγότερο από το 1% των συνολικών τιμών. Αυτά τα αποτελέσματα παρουσιάζονται στην Εικόνα 22.



Εικόνα 22 - Πλήθος αρνητικών τιμών του συντελεστή ισχύος του W4 ανά ημέρα της εβδομάδας και ώρα της ημέρας

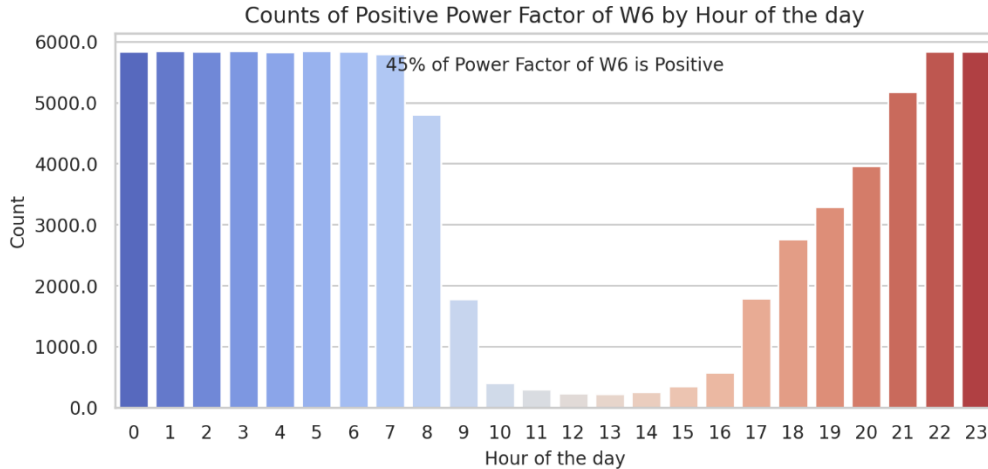
Για τον συντελεστή ισχύος του φωτοβολταϊκού (W6) βλέπουμε, όπως ήταν αναμενόμενο, σταδιακή αύξηση της τιμής του με την πάροδο της ώρας ώσπου να φτάσει στην μέγιστη τιμή του τις πρωινές – μεσημεριανές ώρες και μετά να μειωθεί ξανά σταδιακά μέχρι το μηδέν. Παρατηρούμε ότι υπάρχουν κάποιες θετικές τιμές αμέσως πριν και μετά την σταδιακή αύξηση και μείωση, αντίστοιχα, της τιμής του (Εικόνα 23).

Παρατηρήσαμε, ωστόσο, μέσα από την στατιστική περιγραφή του συντελεστή ισχύος του W6 ότι παρουσιάζει μέγιστη τιμή 1.31, που είναι αδύνατο, καθώς ο Σ.Ι. παίρνει τιμές στο  $[-1,1]$ . Ερευνώντας περαιτέρω το θέμα αυτό, παρατηρήθηκε ότι ο Σ.Ι. υπολογίζεται μεγαλύτερος του 1 σε δύο μετρήσεις. Οι μετρήσεις αυτές, προφανώς οφείλονται σε λάθη στη μέτρηση γι' αυτό και αντικαταστάθηκαν με NaNs.



Εικόνα 23 - Γραφική αναπαράσταση συντελεστή ισχύος του W6 σε διάστημα τριών ημερών

Στην Εικόνα 24 βλέπουμε ότι οι ώρες της ημέρας που εμφανίζονται αυτές οι θετικές τιμές είναι οι απογευματινές – βραδινές ώρες. Οι τιμές αυτές είναι πολύ μικρές, μικρότερες από 1kW και είναι αναμενόμενες, όπως έχει ήδη εξηγηθεί, καθώς αν και το φωτοβολταϊκό δεν λειτουργεί εκείνες τις ώρες, μπορεί να υπάρχει μια μικρή κατανάλωση ισχύος που οφείλεται συνήθως σε εξαρτήματα του φωτοβολταϊκού. Επομένως, έχοντας μια θετική ενεργό και φαινόμενη ισχύ, οδηγούμαστε σε θετικό συντελεστή ισχύος. Για να αποφευχθεί αυτή η κατανάλωση απαιτείται η αποσύνδεση του φωτοβολταϊκού από το δίκτυο, διαδικασία που απαιτεί την εισαγωγή περαιτέρω μηχανισμών και χρειάζεται καθημερινή παρακολούθηση, γι' αυτό και συνήθως δεν προτιμάται.

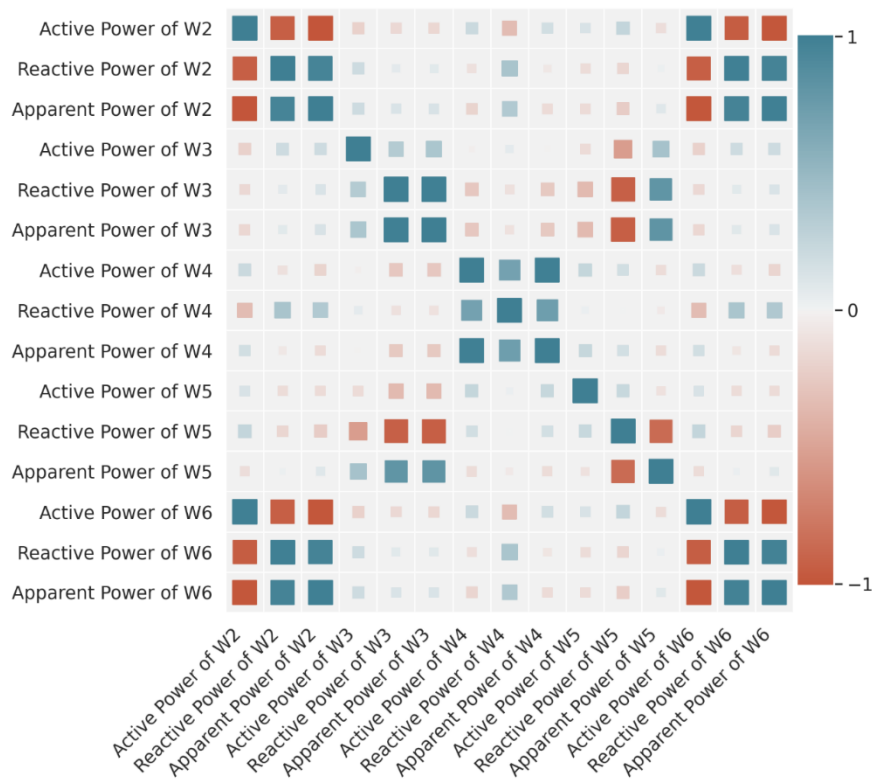


Εικόνα 24 - Πλήθος αρνητικών τιμών συντελεστή ισχύος W6 ανά ώρα της ημέρας

### 3.3.3.5. Συσχετίσεις μεταβλητών

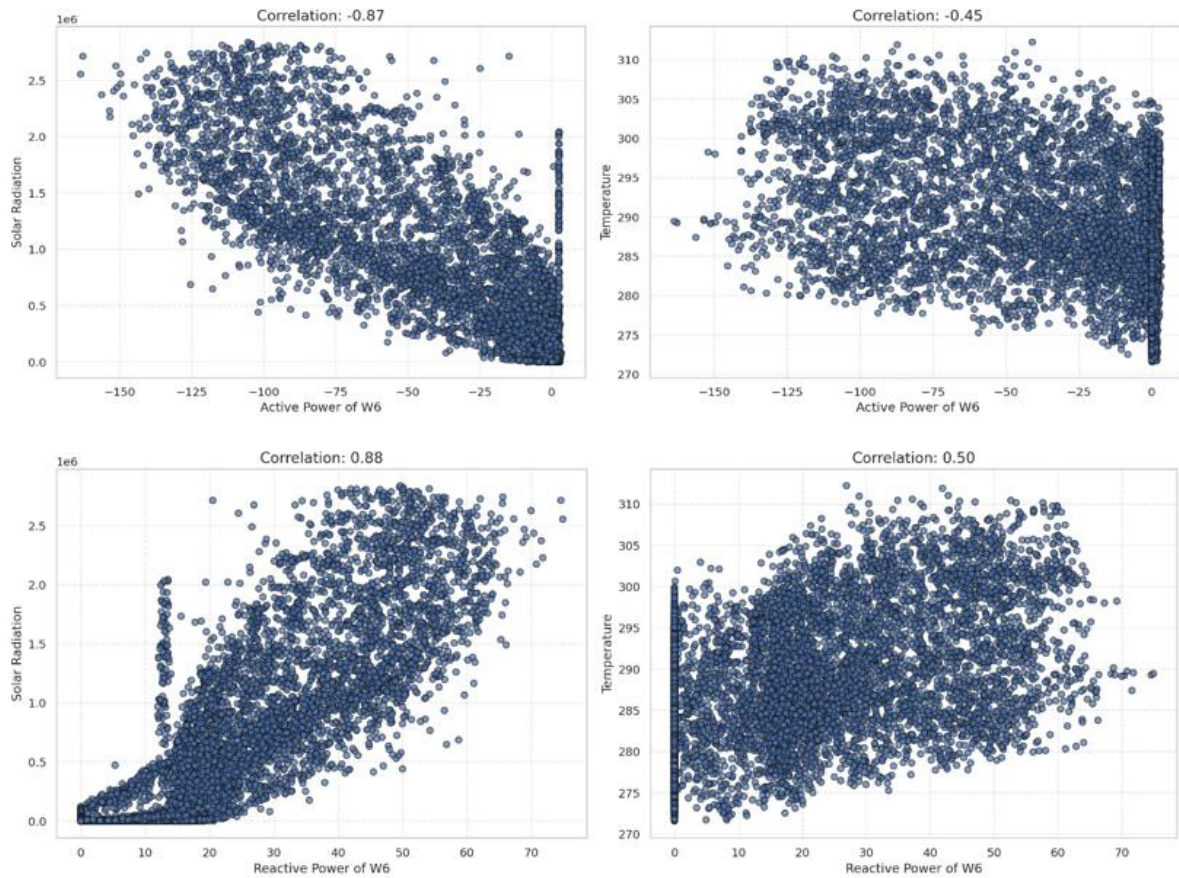
Τέλος σημαντική είναι η διερεύνηση των σχέσεων μεταξύ των μεταβλητών και της συσχέτισής τους. Για αυτόν τον σκοπό δημιουργήθηκαν πίνακες συσχέτισης τόσο για τα επιμέρους δεδομένα όσο και για το σύνολο αυτών (τάση, ρεύμα, ισχύς). Κάθε κελί του πίνακα αναπαριστά τη συσχέτιση μεταξύ δύο μεταβλητών. Το χρώμα και το μέγεθος που έχουν τα τετράγωνα των κελιών αντιπροσωπεύουν την κατεύθυνση και το μέγεθος της συσχέτισης, ώστε να είναι πιο εύκολα διακριτές οι ισχυρές θετικές ή αρνητικές συσχετίσεις. Στην Εικόνα 25 φαίνεται ο πίνακας συσχέτισης για τα δεδομένα της ισχύος που μετρήθηκαν από όλους τους μετρητές.

Όπως παρατηρείται υπάρχει ισχυρή αρνητική συσχέτιση της ενεργούς ισχύος των μετρητών W2 και W6 με τις αντίστοιχες άεργες, καθώς και θετική συσχέτιση της ενεργούς με την άεργο ισχύ του W4. Επιπλέον, έντονη θετική συσχέτιση υπάρχει μεταξύ της άεργου και της φαινόμενης ισχύος του W3 αλλά και έντονη αρνητική συσχέτιση μεταξύ της άεργου και της φαινόμενης ισχύος του W5. Πέρα από αυτά, βλέπουμε ισχυρή θετική συσχέτιση ανάμεσα στην ενεργό, άεργο και φαινόμενη ισχύ του φωτοβολταϊκού και των αντίστοιχων που μετρήθηκαν από τον μετρητή W2. Επιπλέον φαίνεται η άεργος ισχύς του γρήγορου σταθμού φόρτισης (W5) να έχει ισχυρή αρνητική συσχέτιση με την άεργο και την φαινόμενη ισχύ της μπαταρίας ( W3).



Εικόνα 25 - Πίνακας συσχέτισης δεδομένων ισχύος

Ερευνήθηκε, τέλος, η συσχέτιση των μετεωρολογικών δεδομένων με την ισχύ, ενεργό και άεργο των W4 και W6. Μετά από δημιουργία των αντίστοιχων scatterplots που δείχνουν γραφικά την συσχέτιση των μεταβλητών, παρατηρήθηκε η ισχυρή αρνητική (λόγω του πρόσημου της ισχύος) συσχέτιση της ενεργούς ισχύος με την ηλιακή ακτινοβολία και η μέτρια συσχέτισή της με την θερμοκρασία. Αντίστοιχα ήταν τα ευρήματα σχετικά με την άεργο ισχύ του W6, με μοναδική διαφορά το πρόσημο και το ελαφρώς μεγαλύτερο μέγεθος της συσχέτισης. Τα αποτελέσματα αυτά των συσχετίσεων της ισχύος του W6 με τα μετεωρολογικά δεδομένα αναπαρίστανται στην Εικόνα 26.

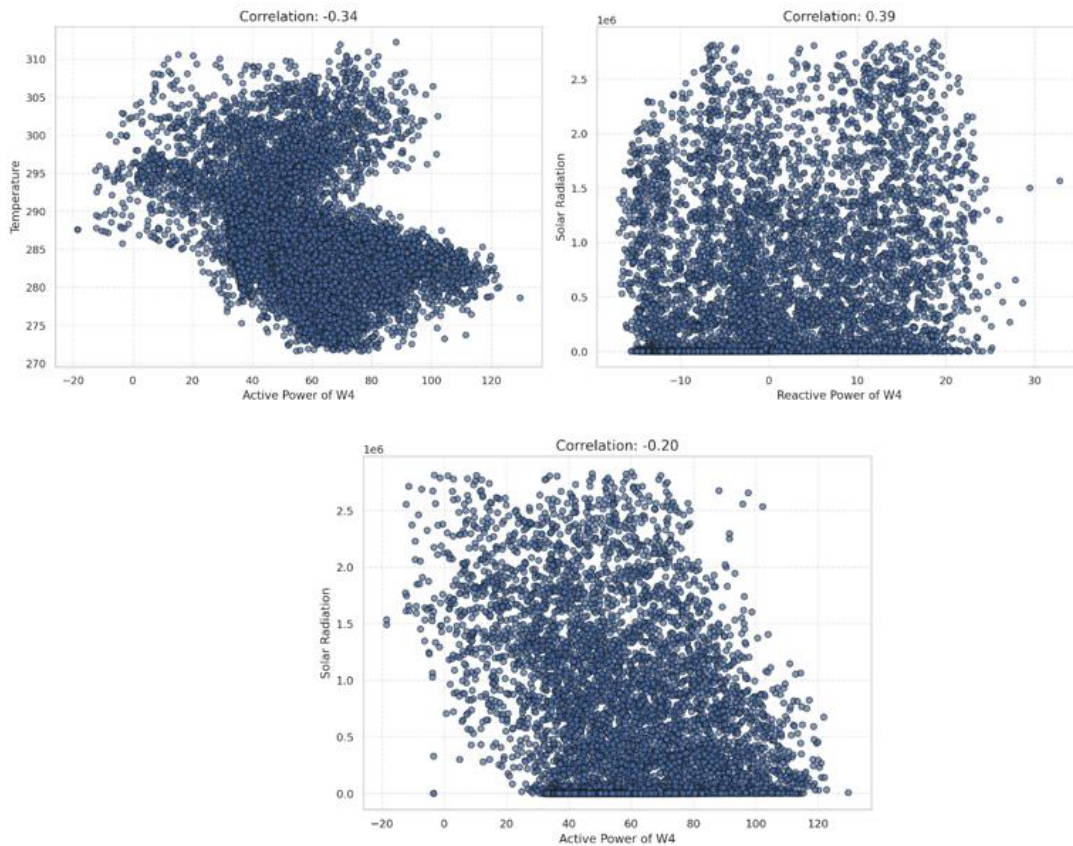


Εικόνα 26 - Scatterplots μετεωρολογικών δεδομένων και δεδομένων ισχύος του W6

Επιπλέον, σχετικά με την ισχύ του W4, παρατηρήθηκε μέτρια θετική και αρνητική συσχέτιση της αέργου και της ενεργού ισχύος της, αντίστοιχα, με την ηλιακή ακτινοβολία. Τέλος αρνητική συσχέτιση παρατηρήθηκε μεταξύ της ενεργούς ισχύος της και της θερμοκρασίας. Τα ευρήματα αυτά παρουσιάζονται στην Εικόνα 27.

Σχετικά με τα δεδομένα της νεφοκάλυψης δεν παρατηρήθηκε κάποια αξιοσημείωτη συσχέτιση ούτε με την ισχύ του W4 ούτε με του W6.





Εικόνα 27 - Scatterplots μετεωρολογικών δεδομένων και δεδομένων ισχύος του W4

### 3.4. Διαδικασία Πρόβλεψης

#### 3.4.1. Επιλογή Συμμεταβλητών

Η διάκριση των συμμεταβλητών, δηλαδή των μεταβλητών που δίνονται στο μοντέλο κατά τη διάρκεια της εκπαίδευσης μαζί με την χρονοσειρά που επιχειρείται να προβλεφθεί, γίνεται σε δύο κατηγορίες: 1) τις παρελθοντικές και 2) τις μελλοντικές συμμεταβλητές.

Παρελθοντικές είναι οι συμμεταβλητές των οποίων οι τιμές είναι γνωστές μέχρι και την χρονική στιγμή  $t$  της πρόβλεψης ενώ μελλοντικές είναι οι συμμεταβλητές των οποίων οι τιμές είναι γνωστές και μετά την χρονική στιγμή  $t$  της πρόβλεψης. Πιο συγκεκριμένα, για μια πρόβλεψη που γίνεται την χρονική στιγμή  $t$  με ορίζοντα πρόβλεψης  $n$ , τότε οι τιμές των μελλοντικών συμμεταβλητών για τις χρονικές στιγμές  $t, t+1, t+2, \dots, t+n$  είναι γνωστές. Παράδειγμα παρελθοντικών και μελλοντικών συμμεταβλητών αποτελούν η πραγματική τιμή της ηλιακής ακτινοβολίας και η πρόβλεψη της ηλιακής ακτινοβολίας, αντίστοιχα.



Τα δεδομένα που χρησιμοποιήθηκαν ως συμμεταβλητές είναι τα μετεωρολογικά δεδομένα και συγκεκριμένα η ηλιακή ακτινοβολία, το ποσοστό της νεφοκάλυψης και η θερμοκρασία του αέρα, τα ημερολογιακά δεδομένα και συγκεκριμένα η ώρα της ημέρας, καθώς και ο συνδυασμός τους.

Τα μετεωρολογικά δεδομένα μπορούν να χρησιμοποιηθούν ως παρελθοντικές συμμεταβλητές αφού αποτελούν πραγματικές τιμές, άγνωστες στο μέλλον, ενώ τα ημερολογιακά δεδομένα ως παρελθοντικές ή μελλοντικές συμμεταβλητές. Επιλέχθηκε, ωστόσο, η χρήση των ημερολογιακών δεδομένων ως παρελθοντικών συμμεταβλητών καθώς τα περισσότερα μοντέλα που χρησιμοποιήθηκαν δεν υποστηρίζουν την χρήση μελλοντικών συμμεταβλητών,

Η αδυναμία ενσωμάτωσης μελλοντικών συμμεταβλητών κάποιων μοντέλων της βιβλιοθήκης Darts, αποδίδεται, κυρίως, στην αρχιτεκτονική τους. Συγκεκριμένα, μελλοντικές συμμεταβλητές δέχονται τα μοντέλα που κάνουν προβλέψεις μία προς μία, συνολικού μήκους όσο ο ορίζοντας πρόβλεψης, όπως είναι για παράδειγμα το LightGBM μοντέλο. Εν αντιθέσει, μοντέλα που πραγματοποιούν προβλέψεις σε blocks, όπως είναι τα μοντέλα NBEATS και TCN που αναλύθηκαν στην θεωρία, δεν δέχονται μελλοντικές μεταβλητές.

#### 3.4.2. Επιλογή και Εκπαίδευση Μοντέλων

Αρχικά, λόγω των πολλών συνεχομένων κενών τιμών, όπως αυτές αναλύθηκαν στο προηγούμενο κεφάλαιο, επιλέξαμε να χρησιμοποιήσουμε την μέθοδο `extract_subseries` της βιβλιοθήκης Darts της Python. Μέσω αυτής, η εκάστοτε χρονοσειρά που πρόκειται να προβλεφθεί χωρίστηκε σε επιμέρους χρονοσειρές, οι οποίες περιλαμβάνουν το πολύ 4 κενές, διαδοχικές ώρες. Επιπλέον οι χρονοσειρές που ήταν πολύ μικρές, λιγότερο από δύο εβδομάδες δεδομένων, απορρίφθηκαν, ενώ ως τελευταία χρονοσειρά επιλέχθηκε εκείνη που περιείχε δεδομένα τουλάχιστον δύο μηνών, ώστε να μπορεί να χρησιμοποιηθεί στη διαδικασία της αξιολόγησης της πρόβλεψης. Στη συνέχεια πραγματοποιήθηκε κάλυψη των κενών που απέμειναν με απλή γραμμική παρεμβολή.

Ο χρονικός ορίζοντας της πρόβλεψης έχει τεθεί σε μία ημέρα (day-ahead forecast). Για την πρόβλεψη χρησιμοποιήθηκαν 3 μοντέλα, το μοντέλο μηχανικής μάθησης LightGBM και τα μοντέλα βαθιάς μάθησης N-BEATS και TCN καθώς έχουν αποδειχτεί ιδιαίτερα αποδοτικά στην βραχυπρόθεσμη πρόβλεψη (short term forecasting) χρονοσειρών. [66][67][68][1]. Επιπλέον, εξετάστηκε η χρήση της τεχνικής Ensemble στο μοντέλο N-BEATS, με χρήση του μοντέλου NaiveEnsemble της βιβλιοθήκης Darts. Δοκιμάστηκαν πολλαπλά παράθυρα προηγούμενης γνώσης (lookback window) για όλα τα μοντέλα, από μια έως δέκα ημέρες, και αποφασίστηκε η μία εβδομάδα ως η πιο αποδοτική επιλογή (7\*24 τιμές). Αντίστοιχα, για το παράθυρο προηγούμενης γνώσης των δεδομένων των συμμεταβλητών, όπου αυτές χρησιμοποιήθηκαν, έπειτα από δοκιμές, επιλέχθηκε ως βέλτιστη τιμή η μια ημέρα (24 τιμές). Επιπλέον, ως σημείο αναφοράς (benchmark) χρησιμοποιήθηκε το Naïve Seasonal μοντέλο, με περίοδο εποχιακότητας τις 24 ώρες.

Οι επιμέρους επιλογές που έγιναν στην εκπαίδευση κάθε μοντέλου αναλύονται στην συνέχεια:

- LightGBM Model

Για την εκπαίδευση του μοντέλου μηχανικής μάθησης (ML) LightGBM δόθηκε ως σύνολο εκπαίδευσης (train set), το σύνολο των subseries που διαχωρίσαμε, εκτός της τελευταίας χρονοσειράς της λίστας των subseries, που είχε μήκος 60 ημερών, δηλαδή  $60 \times 24 = 1440$  data points. Το μοντέλο δοκιμάστηκε τόσο με κανονικοποιημένα δεδομένα όσο και με μη κανονικοποιημένα.

- Neural basis expansion analysis time series forecasting (N-BEATS) Model

Για την εκπαίδευση του μοντέλου βαθιάς μάθησης (DL) N-BEATS δόθηκε ως σύνολο εκπαίδευσης (train set), το σύνολο των subseries που διαχωρίσαμε εκτός της τελευταίας χρονοσειράς της λίστας των subseries. Από την αυτή την τελευταία χρονοσειρά, που είχε μήκος 2 μηνών, ο ένας μήνας δόθηκε ως σύνολο αξιολόγησης (validation set). Το μοντέλο εκπαιδεύτηκε με κανονικοποιημένα δεδομένα καθώς τα νευρωνικά λειτουργούν καλύτερα με δεδομένα που έχουν κανονικοποιηθεί.

Για τη βελτιστοποίηση του μοντέλου πραγματοποιήθηκε hyperparameter tuning, χρησιμοποιώντας τη μέθοδο *gridsearch* της βιβλιοθήκης Darts. Πρόκειται για την διαδικασία στην οποία σε κάθε παράμετρο του μοντέλου που επιλέγουμε, δίνεται μια λίστα με πολλαπλές τιμές και, στη συνέχεια, μέσω του grid search ελέγχεται κάθε πιθανός συνδυασμός αυτών των παραμέτρων και επιλέγεται ο βέλτιστος, δηλαδή αυτός που μειώνει το συγκεκριμένο δείκτη που ορίζουμε κατά την διαδικασία. Λόγω του ότι η διαδικασία είναι πολύ χρονοβόρα, επιλέξαμε να κρατήσουμε σταθερό το πλάτος των layers, width = 64, καθώς και το batch\_size = 1024. Η επιλογή του πλάτους των layers έγινε κρίνοντας από αντίστοιχες έρευνες [67] ενώ η επιλογή του batch\_size προέκυψε δοκιμάζοντας διάφορες επιλογές σε μεμονωμένα μοντέλα. Τέλος, σχετικά με τον αριθμό των epochs που δίνεται για την εκπαίδευση του μοντέλου, επιλέξαμε έναν αρκετά μεγάλο αριθμό, epochs = 500, καθώς με την μέθοδο του Early Stopping που αναλύεται στην συνέχεια, το μοντέλο θα σταματήσει να εκπαιδεύεται αρκετά νωρίτερα.

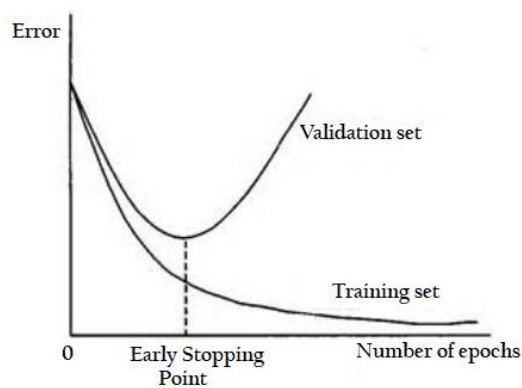
Αυτή η διαδικασία της βελτιστοποίησης του N-BEATS διήρκεσε 390 λεπτά και παρήγαγε τις ακόλουθες τιμές για τις υπερπαραμέτρους του μοντέλου.

Συγκεντρωτικά, οι βέλτιστες τιμές καθώς και η λίστα των τιμών που δόθηκε σε κάθε υπερπάρμετρο κατά την διαδικασία του tuning, παρουσιάζονται στον Πίνακα 2.

<i>Hyperparameter</i>	<i>Value</i>	<i>Grid</i>
<b>Epochs</b>	<b>500</b>	<b>500</b>
<b>Width (d)</b>	<b>64</b>	<b>64</b>
<b>Blocks (R)</b>	<b>5</b>	<b>[3, 4, 5, 6]</b>
<b>Layers (L)</b>	<b>2</b>	<b>[2, 3]</b>
<b>Lookback period (<i>lw</i>, <i>days</i>)</b>	<b>7</b>	<b>7</b>
<b>Stacks</b>	<b>7</b>	<b>[3, 5, 7]</b>
<b>Batch size</b>	<b>1024</b>	<b>1024</b>

Πίνακας 2 - Τιμές υπερπαραμέτρων μοντέλου N-BEATS

Χρησιμοποιήθηκε, επιπλέον, η μέθοδος του Early Stopping. Πρόκειται για μια μέθοδο για την αποφυγή του overfitting του μοντέλου. Ο όρος overfitting περιγράφει ουσιαστικά την περίπτωση που το μοντέλο μαθαίνει υπερβολικά καλά το σύνολο εκπαίδευσης, σχεδόν απομνημονεύοντάς το, με αποτέλεσμα να μη λειτουργεί καλά στο test set, όπου τα δεδομένα δεν είναι πανομοιότυπα με αυτά του train set, αφού δεν έχει καταφέρει τελικά να μάθει μοτίβα και σχήματα. Με την μέθοδο του Early Stopping, ελέγχεται σε κάθε epoch η ακρίβεια του μοντέλου και όταν αυτή σταματήσει να αυξάνεται (δηλαδή το validation loss σταματήσει να μειώνεται) ολοκληρώνεται η εκπαίδευση του μοντέλου[69].



Εικόνα 28- Γραφική εξήγηση μεθόδου Early Stopping

Στη συγκεκριμένη περίπτωση επιλέξαμε να συνεχίσουμε την εκπαίδευση για άλλα 10 epochs αφού βρεθεί το ελάχιστο validation loss, για να είμαστε βέβαιοι ότι δεν πρόκειται να αυξηθεί ξανά η ακρίβεια.

- Naïve Ensemble Model

Για τη σύνθεση του μοντέλου NaiveEnsemble Model, της βιβλιοθήκης Darts, εκπαιδεύτηκαν 10 N-BEATS μοντέλα, με τις υπερπαραμέτρους που αναφέρθηκαν στην ανάλυση του N-BEATS μοντέλου και που προέκυψαν από το hyperparameter tuning, με τη διαφορά ότι σε κάθε μοντέλο δόθηκε διαφορετική τυχαία αρχικοποίηση των βαρών του νευρωνικού (random state) [67]. Συγκεκριμένα δόθηκαν οι παρακάτω τιμές:

```
random_states = [5, 6,7, 8, 10, 12, 14,15, 16, 18]
```

Ως αριθμός εποχών δόθηκε σε κάθε μοντέλο ο βέλτιστος αριθμός που βρέθηκε μέσω της διαδικασίας Early Stopping κατά την εκπαίδευση του απλού μοντέλου N-BEATS. Το Naïve Ensemble Model της βιβλιοθήκης Darts, επιστρέφει ως τελική πρόβλεψη τον μέσο όρο των προβλέψεων των επιμέρους μοντέλων.

- Temporal Convolutional Network (TCN)

Στην περίπτωση του μοντέλου TCN, όπως και στην περίπτωση του μοντέλου βαθιάς μάθησης N-BEATS, δόθηκε για την εκπαίδευση του μοντέλου ως σύνολο εκπαίδευσης η λίστα με τις επιμέρους χρονοσειρές (subseries) εκτός της τελευταίας χρονοσειράς της λίστας. Από την αυτή την τελευταία χρονοσειρά, μήκους 2 μηνών, ο ένας μήνας δόθηκε ως σύνολο αξιολόγησης (validation set). Επιπλέον, χρησιμοποιήθηκε η μέθοδος Early Stopping για την βελτιστοποίηση του αριθμού των εποχών που εκπαιδεύεται το μοντέλο και την αποφυγή του overfitting. Η αρχιτεκτονική του μοντέλου που χρησιμοποιήθηκε είναι αυτή του Πίνακα 3.

***TCN Architecture***

<b><i>Parameter</i></b>	<b><i>Value</i></b>
Dilation base	2
Kernel size	5
Number of filters	5

*Πίνακας 3 -Τιμές παραμέτρων μοντέλου TCN*

### 3.4.3. Αξιολόγηση Μοντέλων

Για την αξιολόγηση των μοντέλων πρόβλεψης υπάρχουν πολλοί δείκτες ακριβείας που μπορούν να χρησιμοποιηθούν, όπως αναλύθηκε στο Κεφάλαιο 2.2.5. Στατιστική Ανάλυση Ακριβείας Προβλέψεων. Ωστόσο, η εύρεση των κατάλληλων δεικτών που ανταποκρίνονται στα δεδομένα

μας και ερμηνεύουν σωστά τα αποτελέσματα των προβλέψεων, είναι μια διαδικασία που απαιτεί προσοχή καθώς επιφυλάσσει κινδύνους για τη σωστή αξιολόγηση των αποτελεσμάτων.

Έρευνες στην αξιολόγηση μοντέλων για την πρόβλεψη της παραγόμενης ισχύος φωτοβολταϊκών και αιολικών πηγών έχουν δείξει ότι δείκτες όπως ο Mean Absolute Percentage Error (MAPE) και ο Root Mean Squared Error (RMSE) δεν είναι κατάλληλοι, καθώς η παραγόμενη ισχύς από αυτές τις πηγές είναι μηδενική – (ή πολύ κοντά στο μηδέν) αρκετές ώρες της ημέρας. Έτσι οι τιμές των δεικτών είναι πολύ υψηλές και αν λαμβάνονταν υπόψη το μοντέλο που αξιολογούν θα κρίνονταν ακατάλληλο, χωρίς αυτό να είναι απαραίτητα αλήθεια [70], [71].

Για την αξιολόγηση των αποτελεσμάτων αποφασίστηκε τελικά να χρησιμοποιηθούν οι στατιστικοί δείκτες MASE και MAE. Σχετικά με τον δείκτη MASE, η επιλογή έγινε λόγω των πλεονεκτημάτων στη χρήση του που αναλύθηκαν στο κεφάλαιο 2.2.5. Στατιστική Ανάλυση Ακριβείας Προβλέψεων όπως το ότι δεν εξαρτάται από την κλίμακα των δεδομένων και μπορεί να χρησιμοποιηθεί σε αρνητικά, ή πολύ κοντά στο μηδέν, δεδομένα, όπως συμβαίνει στην δική μας περίπτωση [72]–[74]. Τέλος, η επιλογή του δείκτη Mean Absolute Error (MAE) έγινε καθώς έχει χρησιμοποιηθεί ευρέως σε αντίστοιχα προβλήματα βραχυπρόθεσμης πρόβλεψης ισχύος, με πολύ καλά αποτελέσματα [75]–[78].

## 4. Αποτελέσματα

Στο παρόν κεφάλαιο παρουσιάζονται τα αποτελέσματα των μοντέλων που χρησιμοποιήθηκαν. Θα ερευνήσουμε το κάθε μοντέλο ξεχωριστά. Κάθε μοντέλο χρησιμοποιήθηκε για να προβλέψει την ενεργό αλλά και την άεργο ισχύ που μετρείται από τους μετρητές W4 και W6.

### 4.1. Αποτελέσματα Μοντέλων

#### 4.1.1. Αποτελέσματα Μοντέλου LightGBM

Αρχικά, εξετάστηκε αν το μοντέλο έχει καλύτερα αποτελέσματα με κανονικοποιημένα, ή όχι, δεδομένα. Δοκιμάστηκαν, στη συνέχεια, να δοθούν στο μοντέλο ως παρελθοντικές συμμεταβλητές τα μετεωρολογικά δεδομένα, δηλαδή τα δεδομένα της θερμοκρασίας, της ακτινοβολίας και της νεφοκάλυψης, τα ημερολογιακά δεδομένα, και συγκεκριμένα η ώρα της ημέρας, καθώς και ο συνδυασμός τους. Τα αποτελέσματα παρατίθενται στους Πίνακες 4, 5, 6 και 7 που ακολουθούν, με καταγραφή των δεικτών MASE και MAE σε κάθε περίπτωση.

<i>Active Power measured by W4</i>			
Normalized Data	Model	MASE	MAE
No	Purely Autoregressive	0.82	8.59
Yes	Purely Autoregressive	0.81	8.56
No	Weather Covariates	0.77	8.15
<b>Yes</b>	<b>Weather Covariates</b>	<b>0.77</b>	<b>8.12</b>
No	Weather & Calendar Covariates	0.81	8.54
Yes	Weather & Calendar Covariates	0.78	8.19
No	Calendar Covariates	0.85	8.96
Yes	Calendar Covariates	0.84	8.81

Πίνακας 4 - Δείκτες ακριβείας πρόβλεψης της ενεργούς ισχύος του φορτίου του παρόχου με το μοντέλο LightGBM, ανά περίπτωση

<i>Reactive Power measured by W4</i>			
Normalized Data	Past Covariates	MASE	MAE
No	Purely Autoregressive	1.09	2.93
Yes	Purely Autoregressive	1.05	2.83
No	Weather Covariates	1.07	2.89
Yes	Weather Covariates	1.07	2.88
No	Weather & Calendar Covariates	1.07	2.88
Yes	Weather & Calendar Covariates	1.07	2.88
No	Calendar Covariates	1.05	2.84
<b>Yes</b>	<b>Calendar Covariates</b>	<b>1.03</b>	<b>2.77</b>

Πίνακας 5 - Δείκτες ακριβείας πρόβλεψης της άεργου ισχύος του φορτίου του παρόχου με το μοντέλο LightGBM, ανά περίπτωση

<i>Active Power measured by W6</i>			
Normalized Data	Past Covariates	MASE	MAE
No	Purely Autoregressive	0.77	6.91
Yes	Purely Autoregressive	0.77	6.91
No	Weather Covariates	0.53	4.76
Yes	Weather Covariates	0.55	4.93
<b>No</b>	<b>Weather &amp; Calendar Covariates</b>	<b>0.50</b>	<b>4.53</b>
Yes	Weather & Calendar Covariates	0.53	4.80
No	Calendar Covariates	0.80	7.17
Yes	Calendar Covariates	0.80	7.17

Πίνακας 6 - Δείκτες ακριβείας πρόβλεψης της ενεργούς ισχύος του φωτοβολταϊκού με το μοντέλο LightGBM ανα περίπτωση

<i>Reactive Power measured by W6</i>			
Normalized Data	Past Covariates	MASE	MAE
No	Purely Autoregressive	0.77	2.51
Yes	Purely Autoregressive	0.79	2.58
No	Weather Covariates	0.57	1.85
<b>Yes</b>	<b>Weather Covariates</b>	<b>0.53</b>	<b>1.72</b>
No	Weather & Calendar Covariates	0.55	1.78
Yes	Weather & Calendar Covariates	0.55	1.80
No	Calendar Covariates	0.77	2.51
Yes	Calendar Covariates	0.77	2.52

Πίνακας 7 - Δείκτες ακριβείας πρόβλεψης της αέργου ισχύος του φωτοβολταϊκού με το μοντέλο LightGBM ανα περίπτωση

Όπως βλέπουμε στους Πίνακες 4 **Error! Reference source not found.** και 5, το μοντέλο στην πρόβλεψη της ισχύος που μετριέται από τον W4, λειτουργεί καλύτερα με μη κανονικοποιημένα δεδομένα, σε όλες τις περιπτώσεις. Το ίδιο ισχύει στην πρόβλεψη της ισχύος του φωτοβολταϊκού, μόνο όμως όταν δε δίνονται παρελθοντικές συμμεταβλητές στο μοντέλο (Πίνακας 6, Πίνακας 7).

Σχετικά με τις παρελθοντικές συμμεταβλητές, παρατηρήθηκε ότι για την ενεργό ισχύ του W4, η προσθήκη τόσο των μετεωρολογικών όσο και του συνδυασμού τους με τα ημερολογιακά δεδομένα στο μοντέλο βελτιώνει την απόδοση του. Επιπλέον τα ημερολογιακά δεδομένα μεμονωμένα φαίνεται να χειροτερεύουν την απόδοση του μοντέλου, τόσο στην περίπτωση των φυσικών όσο και των κανονικοποιημένων δεδομένων. Για την αέργου ισχύ του W4, παρατηρήθηκε ότι η απόδοση του μοντέλου σε όλες τις περιπτώσεις βελτιώνεται με την προσθήκη παρελθοντικών συμμεταβλητών.

Σχετικά με την ισχύ του φωτοβολταϊκού, παρατηρήθηκε ότι το μοντέλο βελτιώνεται αισθητά με την προσθήκη των μετεωρολογικών δεδομένων. Συγκεκριμένα, στην περίπτωση της ενεργούς ισχύος του W6, η προσθήκη των μετεωρολογικών δεδομένων βελτιώνει την μετρική απόδοσης MASE κατά 31.2%, καθώς η τιμή του μειώνεται από 0.77 σε 0.53. Επιπλέον ο συνδυασμός των μετεωρολογικών και των ημερολογιακών δεδομένων φαίνεται να βελτιώνει περαιτέρω την απόδοση του μοντέλου, με μείωση του MASE κατά 5.66%. Το γεγονός αυτό ήταν αναμενόμενο

αν αναλογιστεί κανείς την υψηλή εξάρτηση της παραγόμενης ισχύος του από τα καιρικά φαινόμενα. Τέλος τα ημερολογιακά δεδομένα μεμονωμένα φαίνεται να χειροτερεύουν ή να μη βελτιώνουν την απόδοση του μοντέλου.

Ο χρόνος εκπαίδευσης του μοντέλου αποτελεί ένα από τα σημαντικότερα πλεονεκτήματά του καθώς σε καμία περίπτωση δεν ξεπέρασε τα λίγα δευτερόλεπτα.

Συγκεντρωτικά, σχετικά με τη χρήση του μοντέλου LightGBM παρατηρείται ότι έχει καλή απόδοση στην πρόβλεψη των τιμών της ενεργούς και αέργου ισχύος του W6 αλλά και της ενεργούς ισχύος του W4 ενώ δεν αποδίδει ικανοποιητικά στην πρόβλεψη της αέργου ισχύος του W4, αφού η μετρική απόδοσης MASE είναι, σε όλες τις περιπτώσεις, μεγαλύτερη της μονάδας.

#### 4.1.2. Αποτελέσματα Μοντέλου N-BEATS

Όπως και προηγουμένως, αρχικά εξετάστηκε εάν το μοντέλο λειτουργεί καλύτερα με κανονικοποιημένα ή μη δεδομένα, παρατηρώντας σημαντική βελτίωση της απόδοσής του μοντέλου όταν δέχεται κανονικοποιημένα δεδομένα. Για τον λόγο αυτό, σε όλα τα πειράματα που διενεργήθηκαν και καταγράφονται στους Πίνακες 8 και 9, χρησιμοποιήθηκαν κανονικοποιημένα δεδομένα. Στα πειράματα αυτά, ελέγχθηκε η απόδοση του μοντέλου όταν αυτό εκπαιδεύεται χωρίς αλλά και με παρελθοντικές συμμεταβλητές. Ως τέτοιες επιλέχθηκαν τα μετεωρολογικά και τα ημερολογιακά δεδομένα καθώς και ο συνδυασμός τους. Τα αποτελέσματα φαίνονται στους παρακάτω πίνακες.

Model	Past Covariates	Active Power measured by W4		Reactive Power measured by W4	
		MASE	MAE	MASE	MAE
N-BEATS	Purely Autoregressive	0.81	8.57	1.03	2.78
N-BEATS	Weather Covariates	0.94	9.87	1.13	3.05
N-BEATS	Weather & Calendar Covariates	1	10.58	1.12	3.01
N-BEATS	Calendar Covariates	1.02	10.73	1.12	3.01
<b>Ensemble N-BEATS</b>	<b>Purely Autoregressive</b>	<b>0.77</b>	<b>8.10</b>	<b>1.03</b>	<b>2.77</b>

Πίνακας 8 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W4 με το μοντέλο N-BEATS ανα περίπτωση



Model	Past Covariates	Active Power measured by W6		Reactive Power measured by W6	
		MASE	MAE	MASE	MAE
N-BEATS	Purely Autoregressive	1.04	9.33	1.15	3.75
N-BEATS	Weather Covariates	0.99	8.91	1.28	4.17
N-BEATS	Weather & Calendar Covariates	1.03	9.21	1.37	4.48
N-BEATS	Calendar Covariates	1.26	11.28	1.31	4.29
<b>Ensemble N-BEATS</b>	<b>Purely Autoregressive</b>	<b>0.87</b>	<b>7.83</b>	<b>1.15</b>	<b>3.75</b>

Πίνακας 9 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W6 με το μοντέλο N-BEATS ανά περίπτωση

Όπως παρατηρείται από τους παραπάνω πίνακες, η χειροτέρευση της απόδοσης του μοντέλου είναι σημαντική με την εισαγωγή των παρελθοντικών συμμεταβλητών και ειδικά των ημερολογιακών δεδομένων.

Συγκεκριμένα, στη χρονοσειρά της ενεργούς ισχύος που μετρήθηκε από τον W4, η εισαγωγή των ημερολογιακών δεδομένων αυξάνει τον MASE κατά 25.9%. Για το γεγονός αυτό ευθύνεται, εν μέρει, η αρχιτεκτονική του μοντέλου καθώς μέσω της διαδικασίας εκπαίδευσής του και της μεθόδου backcasting που αναφέρθηκε στην ενότητα 2.3.2. Μοντέλα Μηχανικής και Βαθιάς Μάθησης, το μοντέλο είναι ικανό να αναγνωρίσει τις χρονικές συσχετίσεις. Η αρνητική επίδραση των συμμεταβλητών του καιρού στην απόδοση του μοντέλου, ωστόσο, δεν είναι αναμενόμενη. Παρατηρείται ότι η μόνη περίπτωση στην οποία η προσθήκη των μετεωρολογικών δεδομένων βελτιώνει ελαφρώς την απόδοση του μοντέλου είναι η περίπτωση της ενεργούς ισχύος του W6, με μείωση του MASE κατά, περίπου, 4.8%.

Ο χρόνος εκπαίδευσης του N-BEATS, καθώς πρόκειται για ένα μοντέλο βαθιάς μάθησης με πολύπλοκη αρχιτεκτονική, είναι αρκετά μεγαλύτερος του μοντέλου LightGBM. Κατά μέσο όρο χρειάστηκαν 8 λεπτά για την εκπαίδευση του μοντέλου. Ωστόσο υπήρχαν σημαντικές διαφοροποιήσεις στον χρόνο εκπαίδευσης μεταξύ των ισχύων του W4 και W6. Αυτό οφείλεται στον αριθμό των εποχών που χρειάστηκαν τα αντίστοιχα μοντέλα για την εκπαίδευση αφού για παράδειγμα για την ενεργό ισχύ του W4 χρειάστηκαν 66 epochs μέχρι να βρεθεί το ελάχιστο validation loss ενώ για την ενεργό ισχύ του W6 χρειάστηκαν μόλις 22.

Τέλος, καταγράφηκαν τα αποτελέσματα του Ensemble μοντέλου. Συγκεκριμένα, επιλέχθηκαν τα N-BEATS μοντέλα που είχαν την καλύτερη απόδοση, για κάθε μέγεθος που προβλέπεται, και εκπαιδεύτηκαν δέκα τέτοια ίδια μοντέλα. Επομένως, για την ενεργό και την άεργο ισχύ του W4 επιλέχθηκε να εκπαιδευτεί το μοντέλο χωρίς παρελθοντικές συμμεταβλητές, ενώ για την ενεργό και την άεργο ισχύ του W6 εκπαιδεύτηκαν δύο Ensemble μοντέλα, τα οποία είχαν παρόμοια

αποτελέσματα: ένα με τα μετεωρολογικά δεδομένα ως παρελθοντικές συμμεταβλητές και ένα χωρίς.

Η χρήση της τεχνικής ensemble οδηγεί σε αρκετά καλύτερα αποτελέσματα όσον αφορά τις ενεργές ισχύς και ιδιαίτερα αυτή του W6, καθώς μειώνει τον δείκτη MASE κατά 7.8%.

Ο μέσος χρόνος εκπαίδευσης εδώ είναι τα 45 λεπτά, με μέγιστο χρόνο εκπαίδευσης αυτόν του μοντέλου για την πρόβλεψη της ισχύος του W4, τα 60 λεπτά.

Συγκεντρωτικά, το μοντέλο φαίνεται να αποδίδει ικανοποιητικά στην πρόβλεψη της ενεργούς ισχύος του W4 και του W6 ενώ η απόδοση του δεν είναι καλή στην πρόβλεψη της αέργου ισχύος.

#### 4.1.3. Αποτελέσματα Μοντέλου TCN

Όπως και στην περίπτωση του μοντέλου βαθιάς μάθησης, N-BEATS, έτσι και στην περίπτωση του μοντέλου TCN παρατηρήθηκε, μετά από δοκιμή, ότι αποδίδει καλύτερα με κανονικοποιημένα δεδομένα ισχύος. Γι' αυτό το λόγο στα παρακάτω πειράματα έχουν χρησιμοποιηθεί μόνο κανονικοποιημένα δεδομένα.

Στους Πίνακες 10 και 11, παρατίθενται τα αποτελέσματα των πειραμάτων με το μοντέλο βαθιάς μάθησης, TCN καθώς και οι διάφορες παρελθοντικές συμμεταβλητές που χρησιμοποιήθηκαν.

Past Covariates	Active Power measured by W4		Reactive Power measured by W4	
	MASE	MAE	MASE	MAE
Purely Autoregressive	1.03	10.82	<b>1.06</b>	<b>2.84</b>
<b>Weather Covariates</b>	<b>0.93</b>	<b>9.79</b>	1.11	2.99
Weather & Calendar Covariates	0.97	10.24	1.12	3.01
Calendar Covariates	0.96	10.12	1.10	2.95

Πίνακας 10 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W4 με το μοντέλο TCN ανά περίπτωση

Past Covariates	Active Power measured by W6		Reactive Power measured by W6	
	MASE	MAE	MASE	MAE
Purely Autoregressive	0.92	8.30	0.83	2.70
<b>Weather Covariates</b>	<b>0.74</b>	<b>6.61</b>	<b>0.80</b>	<b>2.60</b>
Weather & Calendar Covariates	0.76	6.83	0.97	3.16
Calendar Covariates	0.81	7.31	0.84	2.77

Πίνακας 11 - Δείκτες ακριβείας πρόβλεψης της ισχύος του W6 με το μοντέλο TCN ανά περίπτωση

Παρατηρείται έντονη θετική επίδραση των παρελθοντικών συμμεταβλητών στο μοντέλο σε όλες τις περιπτώσεις. Πιο ισχυρή φαίνεται, ωστόσο, η επίδραση των μετεωρολογικών δεδομένων, καθώς στην περίπτωση της ενεργούς ισχύος του W4, για παράδειγμα, βελτιώνει τη μετρική απόδοσης MASE κατά 9.8%, αφού από 1.03 μειώνεται σε 0.93, ενώ στην περίπτωση της ενεργούς ισχύος του W6 βελτιώνει αξιοσημείωτα την ίδια μετρική, κατά ποσοστό 19.5%, με μείωσή της τιμής της από 0.92 σε 0.74. Μοναδικές εξαιρέσεις αποτελούν η περίπτωση της αέργου ισχύος του W4, όπου η εισαγωγή των παρελθοντικών συμμεταβλητών φαίνεται να χειροτερεύει ελαφρώς τα αποτελέσματα του μοντέλου και η περίπτωση της αέργου ισχύος του W6 όπου τα ημερολογιακά δεδομένα φαίνεται να χειροτερεύουν το μοντέλο ενώ η θετική επίδραση των μετεωρολογικών δεδομένων είναι μικρή.

Ο χρόνος εκπαίδευσης του μοντέλου ήταν σαφώς μεγαλύτερος από αυτόν του N-BEATS μοντέλου, με μέσο χρόνο εκπαίδευσης τα 20 λεπτά.

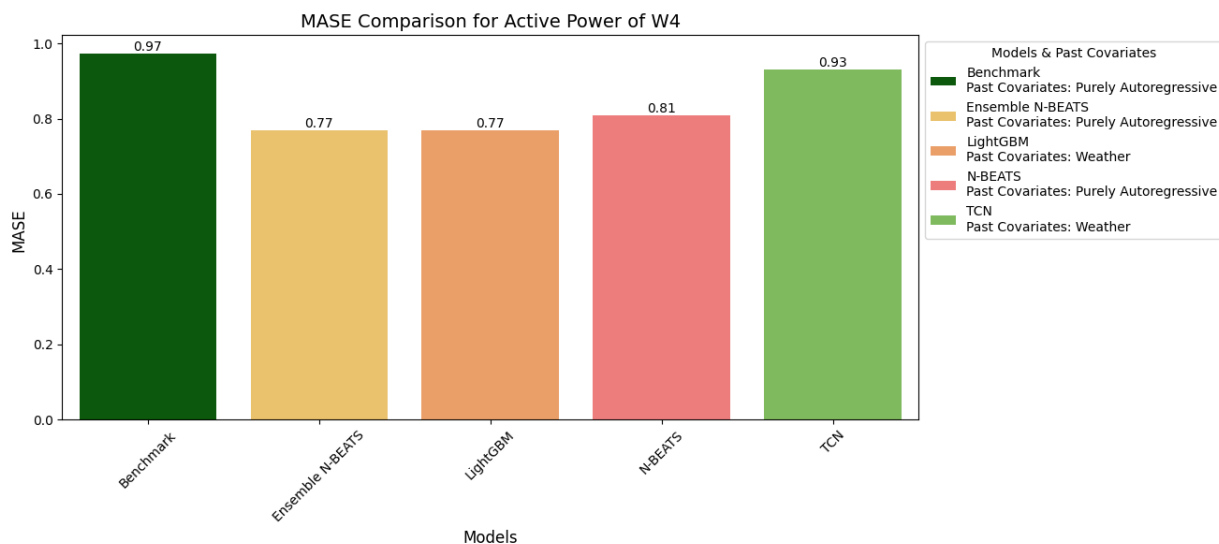
Συγκεντρωτικά, το μοντέλο TCN φαίνεται να αποδίδει ικανοποιητικά σε όλες τις περιπτώσεις, με μοναδική εξαίρεση την άεργο ισχύ του W4 όπου τα αποτελέσματα δείχνουν το μοντέλο να μην λειτουργεί καλά.

## 4.2. Οπτικοποίηση Αποτελεσμάτων – Συγκριτική Ανάλυση

### 4.2.1. Ενεργός Ισχύς W4

Στην ενότητα αυτή θα παρουσιαστεί συγκριτική ανάλυση των αποτελεσμάτων του κάθε μοντέλου που χρησιμοποιήθηκε για την πρόβλεψη της ενεργούς ισχύος που μετριέται από τον μετρητή W4, δηλαδή εκείνης που μετριέται στο ζυγό που είναι συνδεδεμένα το φορτίο τού παρόχου και το φωτοβολταϊκό ισχύος 60kW.

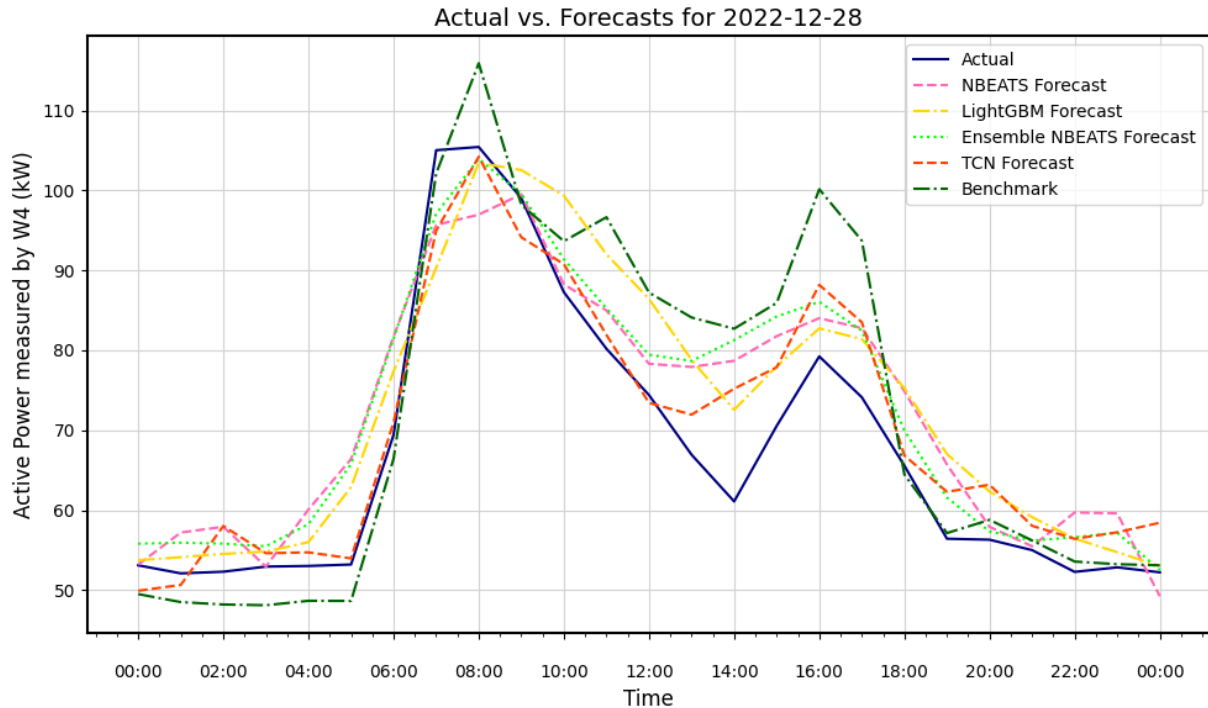
Στην Εικόνα 29 απεικονίζονται οι ελάχιστες τιμές της μετρικής απόδοσης MASE που υπολογίσθηκαν για κάθε μοντέλο καθώς και οι παρελθοντικές συμμεταβλητές που αντιστοιχούν στο μοντέλο με το ελάχιστο MASE.



Εικόνα 29 – Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της ενεργούς ισχύος του W4

Παρατηρείται ότι η απόδοση του μοντέλου TCN δεν είναι τόσο καλή όσο αυτή των υπόλοιπων μοντέλων ενώ καλύτερα αποδίδουν το μοντέλο LightGBM με τα μετεωρολογικά δεδομένα ως συμμεταβλητές και του purely autoregressive Ensemble N-BEATS, με την τιμή του MASE να υπολογίζεται, και στα δύο μοντέλα, σε 0.77.

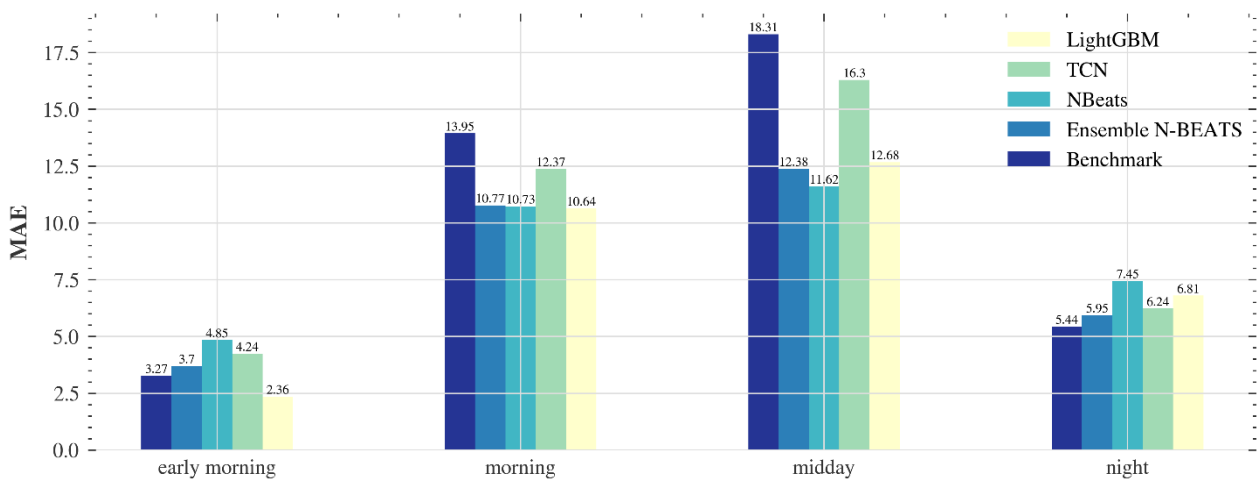
Επιπλέον, στην Εικόνα 30 φαίνεται ένα στιγμιότυπο πρόβλεψης της ενεργούς ισχύος του W4 για μια ημέρα, με τα μοντέλα της Εικόνας 29.



Εικόνα 30 – Στιγμιότυπο πρόβλεψης της ενεργούς ισχύος του W4

Η υστέρηση του μοντέλου TCN παρατηρείται και γραφικά καθώς το μοντέλο είναι ανίκανο να προβλέψει την έντονη πτώση ισχύος κατά τις μεσημεριανές ώρες. Αυτή η αδυναμία ωστόσο παρατηρούμε ότι υπάρχει και στα υπόλοιπα μοντέλα, αν και σε μικρότερο βαθμό.

Για την περαιτέρω διερεύνηση αυτής της αδυναμίας, παρουσιάζεται στην Εικόνα 31 το ραβδόγραμμα των σφαλμάτων των μοντέλων, και συγκεκριμένα της μετρικής απόδοσης MAE, ανάλογα με την ώρα της ημέρας.

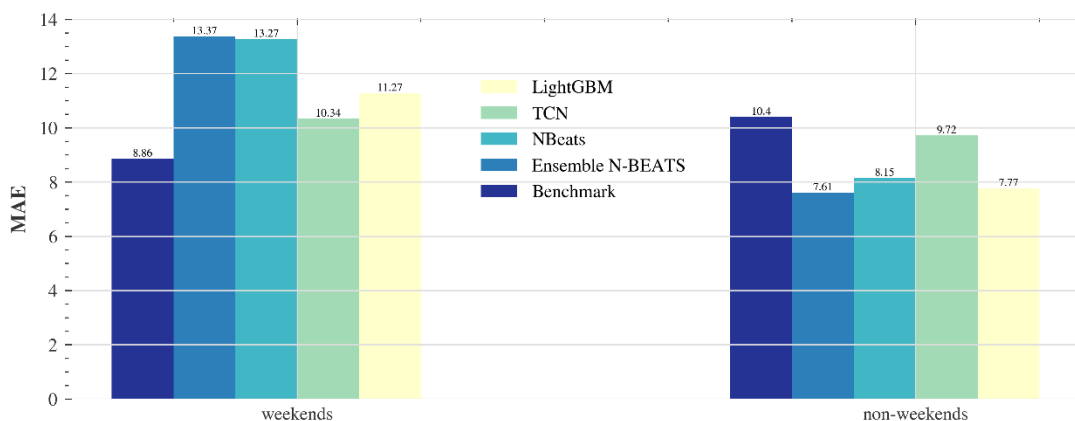


Εικόνα 31 – Σύγκριση του MAE των μοντέλων για την πρόβλεψη της ενεργούς ισχύος του W4 ανά χρονικό διάστημα της ημέρας

Από το παραπάνω ραβδόγραμμα παρατηρείται ότι τα σφάλματα είναι μεγαλύτερα κατά τις πρωινές (06:00 – 12:00) και τις μεσημεριανές – πρώτες απογευματινές (12:00 – 18:00) ώρες. Η αύξηση των σφαλμάτων κατά τις πρωινές ώρες οφείλεται, εν μέρει, στην αύξηση της καταναλισκόμενης ισχύος αφού η ζήτηση ηλεκτρικού φορτίου αυξάνεται σημαντικά εκείνες τις ώρες. Οπότε, με μεγαλύτερες πραγματικές τιμές οδηγούμαστε και σε μεγαλύτερα σφάλματα. Η αύξηση του σφάλματος, ωστόσο, κατά τις μεσημεριανές – πρώτες απογευματινές ώρες, για το μοντέλο TCN είναι αξιοσημείωτη. Επιπλέον παρατηρείται ότι το σφάλμα του μοντέλου ensemble N- BEATS διατηρεί ελαφρώς πιο σταθερές τιμές αυτό το διάστημα σε σχέση με το αντίστοιχο του μοντέλου LightGBM. Τέλος το σφάλμα του Ensemble N – BEATS κατά τη διάρκεια της ημέρας κυμαίνεται σε ένα πιο περιορισμένο διάστημα σε σχέση με το σφάλμα του LightGBM, αφού το πρώτο ανήκει στο διάστημα [3.70 , 10.77] ενώ το δεύτερο στο διάστημα [2.36, 12.68].

Επομένως, τα δύο αποδοτικότερα μοντέλα, εδώ, φαίνεται να είναι το Ensemble N-BEATS και το LightGBM, τα οποία εμφανίζουν ίδια μετρική απόδοσης MASE ενώ το μοντέλο Ensemble N-BEATS έχει ελαφρώς μικρότερο εύρος τιμών κατά τη διάρκεια της ημέρας.

Τέλος, ενδιαφέρον παρουσιάζει το ραβδόγραμμα της Εικόνας 32, το οποίο δείχνει την τιμή του δείκτη MAE ανάλογα με το αν υπολογίζεται το Σαββατοκύριακο ή όχι.

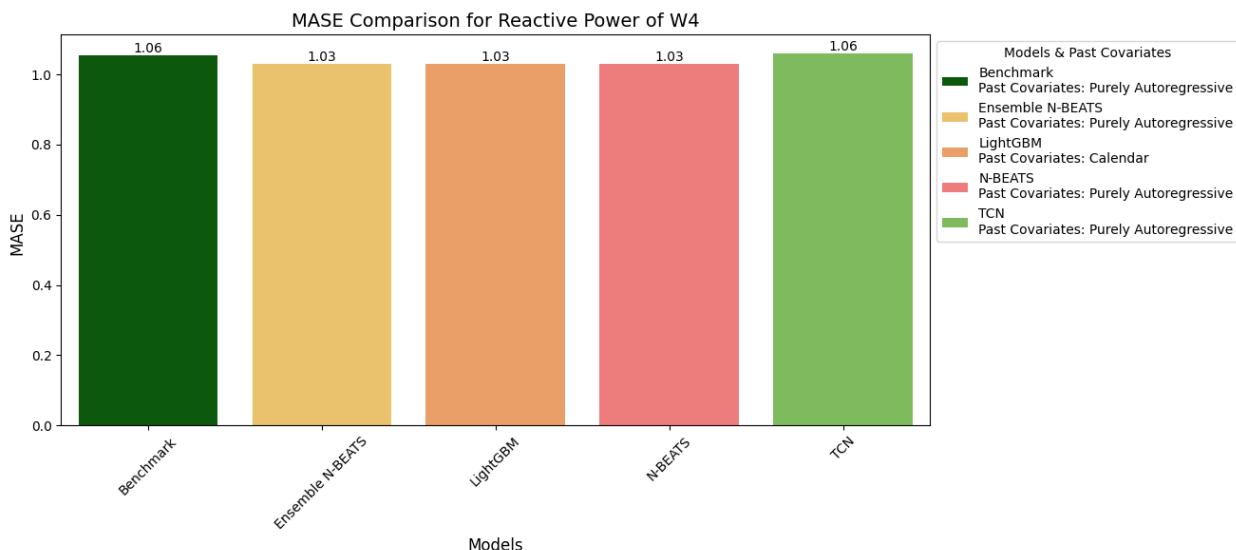


Εικόνα 32 – Σύγκριση του MAE των μοντέλων για την πρόβλεψη της ενεργούς ισχύος του W4 ανα εργάσιμη ή μη ημέρα

Παρατηρείται ότι τα αποτελέσματα των μοντέλων επιδεινώνονται αρκετά κατά το Σαββατοκύριακο, με μοναδική εξαίρεση τα αποτελέσματα του μοντέλου TCN που παραμένουν σχεδόν σταθερά. Πιο αναλυτικά, το σφάλμα του μοντέλου N- BEATS αυξάνεται κατά 62% τα Σαββατοκύριακα ενώ του μοντέλου LightGBM κατά 45%. Το γεγονός αυτό μπορεί να οφείλεται σε αυξημένη κατανάλωση ισχύος κατά τα Σαββατοκύριακα ή ακόμη και σε λάθη που μπορεί να υπάρχουν στις μετρήσεις που συμβαίνουν αυτές τις ημέρες.

#### 4.2.2. Άεργος Ισχύς W4

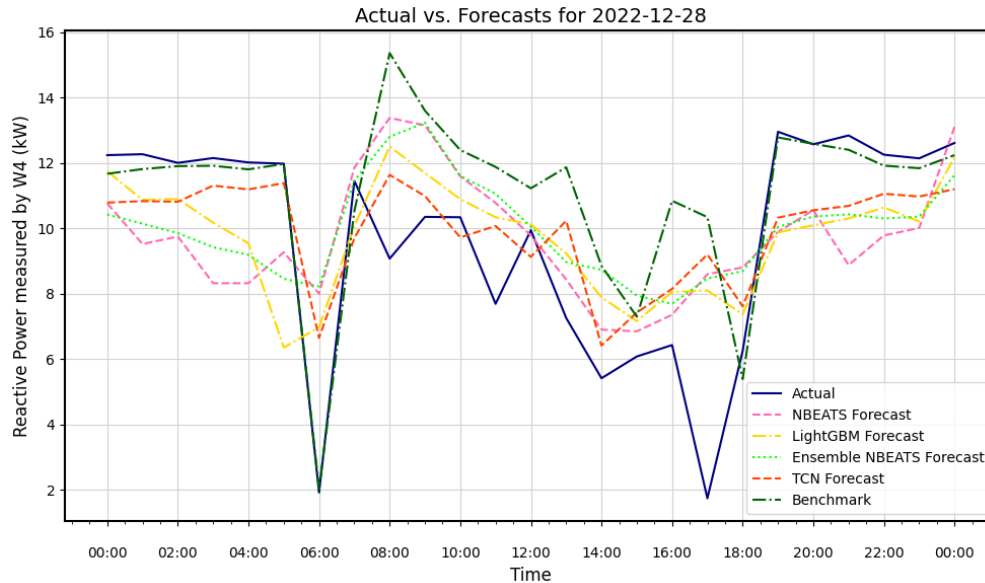
Το ραβδόγραμμα του δείκτη MASE, που υπολογίστηκε για κάθε μοντέλο που προβλέπει την άεργο ισχύ που μετρείται από τον W4 παρουσιάζεται στην Εικόνα 33. Όπως και προηγουμένως, για κάθε μοντέλο έχει επιλεγθεί να απεικονίζεται η περίπτωση των παρελθοντικών συμμεταβλητών για την οποία το MASE που μετρήθηκε ήταν το ελάχιστο.



Εικόνα 33 – Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της άεργου ισχύος του W4

Παρατηρείται ότι η απόδοση των μοντέλων ως προς τον δείκτη MASE είναι ίδια με εξαίρεση την περίπτωση του μοντέλου TCN, το οποίο φαίνεται να αποδίδει ελαφρώς χειρότερα και ίδια με το μοντέλο αναφοράς. Ωστόσο, κρίνοντας από τον δείκτη MASE κανένα από τα μοντέλα δεν έχει ικανοποιητικά αποτελέσματα στην πρόβλεψη καθώς η τιμή του είναι μεγαλύτερη της μονάδας σε όλες τις περιπτώσεις.

Ένα στιγμιότυπο πρόβλεψης μιας ημέρας του Δεκεμβρίου 2022 παρουσιάζεται στην Εικόνα 34.

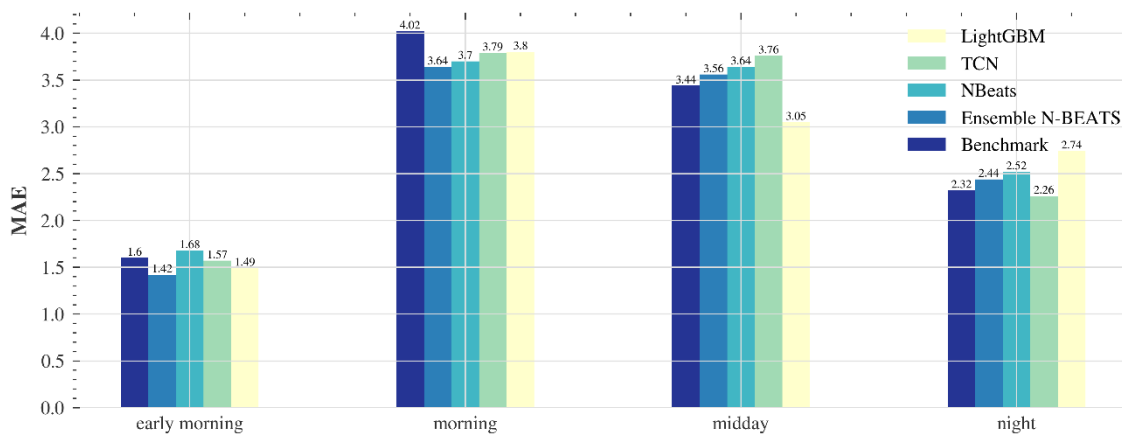


Εικόνα 34 - Στιγμιότυπο πρόβλεψης της αέργου ισχύος του W4

Η μειωμένη προβλεπτική ικανότητα των μοντέλων γίνεται αντιληπτή και από την γραφική απεικόνιση των προβλέψεων τους, καθώς αδυνατούν σε μεγάλο βαθμό να συμβαδίσουν με τα πραγματικά δεδομένα. Αυτό μπορεί να οφείλεται σε πολλούς λόγους συμπεριλαμβανομένου και αυτού της φύσης της χρονοσειράς αφού, όπως φαίνεται, παρουσιάζει πολλές διακυμάνσεις κατά τη διάρκεια της ημέρας. Λόγω των χαρακτηριστικών της αέργου ισχύος, όπως η μεγάλη διασπορά και η μη σαφής κατεύθυνση ροής, είναι δύσκολο να προβλεφθεί η αέργος ισχύς με την ίδια ακρίβεια που προβλέπεται η ενεργός ισχύς. Επιπλέον, στην συγκεκριμένη περίπτωση, η αέργος ισχύς που μετριέται από τον W4 αντιπροσωπεύει την αέργο ισχύ που μετριέται στον ζυγό που είναι συνδεδεμένα τόσο το φορτίο του παρόχου όσο και το φωτοβολταϊκό μέγιστης ισχύος 60 kW γεγονός που επιφέρει περισσότερες εξαρτήσεις που είναι δύσκολο να εντοπιστούν από τα μοντέλα.

Ενδιαφέρον παρουσιάζει και σε αυτή την περίπτωση το ραβδόγραμμα που δείχνει τις τιμές του δείκτη MAE ανάλογα με την ώρα της ημέρας, το οποίο απεικονίζεται στην Εικόνα 35, όπου φαίνεται και πάλι μείωση της απόδοσης των μοντέλων στο διάστημα 06:00 – 18:00. Το μοντέλο LightGBM, ωστόσο, έχει σαφώς βελτιωμένη απόδοση σε σχέση με τα υπόλοιπα μοντέλα στο διάστημα 12:00 – 18:00.





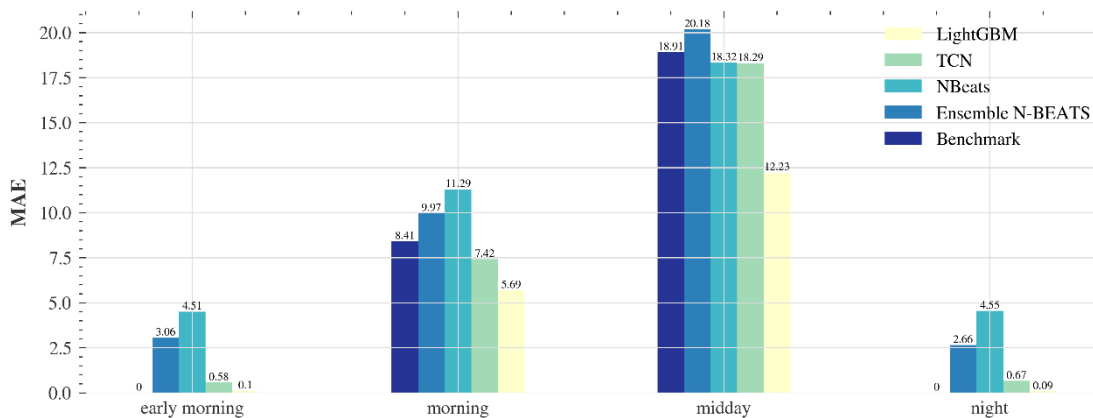
Εικόνα 35 - Σύγκριση του MAE των μοντέλων για την πρόβλεψη της αέργου ισχύος του W4 ανά χρονικό διάστημα της ημέρας

Τέλος της νυχτερινές ώρες (18:00 – 00:00) το μοντέλο TCN φαίνεται να λειτουργεί ελαφρώς καλύτερα ενώ στις υπόλοιπες ώρες της ημέρας το μοντέλο Ensemble N-BEATS με τις παρελθοντικές συμμεταβλητές που αναγράφονται στη Εικόνα 33, έχει την καλύτερη απόδοση με μικρή ωστόσο διαφορά από αυτή των υπολοίπων.

#### 4.2.3. Ενεργός Ισχύς W6

Σε αυτή την υποενότητα θα παρουσιαστούν τα αποτελέσματα των μοντέλων που χρησιμοποιήθηκαν για την πρόβλεψη της ενεργούς ισχύος που μετρείται από τον μετρητή W6, δηλαδή της ενεργούς ισχύος του φωτοβολταϊκού 185kW.

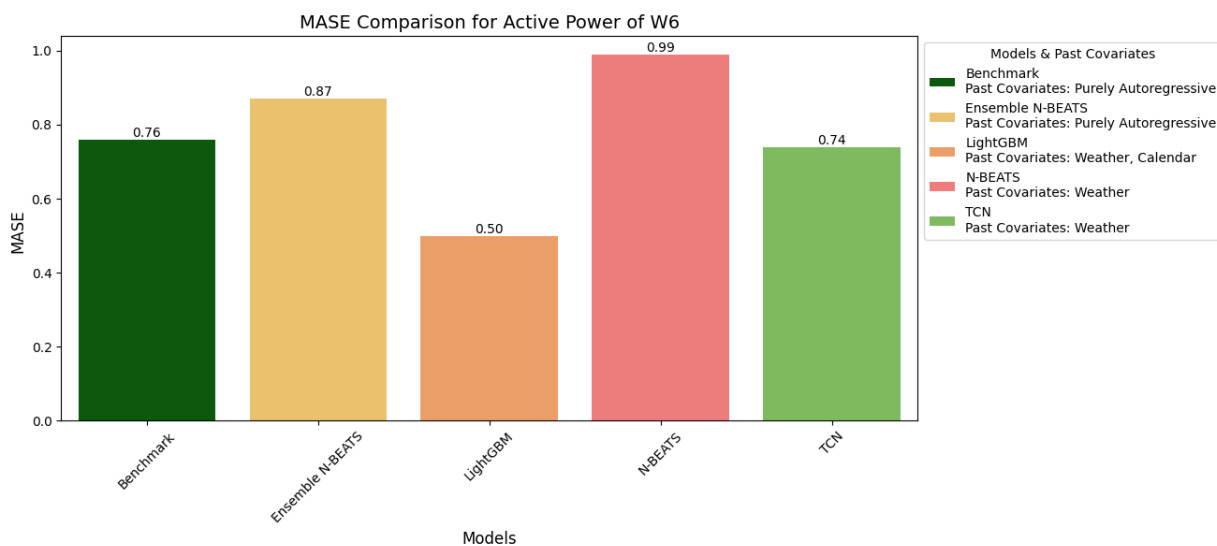
Καθώς η ισχύς του φωτοβολταϊκού εξαρτάται άμεσα από την ώρα της ημέρας, η συγκριτική ανάλυση των αποτελεσμάτων θα ξεκινήσει με την παρουσίαση του ραβδογράμματος που δείχνει την τιμή του MAE ανάλογα με αυτήν (Εικόνα 36).



Εικόνα 36 - Σύγκριση του MAE των μοντέλων για την πρόβλεψη της ενεργούς ισχύος του W6 ανά χρονικό διάστημα της ημέρας

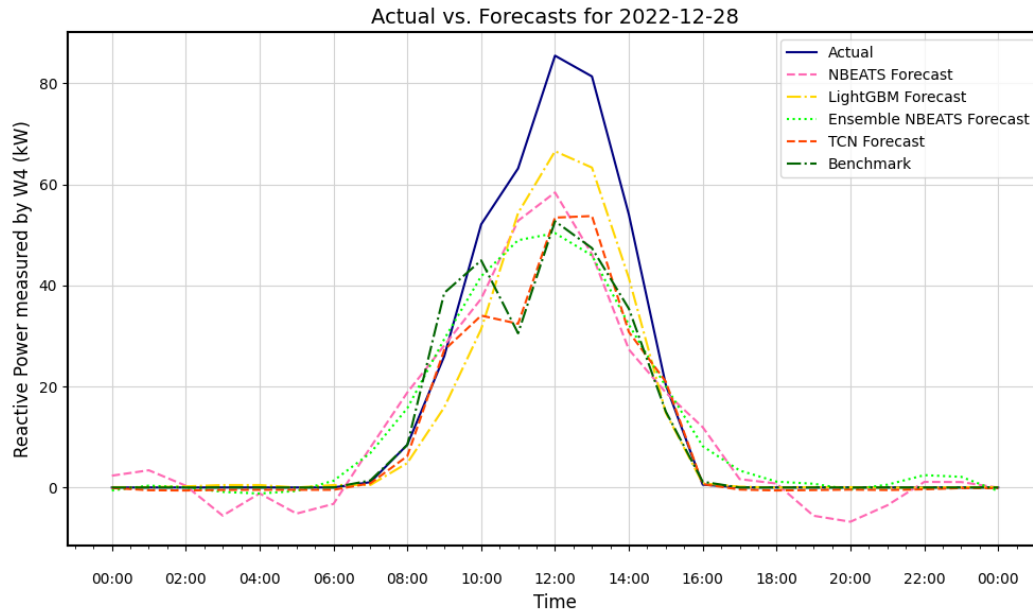
Όπως φαίνεται στην Εικόνα 36, το διάστημα 12:00 – 18:00, δηλαδή το διάστημα κατά το οποίο η ισχύς του φωτοβολταϊκού έχει φτάσει τη μέγιστη τιμή της και σταδιακά ελαττώνεται, το σφάλμα αντίστοιχα αυξάνεται αρκετά ενώ τις νυχτερινές και πρώτες πρωινές ώρες το σφάλμα του μοντέλου TCN και ειδικά του LightGBM είναι σχεδόν μηδενικό.

Από την Εικόνα 37 παρατηρούμε πως τα μοντέλα LightGBM, TCN έχουν πολύ καλή απόδοση ενώ η απόδοση του μοντέλου N-BEATS βελτιώνεται αισθητά με την τεχνική Ensemble, οδηγώντας σε ικανοποιητικά αποτελέσματα. Πιο συγκεκριμένα, ο δείκτης MASE για το μοντέλο LightGBM που εκπαιδεύτηκε με παρελθοντικές συμμεταβλητές τα μετεωρολογικά και τα ημερολογιακά δεδομένα υπολογίστηκε ως 0.5. Η τιμή αυτή υποδεικνύει ότι το μοντέλο έχει διπλασιάσει την ακρίβεια που θα είχε η πρόβλεψη με χρήση του αφελούς μοντέλου (Naïve). Αξιόλογη επιπλέον είναι και η απόδοση του μοντέλου TCN όταν αυτό εκπαιδεύεται με παρελθοντικές συμμεταβλητές τα καιρικά δεδομένα.



Εικόνα 37 -Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της ενεργού ισχύος του W6

Τέλος στην Εικόνα 38 απεικονίζεται ένα στιγμιότυπο πρόβλεψης με παρουσίαση όλων των μοντέλων που αναφέρονται στο παραπάνω ραβδόγραμμα. Η αυξημένη προβλεπτική ικανότητα των μοντέλων γίνεται αντιληπτή και γραφικά, καθώς οι προβλέψεις πλησιάζουν πολύ τις πραγματικές τιμές, με την πρόβλεψη του LightGBM να τις προσεγγίζει σε αρκετά μεγάλο βαθμό.

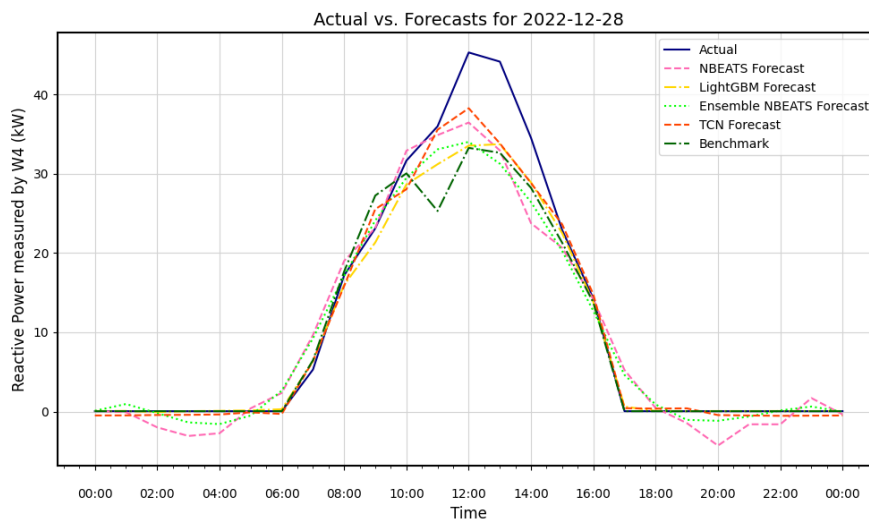


Εικόνα 38 - Στιγμιότυπο πρόβλεψης της ενεργού ισχύος του W6

#### 4.2.4. Άεργος Ισχύς W6

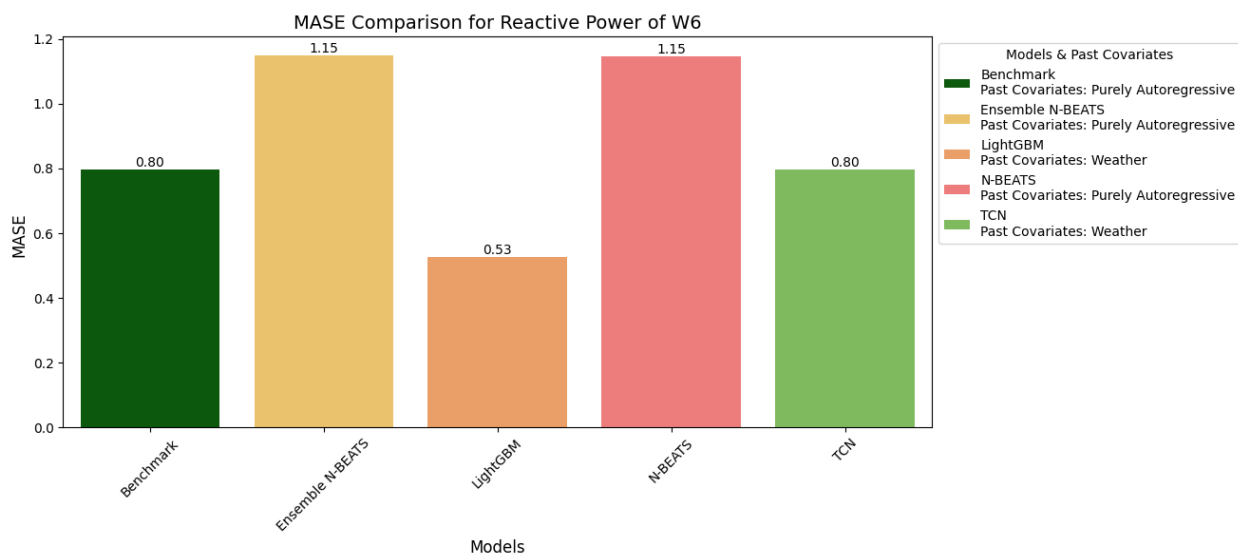
Στο τελευταίο μέρος αυτής της ενότητας, θα αναλύσουμε την απόδοση των μοντέλων που χρησιμοποιήθηκαν για την πρόβλεψη της άεργου ισχύος που μετρείται από τον W6.

Ξεκινώντας με ένα στιγμιότυπο πρόβλεψης, όπως αυτό φαίνεται στην Εικόνα 39, διαπιστώνεται ότι τα μοντέλα έχουν καλή απόδοση, καθώς οι αποκλίσεις των προβλέψεων από τις πραγματικές τιμές των δεδομένων είναι μικρές.



Εικόνα 39 - Στιγμιότυπο πρόβλεψης της ενεργού ισχύος του W6

Πιο αναλυτικά, στην Εικόνα 40 καταγράφονται οι ελάχιστες τιμές του MASE καθώς και τα μοντέλα στον οποίον τις προβλέψεις υπολογίστηκαν.



Εικόνα 40 - Σύγκριση MASE των αποτελεσμάτων των μοντέλων για την πρόβλεψη της αέργου ισχύος του W6

Όπως και στην περίπτωση της ενεργούς ισχύος έτσι κι στην άεργο ισχύ καταγράφεται πολύ καλή απόδοση του μοντέλου LightGBM, όταν αυτό εκπαιδεύεται με τα μετεωρολογικά δεδομένα ως παρελθοντική συμμεταβλητή, στην πρόβλεψη της τιμής της. Το μοντέλο N-BEATS, ωστόσο, σε αυτήν την περίπτωση δεν αποδίδει καλά, ούτε όταν χρησιμοποιείται μόνο του αλλά ούτε με την χρήση της μεθόδου Ensemble.

#### 4.2.5. Βέλτιστα Μοντέλα

Στον Πίνακα 12 που ακολουθεί παρουσιάζονται συνοπτικά τα βέλτιστα μοντέλα που βρέθηκαν για κάθε χρονοσειρά μαζί με δεδομένα για το αν χρησιμοποιήθηκαν παρελθοντικές συμμεταβλητές και κανονικοποιημένα δεδομένα κατά την εκπαίδευση των μοντέλων.

Timeseries	Model	Method	Normalized Data	MASE
Active Power of W4	LightGBM	Weather Covariates	Yes	0.77
Reactive Power of W4	Ensemble N-BEATS	Purely Autoregressive	Yes	1.03
Active Power of W6	LightGBM	Weather & Calendar Covariates	No	0.5
Reactive Power of W6	LightGBM	Weather Covariates	Yes	0.53

Πίνακας 12 - Βέλτιστα μοντέλα για κάθε χρονοσειρά

Παρατηρείται ότι το μοντέλο LightGBM αποδίδει καλύτερα στην πρόβλεψη της ενεργού και αέργου ισχύος που μετριέται από τον W6 καθώς και για την πρόβλεψη της ενεργού ισχύος που μετριέται από τον W4 ενώ σε όλες τις περιπτώσεις που αναφέρθηκαν αναγκαία είναι η χρήση εξωτερικών μεταβλητών. Συγκεκριμένα, στην πρόβλεψη της ενεργού ισχύος του W4 και της αέργου του W6 αρκούν τα μετεωρολογικά δεδομένα ως συμμεταβλητές για την βέλτιστη απόδοση του LightGBM ενώ στην περίπτωση της ενεργού ισχύος του W6 η επιπλέον προσθήκη των χρονικών δεδομένων καλυτερεύουν το μοντέλο. Αυτό δικαιολογείται καθώς η ενεργός ισχύς του φωτοβολταϊκού που είναι συνδεδεμένο στον W6 συσχετίζεται έντονα με την ώρα της ημέρας και επομένως η γνώση των χρονικών μεταβλητών κατά τη διάρκεια της εκπαίδευσης οδηγεί σε καλύτερες προβλέψεις.

Μοναδική περίπτωση που το μοντέλο LightGBM υστερεί είναι η περίπτωση της αέργου ισχύος του W4 στην οποία το ensemble N-BEATS μοντέλο φαίνεται να λειτουργεί καλύτερα, στην purely autoregressive μορφή του.

## 5. Συμπεράσματα

Μέσω των πειραμάτων που πραγματοποιήθηκαν, έγινε προσπάθεια βραχυπρόθεσμης πρόβλεψης και συγκεκριμένα ορίστηκαν ως ορίζοντας πρόβλεψης οι 24 ώρες, της ενεργού και αέργου ισχύος που μετριέται από δύο smart meters τοποθετημένους σ' ένα έξυπνο δίκτυο χαμηλής τάσης στην Ιταλία. Τα μοντέλα που χρησιμοποιήθηκαν ήταν: το μοντέλο μηχανικής μάθησης LighGBM και τα μοντέλα βαθιάς μάθησης TCN και N-BEATS καθώς και ένα Ensemble μοντέλο, συνδυάζοντας δέκα N-BEATS μοντέλα. Τα συμπεράσματα που διεξήχθησαν μέσω της πειραματικής διαδικασίας είναι τα ακόλουθα:

- Η κανονικοποίηση των δεδομένων δεν έχει την ίδια επίδραση στις αποδόσεις των μοντέλων. Συγκεκριμένα, παρατηρήθηκε σημαντική βελτίωση της απόδοσης των μοντέλων βαθιάς μάθησης N-BEATS και TCN όταν αυτά εκπαιδεύονται με κανονικοποιημένα δεδομένα. Αντιθέτως, στο μοντέλο μηχανικής μάθησης LightGBM παρατηρήθηκε ουδέτερη επίδραση καθώς το μοντέλο παρουσίασε παρόμοιες επιδόσεις είτε εκπαιδεύεται με κανονικοποιημένα δεδομένα είτε όχι. Για τα μοντέλα βαθιάς μάθησης υπάρχουν κι άλλες αναφορές στην βιβλιογραφία που καταγράφουν την καλύτερη απόδοσή τους όταν λειτουργούν με κανονικοποιημένα δεδομένα. Το γεγονός αυτό δικαιολογείται από την αρχιτεκτονική των μοντέλων, καθώς μεγάλες τιμές αρχικών δεδομένων, για παράδειγμα, οδηγούν τα μοντέλα να μάθουν μεγάλες τιμές βαρών. Μοντέλα με μεγάλες τιμές βαρών είναι συχνά ασταθή, παρουσιάζουν δηλαδή χαμηλή απόδοση και ευαισθησία στην διαδικασία μάθησης οδηγώντας σε μεγαλύτερα σφάλματα.
- Η προσθήκη των μετεωρολογικών δεδομένων ως παρελθοντικές συμμεταβλητές στην διαδικασία εκπαίδευσης των μοντέλων LightGBM και TCN οδήγησαν τα μοντέλα σε βελτιωμένες αποδόσεις στην πλειονότητα των περιπτώσεων. Πιο συγκεκριμένα, ιδιαίτερα έντονη ήταν η θετική επίδραση των μετεωρολογικών δεδομένων στην πρόβλεψη της ενεργούς ισχύος που μετριέται από τον W6, με μέση τιμή βελτίωσης του δείκτη MASE με μέση τιμή βελτίωσης το 28%. Αντίστοιχη ήταν η θετική επίδραση στην πρόβλεψη των υπόλοιπων ισχύων με εξαίρεση την άεργο ισχύ που μετριέται από τον W4 που η επίδραση των μετεωρολογικών δεδομένων ήταν μικρότερη. Το μοντέλο βαθιάς μάθησης N-BEATS έχει κατά κανόνα, σαφώς καλύτερη απόδοση όταν δεν χρησιμοποιεί τα μετεωρολογικά δεδομένα ως παρελθοντικές συμμεταβλητές. Συγκεκριμένα, η προσθήκη τους στην διαδικασία εκπαίδευσης του μοντέλου προκαλεί επιδείνωση του δείκτη MASE κατά, μέσο όρο, 12%, με εξαίρεση την περίπτωση της πρόβλεψης της ενεργούς ισχύος που μετριέται από τον W6 στην οποία προκαλείται βελτίωση του MASE κατά 4.8%. Η λιγότερο καλή απόδοση του μοντέλου N-BEATS της βιβλιοθήκης darts έχει παρατηρηθεί και σε άλλες εργασίες [1], [67]. Η πολύπλοκη αρχιτεκτονική του μοντέλου σε συνδυασμό με την εφαρμογή του από την συγκεκριμένη βιβλιοθήκη της rython να ευθύνεται γι' αυτό το γεγονός.

- Η προσθήκη των ημερολογιακών δεδομένων ως παρελθοντικές συμμεταβλητές στην διαδικασία εκπαίδευσης του μοντέλου TCN οδήγησε σε βελτίωση της απόδοσής του, μόνο όμως στην περίπτωση της πρόβλεψης της ενεργού ισχύος, με μείωση του δείκτη MASE κατά μέσο όρο 10%. Το μοντέλο LightGBM απέδωσε λιγότερο καλά με την προσθήκη των ημερολογιακών δεδομένων στην πρόβλεψη της ενεργούς ισχύος ενώ στην περίπτωση της αέργου υπήρχε μια βελτίωση της απόδοσης κατά 5%, για την ισχύ που μετριέται από τον W4. Τέλος το μοντέλο N-BEATS δεν επωφελείται καθόλου από τα ημερολογιακά δεδομένα αφού σε όλες τις στην περίπτωση της ενεργούς ισχύος υπάρχει σημαντική αύξηση του δείκτη MASE κατά μέσο όρο 23% και στην περίπτωση της αέργου ισχύος αύξηση του δείκτη MASE κατά μέσο όρο 11%.
- Η τεχνική Naive Ensemble με την με διαφορετική αρχικοποίηση βαρών των επιμέρους μοντέλων και τον συνδυασμό των προβλέψεών τους, οδηγεί το μοντέλο βαθιάς μάθησης N-BEATS, μέσω της εξάλειψης της επίδρασης των επιμέρους βαρών των μοντέλων, σε πιο ακριβείς προβλέψεις. Συγκεκριμένα, το μοντέλο βελτιώθηκε αισθητά όσον αφορά την πρόβλεψη της ενεργούς ισχύος ενώ στην πρόβλεψη της αέργου ισχύος η απόδοση του ήταν σταθερή.

Συνοψίζοντας, για την πρόβλεψη της ισχύος, ενεργούς και αέργου, που μετριέται στον ζυγό που είναι συνδεδεμένα το φορτίο του παρόχου (headquarter load) και το φωτοβολταϊκό μέγιστης ισχύος 60kW, τόσο το μοντέλο LightGBM όσο και το N-BEATS με την τεχνική Naive Ensemble έχουν παρόμοιες επιδόσεις, ωστόσο λόγω του μεγάλου χρόνου εκπαίδευσης που παρουσιάζει το δεύτερο, επιλέγεται το μοντέλο μηχανικής μάθησης LightGBM ως το πιο αποδοτικό. Για την πρόβλεψη της ισχύος, ενεργούς και αέργου, που μετριέται στην έξοδο του φωτοβολταϊκού μέγιστης ισχύος 165kW, το μοντέλο βαθιάς μάθησης TCN καθώς και το μοντέλο μηχανικής μάθησης LightGBM παρουσιάζουν πολύ καλές επιδόσεις, με σαφώς καλύτερη την απόδοση του τελευταίου.

Προτείνεται, λοιπόν για την επίλυση του προβλήματος βραχυπρόθεσμης πρόβλεψης της ισχύος, ενεργούς και αέργου, που μετριέται από έξυπνους μετρητές σε ένα δίκτυο αντίστοιχης τοπολογίας με αυτή της Εικόνας 6, η χρήση του μοντέλου LightGBM καθώς υπερέχει των υπόλοιπων μοντέλων βαθιάς μάθησης που παρουσιάστηκαν. Η υπεροχή του έγκειται τόσο σε επίπεδο επιδόσεων, αφού σε όλες τις περιπτώσεις που εξετάστηκαν τα αποτελέσματα της πρόβλεψης με αυτό το μοντέλο ήταν σαφώς καλύτερα από αυτά των υπολοίπων, όσο και σε επίπεδο υπολογιστικής ισχύος, καθώς ο μικρός χρόνος εκπαίδευσής του αλλά και η δυνατότητα τού μοντέλου να λειτουργεί χωρίς την χρήση GPU αποτελούν σημαντικά πλεονεκτήματα. Είναι αναγκαία ωστόσο η χρήση εξωγενών μεταβλητών και συγκεκριμένα μετεωρολογικών δεδομένων με μετρήσεις της ηλιακής ακτινοβολίας, της θερμοκρασίας και της νεφοκάλυψης, για να επιτευχθούν υψηλές αποδόσεις.

Αξίζει να παρατηρηθεί, ότι κανένα μοντέλο δεν είχε ικανοποιητικές αποδόσεις στην πρόβλεψη της αέργου ισχύος που μετρήθηκε από τον W4, γεγονός που μπορεί να οφείλεται στην έλλειψη δεδομένων που είχαμε και αξίζει να ερευνηθεί περαιτέρω στο μέλλον.

Εν κατακλείδι, έγινε μια προσπάθεια εκτενούς ανάλυσης των δεδομένων, εφαρμογής μοντέλων πρόβλεψης και ερμηνείας των αποτελεσμάτων των προβλέψεων, καταλήγοντας σε ένα μοντέλο που μπορεί να λειτουργήσει αποδοτικά σε πολλαπλά σύνολα δεδομένων που αφορούν χρονοσειρές ηλεκτρικής ενέργειας και συλλέγονται από smart meters τοποθετημένα σε έξυπνα δίκτυα.



## 6. Μελλοντική Εργασία – Προεκτάσεις

Υπό το φως των πορισμάτων που εξήχθησαν από την εργασία αυτή, είναι σημαντικό να αναγνωρίσουμε τους περιορισμούς της τρέχουσας έρευνας και να σκιαγραφήσουμε πιθανές οδούς για περαιτέρω εξερεύνηση και βελτίωση.

Συγκεκριμένα, προέκυψαν περιορισμοί ως προς τη διαθεσιμότητα και την ποιότητα των δεδομένων που χρησιμοποιήθηκαν, καθώς μια σημαντική πρόκληση αποτέλεσαν οι κενές τιμές οι οποίες αφορούσαν αρκετούς μήνες εντός του υπό εξέταση χρονικού πλαισίου. Η ύπαρξη αυτών των κενών τιμών στο σύνολο των δεδομένων υπήρξε σημαντικό εμπόδιο για την επίτευξη ολοκληρωμένων και ακριβών αποτελεσμάτων πρόβλεψης. Τα δεδομένα που λείπουν μπορεί να δημιουργήσουν αβεβαιότητα και να εμποδίσουν την ακριβή μοντελοποίηση των προτύπων χρονοσειρών, οδηγώντας δυνητικά σε μη βέλτιστες προβλέψεις. Η αντιμετώπιση αυτού του περιορισμού και η ανάπτυξη ισχυρότερων μεθοδολογιών για τον χειρισμό των δεδομένων που λείπουν θα είναι ζωτικής σημασίας για μελλοντικές ερευνητικές προσπάθειες με στόχο την ενίσχυση της ακρίβειας και της αξιοπιστίας των προβλέψεων των μοντέλων που χρησιμοποιούνται.

Επιπλέον, είναι σημαντικό να σημειωθεί ότι οι διαθέσιμοι υπολογιστικοί πόροι για αυτήν την έρευνα ήταν περιορισμένοι, ιδιαίτερα λόγω της απουσίας δυνατοτήτων GPU. Ως αποτέλεσμα, επιβλήθηκαν ορισμένοι περιορισμοί καθώς απαιτήθηκε η μείωση του χώρου αναζήτησης παραμέτρων για το μοντέλο N-BEATS κατά την διαδικασία της αναζήτησης πλέγματος (grid search). Παρόλα αυτά, έγιναν προσπάθειες να επιλεγεί ένα εύλογο φάσμα παραμέτρων για εξερεύνηση, επιτρέποντας ολοκληρωμένη αξιολόγηση της απόδοσης του μοντέλου εντός των δεδομένων περιορισμών. Δεδομένης της σημασίας της βελτιστοποίησης των παραμέτρων των μοντέλων, είναι σημαντικό σε μελλοντική εργασία να αντιμετωπιστούν αυτοί οι υπολογιστικοί περιορισμοί και, αξιοποιώντας τις δυνατότητες της GPU, να διεξαχθεί μια πιο ολοκληρωμένη εξερεύνηση των υπερπαραμέτρων του μοντέλου N-BEATS που χρησιμοποιήθηκε.

Ακόμα μια προέκταση της εργασίας είναι αυτή της εξερεύνησης της επίδρασης περισσότερων παρελθοντικών συμμεταβλητών. Συγκεκριμένα, η συμπερίληψη των δεδομένων ρεύματος και τάσης που ήδη έχουμε ως συμμεταβλητών στα μοντέλα πρόβλεψης μπορεί να προσφέρει πολλά πιθανά οφέλη, καθώς επιτρέπει μια πιο ολοκληρωμένη κατανόηση της υποκείμενης δυναμικής του συστήματος. Οι μετρήσεις του ρεύματος και της τάσης παρέχουν πληροφορίες για την ηλεκτρική συμπεριφορά του δικτύου, καταγράφοντας διακυμάνσεις και ανωμαλίες που μπορεί να επηρεάσουν την έξοδο ισχύος. Ενσωματώνοντας αυτές τις μεταβλητές ως παρελθοντικές συμμεταβλητές, τα μοντέλα πρόβλεψης μπορούν να αξιοποιήσουν αυτές τις πρόσθετες πληροφορίες και να κάνουν πιο ακριβείς προβλέψεις. Προτείνεται, λοιπόν σε μελλοντική έρευνα, η διεξαγωγή πειραμάτων για την αξιολόγηση της επίδρασης της προσθήκης των συμμεταβλητών ρεύματος και τάσης στην ακρίβεια πρόβλεψης ισχύος.

Τέλος, σε αυτή την εργασία το πρόβλημα αντιμετωπίστηκε χρησιμοποιώντας univariate timeseries, αφού για κάθε μετρητή είχαμε ξεχωριστές χρονοσειρές για την ενεργό και άεργο ισχύ του και εκπαιδεύτηκαν ξεχωριστά μοντέλα πάνω σε αυτές τις χρονοσειρές. Σε μελλοντική εργασία προτείνεται ωστόσο να εξετασθεί η ακρίβεια των μοντέλων πρόβλεψης χρησιμοποιώντας multivariate timeseries. Δηλαδή, θα μπορούσαμε για κάθε μετρητή να εκπαιδεύσουμε ένα μοντέλο, πάνω σε μια χρονοσειρά με δύο εξαρτημένες μεταβλητές, την ενεργό και την άεργο ισχύ του κάθε μετρητή. Διαφορετικά, θα μπορούσε να εκπαιδευτεί μόνο ένα μοναδικό μοντέλο, πάνω σε μια multivariate χρονοσειρά που θα περιέχει τα δεδομένα της ισχύος κάθε μετρητή W2, W4, δηλαδή 4 μεταβλητές.

Συμπερασματικά, υπάρχουν αρκετές μελλοντικές επεκτάσεις της παρούσας εργασίας οι οποίες αξίζει να ερευνηθούν προκειμένου να ενισχυθεί η αποδοτικότητα των μοντέλων πρόβλεψης που χρησιμοποιήθηκαν για την αποτελεσματικότερη πρόβλεψη πολλαπλών δεδομένων που αφορούν χρονοσειρές ηλεκτρικής ενέργειας.

## 7. Κατάλογος Βιβλιογραφίας

- [1] S. Pelekis *et al.*, “In Search of Deep Learning Architectures for Load Forecasting: A Comparative Analysis and the Impact of the Covid-19 Pandemic on Model Performance,” 2022.
- [2] B. M. Hodge, A. Zeiler, D. Brooks, G. Blau, J. Pekny, and G. Reklatis, “Improved Wind Power Forecasting with ARIMA Models,” *Computer Aided Chemical Engineering*, vol. 29, pp. 1789–1793, Jan. 2011, doi: 10.1016/B978-0-444-54298-4.50136-7.
- [3] W. Li and Z. G. Zhang, “Based on time sequence of ARIMA model in the application of short-term electricity load forecasting,” *ICRCCS 2009 - 2009 International Conference on Research Challenges in Computer Science*, pp. 11–14, 2009, doi: 10.1109/ICRCCS.2009.12.
- [4] Y. Li, Y. Su, and L. Shu, “An ARMAX model for forecasting the power output of a grid connected photovoltaic system,” *Renew Energy*, vol. 66, pp. 78–89, Jun. 2014, doi: 10.1016/J.RENENE.2013.11.067.
- [5] H. T. Yang and Chao-Ming, “Identification of ARMAX model for short term load forecasting: An evolutionary programming approach,” *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 403–408, 1996, doi: 10.1109/59.486125.
- [6] H. Cui and X. Peng, “Short-Term City Electric Load Forecasting with Considering Temperature Effects: An Improved ARIMAX Model,” *Math Probl Eng*, vol. 2015, 2015, doi: 10.1155/2015/589374.
- [7] P. De Falco, L. P. Di Noia, and R. Rizzo, “Exponential Smoothing Model for Photovoltaic Power Forecasting,” *Proceedings of 2021 9th International Conference on Modern Power Systems, MPS 2021*, Jun. 2021, doi: 10.1109/MPS52805.2021.9492538.
- [8] W. R. Christiaanse, “Short-term load forecasting using general exponential smoothing,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 2, pp. 900–911, Mar. 1971, doi: 10.1109/TPAS.1971.293123.
- [9] J. X. Che and J. Z. Wang, “Short-term load forecasting using a kernel-based support vector regression combination model,” *Appl Energy*, vol. 132, pp. 602–609, Nov. 2014, doi: 10.1016/j.apenergy.2014.07.064.
- [10] J. Zeng and W. Qiao, “Support vector machine-based short-term wind power forecasting,” *2011 IEEE/PES Power Systems Conference and Exposition, PSCE 2011*, 2011, doi: 10.1109/PSCE.2011.5772573.
- [11] J. Kou *et al.*, “Photovoltaic power forecasting based on artificial neural network and meteorological data,” *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, 2013, doi: 10.1109/TENCON.2013.6718512.

- [12] A. R. Pazikadin, D. Rifai, K. Ali, M. Z. Malik, A. N. Abdalla, and M. A. Faraj, "Solar irradiance measurement instrumentation and power solar generation forecasting based on Artificial Neural Networks (ANN): A review of five years research trend," *Science of The Total Environment*, vol. 715, p. 136848, May 2020, doi: 10.1016/J.SCITOTENV.2020.136848.
- [13] D. Baczyński and M. Parol, "Influence of artificial neural network structure on quality of short-term electric energy consumption forecast," *IEE Proceedings: Generation, Transmission and Distribution*, vol. 151, no. 2, pp. 241–245, Mar. 2004, doi: 10.1049/IP-GTD:20040070.
- [14] D. Niu, K. Wang, L. Sun, J. Wu, and X. Xu, "Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study," *Appl Soft Comput*, vol. 93, p. 106389, Aug. 2020, doi: 10.1016/J.ASOC.2020.106389.
- [15] S. Ben Taieb and R. J. Hyndman, "A gradient boosting approach to the Kaggle load forecasting competition".
- [16] T. Limouni, R. Yaagoubi, K. Bouziane, K. Guissi, and E. H. Baali, "Accurate one step and multistep forecasting of very short-term PV power using LSTM-TCN model," *Renew Energy*, vol. 205, pp. 1010–1024, Mar. 2023, doi: 10.1016/J.RENENE.2023.01.118.
- [17] W. Sheng, K. Liu, D. Jia, S. Chen, and R. Lin, "Short-Term Load Forecasting Algorithm Based on LST-TCN in Power Distribution Network," *Energies 2022, Vol. 15, Page 5584*, vol. 15, no. 15, p. 5584, Aug. 2022, doi: 10.3390/EN15155584.
- [18] S. H. Rafi, N. Al-Masood, S. R. Deeba, and E. Hossain, "A short-term load forecasting method using integrated CNN and LSTM network," *IEEE Access*, vol. 9, pp. 32436–32448, 2021, doi: 10.1109/ACCESS.2021.3060654.
- [19] L. G. Chen, H. D. Chiang, N. Dong, and R. P. Liu, "Group-based chaos genetic algorithm and non-linear ensemble of neural networks for short-term load forecasting," *IET Generation, Transmission & Distribution*, vol. 10, no. 6, pp. 1440–1447, Apr. 2016, doi: 10.1049/IET-GTD.2015.1068.
- [20] S. Mishra and S. K. Patra, "Short Term Load Forecasting Using Neural Network Trained with Genetic Algorithm & Particle Swarm Optimization," *2008 First International Conference on Emerging Trends in Engineering and Technology*, pp. 606–611, 2008, doi: 10.1109/ICETET.2008.94.
- [21] K. Βourνας and Γ. Κονταξής, *Εισαγωγή στα Συστήματα Ηλεκτρικής Ενέργειας*. 2001.
- [22] R. Peters, "Basics on Base Load: Meeting Ontario's Base Load Electricity Demand with Renewable Power Sources." 2007.
- [23] European SmartGrids Technology Platform, "Strategic Deployment Document for Europe's Electricity Networks of the Future," 2010.

- [24] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid - The new and improved power grid: A survey," *IEEE Communications Surveys and Tutorials*, vol. 14, no. 4, pp. 944–980, 2012, doi: 10.1109/SURV.2011.101911.00087.
- [25] M. Çınar and A. Kaygusuz, "Self-Healing In Smart Grid: A Review," *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, vol. 7, pp. 492–503, May 2018, doi: 10.17798/bitlisfen.460164.
- [26] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, Jan. 2010, doi: 10.1109/MPE.2009.934876.
- [27] S. Pelekis *et al.*, "TARGETED DEMAND RESPONSE FOR FLEXIBLE ENERGY COMMUNITIES USING CLUSTERING TECHNIQUES," 2023.
- [28] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid - The new and improved power grid: A survey," *IEEE Communications Surveys and Tutorials*, vol. 14, no. 4, pp. 944–980, 2012, doi: 10.1109/SURV.2011.101911.00087.
- [29] National Energy Technology Laboratory, "The NETL Modern Grid Strategy Powering our 21st-Century Economy: Advanced Metering Infrastructure ," *U.S. Department of Energy, National Energy Technology Laboratory*. 2008.
- [30] R. Rashed Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on Advanced Metering Infrastructure," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 473–484, Dec. 2014, doi: 10.1016/J.IJEPES.2014.06.025.
- [31] S. Gheorghe, N. Golovanov, G. C. Lazaroiu, and R. Porumb, "Smart Grid, Integration of Renewable Sources and Improvement of Power Quality," *Proceedings - 2017 21st International Conference on Control Systems and Computer, CSCS 2017*, pp. 641–645, Jul. 2017, doi: 10.1109/CSCS.2017.98.
- [32] Φ. Ασημακόπουλος Β. Πετρόπουλος, "Επιχειρησιακές προβλέψεις," *Εκδόσεις Συμμετρία*.
- [33] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. Australia: OTexts, 2018.
- [34] V. Prema and K. U. Rao, "Time series decomposition model for accurate wind speed forecast," *Renewables: Wind, Water, and Solar 2015 2:1*, vol. 2, no. 1, pp. 1–11, Nov. 2015, doi: 10.1186/S40807-015-0018-9.
- [35] S. So, "Why is the sample variance a biased estimator?," 2008.
- [36] J. Panaretos and E. Xekalaki, *Introduction to Statistical Thinking, vol. II (Estimation and Testing of Hypothesis)*. 2000.
- [37] C. Kosma, G. Nikolentzos, N. Xu, and M. Vazirgiannis, "Time Series Forecasting Models Copy the Past: How to Mitigate," 2022.

- [38] S. Makridakis, "Accuracy measures: theoretical and practical concerns," *Int J Forecast*, vol. 9, no. 4, pp. 527–529, 1993, doi: 10.1016/0169-2070(93)90079-3.
- [39] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int J Forecast*, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: 10.1016/J.IJFORECAST.2006.03.001.
- [40] M. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. Tyukov, T. A. Janovsky, and V. A. Kamaev, "A survey of forecast error measures," *World Appl Sci J*, vol. 24, pp. 171–176, Jun. 2013, doi: 10.5829/idosi.wasj.2013.24.itmies.80032.
- [41] G. Wiederhold, J. McCarthy, and E. Feigenbaum, "Arthur Samuel: pioneer in machine learning," *Commun ACM*, vol. 33, no. 11, pp. 137–139, Nov. 1990, Accessed: May 22, 2023. [Online]. Available: <https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=00010782&v=2.1&it=r&id=GALE%7CA9597979&sid=googleScholar&linkaccess=fulltext>
- [42] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," *IBM J Res Dev*, vol. 3, no. 3, pp. 210–229, Jul. 1959, doi: 10.1147/RD.33.0210.
- [43] T. M. Mitchell, *Machine learning.*, vol. 1. New York: McGraw-hill, 2007.
- [44] F. Chollet, *Deep Learning with Python*. Manning, 2017.
- [45] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature* 2017 542:7639, vol. 542, no. 7639, pp. 115–118, Jan. 2017, doi: 10.1038/nature21056.
- [46] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering." pp. 815–823, 2015.
- [47] "Handbook of Natural Language Processing," *Handbook of Natural Language Processing*, Feb. 2010, doi: 10.1201/9781420085938.
- [48] E. Stellwagen and L. Tashman, "ARIMA: The Models of Box and Jenkins," *Foresight: Int. J. Appl. Forecast.*, pp. 28–33, Jul. 2013.
- [49] E. S. Gardner, "Exponential smoothing: The state of the art," *J Forecast*, vol. 4, no. 1, pp. 1–28, Jan. 1985, doi: 10.1002/FOR.3980040103.
- [50] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: Results, findings, conclusion and way forward," *Int J Forecast*, vol. 34, no. 4, pp. 802–808, Oct. 2018, doi: 10.1016/J.IJFORECAST.2018.06.001.
- [51] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M5 competition: Background, organization, and implementation," *Int J Forecast*, vol. 38, no. 4, pp. 1325–1336, Oct. 2022, doi: 10.1016/J.IJFORECAST.2021.07.007.
- [52] T. O. Ayodele, "Types of machine learning algorithms," in *New advances in machine learning* 3, 2010, pp. 19–48.

- [53] A. Jet and H. J. O, "Supervised Machine Learning Algorithms: Classification and Comparison," *International Journal of Computer Trends and Technology*, vol. 48, 2017, doi: 10.14445/22312803/IJCTT-V48P126.
- [54] J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Source: The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001, Accessed: Jun. 22, 2023. [Online]. Available: <http://www.jstor.org>URL:<http://www.jstor.org/stable/2699986>
- [55] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree", Accessed: Jun. 22, 2023. [Online]. Available: <https://github.com/Microsoft/LightGBM>.
- [56] B. N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio Mila, "N-BEATS: NEURAL BASIS EXPANSION ANALYSIS FOR INTERPRETABLE TIME SERIES FORECASTING".
- [57] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," 2018.
- [58] Y. Lin, I. Koprinska, and M. Rana, "Temporal Convolutional Attention Neural Networks for Time Series Forecasting," *Proceedings of the International Joint Conference on Neural Networks*, vol. 2021-July, Jul. 2021, doi: 10.1109/IJCNN52387.2021.9534351.
- [59] A. Herle, J. Channegowda, and D. Prabhu, "A Temporal Convolution Network Approach to State-of-Charge Estimation in Li-ion Batteries," *2020 IEEE 17th India Council International Conference, INDICON 2020*, Dec. 2020, doi: 10.1109/INDICON49873.2020.9342315.
- [60] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 8, no. 4, p. e1249, Jul. 2018, doi: 10.1002/WIDM.1249.
- [61] "<https://www.copernicus.eu/en>."
- [62] W. McKinney, "Data Structures for Statistical Computing in Python," 2010.
- [63] E. Bisong, "Matplotlib and Seaborn," *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, pp. 151–165, 2019, doi: 10.1007/978-1-4842-4470-8\_12.
- [64] A. Jain, "DARTS - Multibody Modeling, Simulation and Analysis Software," *Computational Methods in Applied Sciences*, vol. 53, pp. 433–441, 2020, doi: 10.1007/978-3-030-23132-3\_52/COVER.
- [65] J. Peppanen, X. Zhang, S. Grijalva, and M. J. Reno, "Handling bad or missing smart meter data through advanced data imputation," *2016 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2016*, Dec. 2016, doi: 10.1109/ISGT.2016.7781213.
- [66] A. Bielecki *et al.*, "Short-Term Wind Power Prediction Based on LightGBM and Meteorological Reanalysis," *mdpi.com*, 2022, doi: 10.3390/en15176287.

- [67] S. Pelekis *et al.*, “A comparative assessment of deep learning models for day-ahead load forecasting: Investigating key accuracy drivers,” *ArXiv preprint*, Feb. 2023, Accessed: Jun. 17, 2023. [Online]. Available: <https://arxiv.org/abs/2302.12168v1>
- [68] A. K. Shaikh, A. Nazir, I. Khan, and A. S. Shah, “Short term energy consumption forecasting using neural basis expansion analysis for interpretable time series,” *Scientific Reports* 2022 12:1, vol. 12, no. 1, pp. 1–18, Dec. 2022, doi: 10.1038/s41598-022-26499-y.
- [69] M. Decuyper, M. Stockhoff, S. Vandenberghe, al -, and X. Ying, “An Overview of Overfitting and its Solutions You may also like Artificial neural networks for positioning of gamma interactions in monolithic PET detectors Analysis of overfitting in the regularized Cox model An Overview of Overfitting and its Solutions,” p. 22022, 2019, doi: 10.1088/1742-6596/1168/2/022022.
- [70] H. Lu, X. Ma, K. Huang, and M. Azimi, “Prediction of offshore wind farm power using a novel two-stage model combining kernel-based nonlinear extension of the Arps decline model with a multi-objective grey wolf optimizer,” *Renewable and Sustainable Energy Reviews*, vol. 127, p. 109856, 2020, doi: 10.1016/j.rser.2020.109856.
- [71] M. Krechowicz, A. Krechowicz, L. Lichołai, A. Pawelec, J. Z. Piotrowski, and A. St., “Reduction of the Risk of Inaccurate Prediction of Electricity Generation from PV Farms Using Machine Learning,” 2022, doi: 10.3390/en15114006.
- [72] M. Perera, J. De Hoog, K. Bandara, and S. Halgamuge, “Multi-resolution, multi-horizon distributed solar PV power forecasting with forecast combinations,” *Expert Syst Appl*, vol. 205, p. 117690, Nov. 2022, doi: 10.1016/J.ESWA.2022.117690.
- [73] K. Rosen, C. Angeles-Camacho, V. Elvira, and S. Tulio Guillén-Burguete, “Intra-hour photovoltaic forecasting through a time-varying Markov switching model,” 2023, doi: 10.1016/j.energy.2023.127952.
- [74] K. J. Nam, S. Hwangbo, and C. K. Yoo, “A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea,” *Renewable and Sustainable Energy Reviews*, vol. 122, p. 109725, Apr. 2020, doi: 10.1016/J.RSER.2020.109725.
- [75] Y. Chu, B. Urquhart, S. M. I. Gohari, H. T. C. Pedro, J. Kleissl, and C. F. M. Coimbra, “Short-term reforecasting of power output from a 48 MWe solar PV plant,” *Solar Energy*, vol. 112, pp. 68–77, Feb. 2015, doi: 10.1016/J.SOLENER.2014.11.017.
- [76] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, “Solar photovoltaic generation forecasting methods: A review,” *Energy Convers Manag*, vol. 156, pp. 459–497, Jan. 2018, doi: 10.1016/J.ENCONMAN.2017.11.019.
- [77] Y. Huang and Y. Wu, “Short-Term Photovoltaic Power Forecasting Based on a Novel Autoformer Model,” *Symmetry* 2023, Vol. 15, Page 238, vol. 15, no. 1, p. 238, Jan. 2023, doi: 10.3390/SYM15010238.



- [78] A. Dobbs and A. Florita, "Short-term solar forecasting performance of popular machine learning algorithms," 2017.
- [79] N. Sagar, "Powerline communications systems: overview and analysis," 2011, doi: 10.7282/T3RR1XJ7.