



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

ΟΜΑΔΟΠΟΙΗΣΗ ΟΙΚΙΑΚΩΝ ΚΑΤΑΝΑΛΩΤΩΝ ΓΙΑ ΠΑΡΟΧΗ ΕΥΕΛΙΞΙΑΣ ΣΕ ΠΕΡΙΟΔΟΥΣ ΑΙΧΜΩΝ ΖΗΤΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Στυλιανός Σ. Τζαγκαράκης

Επίβλεψη: Παύλος Σ. Γεωργιλάκης, Καθηγητής Ε.Μ.Π.

Επόπτευση: Χριστόφορος Μένος-Αικατερινιάδης, Υποψήφιος Διδάκτορας Ε.Μ.Π.

Αθήνα, Ιούνιος 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

ΟΜΑΔΟΠΟΙΗΣΗ ΟΙΚΙΑΚΩΝ ΚΑΤΑΝΑΛΩΤΩΝ ΓΙΑ ΠΑΡΟΧΗ ΕΥΕΛΙΞΙΑΣ ΣΕ ΠΕΡΙΟΔΟΥΣ ΑΙΧΜΩΝ ΖΗΤΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Στυλιανός Σ. Τζαγκαράκης

Επίβλεψη: Παύλος Σ. Γεωργιλάκης, Καθηγητής Ε.Μ.Π

Επόπτευση: Χριστόφορος Μένος-Αικατερινιάδης, Υποψήφιος Διδάκτορας Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 27^η Ιουνίου 2025

.....
Παύλος Γεωργιλάκης
Καθηγητής Ε.Μ.Π.

.....
Βασίλειος Νικολαΐδης
Επικ. Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Τσεκούρας
Αναπλ. Καθηγητής Πα.Δ.Α.

Αθήνα, Ιούνιος 2025

.....

Στυλιανός Σ. Τζαγκαράκης

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

Copyright © Στυλιανός Τζαγκαράκης, 2025.

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.

ΠΕΡΙΛΗΨΗ

Η παρούσα Διπλωματική Εργασία εξετάζει την ομαδοποίηση οικιακών καταναλωτών με στόχο την παροχή ευελιξίας σε περιόδους αιχμών ζήτησης. Εστιάζει στην ταξινόμηση καταναλωτών σε ομάδες με παρόμοια πρότυπα συμπεριφοράς, προκειμένου να ενισχυθεί η αποτελεσματικότητα των δράσεων Απόκρισης Ζήτησης. Η ανάλυση βασίζεται σε πραγματικά δεδομένα κατανάλωσης μετρήσεων ανά 30 λεπτά για 5.567 νοικοκυριά στο Λονδίνο για το έτος 2013, όπως δημοσιεύονται από τον διαχειριστή UK Power Networks, ενώ η τελική ανάλυση εφαρμόζεται μετά από κατάλληλη προεπεξεργασία. Υπολογίζονται δείκτες συμπεριφοράς που σχετίζονται με τη δραστηριότητα σε ώρες αιχμής και επιλέγονται τα πλέον διακριτικά χαρακτηριστικά. Ακολουθεί εφαρμογή του αλγορίθμου K-Means για την ομαδοποίηση. Ο αριθμός των συστάδων επιλέγεται βάσει συγκριτικής αξιολόγησης μεθόδων, όπως οι μέθοδοι Elbow, Silhouette Score και Calinski–Harabasz, ώστε να εξασφαλιστεί ικανοποιητικός διαχωρισμός των καταναλωτών. Η μοντελοποίηση γίνεται με χρήση της γλώσσας προγραμματισμού Python.

Τα αποτελέσματα της ανάλυσης αναδεικνύουν δύο συστάδες, συνολικού πλήθους 1.600 καταναλωτών, οι οποίες θεωρούνται καταλληλότερες για συμμετοχή σε δράσεις ευελιξίας. Η επιλογή βασίζεται στην υψηλή κατανάλωση και την έντονη δραστηριότητα κατά τις ώρες αιχμής.

Η αξιολόγηση της μεθοδολογίας πραγματοποιείται με βάση τη δυνατότητα αξιοποίησης των προτεινόμενων συστάδων σε σενάρια ευελιξίας, τα οποία βασίζονται σε πραγματικά αιτήματα καθοδικής ευελιξίας για την περιοχή του Λονδίνου, όπως δημοσιεύονται από τον διαχειριστή UK Power Networks. Επιλέγονται δέκα αντιπροσωπευτικά αιτήματα του έτους 2024, τα οποία προσαρμόζονται χρονικά, προκειμένου να διασφαλιστεί η συμβατότητα με το σύνολο κατανάλωσης του 2013. Εξετάζονται δύο διαφορετικές στρατηγικές επιλογής καταναλωτών για την κάλυψη των αιτημάτων: μία γενική, βασισμένη μόνο στην ιστορική στιγμιαία κατανάλωση κατά την ώρα και μέρα του αιτήματος ευελιξίας, και μία στοχευμένη, βασισμένη στα αποτελέσματα της ομαδοποίησης για τις δύο επικρατούσες συστάδες. Μέσω της σύγκρισης των δύο προσεγγίσεων, αναδεικνύεται η λειτουργική αξία της προτεινόμενης μεθοδολογίας ως εργαλείου υποστήριξης για Φορείς Σωρευτικής Εκπροσώπησης (ΦοΣΕ) Φορτίου στο πλαίσιο προγραμμάτων Απόκρισης Ζήτησης.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Ομαδοποίηση οικιακών καταναλωτών, ώρες αιχμής, παροχή ευελιξίας, έξυπνοι μετρητές, Λονδίνο, UK Power Networks, ΦοΣΕ Φορτίου, αλγόριθμος K-Means, προφίλ κατανάλωσης, προγράμματα Απόκρισης Ζήτησης, αιτήματα ευελιξίας

ABSTRACT

The purpose of this Diploma Thesis is to examine the clustering of residential electricity consumers with the aim of providing flexibility during periods of peak demand. It focuses on classifying consumers into groups with similar behavioral patterns to improve the effectiveness of Demand Response actions. The analysis is based on real smart meter data with 30-minute resolution for 5,567 households in London for the year 2013, as published by the distribution network operator UK Power Networks, while the final analysis is conducted following appropriate data preprocessing. Behavioral indicators related to peak activity are calculated and the most discriminative features are selected. The K-Means algorithm is then applied for clustering. The number of clusters is determined through comparative evaluation of methods such as Elbow, Silhouette Score and Calinski–Harabasz, ensuring a satisfactory separation of consumers. The modeling is implemented using the Python programming language.

The results of the analysis identify two clusters, comprising a total of 1,600 consumers, which are considered the most suitable for participation in flexibility programs. Selection is based on high consumption and frequent activity during peak hours.

The methodology is evaluated based on the potential use of the identified clusters in flexibility scenarios, which rely on actual downward flexibility requests in the London area, as published by UK Power Networks. Ten representative requests from 2024 are selected and temporally adapted to ensure compatibility with the 2013 consumption dataset. Two consumer selection strategies are examined: a general one, based solely on historical instantaneous consumption at the time and day of the flexibility request, and a targeted one, based on the clustering results for the two most prominent groups. By comparing these approaches, the proposed methodology demonstrates its practical value as a supporting tool for Load Aggregators in the context of Demand Response programs.

KEY WORDS

Residential consumer clustering, peak hours, flexibility provision, smart meters, London, UK Power Networks, Load Aggregators, K-Means algorithm, consumption profiles, Demand Response programs, flexibility requests

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2024–2025 υπό την επίβλεψη του κ. Παύλου Γεωργιάκη, καθηγητή της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π. στον οποίο οφείλω ιδιαίτερες ευχαριστίες για την ανάθεσή της, δίνοντάς μου την ευκαιρία να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα. Επίσης, θα ήθελα να ευχαριστήσω θερμά τον υποψήφιο διδάκτορα Χριστόφορο Μένο-Αικατερινιάδη για την πολύτιμη βοήθεια και καθοδήγηση που μου παρείχε σε όλη τη διάρκεια εκπόνησης της εργασίας, καθώς και για τον πολύτιμο χρόνο που μου αφιέρωσε.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου, τη μητέρα μου Δέσποινα, τον πατέρα μου Στυλιανό και τον αδερφό μου Δημήτριο, οι οποίοι αποτελούν πηγή δύναμης σε όλη τη διάρκεια της ζωής μου, αλλά και όλους τους φίλους και συναδέλφους που με στήριξαν σε όλη τη διάρκεια των σπουδών μου.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ	1
1.1 Περιοχή έρευνας – Αντικείμενο της εργασίας	1
1.2 Συνοπτική περιγραφή της μεθοδολογίας	2
1.3 Δομή της εργασίας	2
ΚΕΦΑΛΑΙΟ 2: ΕΥΕΛΙΞΙΑ ΟΙΚΙΑΚΩΝ ΚΑΤΑΝΑΛΩΤΩΝ	5
2.1 Ενεργειακές προκλήσεις και ανάγκη για ευελιξία	5
2.1.1 Μεταβλητότητα παραγωγής και ζήτησης	5
2.1.2 Η συμφόρηση στα ηλεκτρικά δίκτυα και ο ρόλος των διαχειριστών ΣΗΕ	6
2.2 Παροχή ευελιξίας από την πλευρά της ζήτησης	7
2.2.1 Έννοια και σημασία της παροχής ευελιξίας	7
2.2.2 Ρόλος των ΦοΣΕ	8
2.3 Οφέλη των προγραμμάτων διαχείρισης ζήτησης	8
2.4 Προγράμματα Απόκρισης Ζήτησης και στοχευμένη επιλογή καταναλωτών	9
ΚΕΦΑΛΑΙΟ 3: ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΚΑΤΑΝΑΛΩΣΗΣ ΑΠΟ ΕΞΥΠΝΟΥΣ ΜΕΤΡΗΤΕΣ	11
3.1 Γενική μεθοδολογία	11
3.2 Δεδομένα εισόδου	12
3.3 Μέθοδος ανάλυσης	14
3.4 Επεξεργασία δεδομένων με κώδικα Python	15
3.4.1 Ανάγνωση δεδομένων	15
3.4.2 Αναδιαμόρφωση δεδομένων	16
3.4.3 Διαχείριση ελλিপών, αρνητικών και ακραίων τιμών	16
3.4.4 Φιλτράρισμα βάσει πληρότητας	17
3.4.5 Εξαγωγή τελικής μορφής συνόλου δεδομένων	19
3.5 Επικαιρότητα των δεδομένων σε σύγκριση με το σήμερα	20
ΚΕΦΑΛΑΙΟ 4: ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΟΜΑΔΟΠΟΙΗΣΗΣ ΟΙΚΙΑΚΩΝ ΚΑΤΑΝΑΛΩΤΩΝ	23
4.1 Εξαγωγή χαρακτηριστικών εισόδου	23
4.1.1 Συμπεριφορά οικιακών καταναλωτών	23
4.1.2 Χαρακτηριστικά εισόδου	26
4.2 Κανονικοποίηση χαρακτηριστικών εισόδου	28

4.3	Περιγραφή αλγορίθμου K-Means	29
4.4	Προσδιορισμός του βέλτιστου αριθμού συστάδων	31
4.4.1	Μέθοδος Elbow	31
4.4.2	Μέθοδος Silhouette	32
4.4.3	Μέθοδος Calinski-Harabasz	33
ΚΕΦΑΛΑΙΟ 5: ΑΠΟΤΕΛΕΣΜΑΤΑ ΟΜΑΔΟΠΟΙΗΣΗΣ ΚΑΙ ΠΡΑΚΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ		35
5.1	Αξιολόγηση κριτηρίων εισόδου και επιλογή βέλτιστου συνδυασμού	35
5.2	Τελική επιλογή του βέλτιστου αριθμού συστάδων	39
5.3	Πρακτικές Εφαρμογές	51
ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ - ΕΠΕΚΤΑΣΕΙΣ		59
6.1	Σύνοψη της εργασίας και των αποτελεσμάτων της	59
6.2	Πιθανές επεκτάσεις της εργασίας	60
ΚΕΦΑΛΑΙΟ 7: ΒΙΒΛΙΟΓΡΑΦΙΑ		63

ΕΙΣΑΓΩΓΗ

1.1 ΠΕΡΙΟΧΗ ΕΡΕΥΝΑΣ – ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΕΡΓΑΣΙΑΣ

Η περιοχή έρευνας αυτής της διπλωματικής εργασίας επικεντρώνεται στη διαχείριση της ενέργειας και στην παροχή ευελιξίας από οικιακούς καταναλωτές. Η παροχή της ευελιξίας συνδέεται άμεσα με τις εξελίξεις που λαμβάνουν χώρα στο πλαίσιο της ενεργειακής μετάβασης, η οποία επιβάλλει νέες απαιτήσεις στο σύστημα ηλεκτρικής ενέργειας. Η ενεργειακή μετάβαση συνοδεύεται από σημαντικές αλλαγές στο ενεργειακό σύστημα, όπως η αύξηση της διείσδυσης των Ανανεώσιμων Πηγών Ενέργειας (ΑΠΕ) και η ταυτόχρονη μείωση της εξάρτησης από συμβατικές και ελεγχόμενες μονάδες παραγωγής. Παράλληλα, παρατηρείται διαφοροποίηση στις συνήθειες κατανάλωσης, οδηγώντας στην εμφάνιση αιχμών ζήτησης σε ενδοημερήσιο ορίζοντα. Η προσαρμογή της ζήτησης για τη μείωση αυτών των αιχμών κρίνεται αναγκαία, στο πλαίσιο της διασφάλισης της ισορροπίας του συστήματος.

Η ευελιξία ως προς τη ζήτηση αναφέρεται στη δυνατότητα των καταναλωτών να προσαρμόζουν (αλλάζουν/μετατοπίζουν) την ηλεκτρική τους κατανάλωση ως απόκριση σε ένα εξωτερικό αίτημα, όπως είναι η τιμή της ηλεκτρικής ενέργειας, οικονομικά κίνητρα ή τεχνικές απαιτήσεις. Αυτό μπορεί να επιτευχθεί μέσα από κατάλληλα προγράμματα Απόκρισης Ζήτησης, στα οποία μπορούν να συμμετέχουν οι οικιακοί καταναλωτές μετατοπίζοντας την κατανάλωσή τους μακριά από τις περιόδους αιχμής. Τα νοικοκυριά μπορούν να διαδραματίσουν σημαντικό ρόλο στην παροχή ευελιξίας και στην αποσυμφόρηση του δικτύου, μέσω τεχνολογιών που επιτρέπουν μεγαλύτερο έλεγχο στην κατανάλωση μεμονωμένων συσκευών είτε αθροιστικά, σε επίπεδο οικίας. Βελτιστοποιώντας τα προγράμματα λειτουργίας ενεργοβόρων ηλεκτρικών συσκευών και αξιοποιώντας την παραγωγή πράσινης ενέργειας και την αποθήκευση, μπορούν να συμβάλλουν τόσο στη μείωση του κόστους, όσο και στον περιορισμό των περιβαλλοντικών επιπτώσεων [1].

Η παρούσα διπλωματική εργασία ασχολείται με την ομαδοποίηση οικιακών καταναλωτών με χρήση δεδομένων από έξυπνους μετρητές, με στόχο τον εντοπισμό των καταλληλότερων ομάδων για συμμετοχή σε προγράμματα ευελιξίας. Μέσα από ανάλυση δεδομένων πραγματικής κατανάλωσης προσδιορίζονται προφίλ κατανάλωσης, τα οποία παρουσιάζουν έντονη δραστηριότητα κατά τις ώρες αιχμής, με στόχο να αξιοποιηθούν κατάλληλα από τους Φορείς Σωρευτικής Εκπροσώπησης (ΦοΣΕ) Φορτίου για παροχή ευελιξίας. Οι ΦοΣΕ Φορτίου, οι οποίοι λειτουργούν ως σύνδεσμος μεταξύ των καταναλωτών και της αγοράς ηλεκτρικής ενέργειας, συγκεντρώνουν το συνολικό ηλεκτρικό φορτίο πολλών νοικοκυριών. Μέσω αυτής της συγκεντρωτικής προσέγγισης, έχουν την ικανότητα να συμμετέχουν σε αγορές ενέργειας και να παρέχουν υπηρεσίες ευελιξίας σε ολόκληρο το ηλεκτρικό δίκτυο, βοηθώντας στην εξισορρόπηση της παραγωγής και της ζήτησής του. Για παράδειγμα, μπορούν να μειώνουν τη ζήτηση κατά τις ώρες αιχμής ή να αυξήσουν σημαντικά την κατανάλωση κατά περιόδους υψηλής παραγωγής από ΑΠΕ.

1.2 ΣΥΝΟΠΤΙΚΗ ΠΕΡΙΓΡΑΦΗ ΤΗΣ ΜΕΘΟΔΟΛΟΓΙΑΣ

Η μεθοδολογία που ακολουθείται σε αυτήν τη διπλωματική εργασία περιλαμβάνει την επεξεργασία και ανάλυση δεδομένων κατανάλωσης από έξυπνους μετρητές για 5.567 οικιακούς καταναλωτές. Τα δεδομένα αυτά προέρχονται από τον διαχειριστή UK Power Networks του Λονδίνου και αφορούν τη χρονική περίοδο 2011–2014. Η αρχική επεξεργασία περιλαμβάνει την αναδιαμόρφωση των δεδομένων, τη διαχείριση ελλিপών και ακραίων τιμών, καθώς και τον υπολογισμό χαρακτηριστικών που αντιπροσωπεύουν την καταναλωτική συμπεριφορά χρηστών σε ώρες αιχμής.

Στη συνέχεια, εφαρμόζεται ο αλγόριθμος ομαδοποίησης K-Means για την ανάλυση συστάδων καταναλωτών βάσει των επιλεγμένων χαρακτηριστικών. Η διαδικασία περιλαμβάνει την κανονικοποίηση των δεδομένων, την επιλογή του βέλτιστου αριθμού συστάδων μέσω μετρικών αξιολόγησης (Elbow, Silhouette, Calinski-Harabasz), καθώς και την ερμηνεία των αποτελεσμάτων ως προς την καταναλωτική συμπεριφορά κάθε συστάδας.

Τέλος, πραγματοποιείται κατηγοριοποίηση των συστάδων με βάση την ικανότητα για παροχή ευελιξίας μέσω διαχείρισης του φορτίου οικιακών καταναλωτών και γίνεται σύγκριση της ανταπόκρισης των καταναλωτών σε πραγματικά προγράμματα ευελιξίας, με και χωρίς ομαδοποίηση.

1.3 ΔΟΜΗ ΤΗΣ ΕΡΓΑΣΙΑΣ

Η παρούσα διπλωματική εργασία οργανώνεται σε επτά κεφάλαια, καθένα από τα οποία καλύπτει συγκεκριμένες πτυχές της έρευνας και των ευρημάτων σχετικά με την ομαδοποίηση οικιακών καταναλωτών για παροχή ευελιξίας σε περιόδους αιχμών ζήτησης.

- ✓ **Κεφάλαιο 2:** Αναλύεται η έννοια της ευελιξίας από την πλευρά των οικιακών καταναλωτών, οι ενεργειακές προκλήσεις και ο ρόλος των ΦοΣΕ για στοχευμένη επιλογή των καταναλωτών.
- ✓ **Κεφάλαιο 3:** Περιγράφεται η γενική μεθοδολογία της διπλωματικής εργασίας. Παρουσιάζονται τα δεδομένα κατανάλωσης από έξυπνους μετρητές που χρησιμοποιούνται στην ανάλυση, καθώς και η διαδικασία επεξεργασίας των δεδομένων με Python, περιλαμβάνοντας τα στάδια καθαρισμού και διαμόρφωσής τους.
- ✓ **Κεφάλαιο 4:** Περιγράφονται, σε θεωρητικό επίπεδο, οι κύριες τεχνικές που χρησιμοποιούνται στην ανάλυση της ομαδοποίησης. Παρουσιάζεται η εξαγωγή χαρακτηριστικών εισόδου, η κανονικοποίησή τους, η περιγραφή του αλγορίθμου K-Means και ο προσδιορισμός του βέλτιστου αριθμού συστάδων μέσω κατάλληλων μεθόδων.
- ✓ **Κεφάλαιο 5:** Αναλύεται η αξιολόγηση των κριτηρίων εισόδου και η επιλογή βέλτιστου συνδυασμού. Πραγματοποιείται τελική επιλογή του βέλτιστου αριθμού συστάδων και παρουσιάζονται πρακτικές εφαρμογές.

- ✓ **Κεφάλαιο 6:** Γίνεται η σύνοψη της παρούσας διπλωματικής εργασίας, καταγράφονται τα συνολικά συμπεράσματά της και προτείνονται επεκτάσεις της.
- ✓ **Κεφάλαιο 7:** Παρουσιάζεται η βιβλιογραφία που χρησιμοποιήθηκε σε όλη την διπλωματική εργασία.

ΕΥΕΛΙΞΙΑ ΟΙΚΙΑΚΩΝ ΚΑΤΑΝΑΛΩΤΩΝ

2.1 ΕΝΕΡΓΕΙΑΚΕΣ ΠΡΟΚΛΗΣΕΙΣ ΚΑΙ ΑΝΑΓΚΗ ΓΙΑ ΕΥΕΛΙΞΙΑ

Τα τελευταία χρόνια, το ενεργειακό σύστημα έχει αρχίσει να αλλάζει σημαντικά. Η υψηλή διείσδυση των ΑΠΕ και η αυξανόμενη παρουσία μεταβλητών φορτίων στα σύγχρονα δίκτυα, ενισχύουν τη στοχαστικότητα της παραγωγής και την αστάθεια της ζήτησης. Αυτό δημιουργεί προβλήματα συμφόρησης και ανισορροπίας μεταξύ προσφοράς και ζήτησης στο σύστημα ηλεκτροδότησης.

2.1.1 Μεταβλητότητα παραγωγής και ζήτησης

Η ραγδαία αυξανόμενη ενσωμάτωση αποκεντρωμένων ενεργειακών πόρων, όπως τα φωτοβολταϊκά συστήματα, τα συστήματα αποθήκευσης ενέργειας και τα ηλεκτρικά οχήματα, σε συνδυασμό με τη χρήση φορτίων που διασυνδέονται μέσω ηλεκτρονικών ισχύος, όπως οι αντλίες θερμότητας, έχει αυξήσει σημαντικά την πολυπλοκότητα των ηλεκτρικών δικτύων διανομής χαμηλής τάσης. Αυτό το νέο σύνολο συσκευών και οικιακών φορτίων μπορεί να μεταβάλλει τις καταναλωτικές συνήθειες και, κατά συνέπεια, να επηρεάσει τα τυπικά πρότυπα κατανάλωσης. Ως εκ τούτου, η ισορροπία, σχεδόν σε πραγματικό χρόνο, μεταξύ διακοπτόμενης παραγωγής και απρόβλεπτης ζήτησης, καθώς και ο έλεγχος τάσης, η ρύθμιση συχνότητας και η παρακολούθηση της ποιότητας ισχύος καθίστανται όλο και πιο κρίσιμες για τη διατήρηση της ευστάθειας και της αξιοπιστίας του συστήματος ηλεκτρικής ενέργειας.

Παραδοσιακά, η ευθύνη για την αντιμετώπιση αυτών των προκλήσεων βάραινε τους διαχειριστές του συστήματος, οι οποίοι ενίσχυαν το δίκτυο για να ανταποκριθούν. Ωστόσο, στην εποχή των έξυπνων δικτύων διανομής ηλεκτρικής ενέργειας, οι ροές ισχύος είναι πλέον αμφίδρομες και οι καταναλωτές μπορούν να συμμετέχουν ενεργά στη διαμόρφωση ενός πιο αξιόπιστου, αποδοτικού και εύελικτου ενεργειακού συστήματος. Μία από τις πιο υποσχόμενες λύσεις στο επίπεδο της παροχής ευελιξίας είναι η διαχείριση της ζήτησης. Η μείωση των αιχμών ζήτησης, η μετατόπιση φορτίου από ώρες αιχμής σε ώρες χαμηλής ζήτησης και η αύξηση ή μείωση της κατανάλωσης σε ωριαία βάση είναι μερικά από τα οφέλη που μπορεί να προσφέρει [2]. Επιπλέον, η δυσκολία διατήρησης της ισορροπίας μεταξύ προσφοράς και ζήτησης σε πραγματικό χρόνο οδηγεί συχνά στην ανάγκη προσθήκης ή απόσυρσης συμβατικών μονάδων παραγωγής για την κάλυψη των αιχμών φορτίου, δεδομένου ότι η αποθήκευση μεγάλων ποσοτήτων ενέργειας παραμένει έως σήμερα περιορισμένη. Στο πλαίσιο αυτό, τα ευφυή δίκτυα ηλεκτρικής ενέργειας μπορούν να συνεισφέρουν καθοριστικά, επιτρέποντας την έξυπνη διαχείριση της ζήτησης, βασισμένα σε δεδομένα που συλλέγονται σε πραγματικό χρόνο από έξυπνους μετρητές και συσκευές Δικτύου των Πραγμάτων (Internet of Things – IoT). Πέρα από αυτό, η αξιοποίηση τεχνολογιών απομακρυσμένου ελέγχου σε οικιακές συσκευές και ενεργειακούς πόρους, όπως ηλεκτρικά οχήματα και μπαταρίες, καθιστά δυνατή την καλύτερη προσαρμογή της

κατανάλωσης στις διακυμάνσεις της παραγωγής. Η στοχευμένη αυτή ευελιξία οδηγεί σε πιο αποδοτική ένταξη των ΑΠΕ στο σύστημα, ενισχύοντας την αξιοπιστία του δικτύου και περιορίζοντας την εξάρτηση από ορυκτά καύσιμα.

2.1.2 Η συμφόρηση στα ηλεκτρικά δίκτυα και ο ρόλος των διαχειριστών ΣΗΕ

Η συμφόρηση στα ηλεκτρικά δίκτυα αποτελεί ένα από τα πλέον κρίσιμα ζητήματα στη σύγχρονη λειτουργία των συστημάτων ηλεκτρικής ενέργειας (ΣΗΕ). Ορίζεται ως η κατάσταση κατά την οποία η διαθέσιμη ικανότητα μεταφοράς του δικτύου δεν επαρκεί για να εξυπηρετήσει τη ζήτηση ή να μεταφέρει την παραγόμενη ενέργεια από τις περιοχές παραγωγής προς τα κέντρα κατανάλωσης. Η συμφόρηση εκδηλώνεται όταν οι φυσικές ροές ισχύος υπερβαίνουν τα επιτρεπόμενα όρια ενός ή περισσότερων στοιχείων του συστήματος, όπως γραμμές μεταφοράς ή μετασχηματιστές, με αποτέλεσμα να απειλείται η ασφάλεια και η ευστάθεια του δικτύου [3].

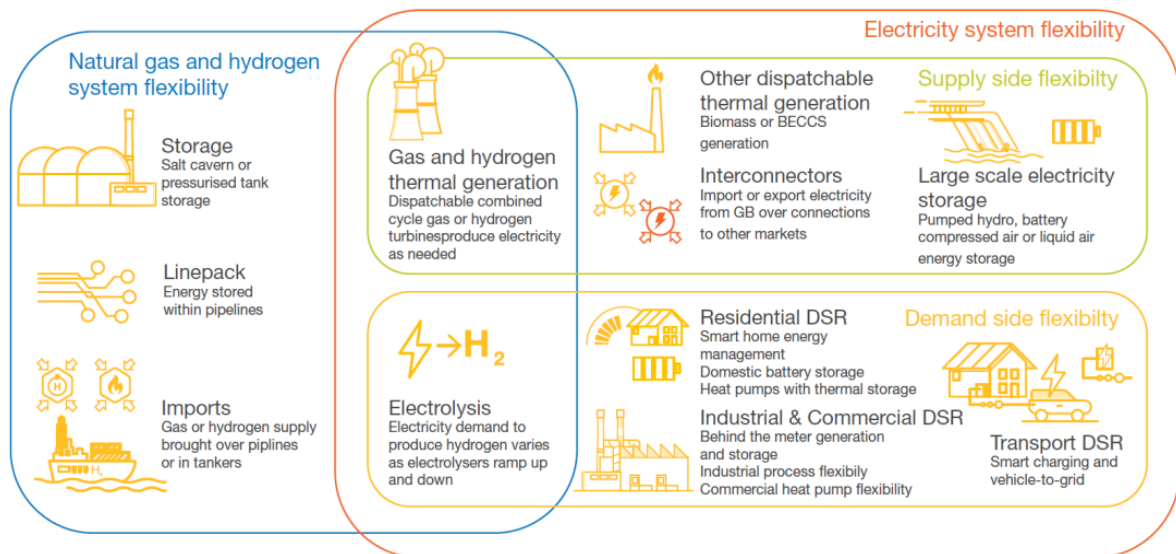
Τα αίτια που προκαλούν συμφόρηση είναι ποικίλα και συνδέονται τόσο με τη διαμόρφωση της παραγωγής, όσο και με τις καταναλωτικές συνήθειες. Μεταξύ των βασικότερων παραγόντων περιλαμβάνονται η αυξημένη ζήτηση ηλεκτρικής ενέργειας, ιδιαίτερα σε ώρες αιχμής, η γεωγραφικά ασύμμετρη κατανομή της παραγωγής και της κατανάλωσης και η εντατική διείσδυση των ΑΠΕ στα δίκτυα. Οι ΑΠΕ, όπως η αιολική και η ηλιακή ενέργεια, συχνά εγκαθίστανται σε απομακρυσμένες περιοχές, όπου το υφιστάμενο δίκτυο δεν επαρκεί για την ασφαλή και αδιάλειπτη απορρόφηση της παραγόμενης ενέργειας, οδηγώντας σε φαινόμενα συμφόρησης, ακόμη και σε περιόδους χαμηλής ζήτησης [3]. Επιπλέον, δημιουργούν νέες προκλήσεις στη λειτουργία των δικτύων, λόγω της στοχαστικής φύσης της παραγωγής τους. Για παράδειγμα, η παραγωγή από φωτοβολταϊκά εξαρτάται από τις καιρικές συνθήκες και μπορεί να επηρεαστεί σημαντικά υπό δυσμενείς συνθήκες, όπως η συννεφιά. Τέτοιες διαταραχές είναι σημαντικά επικίνδυνες, όπως αποτέλεσε το εκτεταμένο μπλακάουτ που σημειώθηκε στην Ιβηρική Χερσόνησο στις 28 Απριλίου 2025. Κατά το συγκεκριμένο περιστατικό, το διασυνδεδεμένο σύστημα Ισπανίας–Πορτογαλίας αποσυνδέθηκε από το Ευρωπαϊκό δίκτυο, έπειτα από απώλεια περίπου 15 GW παραγωγής, η οποία συνέβη εντός ελάχιστων δευτερολέπτων. Η υψηλή διείσδυση ΑΠΕ, σε συνδυασμό με τη μειωμένη περιστρεφόμενη αδράνεια και την ταυτόχρονη ενεργοποίηση συστημάτων προστασίας λόγω ταλαντώσεων στη συχνότητα, πιθανόν να ήταν οι λόγοι που οδήγησαν σε αποσταθεροποίηση και κατάρρευση του συστήματος. Το συμβάν αυτό αναδεικνύει έμπρακτα τον επιχειρησιακό κίνδυνο που προκαλείται από τη συνύπαρξη αυξημένης μεταβλητότητας παραγωγής και ανεπαρκών εφεδρειών στα σύγχρονα ηλεκτρικά δίκτυα [4]. Όλα τα παραπάνω ενδέχεται να έχουν σημαντικές επιπτώσεις στη λειτουργία του συστήματος, όπως η αναγκαστική μείωση της παραγωγής από ΑΠΕ, η αύξηση του λειτουργικού κόστους λόγω ενεργοποίησης ακριβότερων μονάδων παραγωγής, καθώς και ο κίνδυνος πρόκλησης τεχνικών προβλημάτων ή διακοπών ρεύματος σε περιπτώσεις παρατεταμένης υπερφόρτωσης.

Επιπλέον, η εμφάνιση συμφόρησης στη μεταφορά μπορεί να οδηγήσει σε δημιουργία μονοπωλιακής ισχύος για ορισμένους συμμετέχοντες στην αγορά ή ακόμη και να θέσει σε κίνδυνο την ευστάθεια του συστήματος. Επομένως, οι προληπτικές ή διορθωτικές ενέργειες εκ μέρους των διαχειριστών και των ρυθμιστικών αρχών καθίστανται απαραίτητες, τόσο για την άρση της συμφόρησης, όσο και για τη μείωση των λειτουργικών και οικονομικών κινδύνων του συστήματος [3],[5].

Ο ρόλος των διαχειριστών των συστημάτων μεταφοράς είναι καθοριστικός στην πρόληψη και αντιμετώπιση της συμφόρησης. Μέσω στρατηγικών όπως η ανακατανομή φορτίου, η επενδυτική ενίσχυση της υποδομής και η αξιοποίηση έξυπνων τεχνολογιών και αποθηκευτικών λύσεων, οι διαχειριστές στοχεύουν στη διατήρηση της ευστάθειας και της αξιοπιστίας του δικτύου. Παράλληλα, αναδεικνύεται και η σημασία της ευελιξίας στη ζήτηση ως ένα επιπλέον εργαλείο δυναμικής διαχείρισης της συμφόρησης.

2.2 ΠΑΡΟΧΗ ΕΥΕΛΙΞΙΑΣ ΑΠΟ ΤΗΝ ΠΛΕΥΡΑ ΤΗΣ ΖΗΤΗΣΗΣ

2.2.1 Έννοια και σημασία της παροχής ευελιξίας



Σχήμα 2.1: Ευελιξία από την πλευρά της ζήτησης [6]

Η ευελιξία από την πλευρά της ζήτησης ορίζεται ως η τροποποίηση της προβλεπόμενης κατανάλωσης ή παραγωγής από τους τελικούς χρήστες, ως απόκριση σε οικονομικά ή μη οικονομικά κίνητρα και/ή σε δυναμικά σήματα τιμολόγησης ηλεκτρικής ενέργειας, είτε με στόχο τη μείωση του ενεργειακού κόστους, είτε την παροχή υπηρεσιών προς το δίκτυο [7]. Στο πλαίσιο αυτό, η ευελιξία των οικιακών καταναλωτών αναδεικνύεται σε καθοριστικό παράγοντα για την αντιμετώπιση των σημερινών προκλήσεων των δικτύων διανομής. Η προσαρμογή της ζήτησης ενέργειας συμβάλλει στον περιορισμό των αιχμών κατανάλωσης, στην ενίσχυση της ευστάθειας του συστήματος και στην αποτελεσματικότερη αξιοποίηση των διαθέσιμων ενεργειακών πόρων. Η αξιοποίηση σύγχρονων τεχνολογιών διαχείρισης ενέργειας ενισχύει την ικανότητα των καταναλωτών να συμβάλλουν ενεργά στη διαμόρφωση ενός πιο βιώσιμου και αποδοτικού ενεργειακού περιβάλλοντος.

2.2.2 Ρόλος των ΦοΣΕ

Οι Φορείς Σωρευτικής Εκπροσώπησης (ΦοΣΕ) Παραγωγής είναι εξειδικευμένες οντότητες που αναλαμβάνουν, εκ μέρους παραγωγών ΑΠΕ, την εκπροσώπησή τους στις αγορές ηλεκτρικής ενέργειας. Αναλαμβάνουν λειτουργίες, όπως η πρόβλεψη παραγωγής, η υποβολή προσφορών, η διαχείριση αποκλίσεων και η υποστήριξη της συμμόρφωσης των παραγωγών με τις απαιτήσεις της αγοράς. Παράλληλα, παρέχουν συμβουλευτικές και τεχνοοικονομικές υπηρεσίες, διευκολύνοντας τη λειτουργική και οικονομική ένταξη των ΑΠΕ στην αγορά. Μέσω των τεχνολογιών διαχείρισης ενέργειας, μπορούν να ενισχύσουν την ευελιξία και την αποδοτικότητα του συστήματος.

Οι Φορείς Σωρευτικής Εκπροσώπησης Φορτίου αποτελούν νομικές οντότητες, οι οποίες συγκεντρώνουν και διαχειρίζονται φορτία από πλήθος τελικών καταναλωτών, εκπροσωπώντας τα συλλογικά στις αγορές ηλεκτρικής ενέργειας. Οι εν λόγω φορείς δραστηριοποιούνται με σκοπό την παροχή ευελιξίας στο σύστημα μέσω της μείωσης ή της μετατόπισης της κατανάλωσης, ως απόκριση σε οικονομικά ή τεχνικά σήματα του δικτύου. Σε αντίθεση με τους ΦοΣΕ Παραγωγής, οι ΦοΣΕ Φορτίου δεν εκπροσωπούν μονάδες παραγωγής, αλλά καταναλωτές ή ομάδες καταναλωτών. Ο ρόλος τους εστιάζει στη διαχείριση του φορτίου που συγκεντρώνουν, ώστε να παρέχουν υπηρεσίες εξισορρόπησης, ευελιξίας και απόκρισης ζήτησης, κυρίως σε περιόδους αυξημένης ή στοχαστικής ζήτησης. Μέσω της συστηματικής αξιοποίησης αυτών των φορτίων, έχουν τη δυνατότητα να συμμετέχουν αυτόνομα στην αγορά εξισορρόπησης, ενισχύοντας την αποδοτικότητα και την ευστάθεια του συστήματος. Στο πλαίσιο των οικιακών καταναλωτών, οι ΦοΣΕ καλούνται να αναλάβουν κρίσιμο ρόλο στην εξισορρόπηση του συστήματος. Ένας τρόπος είναι η ενεργοποίηση των οικιακών καταναλωτών, δηλαδή η δυνατότητα να προσαρμόσουν την κατανάλωσή τους κατά την αιχμή ζήτησης, ανάλογα με τις ανάγκες του συστήματος. Βασική προϋπόθεση αποτελεί η κατανόηση της καταναλωτικής τους συμπεριφοράς και η ομαδοποίηση των χρηστών με βάση την ικανότητά τους να παρέχουν ευελιξία.

2.3 ΟΦΕΛΗ ΤΩΝ ΠΡΟΓΡΑΜΜΑΤΩΝ ΔΙΑΧΕΙΡΙΣΗΣ ΖΗΤΗΣΗΣ

Οι αιχμές κατανάλωσης, κατά τις οποίες η ζήτηση ηλεκτρικής ενέργειας είναι ιδιαίτερα αυξημένη, ενδέχεται να οδηγήσουν σε υπερφόρτωση των δικτύων διανομής, προκαλώντας καταπόνηση του δικτύου, αύξηση του κόστους παραγωγής και, σε ακραίες περιπτώσεις, διακοπή του ηλεκτρικού ρεύματος. Από την πλευρά των καταναλωτών, η δυνατότητα ευελιξίας κατά τις ώρες αιχμών, μπορεί να συνεισφέρει ουσιαστικά στη διαχείριση αυτών, καθώς οι ίδιοι μπορούν να προσαρμόζουν την κατανάλωσή τους με βάση είτε την τιμή της ηλεκτρικής ενέργειας, είτε τα κίνητρα που προσφέρονται μέσω προγραμμάτων διαχείρισης ενέργειας. Με αυτόν τον τρόπο, είναι εφικτή η μείωση της ζήτησης κατά τις κρίσιμες χρονικές περιόδους, αλλά και η μετατόπιση της κατανάλωσης σε ώρες που η παραγωγή υπερτερεί της ζήτησης. Η προσαρμογή αυτή μπορεί να πραγματοποιηθεί είτε χειροκίνητα, είτε μέσω αυτοματοποιημένων τεχνολογιών. Το αποτέλεσμα είναι η αποσυμφόρηση των δικτύων διανομής και η ταυτόχρονη μείωση του κόστους παραγωγής, καθώς περιορίζεται η ανάγκη για ενεργοποίηση νέων ακριβότερων μονάδων ή για επενδύσεις στο δίκτυο. Επιπλέον, οι καταναλωτές μπορούν να συμβάλλουν στη βελτίωση της αποδοτικότητας, της αξιοπιστίας και της ευελιξίας των ενεργειακών δικτύων, ενεργώντας ταυτόχρονα ως καταναλωτές και παραγωγοί ηλεκτρικής ενέργειας.

Αυτό επιτυγχάνεται μέσω της δυνατότητας αγοράς ή πώλησης ενέργειας, την οποία έχουν παράγει ή αποθηκεύσει, προς τον εκάστοτε πάροχο. Πέρα από τη συνεισφορά των δραστηριοτήτων αυτών προς όφελος του δικτύου, οι καταναλωτές επωφελούνται σημαντικά σε ατομικό επίπεδο από τη συμμετοχή τους σε σχήματα ευέλικτης διαχείρισης της κατανάλωσης.

Η ένταξη των καταναλωτών σε προγράμματα διαχείρισης ενέργειας ή δυναμικής τιμολόγησης επιτρέπει τη μείωση του ενεργειακού τους κόστους, καθώς τους δίνεται η δυνατότητα να προσαρμόζουν την κατανάλωσή τους βάσει χρονικών ζωνών ή τιμών που μεταβάλλονται σε πραγματικό χρόνο. Οι χρήστες επιβραβεύονται μέσω προγραμμάτων που βασίζονται σε κίνητρα, ενισχύοντας έτσι τη συμμετοχή τους και ενθαρρύνοντας την υιοθέτηση ευέλικτης συμπεριφοράς. Η ενσωμάτωση Συστημάτων Οικιακής Διαχείρισης Ενέργειας (Home Energy Management Systems – HEMS) καθιστά δυνατή την αυτόματη και βέλτιστη προγραμματισμένη λειτουργία των φορτίων, μειώνοντας την ανάγκη για ανθρώπινη παρέμβαση και αυξάνοντας την ενεργειακή αποδοτικότητα του νοικοκυριού. Παράλληλα, δίνεται στους καταναλωτές η δυνατότητα να αναλάβουν ενεργό ρόλο στη διαχείριση της ενέργειας που καταναλώνουν ή παράγουν, ενισχύοντας έτσι τη γνώση και τη θέση τους σε ενεργειακό επίπεδο [8].

2.4 ΠΡΟΓΡΑΜΜΑΤΑ ΑΠΟΚΡΙΣΗΣ ΖΗΤΗΣΗΣ ΚΑΙ ΣΤΟΧΕΥΜΕΝΗ ΕΠΙΛΟΓΗ ΚΑΤΑΝΑΛΩΤΩΝ

Για την αντιμετώπιση προκλήσεων, όπως η μείωση των αιχμών κατανάλωσης και η εξομάλυνση της ημερήσιας καμπύλης φορτίου, έχει προταθεί ως λύση η εφαρμογή προγραμμάτων Απόκρισης Ζήτησης (AZ). Η απόκριση ζήτησης αποτελεί την προέκταση της διαχείρισης από την πλευρά της ζήτησης στο πλαίσιο ανταγωνιστικών αγορών ηλεκτρικής ενέργειας. Μέσω σημάτων τιμών και μηχανισμών κινήτρων, καθίσταται κρίσιμη τόσο για την αξιοπιστία του ηλεκτρικού συστήματος, όσο και για την αποδοτικότητα της αγοράς. Ανάλογα με το κίνητρο ενεργοποίησης της, τα προγράμματα απόκρισης ζήτησης στις αγορές ηλεκτρικής ενέργειας διακρίνονται σε δύο κατηγορίες: προγράμματα βάσει τιμής και προγράμματα βάσει κινήτρων [9].

Η στοχευμένη επιλογή των καταναλωτών αποτελεί κρίσιμο στάδιο για την αποτελεσματική εφαρμογή προγραμμάτων απόκρισης ζήτησης. Οι καταναλωτές ενθαρρύνονται να συμμετάσχουν μέσω τιμολογιακών σχημάτων που βασίζονται στον χρόνο, με την επιλογή τους να εξαρτάται συνήθως από την ιστορική συμπεριφορά κατανάλωσης ενέργειας. Επομένως, είναι προτιμότερο να διασφαλίζεται ότι οι επιλεγμένοι χρήστες παρουσιάζουν συνεπή πρότυπα χρήσης και είναι πιθανό να καταναλώνουν ενέργεια σύμφωνα με τις προσδοκίες του παρόχου κατά τις ώρες ενεργοποίησης των προγραμμάτων. Ένα από τα πλεονεκτήματα της προσέγγισης αυτής, είναι ότι τα κίνητρα στοχεύουν αποκλειστικά σε εκείνους τους καταναλωτές των οποίων η αναμενόμενη συμπεριφορά οδηγεί σε μείωση του κόστους για τον προμηθευτή. Καταναλωτές με απρόβλεπτη ή τυχαία συμπεριφορά ενδέχεται να εμποδίσουν την επίτευξη των στόχων των προγραμμάτων απόκρισης ζήτησης από την πλευρά των παρόχων.

Η στοχευμένη παροχή κινήτρων μειώνει τα κόστη που επιβαρύνουν τον προμηθευτή, περιορίζει τις αιχμές στη ζήτηση και ενδέχεται να οδηγήσει σε υψηλότερα ποσοστά διατήρησης των καταναλωτών. Ταυτόχρονα, μειώνονται οι παρεμβάσεις στην καθημερινότητα των καταναλωτών, περιορίζεται η εμφάνιση φαινομένων «κόπωσης

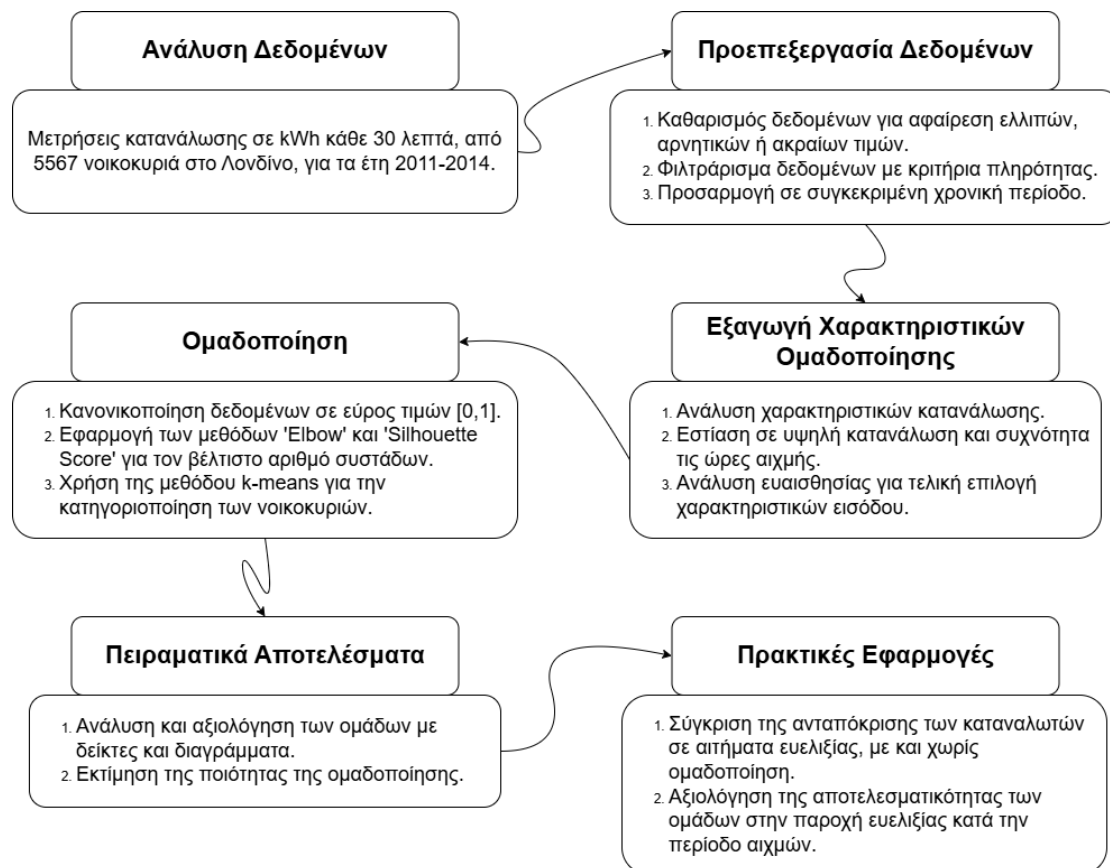
απόκρισης ζήτησης» και αποφεύγεται η χορήγηση κινήτρων σε χρήστες των οποίων η τελική συμπεριφορά αυξάνει το λειτουργικό κόστος του προμηθευτή. Επιπλέον, προτείνεται η στόχευση υποσύνολου καταναλωτών που είναι ταυτόχρονα σχετικοί, δηλαδή είναι πιθανό να χρησιμοποιούν συσκευές μετατοπίσιμου φορτίου, όπως ηλεκτρικά οχήματα και αντλίες θερμότητας κατά τις ώρες αιχμής, και είναι πρόθυμοι να μειώσουν την κατανάλωσή τους (πιθανό να αντιδράσουν θετικά σε αίτημα μείωσης από τον πάροχο ή τον ΦοΣΕ).

Ωστόσο, οι παραδοσιακές τεχνικές ομαδοποίησης, αν και εντοπίζουν ομάδες καταναλωτών σε προγράμματα AZ, δεν λαμβάνουν συχνά υπόψη την ιστορική συνέπεια της συμπεριφοράς ενός καταναλωτή σε βάθος χρόνου. Η ποσοτικοποίηση της συνέπειας στην κατανάλωση είναι κρίσιμη, καθώς υποδεικνύει την πιθανότητα ένας καταναλωτής να ακολουθήσει το ιστορικό του πρότυπο κατά την ημέρα και ώρα εφαρμογής του προγράμματος AZ. Για τον λόγο αυτό, στη βιβλιογραφία έχουν προταθεί μεθοδολογίες για την επιλογή καταναλωτών που παρουσιάζουν αιχμές κατανάλωσης την επιθυμητή χρονική στιγμή, έναντι εκείνων που επιδεικνύουν συνεπή καταναλωτική συμπεριφορά [10],[11]. Η ομαδοποίηση των καταναλωτών και η στοχευμένη επιλογή κατάλληλων ομάδων για την ένταξή τους σε προγράμματα απόκρισης ζήτησης αποτελεί αρμοδιότητα και ευθύνη των ΦοΣΕ.

ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΚΑΤΑΝΑΛΩΣΗΣ ΑΠΟ ΕΞΥΠΝΟΥΣ ΜΕΤΡΗΤΕΣ

3.1 ΓΕΝΙΚΗ ΜΕΘΟΔΟΛΟΓΙΑ

Η παρούσα διπλωματική εργασία ακολουθεί μεθοδολογία ανάλυσης που περιλαμβάνει έξι βασικά στάδια, τα οποία συνοψίζονται στο Σχήμα 3.1 με τη μορφή διαγράμματος ροής. Στο πρώτο στάδιο πραγματοποιείται η επιλογή και συγκέντρωση των δεδομένων κατανάλωσης μέσω έξυπνων μετρητών. Στη συνέχεια, ακολουθεί η προεπεξεργασία των δεδομένων με στόχο τη βελτίωση της ποιότητάς τους. Περιλαμβάνεται ο καθαρισμός από ελλείψεις, αρνητικές ή ακραίες τιμές, η εφαρμογή φίλτρων πληρότητας και η επιλογή της χρονικής περιόδου μελέτης. Το τρίτο στάδιο αφορά την εξαγωγή χαρακτηριστικών εισόδου που σχετίζονται με τη συμπεριφορά κατανάλωσης κατά τις ώρες αιχμής, όπως υψηλή κατανάλωση και συχνότητα, καθώς γίνεται και μία ανάλυση ευαισθησίας για την τελική επιλογή αυτών. Ακολουθεί η διαδικασία ομαδοποίησης, κατά την οποία τα δεδομένα κανονικοποιούνται σε ενιαία κλίμακα τιμών. Εφαρμόζονται κάποιες μέθοδοι για τον προσδιορισμό του βέλτιστου αριθμού συστάδων, ενώ η τελική κατάταξη των καταναλωτών υλοποιείται με χρήση τεχνικής ανάλυσης συστάδων. Τα αποτελέσματα της ομαδοποίησης αξιολογούνται στο πέμπτο στάδιο, μέσω διαγραμμάτων και δεικτών, ώστε να εκτιμηθεί η αποτελεσματικότητα της μεθόδου. Τέλος, παρουσιάζονται πρακτικές εφαρμογές που αποσκοπούν στην αξιολόγηση της ανταπόκρισης των καταναλωτών σε αιτήματα ευελιξίας. Η ανάλυση πραγματοποιείται τόσο σε επίπεδο μεμονωμένων χρηστών, όσο και σε επίπεδο συστάδων, ώστε να διερευνηθεί η συμβολή της ομαδοποίησης στη βελτίωση της διαχειρισσιμότητας και της αποδοτικότητας των μηχανισμών ευελιξίας. Η μεθοδολογική προσέγγιση και η υλοποίηση όλων των σταδίων αυτής, στην παρούσα διπλωματική εργασία, στηρίζονται σε πρακτικές της διεθνούς βιβλιογραφίας [12]. Αξίζει να σημειωθεί ότι όλα τα στάδια ανάλυσης είναι προσαρμοσμένα στις ιδιαιτερότητες του συνόλου δεδομένων, το οποίο παρουσιάζεται στην Ενότητα 3.2.



Σχήμα 3.1: Διάγραμμα ροής της μεθοδολογίας

3.2 ΔΕΔΟΜΕΝΑ ΕΙΣΟΔΟΥ

Η ραγδαία ανάπτυξη των τεχνολογιών έξυπνης μέτρησης και η ευρεία εφαρμογή τους σε οικιακό επίπεδο, καθιστούν δυνατή τη συλλογή λεπτομερών δεδομένων κατανάλωσης σε πραγματικό χρόνο. Τα δεδομένα αυτά προσφέρουν νέες δυνατότητες για την κατανόηση της συμπεριφοράς των καταναλωτών και την ανάπτυξη στρατηγικών ευέλικτης διαχείρισης της ζήτησης, ιδιαίτερα σε συνθήκες αιχμών.

Στην παρούσα εργασία, χρησιμοποιήθηκε το σύνολο δεδομένων “SmartMeter Energy Use Data in London Households” [13]. Αποτελεί ένα πλήρες σύνολο δεδομένων που δημιουργήθηκε στο πλαίσιο του έργου Low Carbon London και υλοποιήθηκε υπό την εποπτεία του διαχειριστή δικτύου διανομής ηλεκτρικής ενέργειας UK Power Networks. Το σύνολο δεδομένων αυτό περιλαμβάνει μετρήσεις κατανάλωσης ηλεκτρικής ενέργειας μέσω έξυπνων μετρητών, οι οποίες προέρχονται από δείγμα νοικοκυριών του Λονδίνου. Οι καταναλωτές που συμμετείχαν στην πιλοτική δράση επιλέχθηκαν, έτσι ώστε να αποτελούν αντιπροσωπευτικό δείγμα του πληθυσμού του ευρύτερου Λονδίνου. Το εν λόγω σύνολο δεδομένων παρέχει λεπτομερείς πληροφορίες σχετικά με τις ενεργειακές συνθήκες των νοικοκυριών, επιτρέποντας τη διερεύνηση της ευελιξίας στη ζήτηση και στη βελτιστοποίηση της κατανάλωσης ηλεκτρικής ενέργειας σε οικιακό επίπεδο. Το αρχείο δεδομένων είναι ελεύθερα προσβάσιμο μέσω της επίσημης διαδικτυακής πύλης του Δήμου του Λονδίνου και ακολουθεί τις αρχές FAIR (Findable, Accessible, Interoperable, Reusable)¹, διασφαλίζοντας

¹ FAIR: <https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households>

τη δυνατότητα εύρεσης, πρόσβασης και επαναχρησιμοποίησης των δεδομένων με τυποποιημένο και τεκμηριωμένο τρόπο.

Η ανάλυση στο πλαίσιο αυτής της διπλωματικής εργασίας επικεντρώνεται στην επεξεργασία δεδομένων, με στόχο την εξαγωγή ενός τελικού συνόλου δεδομένων, το οποίο θα χρησιμοποιηθεί ως βάση για την ομαδοποίηση οικιακών καταναλωτών για ένα ΦοΣΕ Φορτίου. Η χρήση ενός τέτοιου πραγματικού συνόλου δεδομένων αποτελεί πολύτιμο εργαλείο για την κατανόηση της ενεργειακής κατανάλωσης σε νοικοκυριά, δίνοντας τη δυνατότητα στους ερευνητές να προχωρήσουν σε αναλύσεις που αποσκοπούν στη διαχείριση της ενέργειας και παροχή ευελιξίας στο δίκτυο. Η κατανόηση της φύσης και της δομής του διαθέσιμου συνόλου δεδομένων αποτελεί βασικό βήμα για την ορθή εφαρμογή της μεθοδολογίας. Οι μετρήσεις προέρχονται από 5.567 οικιακούς καταναλωτές και αφορούν χρονοσειρές κατανάλωσης ηλεκτρικής ενέργειας μέσω έξυπνων μετρητών με ενδείξεις ανά 30 λεπτά, για τη χρονική περίοδο από Νοέμβριο 2011 έως Φεβρουάριο 2014. Το αρχείο του είναι αποθηκευμένο σε μορφή CSV, έχει μέγεθος περίπου 10 GB σε αποσυμπίεσμένη μορφή και περιλαμβάνει πάνω από 167 εκατομμύρια εγγραφές. Κάθε εγγραφή περιλαμβάνει τις εξής πληροφορίες:

- **LCLid**: μοναδικός αναγνωριστικός αριθμός κάθε καταναλωτή, ο οποίος επιτρέπει την παρακολούθηση των μετρήσεων ανά χρήση.
- **stdorToU**: κατηγορία τιμολογιακού καθεστώτος. Η τιμή “Std” δηλώνει σταθερό τιμολόγιο, ενώ “dToU” δηλώνει δυναμικό τιμολόγιο βάσει χρόνου χρήσης.
- **DateTime**: χρονική σήμανση (timestamp) που καταγράφει την ημερομηνία και ώρα της μέτρησης.
- **kWh/hh (per half hour)**: κατανάλωση ηλεκτρικής ενέργειας, εκφρασμένη σε κιλοβατώρες ανά μισή ώρα.

Το σύνολο δεδομένων περιλαμβάνει δύο κατηγορίες καταναλωτών:

1. Περίπου 1.100 καταναλωτές συμμετείχαν κατά το ημερολογιακό έτος 2013 σε πιλοτικό πρόγραμμα δυναμικής τιμολόγησης βάσει χρόνου χρήσης (dToU – Dynamic Time of Use). Οι χρεώσεις ανακοινώνονταν μία ημέρα νωρίτερα, μέσω της οθόνης του έξυπνου μετρητή ή με γραπτό μήνυμα στο κινητό τους τηλέφωνο. Οι καταναλωτές λάμβαναν σήματα με τιμή χαμηλή, κανονική ή υψηλή, αντίστοιχα, ανάλογα με την ώρα της ημέρας.
2. Οι υπόλοιποι περίπου 4.500 καταναλωτές ακολούθησαν σταθερό τιμολόγιο, με ενιαία χρέωση.

Τα τιμολογιακά σχήματα σχεδιάστηκαν, ώστε να προσομοιώνουν μελλοντικά σενάρια λειτουργίας του συστήματος, όπως η διαχείριση της αυξημένης παραγωγής από ΑΠΕ και η αποσυμφόρηση του δικτύου σε περιόδους αιχμών. Στον Πίνακα 3.1 παρουσιάζεται συνοπτικά το σύνολο δεδομένων “SmartMeter Energy Use Data in London Households”.

Πίνακας 3.1: Περιγραφή του συνόλου δεδομένων

Μεταβλητή	Τύπος Δεδομένων	Περιγραφή	Παράδειγμα
LCLid	String	Μοναδικός αναγνωριστικός αριθμός κάθε καταναλωτή	MAC000295
stdorToU	String	Τύπος τιμολογίου (Std ή dToU) – αντιστοιχεί σε περίπου 4.500 και 1.100 νοικοκυριά αντίστοιχα	Std
DateTime	DateTime	Δεδομένα χρονοσειράς που συλλέχθηκαν από τον Νοέμβριο 2011 έως τον Φεβρουάριο 2014	02/05/2012 1:00:00
kWh/hh (per half hour)	Float	kWh (κιλοβατώρα): αντιστοιχεί στη συνολική ποσότητα ενέργειας που καταναλώθηκε hh (half-hourly): αντιστοιχεί σε μετρήσεις ανά 30 λεπτά από τους έξυπνους μετρητές, οι οποίες αποστέλλονται στον προμηθευτή ενέργειας	0,054

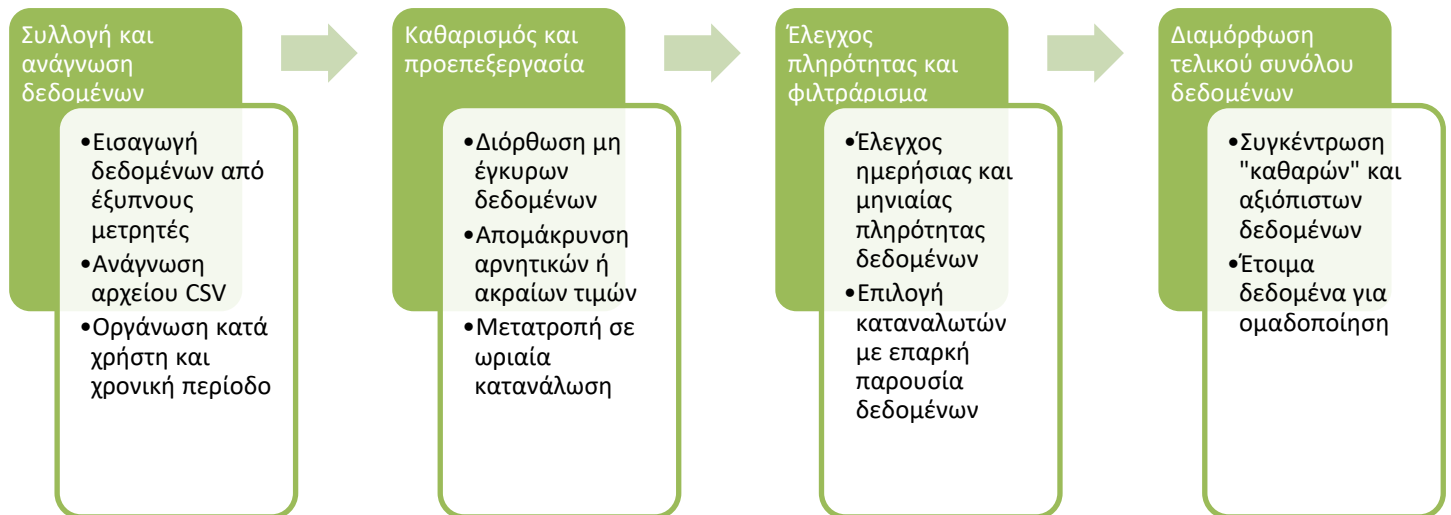
3.3 ΜΕΘΟΔΟΣ ΑΝΑΛΥΣΗΣ

Η μέθοδος ανάλυσης βασίζεται στην επεξεργασία χρονοσειριακών δεδομένων κατανάλωσης ηλεκτρικής ενέργειας, με στόχο τη δημιουργία ενός νέου αξιόπιστου και κατάλληλου προς ανάλυση συνόλου δεδομένων, το οποίο θα χρησιμοποιηθεί για να γίνει η ομαδοποίηση των οικιακών καταναλωτών για έναν ΦοΣΕ. Το αρχικό αρχείο είναι ενιαίο και περιλαμβάνει μετρήσεις από χιλιάδες καταναλωτές για χρονική περίοδο περίπου δύο ετών. Λόγω του μεγάλου όγκου του αρχείου, η ανάλυση απαιτεί συστηματική διαχείριση. Η μεθοδολογία ανάλυσης περιλαμβάνει τα εξής βασικά στάδια:

- **Συλλογή και ανάγνωση δεδομένων:** Τα δεδομένα συλλέγονται από τους έξυπνους μετρητές που καταγράφουν την ενεργειακή κατανάλωση σε πραγματικό χρόνο. Η συλλογή αυτών των δεδομένων πραγματοποιείται σε μορφή αρχείων CSV, τα οποία επιτρέπουν την εύκολη αποθήκευση και ανάγνωση των δεδομένων για περαιτέρω ανάλυση. Η οργάνωση προσαρμόζεται σε συγκεκριμένη χρονική περίοδο και πραγματοποιείται βάσει του μοναδικού αναγνωριστικού κάθε χρήστη.
- **Καθαρισμός και βασική επεξεργασία:** Πραγματοποιείται διόρθωση μη έγκυρων μετρήσεων, όπως ελλειπίς, αρνητικές ή ακραίες τιμές κατανάλωσης και αναγωγή των μετρήσεων κατανάλωσης σε ωριαία συχνότητα.
- **Έλεγχος πληρότητας και φιλτράρισμα:** Οι καταναλωτές αξιολογούνται ως προς την πληρότητα των διαθέσιμων μετρήσεων βάσει ημερήσιων και μηνιαίων ορίων. Διατηρούνται μόνο όσοι πληρούν συγκεκριμένα κριτήρια για το επιλεγμένο χρονικό διάστημα.

- **Διαμόρφωση τελικού συνόλου δεδομένων:** Τα καθαρισμένα και φιλτραρισμένα δεδομένα συγκεντρώνονται σε ένα τελικό ενιαίο dataset, κατάλληλο για την ομαδοποίηση των οικιακών καταναλωτών.

Τα παραπάνω βήματα υλοποιούνται με χρήση της γλώσσας προγραμματισμού Python, αξιοποιώντας βιβλιοθήκες όπως οι Pandas και Numpy, που παρέχουν τη δυνατότητα για διαχείριση και επεξεργασία μεγάλων συνόλων δεδομένων [14],[15],[16]. Στο Σχήμα 3.2 παρουσιάζεται συνοπτικά η μέθοδος ανάλυσης των δεδομένων.



Σχήμα 3.2: Επεξεργασία δεδομένων μέσω της Python

3.4 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ ΜΕ ΚΩΔΙΚΑ PYTHON

3.4.1 Ανάγνωση δεδομένων

Η διαδικασία ανάλυσης ξεκινά με την ανάγνωση του ενιαίου αρχείου δεδομένων, το οποίο είναι αποθηκευμένο σε μορφή CSV. Το αρχείο είναι μεγάλου μεγέθους και περιέχει εκατομμύρια εγγραφές, με μετρήσεις σε διαστήματα μισής ώρας. Η ανάγνωσή του πραγματοποιείται με χρήση της βιβλιοθήκης Pandas, η οποία παρέχει αποδοτικές μεθόδους εισαγωγής και χειρισμού δεδομένων μεγάλης κλίμακας. Στο πλαίσιο της παρούσας διπλωματικής εργασίας, στόχος είναι η ανάλυση του έτους 2013, καθώς είναι το πιο πρόσφατο και πλήρες έτος, από τα διαθέσιμα του συνόλου δεδομένων, γνωρίζοντας ότι το 2014 έχει μόνο 2 μήνες εγγραφών. Για τον σκοπό αυτό, εφαρμόζεται επαναληπτική διαδικασία που επιτρέπει την επεξεργασία των δεδομένων κάθε μήνα του έτους 2013. Η επιμέρους επεξεργασία υλοποιείται με τη βοήθεια ειδικών συναρτήσεων, οι οποίες περιγράφονται αναλυτικά στις επόμενες υποενότητες.

3.4.2 Αναδιαμόρφωση δεδομένων

Στο στάδιο αυτό πραγματοποιείται αναδιαμόρφωση της δομής των εγγραφών, ώστε να καταστούν κατάλληλες για περαιτέρω ανάλυση και φιλτράρισμα. Η επεξεργασία αυτή υλοποιείται μέσω βοηθητικής συνάρτησης, η οποία καλείται εντός της επαναληπτικής διαδικασίας. Αρχικά η στήλη DateTime, η οποία περιέχει τη χρονική σήμανση κάθε μέτρησης, μετατρέπεται σε αντικείμενο τύπου ημερομηνίας και ώρας. Από τη στήλη αυτή εξάγονται επιμέρους χρονικά χαρακτηριστικά, όπως το έτος (Year), ο μήνας (Month), η ημέρα (Day) και η ώρα (Hour). Επιπλέον, προκειμένου να διευκολυνθεί η ανάλυση της συμπεριφοράς των καταναλωτών, οι τιμές κατανάλωσης που καταγράφονται ανά μισή ώρα αναγάζονται σε ωριαία συχνότητα ισχύος, δηλαδή σε kWh/h. Με αυτόν τον τρόπο, εξασφαλίζεται η ομοιογένεια του συνόλου δεδομένων και μειώνεται αρκετά η πολυπλοκότητα της ανάλυσης. Η έξοδος του σταδίου αυτού είναι ένα πλήρως τυποποιημένο υποσύνολο δεδομένων για κάθε μήνα, με συνεπή χρονική δομή και οργανωμένη πληροφορία, έτοιμο για τη διαχείριση ελλিপών, αρνητικών και ακραίων τιμών που ακολουθεί.

Αξίζει να σημειωθεί ότι η παρούσα διπλωματική εργασία δεν εξετάζει την τιμολογιακή πολιτική, όπως αυτή αποτυπώνεται στη στήλη stdorToU, δεδομένου ότι ο στόχος επικεντρώνεται στην ομαδοποίηση καταναλωτών με σκοπό τη μείωση της ζήτησης στο πλαίσιο προγραμμάτων Απόκρισης Ζήτησης από τους ΦοΣΕ Φορτίου. Επειδή ο τύπος τιμολόγησης δεν είναι πάντα διαθέσιμος, η ανάλυση βασίζεται αποκλειστικά σε ιστορικά δεδομένα κατανάλωσης, ώστε να είναι εφαρμόσιμη και σε περιπτώσεις όπου η πληροφορία τιμολόγησης απουσιάζει. Τελικά, η κάθε εγγραφή αναδομείται και εκφράζεται με τη μορφή επτά διακριτών πεδίων:

- μοναδικό αναγνωριστικό καταναλωτή (LCLid)
- έτος (Year)
- μήνας (Month)
- ημέρα (Day)
- ώρα (Hour)
- κατανάλωση ενεργού ισχύος σε κιλοβάτ (kW)

3.4.3 Διαχείριση ελλিপών, αρνητικών και ακραίων τιμών

Σε αυτό το στάδιο εφαρμόζεται βασικός καθαρισμός των δεδομένων κατανάλωσης, με σκοπό την εξάλειψη τιμών που είναι μη έγκυρες ή ενδέχεται να αλλοιώσουν τα αποτελέσματα της ανάλυσης. Αρχικά, οι μη αριθμητικές ή κενές εγγραφές στη στήλη κατανάλωσης μετατρέπονται σε ειδικές τιμές NaN (Not a Number), μέσω της κατάλληλης εντολής για μετατροπή σε αριθμητικό τύπο. Οι τιμές αυτές επισημαίνονται ως ελλειπείς και λαμβάνονται υπόψη στα επόμενα στάδια φιλτραρίσματος. Στη συνέχεια, εντοπίζονται οι αρνητικές τιμές στη στήλη kW, οι οποίες δεν θεωρούνται φυσιολογικές για δεδομένα ηλεκτρικής κατανάλωσης. Οι τιμές αυτές τροποποιούνται σε μηδενικές μέσω ειδικής βοηθητικής συνάρτησης, η οποία ελέγχει κάθε εγγραφή και αντικαθιστά κάθε τιμή μικρότερη του μηδενός με μηδέν. Επιπλέον, αξιοποιείται συνάρτηση εντοπισμού ακραίων τιμών σε ημερήσια βάση. Για κάθε χρήστη υπολογίζεται η συνολική ημερήσια κατανάλωση και επισημαίνονται οι ημέρες, κατά τις οποίες το άθροισμα υπερβαίνει το προκαθορισμένο όριο των 200 kWh, το οποίο ορίστηκε έπειτα από ανάλυση της συμπεριφοράς κατανάλωσης όλων των καταναλωτών. Οι τιμές αυτές καταγράφονται ξεχωριστά, ώστε να είναι δυνατή η

αξιολόγηση και η εξαίρεσή τους από την τελική ανάλυση, εφόσον αυτό κρίνεται αναγκαίο. Παράλληλα, εφαρμόζεται μηχανισμός διασφάλισης της συνέπειας συμμετοχής των καταναλωτών στο χρονικό διάστημα που εξετάζεται. Κατά την επεξεργασία των δεδομένων, επιλέγονται εκείνοι οι χρήστες που παρουσιάζουν σταθερή παρουσία στα υποσύνολα των δεδομένων ανά μήνα, ώστε να εξασφαλίζεται η ομοιογένεια και η αξιοπιστία του τελικού πληθυσμού προς ανάλυση. Η συνδυαστική διαχείριση των παραπάνω τιμών συμβάλλει στη δημιουργία ενός καθαρού και αξιόπιστου συνόλου δεδομένων, το οποίο μπορεί να υποστηρίξει τις επόμενες φάσεις της ανάλυσης. Στα Σχήματα 3.3 και 3.4, παρουσιάζεται ο κώδικας των βοηθητικών συναρτήσεων που χρησιμοποιήθηκαν σε αυτό το στάδιο της μεθοδολογίας.

```
# FUNCTION 2

def negative_values(df):
    # Identify negative values in the "KW" column
    negative_values = df[df["KW"] < 0]
    # print(f"Number of negative values: {len(negative_values)}")

    # Replace negative values with zero
    df["KW"] = df["KW"].apply(lambda x: max(x, 0))

    # Verify that all negative values have been corrected
    negative_values_after = df[df["KW"] < 0]

    # if len(negative_values) > 0:
    print(f"Number of negative values after correction: {len(negative_values_after)}")

    return df
```

Σχήμα 3.3: Βοηθητική συνάρτηση διαχείρισης αρνητικών τιμών

```
# FUNCTION 6

def outliers(df, threshold):
    # Calculate daily consumption per household
    daily_consumption = df.groupby(["ICLId", "Year", "Month", "Day", as_index=False]).agg({"KW": "sum"})
    daily_consumption.rename(columns={"KW": "Daily_KW"}, inplace=True)

    # Identify days where the daily consumption exceeds the threshold
    outliers = daily_consumption[daily_consumption["Daily_KW"] > threshold]

    # Return results only if outliers are detected
    if not outliers.empty:
        print(f"Number of days exceeding {threshold} KW: {len(outliers)}")
        return df, outliers

    # Print message if no outliers are found
    print("No outliers detected.")
    return df
```

Σχήμα 3.4: Βοηθητική συνάρτηση διαχείρισης ακραίων τιμών

3.4.4 Φιλτράρισμα βάσει πληρότητας

Στο τελικό στάδιο επεξεργασίας των δεδομένων, εφαρμόζεται μια διαδικασία φιλτραρίσματος βάσει προκαθορισμένων ορίων πληρότητας, με σκοπό την επιλογή μόνο των καταναλωτών που διαθέτουν επαρκή χρονική κάλυψη στο σύνολο των μετρήσεων. Η

προσέγγιση αυτή διασφαλίζει ότι οι τελικές αναλύσεις θα βασίζονται σε πλήρη και συμπαγή δεδομένα. Η διαδικασία περιλαμβάνει διαδοχικούς ελέγχους σε τρία επίπεδα:

- Σε **ημερήσια βάση**, επιβάλλεται όριο 20 ωρών ανά ημέρα τουλάχιστον. Καταναλωτές με ημιτελή ημερήσια καταγραφή απορρίπτονται, καθώς οι μετρήσεις τους θεωρούνται ανεπαρκείς για περαιτέρω ανάλυση.
- Σε **μηνιαία βάση**, ελέγχεται αν κάθε χρήστης διαθέτει δεδομένα για τουλάχιστον 28 ημέρες ανά μήνα. Όσοι καταναλωτές δεν πληρούν αυτό το κριτήριο, αποκλείονται από το σύνολο για τον αντίστοιχο μήνα.
- Σε **ετήσια βάση**, διατηρούνται μόνο όσοι καταναλωτές πληρούν τα παραπάνω κριτήρια για όλους τους μήνες του έτους. Με αυτόν τον τρόπο, επιλέγεται ένα σύνολο χρηστών με συνεπή παρουσία σε όλη τη διάρκεια της περιόδου μελέτης.

Τα συγκεκριμένα κατώφλια επιλέγονται με σκοπό να επιτευχθεί μια ισορροπία μεταξύ ποιότητας και πληρότητας του συνόλου δεδομένων. Το όριο των 28 ημερών τον μήνα επιτρέπει μικρές απώλειες χωρίς να αλλοιώνεται η συνολική εικόνα κατανάλωσης, ενώ τα 20 ωριαία σημεία ανά ημέρα καλύπτουν περίπου το 80% των μετρήσεων ημερησίως. Η απαίτηση για 12 μήνες συμμετοχής διασφαλίζει τη σύγκριση μεταξύ πλήρων ετήσιων προφίλ. Σε κάθε στάδιο, οι καταναλωτές που δεν ικανοποιούν τα αντίστοιχα κριτήρια αποκλείονται από την ανάλυση. Το τελικό σύνολο δεδομένων που προκύπτει από αυτή τη διαδικασία, περιλαμβάνει μόνο εκείνους τους καταναλωτές των οποίων οι χρονοσειρές είναι συνεχείς, πλήρεις και αξιόπιστες. Η προσέγγιση αυτή βελτιστοποιεί την ποιότητα των αποτελεσμάτων και ενισχύει την εγκυρότητα των αποτελεσμάτων που θα εξαχθούν στα επόμενα στάδια. Τα Σχήματα 3.5, 3.6 και 3.7, παρουσιάζουν τον κώδικα των βοηθητικών συναρτήσεων που χρησιμοποιήθηκαν σε αυτό το στάδιο της μεθοδολογίας.

```
# FUNCTION 4

def missing_hours(df, threshold, eligible_list):
    # Filter only valid values (values > 0)
    valid_df = df[df["KW"] > 0]

    # Group by consumer and day, count valid hourly entries
    valid_hours_per_day = valid_df.groupby(["LCLid", "Day"])["Hour"].nunique()

    # Identify consumers meeting the minimum valid hours threshold
    valid_hours = valid_hours_per_day.groupby("LCLid").min()
    eligible_ids = valid_hours[valid_hours >= threshold].index

    # Update the eligible consumers list
    eligible_list = list(set(eligible_list).union(set(eligible_ids)))

    # Filter the DataFrame
    df = df[df["LCLid"].isin(eligible_ids)].reset_index(drop=True)

    return df, eligible_list
```

Σχήμα 3.5: Βοηθητική συνάρτηση με φίλτρο σε ημερήσια βάση

```
# FUNCTION 3

def missing_days(df, threshold, eligible_list):
    # Filter only valid values (values > 0)
    valid_df = df[df["KW"] > 0]

    # Calculate days with valid data per consumer
    days_per_consumer = valid_df.groupby("LCLid")["Day"].nunique()

    # Filter consumers with at least the threshold number of days
    eligible_ids = days_per_consumer[days_per_consumer >= threshold].index

    # Update the eligible consumers list
    eligible_list = list(set(eligible_list).union(set(eligible_ids)))

    # Filter the DataFrame
    df = df[df["LCLid"].isin(eligible_ids)]

    return df, eligible_list
```

Σχήμα 3.6: Βοηθητική συνάρτηση με φίλτρο σε μηνιαία βάση

```
# FUNCTION 5

def missing_months(df, threshold, eligible_list):
    # Filter only valid values (values > 0)
    valid_df = df[df["KW"] > 0]

    # Count unique months with valid data per consumer
    monthly_counts = valid_df.groupby("LCLid")["Month"].nunique()

    # Filter consumers with at least the threshold number of months
    eligible_ids = monthly_counts[monthly_counts >= threshold].index

    # Update the eligible consumers list
    eligible_list = list(set(eligible_list).intersection(set(eligible_ids)))

    # Filter the DataFrame
    df = df[df["LCLid"].isin(eligible_list)].reset_index(drop=True)

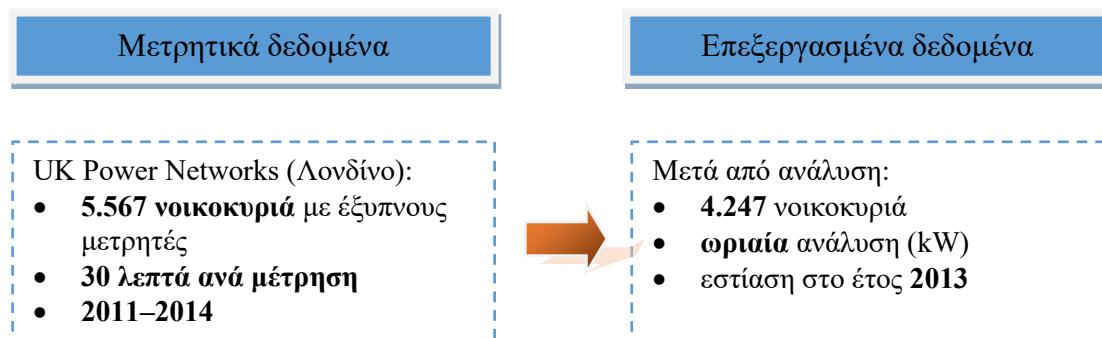
    return df, eligible_list
```

Σχήμα 3.7: Βοηθητική συνάρτηση με φίλτρο σε ετήσια βάση

3.4.5 Εξαγωγή τελικής μορφής συνόλου δεδομένων

Εφόσον ολοκληρωθούν όλα τα προηγούμενα στάδια, συγκεντρώνεται το τελικό σύνολο δεδομένων που ικανοποιεί τα προκαθορισμένα ποιοτικά κριτήρια. Λόγω του ιδιαίτερα μεγάλου όγκου των αρχικών δεδομένων, τα οποία ξεπερνούσαν τις 167 εκατομμύρια εγγραφές, η αρχική επεξεργασία δεν ήταν εφικτό να πραγματοποιηθεί συνολικά. Για τον λόγο αυτό, ακολουθήθηκε σταδιακή προσέγγιση ανά μήνα. Στο τελικό στάδιο, τα επεξεργασμένα υποσύνολα συγχωνεύονται και δημιουργείται ένας ενιαίος πίνακας, ο οποίος αποθηκεύεται σε μορφή αρχείου CSV. Το τελικό αρχείο περιλαμβάνει περίπου 37

εκατομμύρια εγγραφές και προέρχεται από **4.247 καταναλωτές** που πληρούν στο σύνολό τους τα κριτήρια πληρότητας για το έτος **2013**. Η σημαντική μείωση του όγκου δεδομένων, βελτιστοποιεί τη διαχείριση του συνόλου δεδομένων και επιτρέπει την αποτελεσματική εκτέλεση των επόμενων σταδίων ανάλυσης, με μικρότερες απαιτήσεις σε μνήμη και χρόνο υπολογισμού. Ως αποτέλεσμα, το τελικό σύνολο δεδομένων είναι πλέον συμπαγές, καθαρό και έτοιμο προς αξιοποίηση για τη μετέπειτα διαδικασία ομαδοποίησης και μελέτης των καταναλωτικών προφίλ. Στο Σχήμα 3.8 παρουσιάζονται συνοπτικά τα αρχικά και τα επεξεργασμένα δεδομένα.



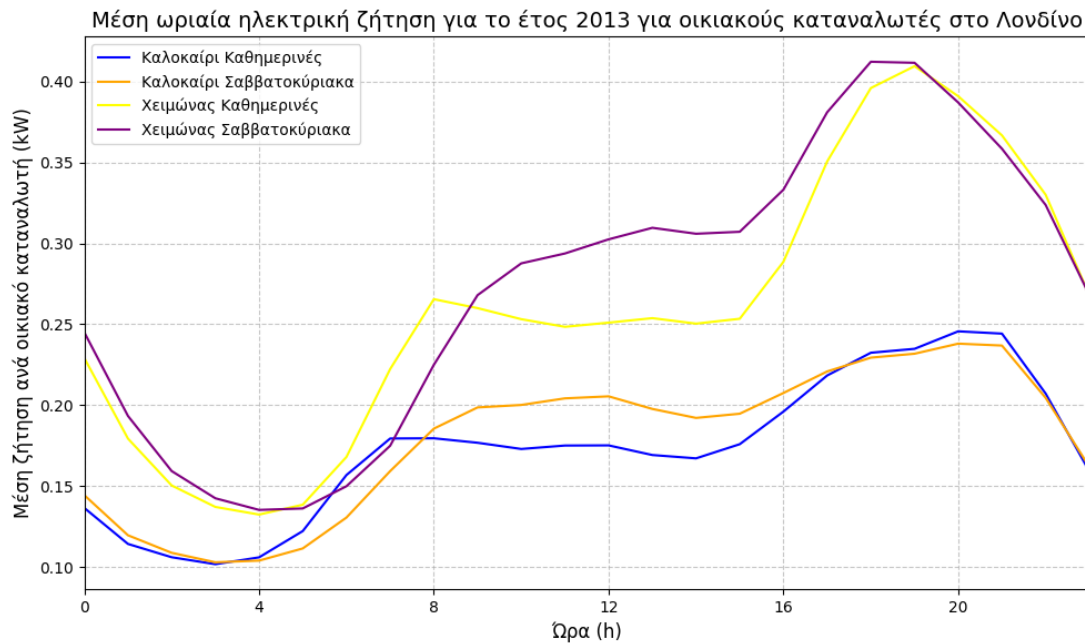
Σχήμα 3.8: Ανάλυση δεδομένων

3.5 ΕΠΙΚΑΙΡΟΤΗΤΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΣΕ ΣΥΓΚΡΙΣΗ ΜΕ ΤΟ ΣΗΜΕΡΑ

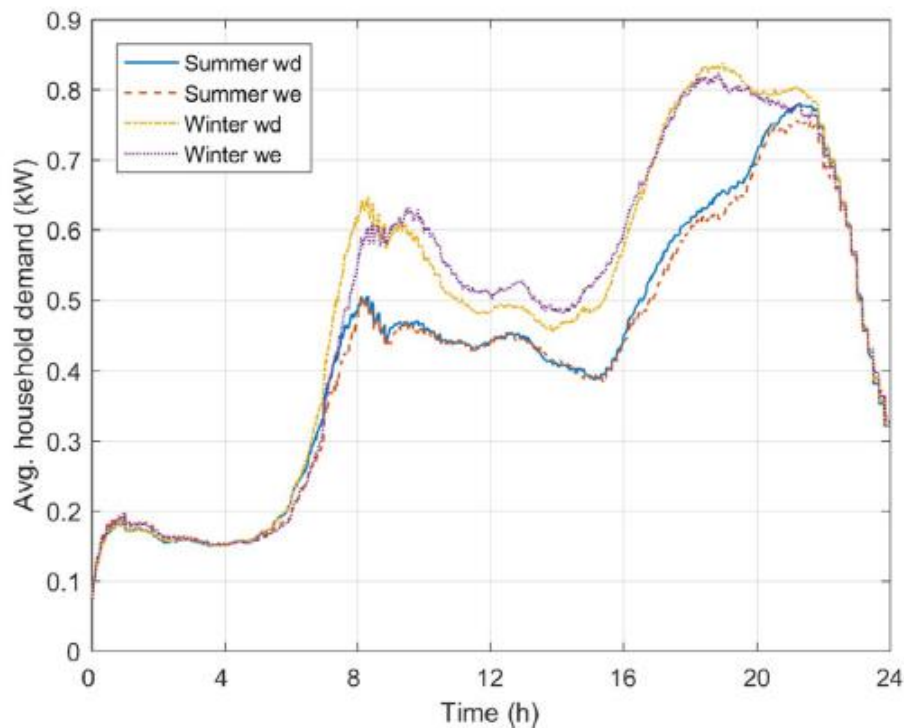
Η παραπάνω ανάλυση βασίζεται σε δεδομένα κατανάλωσης ηλεκτρικής ενέργειας του έτους 2013. Αν και τα δεδομένα αυτά δεν είναι πρόσφατα, επιλέγονται για ανάλυση, καθώς αποτελούν το πιο πρόσφατο, πλήρες και αξιόπιστο διαθέσιμο σύνολο δεδομένων. Η απουσία πιο πρόσφατου συνόλου με αντίστοιχη ποιότητα και πληρότητα, δικαιολογεί την αξιοποίησή τους για τη διερεύνηση προτύπων κατανάλωσης και την εφαρμογή μεθοδολογιών ομαδοποίησης.

Προκειμένου να αξιολογηθεί η ερμηνεία των προτύπων κατανάλωσης του 2013 με τη σημερινή καταναλωτική συμπεριφορά, παρουσιάζεται συγκριτική ανάλυση με πιο πρόσφατα διαθέσιμα στοιχεία. Συγκεκριμένα, γίνεται σύγκριση με ημερήσιες καμπύλες κατανάλωσης για το έτος 2018, οι οποίες έχουν δημοσιευτεί στη βιβλιογραφία και αφορούν οικιακούς καταναλωτές στο Λονδίνο. Στο Σχήμα 3.9 απεικονίζεται η μέση ημερήσια καμπύλη κατανάλωσης για το έτος 2013, ανά εποχή και τύπο ημέρας (καθημερινές/Σαββατοκύριακα). Διακρίνονται δύο ζώνες αιχμής: μία κατά τις πρωινές ώρες (περίπου 7:00–9:00) και μία εντονότερη κατά τις βραδινές (περίπου 17:00–22:00), με υψηλότερες τιμές τη χειμερινή περίοδο. Αντίστοιχα, στο Σχήμα 3.10 παρουσιάζεται η καμπύλη κατανάλωσης του 2018 [17]. Παρατηρείται διατήρηση του ίδιου μοτίβου αιχμών, ωστόσο καταγράφονται αυξημένες τιμές ζήτησης και μεγαλύτερη καταναλωτική δραστηριότητα σε όλη τη διάρκεια της ημέρας. Οι διαφοροποιήσεις αυτές αποδίδονται πιθανώς στη διείσδυση νέων τεχνολογιών, στη χρήση περισσότερων ηλεκτρικών φορτίων και στην εν γένει εξέλιξη των προτύπων διαβίωσης, όπως για παράδειγμα η αυξημένη εξάρτηση από την ηλεκτρική ενέργεια. Συμπερασματικά, αν και τα δεδομένα του 2013 παρουσιάζουν ορισμένες ποσοτικές αποκλίσεις σε σχέση με τα νεότερα, τα ποιοτικά χαρακτηριστικά της ημερήσιας κατανάλωσης παραμένουν σε μεγάλο

βαθμό αντίστοιχα. Το γεγονός αυτό ενισχύει την αξιοπιστία των αποτελεσμάτων που εξάγονται από το επιλεγμένο σύνολο δεδομένων.



Σχήμα 3.9: Μέση ημερήσια καμπύλη κατανάλωσης για το έτος 2013 για οικιακούς καταναλωτές στο Λονδίνο



Σχήμα 3.10: Μέση ημερήσια καμπύλη κατανάλωσης για έτος 2018 για οικιακούς καταναλωτές στο Λονδίνο [17]

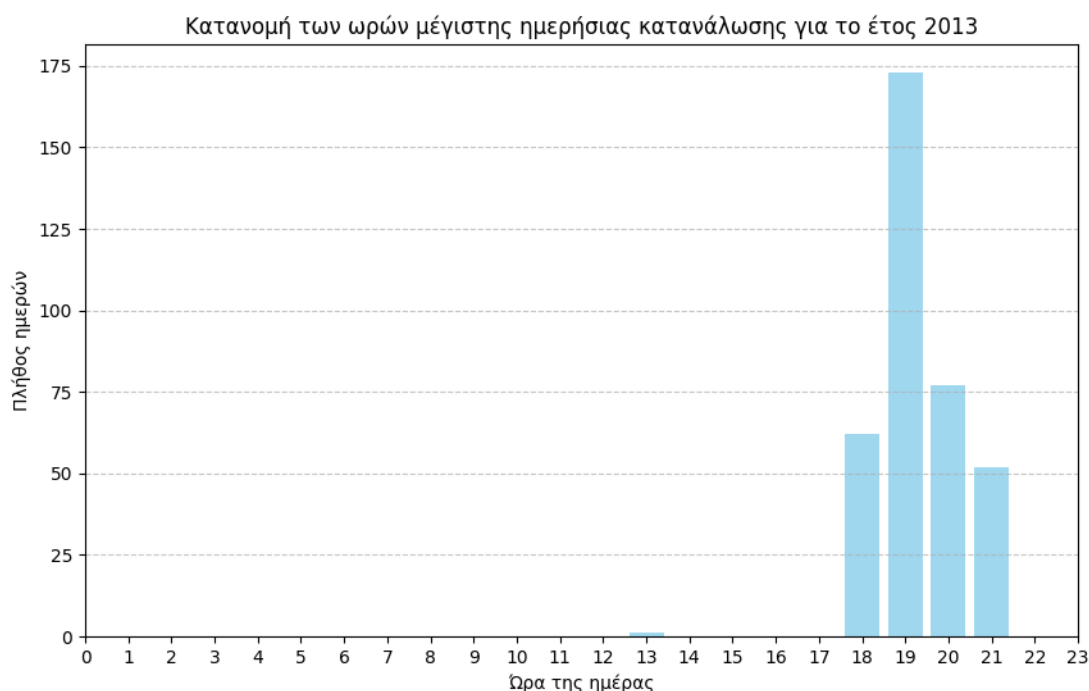
ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΟΜΑΔΟΠΟΙΗΣΗΣ ΟΙΚΙΑΚΩΝ ΚΑΤΑΝΑΛΩΤΩΝ

4.1 ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΕΙΣΟΔΟΥ

Η ομαδοποίηση οικιακών καταναλωτών αποτελεί μια τεχνική που επιτρέπει την ανάλυση των προτύπων κατανάλωσης σε επιμέρους ομάδες με κοινά χαρακτηριστικά. Μέσω της ομαδοποίησης, μπορούν να εντοπιστούν κατηγορίες χρηστών με παρόμοια συμπεριφορά ως προς τη χρήση ηλεκτρικής ενέργειας. Η τεχνική ανάλυσης συστάδων μπορεί να αξιοποιηθεί για στοχευμένα σχήματα διαχείρισης ζήτησης, προγράμματα ευελιξίας ή και δυναμική τιμολόγηση.

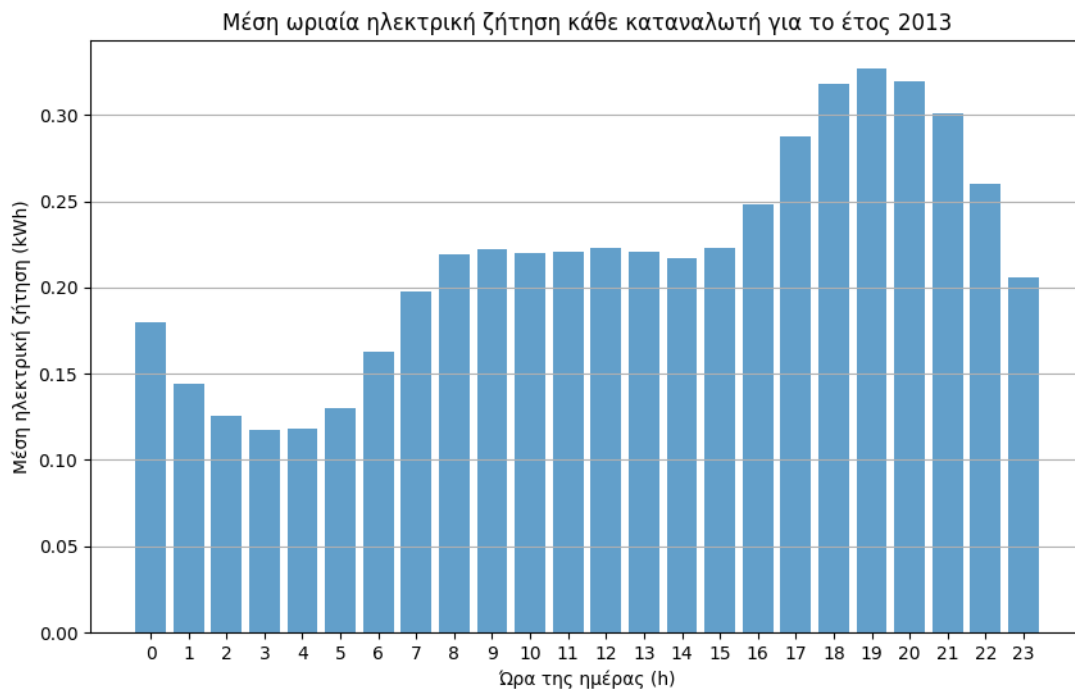
4.1.1 Συμπεριφορά οικιακών καταναλωτών

Για την εφαρμογή μεθόδων ομαδοποίησης απαιτείται, σε πρώτο στάδιο, η εξαγωγή χαρακτηριστικών, τα οποία θα εισαχθούν στο μοντέλο ως βάση για την κατηγοριοποίηση των καταναλωτών. Τα χαρακτηριστικά αυτά συνιστούν το βασικό κριτήριο πάνω στο οποίο θα βασιστεί η διαδικασία του αλγορίθμου για να εντοπίσει πρότυπα και ομοιότητες μεταξύ διαφορετικών καταναλωτικών προφίλ. Επομένως, η επιλογή τους πρέπει να είναι στοχευμένη και να βασίζεται σε κριτήρια που δίνουν ξεχωριστή και ουσιώδη πληροφορία κατά τη μελέτη συμπεριφοράς. Στο πλαίσιο της παρούσας διπλωματικής εργασίας, η ομαδοποίηση επικεντρώνεται στην κατανάλωση ηλεκτρικής ενέργειας κατά τις ώρες αιχμής, όπου η υψηλή κατανάλωση επιβαρύνει σημαντικά το δίκτυο. Ως πρώτο βήμα, είναι απαραίτητο να προσδιοριστούν οι ώρες αιχμής του υπό μελέτη συνόλου δεδομένων, δηλαδή οι χρονικές περίοδοι κατά τις οποίες οι καταναλωτές παρουσιάζουν την υψηλότερη τιμή της ημερήσιας κατανάλωσής τους. Στο Σχήμα 4.1 απεικονίζεται η κατανομή των ωρών αιχμής της αθροιστικής ημερήσιας κατανάλωσης όλων των καταναλωτών για το έτος 2013, η οποία εκφράζει τον αριθμό των ημερών κατά τις οποίες κάθε ώρα της ημέρας αναγνωρίστηκε ως η ώρα μέγιστης ζήτησης. Χρησιμοποιήθηκε η βάση δεδομένων “SmartMeter Energy Use Data in London Households” που περιγράφηκε στο Κεφάλαιο 3 [13]. Από το Σχήμα 4.1 προκύπτει ότι οι ώρες αιχμής εντοπίζονται κυρίως κατά τις απογευματινές και βραδινές ώρες. Το φαινόμενο αυτό συνδέεται με την επιστροφή των καταναλωτών στις οικίες τους μετά την εργασία ή τις καθημερινές τους δραστηριότητες και αντιστοιχεί στη συνηθισμένη καθημερινή ρουτίνα. Οι πληροφορίες αυτές κρίνονται ιδιαίτερα χρήσιμες, καθώς η χρονική ζώνη μεταξύ 18:00 και 22:00 επιλέγεται ως περίοδος αιχμής για την παρούσα ανάλυση. Οι δείκτες κατανάλωσης που θα εξαχθούν από το χρονικό αυτό εύρος θα αποτελέσουν τα χαρακτηριστικά εισόδου για το μοντέλο ομαδοποίησης.

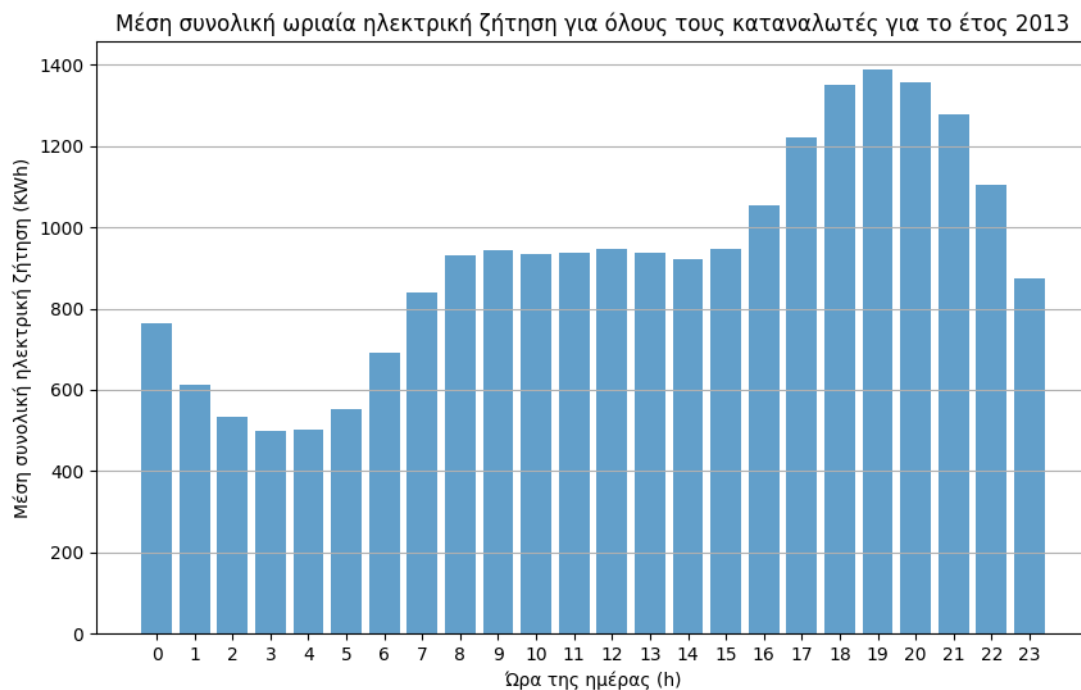


Σχήμα 4.1: Κατανομή των ωρών μέγιστης ημερήσιας κατανάλωσης για το έτος 2013, για το αθροιστικό φορτίο όλων των καταναλωτών

Προκειμένου να αποκτηθεί μία πλήρης εικόνα για τη συνολική καταναλωτική συμπεριφορά των νοικοκυριών, εξετάζονται δύο βασικές καμπύλες ζήτησης για το έτος 2013. Η πρώτη παρουσιάζεται στο Σχήμα 4.2 και απεικονίζει τη μέση ωριαία κατανάλωση ανά καταναλωτή, εκφρασμένη ως ποσοστό της μέσης ημερήσιας κατανάλωσης. Η δεύτερη καμπύλη παρουσιάζεται στο Σχήμα 4.3 και αφορά τη μέση ωριαία ηλεκτρική ζήτηση των καταναλωτών για το έτος 2013. Συνολικά, οι δύο καμπύλες καταδεικνύουν ένα σταθερό και επαναλαμβανόμενο μοτίβο ζήτησης με αιχμές τις απογευματινές και βραδινές ώρες. Η ανάλυση αυτή ενισχύει την κατανόηση της χρονικής κατανομής της ζήτησης και επιβεβαιώνει τη σκοπιμότητα εστίασης της ομαδοποίησης στις ώρες 18:00 έως 22:00, όπου οι διακυμάνσεις της κατανάλωσης αποκτούν ιδιαίτερη σημασία για τη διαχείριση του φορτίου. Όλες οι καμπύλες του μοντέλου, δημιουργούνται με χρήση της γλώσσας προγραμματισμού Python, αξιοποιώντας τη βιβλιοθήκη Matplotlib, η οποία παρέχει εργαλεία για την παραγωγή διαγραμμάτων υψηλής ακρίβειας και την αποτελεσματική οπτικοποίηση δεδομένων [18].



Σχήμα 4.2: Μέση ωριαία κατανάλωση για το έτος 2013 ανά καταναλωτή, συμπεριλαμβάνοντας όλο το πλήθος των καταναλωτών ανεξαρτήτως κατανάλωσης



Σχήμα 4.3: Μέση ωριαία συνολική κατανάλωση για το έτος 2013, για όλους τους καταναλωτές (πλήθος καταναλωτών = 4.247).

4.1.2 Χαρακτηριστικά εισόδου

Οι μετρικές που υπολογίζονται, αποτυπώνουν τόσο την ποσότητα, όσο και τη συχνότητα κατανάλωσης, καθώς και τη χρονική της μεταβλητότητα, παρέχοντας πολύπλευρη εικόνα του προφίλ κάθε νοικοκυριού. Αναλυτικά, τα χαρακτηριστικά που εξάγονται για κάθε καταναλωτή είναι τα εξής:

- **Πλήθος ημερών με κατανάλωση μεγαλύτερη από 3 kWh κατά τις ώρες αιχμής:** Το χαρακτηριστικό αυτό καταγράφει τον αριθμό των ημερών όπου η κατανάλωση υπερβαίνει τις 3 kWh κατά τις ώρες αιχμής. Αποτελεί δείκτη υψηλής κατανάλωσης κατά τις κρίσιμες χρονικές περιόδους και σχετίζεται με την επιβάρυνση που προκαλείται στο δίκτυο.
- **Πλήθος ημερών με κατανάλωση μεγαλύτερη από 1 kWh κατά τις ώρες αιχμής:** Η μεταβλητή αυτή εκφράζει τη συχνότητα με την οποία ένας καταναλωτής παρουσιάζει ενεργή καταναλωτική δραστηριότητα στις ώρες αιχμής, ακόμη και αν η κατανάλωση δεν είναι ιδιαίτερα υψηλή.
- **Ποσοστό μέσης κατανάλωσης τις ώρες αιχμής ως προς τη μέση συνολική ημερήσια κατανάλωση:** Εκφράζει τη σχετική συμβολή των ωρών αιχμής στη συνολική ημερήσια κατανάλωση, παρέχοντας ένδειξη για το κατά πόσο ο καταναλωτής συγκεντρώνει τη ζήτησή του στις περιόδους αιχμών.
- **Μέση τιμή μεγίστων κατανάλωσης κατά τις ώρες αιχμής:** Αντιπροσωπεύει τη μέση τιμή των μεγίστων ωριαίας κατανάλωσης εντός των ωρών αιχμής και σχετίζεται με την ένταση των στιγμιαίων φορτίων κάθε νοικοκυριού.
- **Διακύμανση και εύρος τιμών κατανάλωσης εντός της περιόδου αιχμών:** Οι δείκτες αυτοί καταγράφουν τη σταθερότητα ή μεταβλητότητα του καταναλωτικού προφίλ, προσφέροντας πληροφορία για την ομοιομορφία ή τις αποκλίσεις της κατανάλωσης εντός της περιόδου αιχμών.
- **Αυτοσυσχέτιση:** Εξετάζει τον βαθμό επαναληψιμότητας της καταναλωτικής συμπεριφοράς από ημέρα σε ημέρα. Υψηλή αυτοσυσχέτιση υποδηλώνει σταθερά και προβλέψιμα πρότυπα.
- **Μέση σχετική τυπική απόκλιση:** Αποτελεί κανονικοποιημένο μέτρο της διασποράς της κατανάλωσης, επιτρέποντας τη σύγκριση της μεταβλητότητας μεταξύ καταναλωτών με διαφορετικά επίπεδα ζήτησης.
- **Καθημερινές και Σαββατοκύριακα:** Κάθε χαρακτηριστικό υπολογίζεται διακριτά για Καθημερινές και για Σαββατοκύριακα, προκειμένου να ληφθεί υπόψη η διαφορά στην καταναλωτική συμπεριφορά μεταξύ εργασιμων και μη εργασιμων ημερών.

Αξίζει να σημειωθεί ότι, εφόσον ένας καταναλωτής πληροί το πρώτο κριτήριο (κατανάλωση μεγαλύτερη από 3 kWh κατά τις ώρες αιχμής), πληροί αναγκαστικά και το δεύτερο (κατανάλωση μεγαλύτερη από 1 kWh κατά τις ώρες αιχμής), καθώς το δεύτερο κριτήριο είναι υποσύνολο του πρώτου κριτηρίου. Ωστόσο, η επιλογή των δύο αυτών κριτηρίων εξυπηρετεί διαφορετικούς σκοπούς: το κατώφλι των 3 kWh αποτελεί δείκτη υψηλής κατανάλωσης σε συνθήκες αιχμής και σχετίζεται με την επιβάρυνση του δικτύου, ενώ το κατώφλι της 1 kWh αποτυπώνει τη συχνότητα με την οποία ο καταναλωτής παρουσιάζει ενεργή παρουσία στο χρονικό διάστημα αιχμής, ανεξαρτήτως της έντασης της κατανάλωσης. Συνεπώς, δεν αναμένεται υψηλή αλληλοσυσχέτιση μεταξύ των δύο κριτηρίων (μεταβλητών), καθώς ένας καταναλωτής μπορεί να εμφανίζει συχνή αλλά ήπια κατανάλωση

(πληρώντας μόνο το δεύτερο κριτήριο). Αντίθετα, η ικανοποίηση του πρώτου κριτηρίου υποδηλώνει ταυτόχρονη ικανοποίηση και του δεύτερου.

Η παραπάνω επιλογή χαρακτηριστικών αποσκοπεί στην ποσοτική και ποιοτική κατανόηση καταναλωτικών συμπεριφορών κατά τις ώρες αιχμής, υποστηρίζοντας έτσι τη δημιουργία διαχωρίσιμων ομάδων καταναλωτών μέσω της ομαδοποίησης. Τα παραπάνω χαρακτηριστικά παρουσιάζονται συνοπτικά στον Πίνακα 4.1.

Πίνακας 4.1: Περιγραφή των χαρακτηριστικών εισόδου για το μοντέλο ομαδοποίησης

Χαρακτηριστικό εισόδου	Περιγραφή	Μαθηματική διατύπωση
Ημέρες με > 3 kWh	Πλήθος ημερών με κατανάλωση > 3 kWh κατά τις ώρες 18:00–22:00	$N_3 = \{d \in [1, D]: \sum_{t=18}^{21} E_{d,t} > 3\} $
Ημέρες με > 1 kWh	Πλήθος ημερών με κατανάλωση > 1 kWh κατά τις ώρες 18:00–22:00	$N_1 = \{d \in [1, D]: \sum_{t=18}^{21} E_{d,t} > 1\} $
Συνεισφορά αιχμής στη μέση ημερήσια κατανάλωση	Μέση κατανάλωση της περιόδου αιχμής ως ποσοστό της μέσης συνολικής ημερήσιας κατανάλωσης	$P = \frac{\sum_{t=18}^{21} E_{d,t}}{\sum_{t=0}^{23} E_{d,t}} \times 100\%$
Μέγιστα κατανάλωσης	Μέση τιμή μέγιστων ωριαίων τιμών κατά τις ώρες 18:00–22:00	$M = \frac{1}{ D } \sum_{d \in D} \max_{t \in T_{peak}} E_{d,t}$
Διακύμανση και εύρος τιμών	Στατιστική μεταβλητότητα κατά τις ώρες 18:00–22:00	$\sigma^2 = Var(\{E_{d,t}: t \in T_{peak}\})$ $R = \max_{t \in T_{peak}} E_{d,t} - \min_{t \in T_{peak}} E_{d,t}$
Αυτοσυσχέτιση	Συσχέτιση καταναλωτικής συμπεριφοράς μεταξύ διαδοχικών ημερών κατά τις ώρες 18:00–22:00	$\rho = Corr(\{E_{d,t}, E_{d,t-1}: t \in T_{peak}\})$
Μέση σχετική τυπική απόκλιση	Σύγκριση της σχετικής μεταβλητότητας των προφίλ μεταξύ καταναλωτών με διαφορετικά επίπεδα κατανάλωσης κατά τις ώρες 18:00–22:00	$MRSD = \frac{\sigma(\sum_{t=18}^{21} E_{d,t})}{\mu(\sum_{t=18}^{21} E_{d,t})}$
Καθημερινές και Σαββατοκύριακα	Διάκριση κάθε χαρακτηριστικού εισόδου μεταξύ καθημερινών και σαββατοκύριακου, με ξεχωριστό υπολογισμό για τις δύο κατηγορίες, κατά τις ώρες 18:00–22:00	$N_3^{we,wd}, N_3^{we,wd}, P^{we,wd}, M^{we,wd}, \sigma^{2we,wd}, R^{we,wd}, \rho^{we,wd}, MRSD^{we,wd}$

Οι μεταβλητές και σύμβολα που χρησιμοποιούνται στις μαθηματικές διατυπώσεις των χαρακτηριστικών ορίζονται ως εξής:

- $d \in D$: δείκτης ημέρας εντός του χρονικού ορίζοντα ανάλυσης (π.χ. 1–365 για το έτος 2013)
- $t \in T$: δείκτης ώρας της ημέρας (0–23)
- $T_{peak} = \{18,19,20,21\}$: σύνολο ωρών αιχμής
- $E_{d,t}$: τιμή κατανάλωσης ενέργειας (kWh) την ημέρα d και ώρα t
- N_3, N_1 : αριθμός ημερών με συνολική κατανάλωση στις ώρες αιχμής μεγαλύτερη από 3 kWh και 1 kWh, αντίστοιχα
- P : ποσοστό της ημερήσιας κατανάλωσης που αντιστοιχεί στις ώρες αιχμής
- M : μέση τιμή των ημερήσιων μέγιστων τιμών κατανάλωσης κατά τις ώρες αιχμής
- σ^2, R : διακύμανση και εύρος κατανάλωσης κατά τις ώρες αιχμής
- ρ : συντελεστής αυτοσυσχέτισης εντός των ωρών αιχμής
- μ, σ : μέση τιμή και τυπική απόκλιση
- $MRSD$: μέση σχετική τυπική απόκλιση
- Οι δείκτες wd και we αντιστοιχούν σε καθημερινές (weekdays) και Σαββατοκύριακα (weekends), αντίστοιχα

4.2 ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΕΙΣΟΔΟΥ

Για την εφαρμογή του αλγορίθμου ομαδοποίησης, καθίσταται απαραίτητη η μετατροπή όλων των χαρακτηριστικών εισόδου στην ίδια αριθμητική κλίμακα. Τα δεδομένα που προκύπτουν από διαφορετικές μονάδες μέτρησης ή υπολογισμούς (όπως ποσοστά, αριθμός ημερών ή τιμές ενέργειας σε kWh) ενδέχεται να παρουσιάζουν σημαντικές διαφοροποιήσεις ως προς το εύρος τιμών. Η ανομοιογένεια αυτή μπορεί να επηρεάσει τη διαδικασία, καθώς χαρακτηριστικά με μεγαλύτερα αριθμητικά μεγέθη τείνουν να έχουν δυσανάλογη επίδραση στη λειτουργία του αλγορίθμου, οδηγώντας σε μη αντιπροσωπευτικά αποτελέσματα. Για την αντιμετώπιση του προβλήματος αυτού, εφαρμόζεται κανονικοποίηση τύπου Min-Max [19], η οποία μετασχηματίζει τις τιμές κάθε χαρακτηριστικού, ώστε να κινούνται εντός του επιθυμητού εύρους τιμών [0,1]. Ο μετασχηματισμός πραγματοποιείται σύμφωνα με την παρακάτω μαθηματική εξίσωση (4.1) [19]:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4.1)$$

όπου:

X : η αρχική τιμή του χαρακτηριστικού

X_{min} : η ελάχιστη τιμή του χαρακτηριστικού X

X_{max} : η μέγιστη τιμή του χαρακτηριστικού X

Η μέθοδος αυτή διατηρεί τις σχετικές αποστάσεις μεταξύ των αρχικών τιμών και είναι ιδιαίτερα χρήσιμη σε αλγορίθμους όπως ο K-Means, στους οποίους η απόσταση μεταξύ των παρατηρήσεων αποτελεί βασικό μηχανισμό λειτουργίας. Τα πλεονεκτήματα της κανονικοποίησης Min-Max είναι τα εξής:

- **Απλότητα και ταχύτητα υπολογισμού**, με εύκολη κατανόηση και εφαρμογή ακόμη και σε μεγάλα σύνολα δεδομένων.
- **Διατήρηση των αναλογιών των αρχικών τιμών**, χωρίς παραμόρφωση των σχέσεων μεταξύ των δεδομένων.
- **Κατάλληλη για αλγορίθμους ευαίσθητους στην κλίμακα**, όπως ο K-Means.

Συνοψίζοντας, η κανονικοποίηση Min-Max αποτελεί βασικό εργαλείο στην προεπεξεργασία των δεδομένων, ιδιαίτερα χρήσιμο σε περιπτώσεις όπου απαιτείται ενιαία αριθμητική κλίμακα μεταξύ των χαρακτηριστικών. Με αυτό τον τρόπο συμβάλλει στη βελτίωση της ακρίβειας και της ερμηνευσιμότητας των αποτελεσμάτων.

4.3 ΠΕΡΙΓΡΑΦΗ ΑΛΓΟΡΙΘΜΟΥ K-MEANS

Ο αλγόριθμος K-Means αποτελεί μια από τις πιο διαδεδομένες τεχνικές μη επιβλεπόμενης μάθησης για την ανάλυση συστάδων. Στόχος του είναι η ομαδοποίηση ενός συνόλου δεδομένων σε συγκεκριμένο αριθμό συστάδων, με βάση την ομοιότητα των παρατηρήσεων. Ο αλγόριθμος αυτός είναι ιδιαίτερα χρήσιμος σε εφαρμογές, όπως η ανάλυση καταναλωτικής συμπεριφοράς, η τμηματοποίηση πελατών, η ομαδοποίηση χρονοσειρών με δεδομένα από έξυπνους μετρητές. Κύριος στόχος του K-Means είναι η ελαχιστοποίηση της ενδοσυσταδικής διασποράς, δηλαδή η μείωση της συνολικής απόστασης των σημείων από τα κέντρα των συστάδων τους [20],[21]. Η αντικειμενική συνάρτηση που ελαχιστοποιείται εκφράζεται από την εξίσωση (4.2) [20]:

$$J = \sum_{j=1}^k \sum_{i=1}^c (\|x_i - c_j\|)^2 \quad (4.2)$$

όπου:

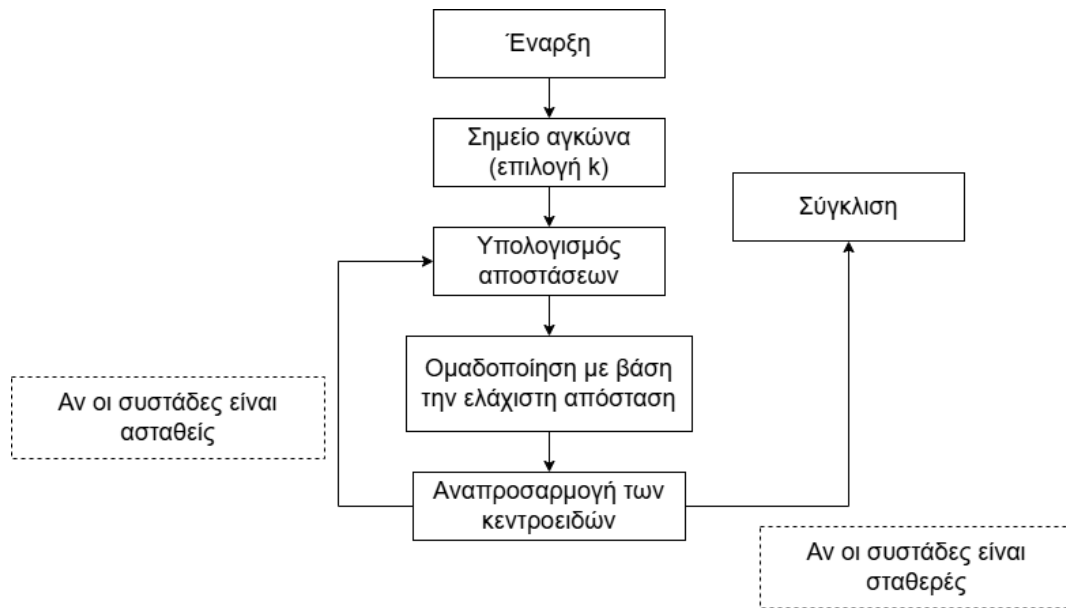
x_i : το σημείο της παρατήρησης i

c_j : το κέντρο της συστάδας j

k : ο αριθμός συστάδων

$\|x_i - c_j\|$: η Ευκλείδεια απόσταση μεταξύ σημείου και κέντρου

Η λειτουργία του αλγορίθμου βασίζεται σε επαναληπτική διαδικασία που περιλαμβάνει αρχικά τον καθορισμό του αριθμού συστάδων k , είτε εμπειρικά, είτε με χρήση μεθόδων, όπως η Elbow method ή η Silhouette Score, οι οποίες περιγράφονται αναλυτικά στην Ενότητα 4.4. Στη συνέχεια, επιλέγονται τυχαία k αρχικά σημεία ως κέντρα και για κάθε σημείο του συνόλου δεδομένων υπολογίζεται η Ευκλείδεια απόστασή του από όλα τα κέντρα. Κάθε σημείο ανατίθεται στο πλησιέστερο κέντρο και τα κέντρα επανυπολογίζονται ως ο μέσος όρος των σημείων της συστάδας. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να επιτευχθεί σύγκλιση, δηλαδή έως ότου τα κέντρα σταθεροποιηθούν, δηλαδή να μην αλλάζουν ή η μεταβολή τους να είναι πολύ μικρή [20]. Η συνολική διαδικασία απεικονίζεται στο Σχήμα 4.4.



Σχήμα 4.4: Διάγραμμα ροής του αλγορίθμου K-Means

Ο αλγόριθμος K-Means συγκεντρώνει πλήθος πλεονεκτημάτων που δικαιολογούν την ευρεία του χρήση. Πρόκειται για μια υπολογιστικά αποδοτική και εύκολα υλοποιήσιμη μέθοδο, η οποία ανταποκρίνεται ικανοποιητικά ακόμη και σε δεδομένα υψηλής διαστασιμότητας ή μεγάλου όγκου. Χάρη στην απλότητά του και την ταχύτητα εκτέλεσής του, ο αλγόριθμος αυτός εφαρμόζεται επιτυχώς σε εφαρμογές πραγματικού χρόνου και επιτρέπει την επεκτασιμότητα σε ακόμη μεγαλύτερα σύνολα. Επιπλέον, προσφέρει ευελιξία ως προς τη χρήση διαφορετικών μετρικών απόστασης και στρατηγικών αρχικοποίησης, ενώ επιδεικνύει καλή προσαρμοστικότητα σε δεδομένα με γεωμετρική ή χρονική δομή. Η ικανότητά του να διατηρεί πληροφορίες σχετικές με τη μορφή των δεδομένων καθιστά τον K-Means ιδιαίτερα κατάλληλο για την ανάλυση χρονοσειρών και την εξαγωγή προφίλ κατανάλωσης ηλεκτρικής ενέργειας.

Παρότι ιδιαίτερα αποτελεσματικός, ο αλγόριθμος εμφανίζει και ορισμένους περιορισμούς. Η ποιότητα της ομαδοποίησης εξαρτάται σε σημαντικό βαθμό από την αρχική επιλογή των κέντρων, καθώς ενδέχεται να οδηγήσει σε τοπικά ελάχιστα της αντικειμενικής συνάρτησης. Επίσης, απαιτείται εκ των προτέρων ορισμός του αριθμού συστάδων, γεγονός που δεν είναι πάντα εύκολο ή αυτονόητο. Τέλος, η παρουσία ακραίων τιμών μπορεί να επηρεάσει αρνητικά τη διαδικασία σύγκλισης και τη δομή των ομάδων. Στην παρούσα ανάλυση, τα ζητήματα αυτά αντιμετωπίζονται κατάλληλα, με χρήση τεχνικών για τη βελτιστοποίηση της αρχικοποίησης, καθώς και με εφαρμογή μεθόδων αξιολόγησης του αριθμού συστάδων, που περιγράφονται στην Ενότητα 4.4 και κανονικοποίησης των δεδομένων, ώστε να περιοριστούν οι επιδράσεις απόκλισης ή ανομοιογένειας στα χαρακτηριστικά [21]. Ειδικότερα, η διαδικασία αρχικοποίησης ενισχύεται με τη στρατηγική 'k-means++', η οποία επιλέγει με ευφυή τρόπο τα αρχικά κέντρα των συστάδων για τη βελτίωση της ταχύτητας και της ποιότητας της σύγκλισης. Για τη διασφάλιση της ευρωστίας και την αποφυγή τοπικών βέλτιστων, ο αλγόριθμος εκτελείται 10 φορές με διαφορετικά αρχικά κέντρα, επιλέγοντας ως τελική την ομαδοποίηση με τη χαμηλότερη αδράνεια. Ο μέγιστος αριθμός επαναλήψεων σε κάθε εκτέλεση έχει οριστεί σε 450, ενώ η χρήση μιας

προκαθορισμένης τιμής τυχαιότητας εγγυάται την πλήρη αναπαραγωγιμότητα των αποτελεσμάτων της ανάλυσης.

Συμπερασματικά, ο αλγόριθμος K-Means κρίνεται κατάλληλος για την παρούσα μελέτη λόγω των προαναφερθέντων ιδιοτήτων του. Σε συνδυασμό με κατάλληλη κανονικοποίηση των χαρακτηριστικών και την εφαρμογή τεχνικών επιλογής του αριθμού συστάδων, δύναται να αποδώσει αξιόπιστα και ερμηνεύσιμα αποτελέσματα ομαδοποίησης, συμβάλλοντας στη διακριτή ταξινόμηση οικιακών καταναλωτών με βάση τα ενεργειακά τους προφίλ.

4.4 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΟΥ ΒΕΛΤΙΣΤΟΥ ΑΡΙΘΜΟΥ ΣΥΣΤΑΔΩΝ

Για τον προσδιορισμό του αριθμού των συστάδων k , θα χρησιμοποιηθούν οι μέθοδοι Elbow, Silhouette και Calinski-Harabasz, οι οποίες προσφέρουν διαφορετικές προσεγγίσεις αξιολόγησης της ποιότητας ομαδοποίησης. Ο ακριβής καθορισμός του k είναι κρίσιμος για την επίτευξη διακριτών και ομοιογενών συστάδων, καθώς επηρεάζει άμεσα τόσο τη συνοχή εντός των ομάδων όσο και τη διαφοροποίηση μεταξύ τους. Οι παραπάνω μέθοδοι βασίζονται σε μετρικές που αποτιμούν είτε τη γεωμετρική εγγύτητα των σημείων εντός κάθε συστάδας, είτε την απόσταση μεταξύ διαφορετικών συστάδων, επιτρέποντας έτσι την αντικειμενική επιλογή της καταλληλότερης τιμής για το k .

4.4.1 Μέθοδος Elbow

Η μέθοδος Elbow αποτελεί μία από τις πιο διαδεδομένες τεχνικές για τον προσδιορισμό του κατάλληλου αριθμού συστάδων κατά την εφαρμογή αλγορίθμων ομαδοποίησης, όπως στην συγκεκριμένη διπλωματική εργασία ο K-Means. Η βασική της αρχή στηρίζεται στην παρατήρηση της μεταβολής ενός μέτρου ενδοσυσταδικής διασποράς, καθώς αυξάνεται ο αριθμός των συστάδων. Για κάθε τιμή του αριθμού συστάδων k , ο αλγόριθμος εκτελείται και υπολογίζεται το άθροισμα των τετραγώνων των αποστάσεων όλων των σημείων από το κέντρο της συστάδας τους [21]. Η ποσότητα αυτή, γνωστή ως ενδοσυσταδική διακύμανση ή συνολικό σφάλμα (δείκτης WCSS – Within Cluster Sum of Squares), τείνει να μειώνεται όσο αυξάνεται ο αριθμός συστάδων, καθώς τα σημεία κατανέμονται σε περισσότερες και μικρότερες ομάδες [22]. Ο υπολογισμός του σφάλματος πραγματοποιείται σύμφωνα με την εξίσωση (4.3) [22]:

$$WCSS = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (4.3)$$

όπου:

k : ο αριθμός των συστάδων

C_i : το σύνολο των σημείων που ανήκουν στη συστάδα i

μ_i : το κέντρο/μέσος όρος της συστάδας

x : σημείο που ανήκει στη συστάδα i

$\|x - \mu_i\|$: η Ευκλείδεια απόσταση μεταξύ του σημείου και του κέντρου

Κατά την υλοποίηση της μεθόδου, κατασκευάζεται διάγραμμα με την τιμή του δείκτη WCSS σε συνάρτηση με το k . Το διάγραμμα αυτό εμφανίζει μία έντονη κάμψη, η οποία μοιάζει με “γόνατο”. Εκείνο το σημείο είναι αυτό στο οποίο η μείωση του σφάλματος παύει να είναι σημαντική. Η τιμή του k που αντιστοιχεί στο σημείο της κάμψης θεωρείται η βέλτιστη, καθώς επιτυγχάνεται ικανοποιητική ομαδοποίηση χωρίς περιττή πολυπλοκότητα [21].

Η μέθοδος βασίζεται κυρίως σε οπτική αξιολόγηση, γεγονός που απαιτεί κρίση και εμπειρία στην ερμηνεία του διαγράμματος. Η μέθοδος Elbow προσφέρει έναν απλό και πρακτικό μηχανισμό για την επιλογή του αριθμού συστάδων, ιδιαίτερα χρήσιμο κατά τα αρχικά στάδια ανάλυσης δεδομένων [22].

4.4.2 Μέθοδος Silhouette

Η μέθοδος Silhouette αξιολογεί την ποιότητα της ομαδοποίησης για δεδομένο αριθμό συστάδων, υπολογίζοντας τον βαθμό στον οποίο κάθε παρατήρηση είναι σωστά τοποθετημένη εντός της συστάδας της σε σύγκριση με άλλες συστάδες. Ο δείκτης Silhouette Score συνδυάζει την ενδοσυσταδική συνοχή και τον διασυσταδικό διαχωρισμό, προσφέροντας μία ποσοτική εκτίμηση της καταλληλότητας της ταξινόμησης κάθε σημείου [23]. Ο υπολογισμός του δείκτη πραγματοποιείται σύμφωνα με την εξίσωση (4.4) [23]:

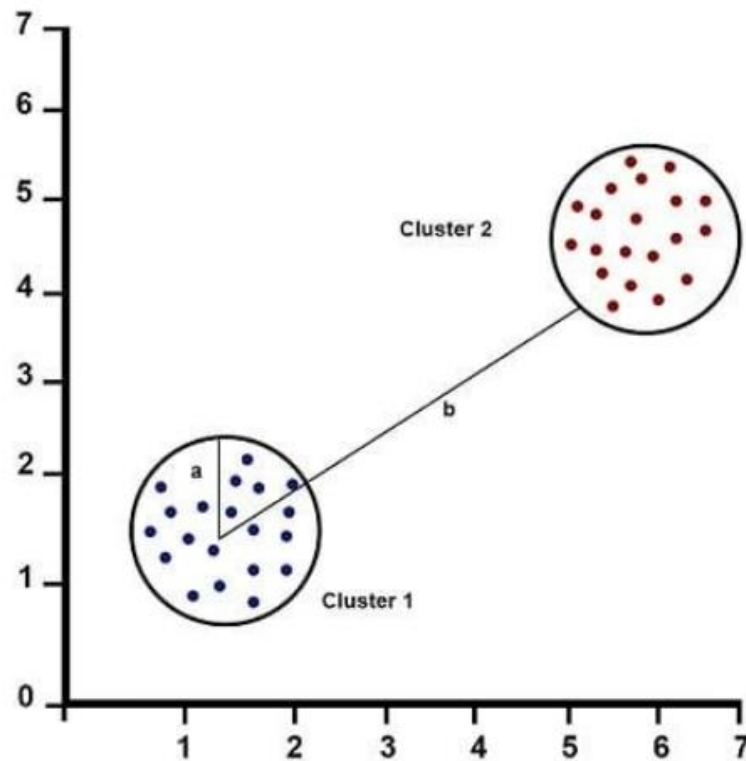
$$s_i = \frac{(b_i - a_i)}{\max(b_i, a_i)} \quad (4.4)$$

όπου:

a_i : η μέση απόσταση του σημείου i από όλα τα άλλα σημεία της ίδιας συστάδας

b_i : η ελάχιστη μέση απόσταση του σημείου i από όλα τα σημεία των υπόλοιπων συστάδων

Η λογική υπολογισμού των ποσοτήτων a_i και b_i αποτυπώνεται στο Σχήμα 4.5. Το σημείο αναφοράς βρίσκεται στη συστάδα 1 και απέχει κατά a_i από τα υπόλοιπα σημεία της ίδιας συστάδας, ενώ η ποσότητα b_i εκφράζει τη μέση απόστασή του από τα σημεία της πλησιέστερης διαφορετικής συστάδας 2.



Σχήμα 4.5: Οπτική απεικόνιση των αποστάσεων a_i , b_i για τον υπολογισμό του δείκτη s_i [22]

Ο μέσος όρος όλων των s_i για ένα δεδομένο k παρέχει ένα συνολικό σκορ (Silhouette Score), βάσει του οποίου μπορεί να επιλεγεί ο αριθμός συστάδων που μεγιστοποιεί τη συνολική ποιότητα ομαδοποίησης. Κυμαίνεται στο διάστημα $[-1,1]$, με τιμές κοντά στο 1 να υποδηλώνουν ότι το σημείο βρίσκεται καλά ενσωματωμένο στη συστάδα του και μακριά από τις υπόλοιπες. Τιμές κοντά στο 0 υποδηλώνουν σημεία που βρίσκονται στα όρια μεταξύ συστάδων, ενώ τιμές κοντά στο -1 δείχνουν πιθανή λανθασμένη ταξινόμηση του σημείου [23].

Σε αντίθεση με τη μέθοδο Elbow, η οποία εστιάζει αποκλειστικά στη μείωση της ενδοσυσταδικής διασποράς (μέσω του δείκτη WCSS), η μέθοδος Silhouette συνδυάζει τη συνεκτικότητα και τη διακριτότητα των ομάδων, καθιστώντας την πιο ισορροπημένη ως προς την αξιολόγηση. Επιπλέον, ενώ το διάγραμμα Elbow απαιτεί υποκειμενική οπτική ερμηνεία, η Silhouette παρέχει συγκρίσιμη και αντικειμενική τιμή για κάθε πιθανό k , καθιστώντας την ιδιαίτερα χρήσιμη όταν δεν υπάρχει σαφής “κάμψη” στο διάγραμμα WCSS [22].

4.4.3 Μέθοδος Calinski-Harabasz

Ο δείκτης Calinski-Harabasz, γνωστός και ως Κριτήριο Αναλογίας Διασποράς, αποτελεί εσωτερικό μέτρο αξιολόγησης της ποιότητας ομαδοποίησης, βασιζόμενο αποκλειστικά στα χαρακτηριστικά του συνόλου δεδομένων και στα αποτελέσματα της ομαδοποίησης. Ο δείκτης αυτός συγκρίνει τη διασπορά μεταξύ των συστάδων με τη διασπορά εντός των συστάδων, προσφέροντας μια ποσοτική εκτίμηση της ποιότητας της ομαδοποίησης [24]. Ο υπολογισμός του δείκτη πραγματοποιείται σύμφωνα με την εξίσωση (4.5) [24],[25]:

$$CH = \frac{\frac{BCSS}{k-1}}{\frac{WCSS}{n-k}} = \frac{1}{WCBCR} \times \frac{n-k}{k-1} \quad (4.5)$$

όπου:

CH : ο δείκτης Calinski–Harabasz

k : ο αριθμός των συστάδων

n : το συνολικό πλήθος των παρατηρήσεων

$BCSS$ (Between-Cluster Sum of Squares): το άθροισμα των τετραγώνων των αποστάσεων μεταξύ των κέντρων των συστάδων και του συνολικού κέντρου των δεδομένων, σταθμισμένο με το πλήθος των σημείων κάθε συστάδας

$WCSS$ (Within-Cluster Sum of Squares): το άθροισμα των τετραγώνων των αποστάσεων μεταξύ των σημείων κάθε συστάδας και του αντίστοιχου κέντρου της

$WCBCR$ (Within-Cluster Sum of Squares to Between-Cluster Variation): δείκτης λόγου του αθροίσματος των τετραγώνων των αποστάσεων κάθε διανύσματος εισόδου από το κέντρο της ομάδας που ανήκει προς τη διασπορά μεταξύ των κέντρων των ομάδων

Ουσιαστικά, ο δείκτης Calinski–Harabasz αξιολογεί την ποιότητα της ομαδοποίησης λαμβάνοντας υπόψη τόσο τη συνοχή εντός των συστάδων όσο και τη διακριτότητα μεταξύ τους. Υψηλές τιμές του δείκτη υποδηλώνουν ότι οι συστάδες είναι συμπαγείς και καλά διαχωρισμένες, ενώ χαμηλές τιμές υποδηλώνουν το αντίθετο. Κατά την εφαρμογή της μεθόδου, ο δείκτης υπολογίζεται για διάφορες τιμές του k , με αυτή που μεγιστοποιεί τον δείκτη να θεωρείται ως η βέλτιστη για την ομαδοποίηση των δεδομένων [24].

Η μέθοδος Calinski–Harabasz προσφέρει ένα αντικειμενικό και υπολογιστικά αποδοτικό μέτρο για την αξιολόγηση της ποιότητας της ομαδοποίησης, καθιστώντας την ιδιαίτερα χρήσιμη σε περιπτώσεις όπου δεν υπάρχουν προκαθορισμένες ετικέτες ή εξωτερικά κριτήρια αξιολόγησης [24].

ΑΠΟΤΕΛΕΣΜΑΤΑ ΟΜΑΔΟΠΟΙΗΣΗΣ ΚΑΙ ΠΡΑΚΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

Η ομαδοποίηση οικιακών καταναλωτών αποτελεί μια τεχνική που επιτρέπει την ανάλυση των προτύπων κατανάλωσης σε επιμέρους ομάδες με κοινά χαρακτηριστικά. Μέσω της ομαδοποίησης, μπορούν να εντοπιστούν κατηγορίες χρηστών με παρόμοια συμπεριφορά ως προς τη χρήση ηλεκτρικής ενέργειας. Η τεχνική ανάλυσης συστάδων μπορεί να αξιοποιηθεί για στοχευμένα σχήματα διαχείρισης ζήτησης, προγράμματα ευελιξίας ή και δυναμική τιμολόγηση.

Το παρόν κεφάλαιο επικεντρώνεται στην παρουσίαση των αποτελεσμάτων που προκύπτουν από την εφαρμογή της μοντελοποίησης της ομαδοποίησης που αναλύθηκε θεωρητικά στο Κεφάλαιο 4. Ειδικότερα, εξετάζονται οι συνδυασμοί χαρακτηριστικών εισόδου, οι τιμές των δεικτών αξιολόγησης για διαφορετικούς αριθμούς συστάδων και η επιλογή του τελικού μοντέλου ομαδοποίησης. Επιπλέον, αναδεικνύονται ενδεικτικές πρακτικές εφαρμογές των αποτελεσμάτων, οι οποίες μπορούν να αξιοποιηθούν από ΦοΣΕ Φορτίου στο πλαίσιο διαχείρισης ευελιξίας και ζήτησης.

5.1 ΑΞΙΟΛΟΓΗΣΗ ΚΡΙΤΗΡΙΩΝ ΕΙΣΟΔΟΥ ΚΑΙ ΕΠΙΛΟΓΗ ΒΕΛΤΙΣΤΟΥ ΣΥΝΔΥΑΣΜΟΥ

Η επιλογή των κατάλληλων χαρακτηριστικών εισόδου αποτελεί καθοριστικό παράγοντα για την ποιότητα της ομαδοποίησης, καθώς επηρεάζει τόσο τη συνοχή εντός των συστάδων όσο και τη διακριτότητα μεταξύ τους. Στο πλαίσιο αυτό, πραγματοποιείται ανάλυση ευαισθησίας με στόχο την αποτίμηση της συνεισφοράς κάθε χαρακτηριστικού στην απόδοση του αλγορίθμου. Η διαδικασία αξιολόγησης περιλαμβάνει μετρικές ποιότητας ομαδοποίησης και ανάλυση συσχέτισης μεταξύ των μεταβλητών, με σκοπό την επιλογή ενός τελικού συνδυασμού, ο οποίος είναι ταυτόχρονα αποτελεσματικός και ερμηνεύσιμος. Η αρχική προσέγγιση εστιάζει σε τρία επιλεγμένα χαρακτηριστικά που θεωρούνται αντιπροσωπευτικά της καταναλωτικής συμπεριφοράς κατά τις ώρες αιχμής, τα οποία είναι τα εξής:

- **Ημέρες με κατανάλωση μεγαλύτερη από 3 kWh (Ημέρες με υψηλή κατανάλωση - N_3)**
- **Ημέρες με κατανάλωση μεγαλύτερη από 1 kWh (Συχνότητα - N_1)**
- **Ποσοστό μέσης κατανάλωσης τις ώρες αιχμής ως προς τη μέση συνολική ημερήσια κατανάλωση (Ποσοστό ωρών αιχμής - PAR)**

Αξίζει να σημειωθεί πως τα παραπάνω τρία χαρακτηριστικά θα αναφέρονται στη συνέχεια με τις αντίστοιχες συντομογραφίες τους: N_3 (Ημέρες με υψηλή κατανάλωση), N_1 (Συχνότητα) και PAR (Ποσοστό ωρών αιχμής). Για τον συνδυασμό των παραπάνω

κριτηρίων, εφαρμόζεται ο αλγόριθμος K-Means και υπολογίζεται ο δείκτης Silhouette Score. Τα αποτελέσματα συγκεντρώνονται στο Σχήμα 5.1, στον οποίο αποτυπώνεται η ικανότητα του συνδυασμού να διαμορφώνει διακριτές ομάδες.

Χαρακτηριστικά	Silhouette Score		
	Βέλτιστο k		
	k=3	k=4	k=5
Ημέρες με κατανάλωση > 3 kWh	0.455	0.476	0.415
Ημέρες με κατανάλωση > 1 kWh			
Ποσοστό αιχμής / ημερήσιας			

Σχήμα 5.1: Δείκτης Silhouette Score των αρχικών χαρακτηριστικών εισόδου

Όπως παρατηρείται στο Σχήμα 5.1, ο συνδυασμός των επιλεγμένων χαρακτηριστικών παρουσιάζει, για $k = 4$, τη μέγιστη τιμή του δείκτη Silhouette, ίση με 0,476, υποδεικνύοντας ικανοποιητικό διαχωρισμό μεταξύ των συστάδων. Η τιμή αυτή κρίνεται επαρκώς υψηλή για τα δεδομένα της ανάλυσης, επιβεβαιώνοντας τη συνεισφορά των συγκεκριμένων μεταβλητών στην ποιότητα της ομαδοποίησης.

Στη συνέχεια εξετάζεται η επίδραση της προσθήκης των υπόλοιπων πέντε χαρακτηριστικών, τα οποία είναι τα εξής:

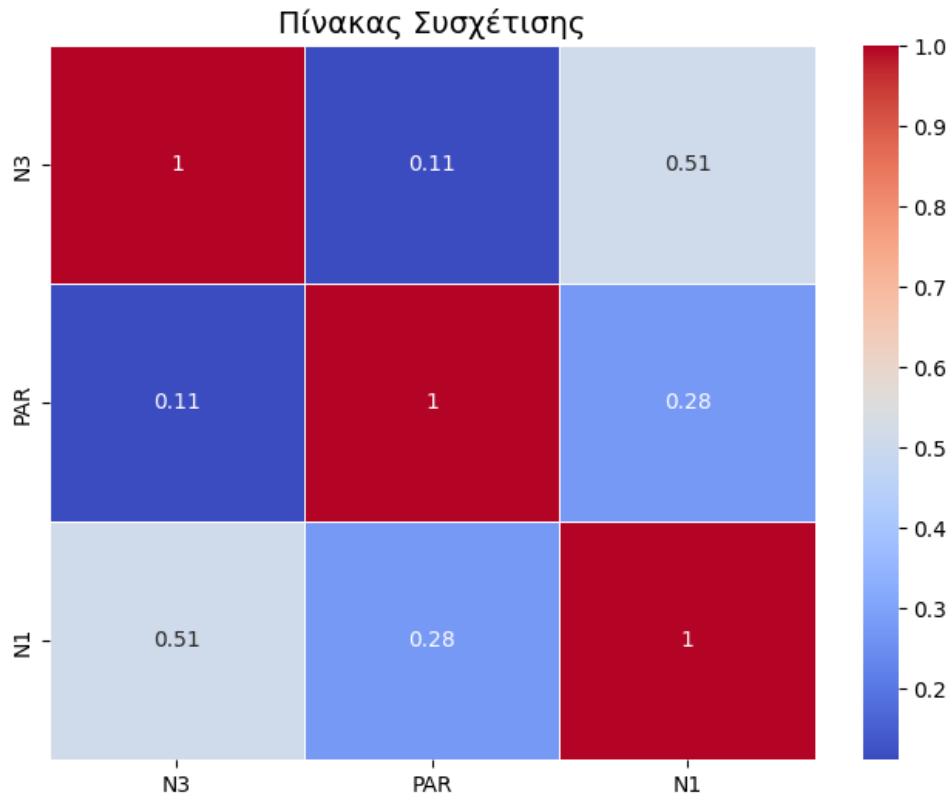
- Μέση τιμή μεγίστων κατανάλωσης κατά τις ώρες αιχμής
- Διακύμανση και εύρος τιμών κατανάλωσης κατά τις ώρες αιχμής
- Μέση σχετική τυπική απόκλιση κατά τις ώρες αιχμής
- Αυτοσυσχέτιση κατά τις ώρες αιχμής
- Καθημερινές και Σαββατοκύριακα κατά τις ώρες αιχμής

Η αναλυτική επεξήγηση όλων των κριτηρίων εισόδου έχει γίνει στην υποενότητα 4.1.2 του Κεφαλαίου 4. Η προσθήκη των χαρακτηριστικών πραγματοποιείται σταδιακά, μεμονωμένα και σε συνδυασμό, προκειμένου να υπολογιστεί η μεταβολή του δείκτη Silhouette και να εκτιμηθεί η επίδραση κάθε μεταβλητής στην ποιότητα της ομαδοποίησης. Τα αποτελέσματα συγκεντρώνονται στο Σχήμα 5.2.

Χαρακτηριστικά	Silhouette Score		
	Βέλτιστο k		
	k=3	k=4	k=5
Μέση τιμή μεγίστων κατανάλωσης	0.333	0.361	0.366
Διακύμανση / Εύρος τιμών	0.326	0.339	0.328
Μέση Σχετική Τυπική Απόκλιση	0.322	0.333	0.35
Αυτοσυσχέτιση	0.362	0.301	0.301
Καθημερινές - Σαββατοκύριακο	0.339	0.351	0.288
Όλα Μαζί	0.411	0.306	0.285

Σχήμα 5.2: Δείκτης Silhouette Score για συνδυασμούς με προσθήκη επιπλέον χαρακτηριστικών

Όπως καταγράφεται στο Σχήμα 5.2, η προσθήκη καθενός από τα επιπλέον χαρακτηριστικά οδηγεί σε σημαντική μείωση της τιμής του δείκτη Silhouette σε όλες τις περιπτώσεις, με τιμές που κυμαίνονται μεταξύ 0,301 και 0,361 για $k = 4$. Πολύ χαμηλή τιμή του δείκτη Silhouette προκύπτει όταν χρησιμοποιούνται όλα τα χαρακτηριστικά μαζί, με τον δείκτη να μειώνεται στο 0.306, γεγονός που επιβεβαιώνει την υποβάθμιση της ποιότητας της ομαδοποίησης λόγω πλεονάζουσας πληροφορίας. Πριν από οποιαδήποτε ανάλυση, πραγματοποιείται κανονικοποίηση των χαρακτηριστικών με τη μέθοδο Min-Max, ώστε να κλιμακωθούν τα χαρακτηριστικά στο διάστημα $[0,1]$. Η διαδικασία αυτή, η οποία έχει παρουσιαστεί αναλυτικά στην Ενότητα 4.2 του Κεφαλαίου 4, εξασφαλίζει τη συγκρισιμότητα των μεταβλητών μεταξύ τους, αποτρέποντας το ενδεχόμενο κάποια χαρακτηριστικά να επηρεάζουν δυσανάλογα τα αποτελέσματα λόγω διαφορετικής κλίμακας. Η κανονικοποίηση εφαρμόζεται στον αρχικό συνδυασμό των τριών χαρακτηριστικών. Η παραπάνω ανάλυση ενισχύεται από την αξιολόγηση συσχέτισης που απεικονίζεται στο Σχήμα 5.3. Το διάγραμμα παρουσιάζει τον πίνακα συσχέτισης που κατασκευάζεται για τα τρία επιλεγμένα χαρακτηριστικά.



Σχήμα 5.3: Πίνακας Συσχέτισης μεταξύ των τελικών χαρακτηριστικών εισόδου

Οι τιμές συσχέτισης μεταξύ των μεταβλητών εμφανίζονται μέτριες έως χαμηλές, στοιχείο που επιβεβαιώνει τη συμπληρωματικότητά τους και ενισχύει την επιλογή τους ως κατάλληλη βάση για τη διαδικασία ομαδοποίησης. Ως τελικό αποτέλεσμα, επιλέγεται ένας συνδυασμός τριών χαρακτηριστικών που παρουσιάζει υψηλή απόδοση βάσει του δείκτη Silhouette και χαμηλή συσχέτιση μεταξύ των μεταβλητών του. Η επιλογή βασίζεται στην ικανότητα των μεταβλητών να εκφράζουν τόσο τη συχνότητα εμφάνισης κατανάλωσης κατά τις ώρες αιχμής, όσο και την ένταση της κατανάλωσης. Κατά συνέπεια, η ομαδοποίηση που προκύπτει βασίζεται σε δύο θεμελιώδεις διαστάσεις της καταναλωτικής συμπεριφοράς: το ύψος και τη χρονική κατανομή της ζήτησης. Η κατηγοριοποίηση αυτή κρίνεται ιδιαίτερα χρήσιμη για την επιχειρησιακή αξιοποίηση των αποτελεσμάτων από ένα ΦοΣΕ Φορτίου, στο πλαίσιο της διαχείρισης ζήτησης. Για την κανονικοποίηση των χαρακτηριστικών και τον υπολογισμό του δείκτη Silhouette αξιοποιείται η βιβλιοθήκη Scikit-learn [26] της γλώσσας προγραμματισμού Python, ενώ για την οπτικοποίηση των σχέσεων μεταξύ των χαρακτηριστικών χρησιμοποιούνται οι βιβλιοθήκες Seaborn [27] και Matplotlib [18]. Στα Σχήματα 5.4 και 5.5, παρουσιάζεται ο κώδικας που χρησιμοποιήθηκε σε αυτό το στάδιο της μεθοδολογίας.

```
# Silhouette Score
from sklearn.metrics import silhouette_score

# Store metrics
silhouette_score_value = silhouette_score(normalized_features, cluster_labels)
silhouette_scores.append(silhouette_score_value)

# Print silhouette score for each k
print(f"Silhouette Score for k={k}: {silhouette_score_value:.4f}")
```

Σχήμα 5.4: Υπολογισμός σε Python του Silhouette Score, για διάφορα k

```
# Correlation Heatmap

# Feature abbreviations
feature_names = {
    "number_of_peak_days_with_high_consumption": "N3",
    "frequency": "N1",
    "percentage_peak_hours": "PAR"
}

# Rename the feature columns using the defined abbreviations
features_renamed = features.rename(columns=feature_names)

# Ensure 'Cluster' column is not present
features_renamed = features_renamed.loc[:, ~features_renamed.columns.str.contains('Cluster', case=False)]

# Compute the correlation matrix
corr_matrix = features_renamed.corr()

# Create the heatmap
plt.figure(figsize=(8, 6))
sns.heatmap(
    corr_matrix,
    annot=True,
    cmap="coolwarm",
    fmt=".2g",
    linewidths=0.5
)

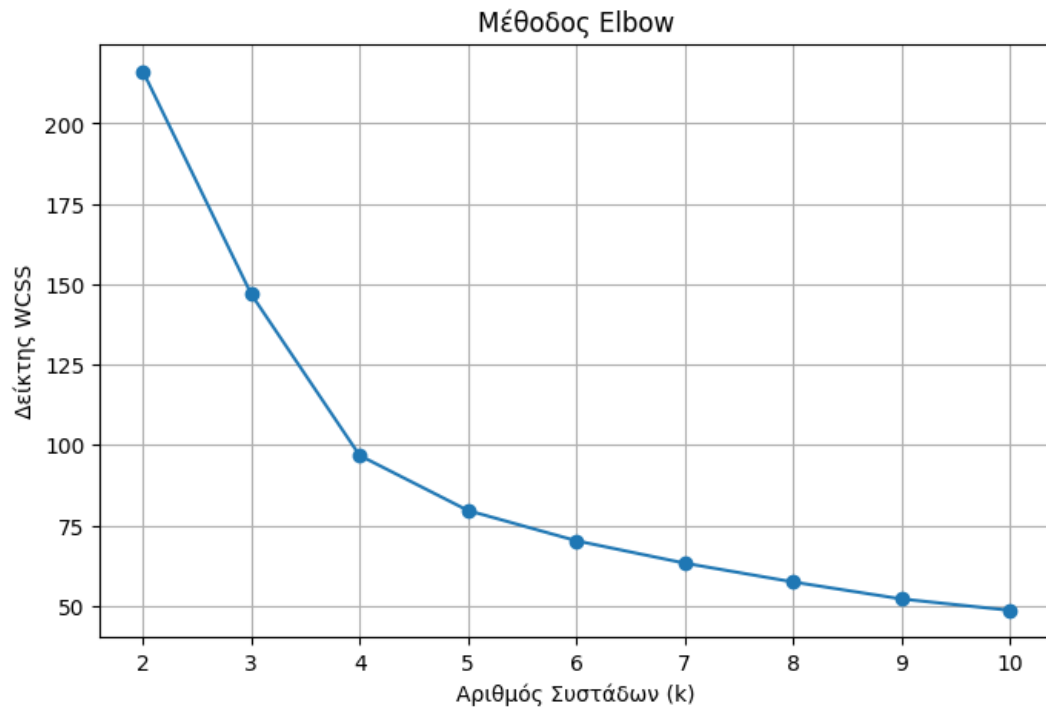
# Add title
plt.title("Πίνακας Συσχέτισης", fontsize=14)

# Show the heatmap
plt.show()
```

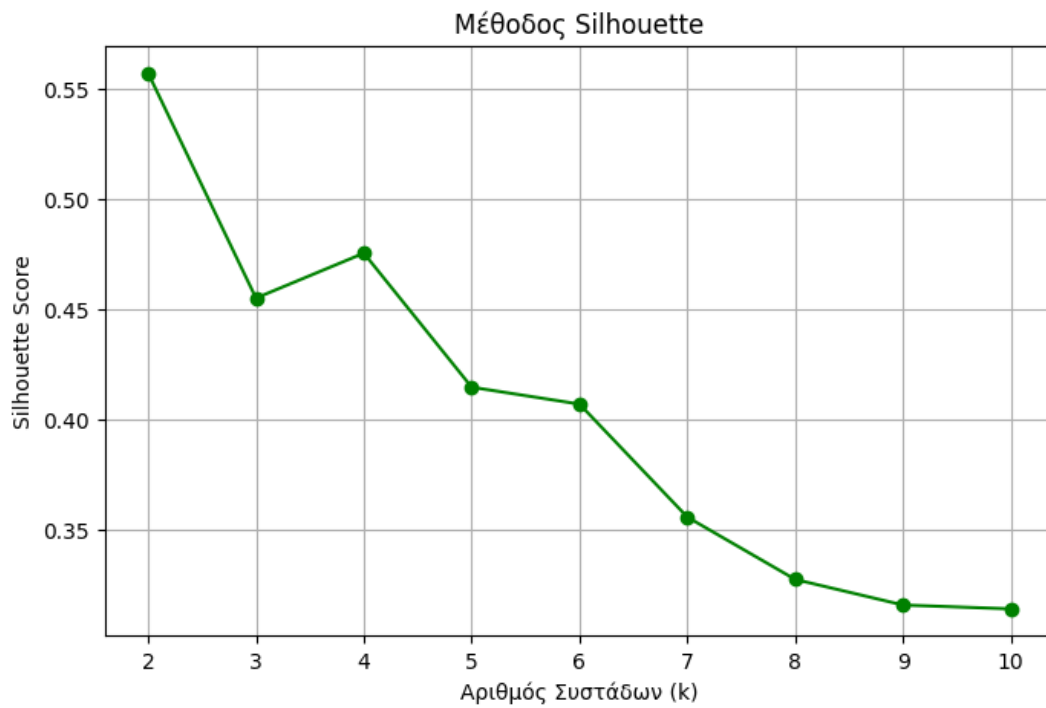
Σχήμα 5.5: Υπολογισμός σε Python του Πίνακα Συσχέτισης

5.2 ΤΕΛΙΚΗ ΕΠΙΛΟΓΗ ΤΟΥ ΒΕΛΤΙΣΤΟΥ ΑΡΙΘΜΟΥ ΣΥΣΤΑΔΩΝ

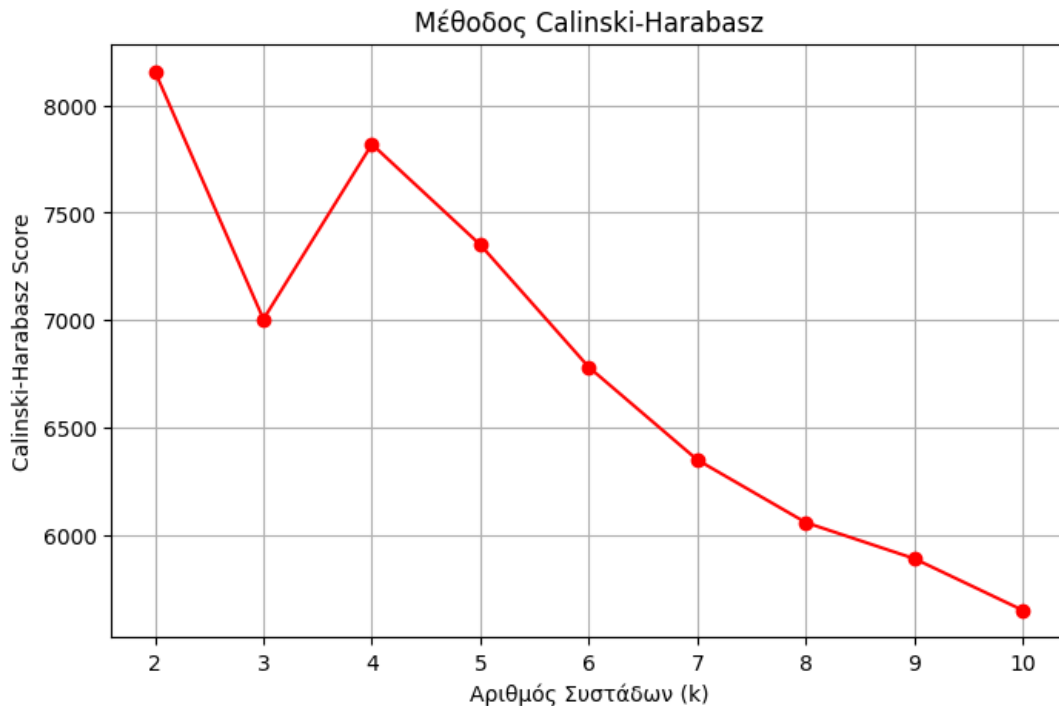
Στο παρόν στάδιο αξιολογείται πρακτικά ο αριθμός των συστάδων που οδηγεί στην καλύτερη ομαδοποίηση, με βάση τα τρία χαρακτηριστικά εισόδου που επιλέχθηκαν στην Ενότητα 5.1. Η διαδικασία βασίζεται στην εφαρμογή των μεθόδων Elbow, Silhouette Score και Calinski–Harabasz, οι οποίες αναλύθηκαν θεωρητικά στην Ενότητα 4.4 του Κεφαλαίου 4. Τα τελικά αποτελέσματα των αντίστοιχων δεικτών απεικονίζονται στα Σχήματα 5.6, 5.7 και 5.8.



Σχήμα 5.6: Μέθοδος Elbow για την επιλογή βέλτιστου αριθμού συστάδων



Σχήμα 5.7: Τιμές σκορ Silhouette για διαφορετικά k



Σχήμα 5.8: Τιμές σκορ Calinski–Harabasz για διαφορετικά k

Από το Σχήμα 5.6 παρατηρείται μια έντονη κάμψη στην τιμή $k = 4$, γεγονός που αποτελεί ένδειξη για την ύπαρξη «γονάτου» στη μέθοδο Elbow και υποδηλώνει σημείο, πέρα από το οποίο η μείωση του ενδοσυσταδικού σφάλματος (WCSS) δεν είναι πλέον σημαντική. Παράλληλα, από το Σχήμα 5.7 διαπιστώνεται ότι για $k = 4$ ο δείκτης Silhouette παίρνει τιμή 0.48, η οποία, αν και όχι η μέγιστη, είναι επαρκώς υψηλή και συνδυάζει συνοχή και διαχωρισιμότητα. Στο Σχήμα 5.8 ο δείκτης Calinski–Harabasz παρουσιάζει επίσης υψηλή τιμή για $k = 4$, η οποία είναι αρκετά κοντά στη μέγιστη. Τελικά, η συνεκτίμηση των τριών δεικτών συγκλίνει στην επιλογή της τιμής $k = 4$ ως βέλτιστης για την ομαδοποίηση. Παρότι δεν αποτελεί το απόλυτο μέγιστο όλων των δεικτών, εντοπίζεται ως η πλέον κατάλληλη τιμή λόγω της συμβατότητάς της με ταυτόχρονες υψηλές αποδόσεις στους διαφορετικούς δείκτες αξιολόγησης.

Στο επόμενο στάδιο, εφαρμόζεται ο αλγόριθμος K-Means, ο οποίος αναλύθηκε θεωρητικά στην Ενότητα 4.3 του Κεφαλαίου 4, με αριθμό συστάδων $k = 4$, όπως προέκυψε από την προηγούμενη αξιολόγηση. Στο Σχήμα 5.9, παρουσιάζεται ο κώδικας που χρησιμοποιήθηκε για την υλοποίηση του αλγορίθμου.

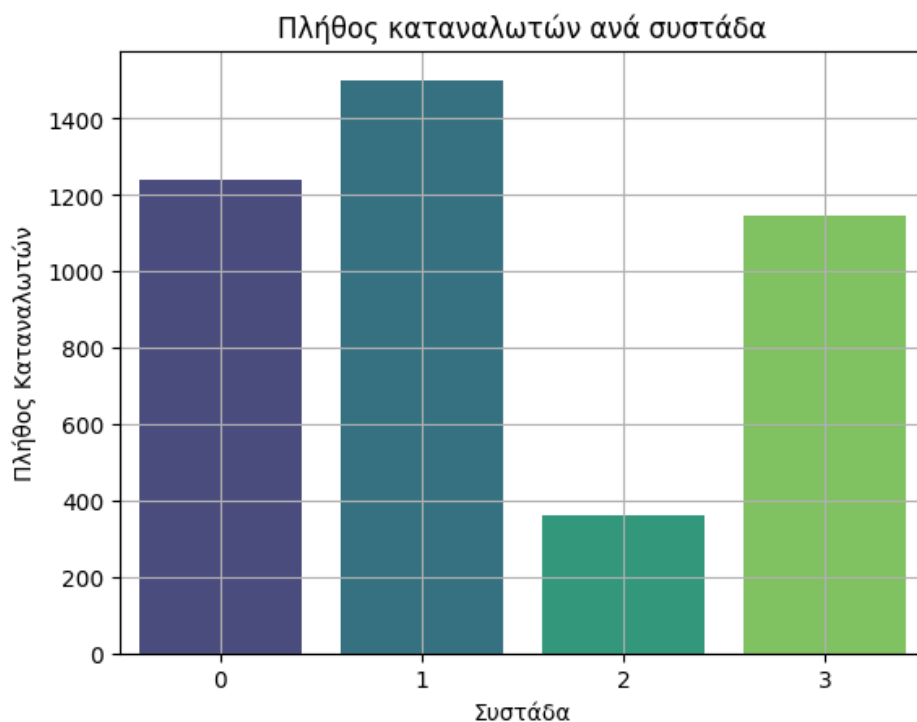
```
# K-means method
from sklearn.cluster import KMeans

# Apply K-Means with optimal k
optimal_k = 4
kmeans = KMeans(n_clusters=optimal_k, random_state=42)
clusters = kmeans.fit_predict(normalized_features)

# Add the cluster labels to the DataFrame
features['Cluster'] = clusters
```

Σχήμα 5.9: Υλοποίηση σε Python του αλγορίθμου K-Means

Η κατανομή των οικιακών καταναλωτών στις τελικές τέσσερις συστάδες αποτυπώνεται στο Σχήμα 5.10. Το διάγραμμα παρουσιάζει τον αριθμό των νοικοκυριών που ταξινομούνται σε κάθε ομάδα, προσφέροντας ένδειξη για την πληθυσμιακή κατανομή των συστάδων.

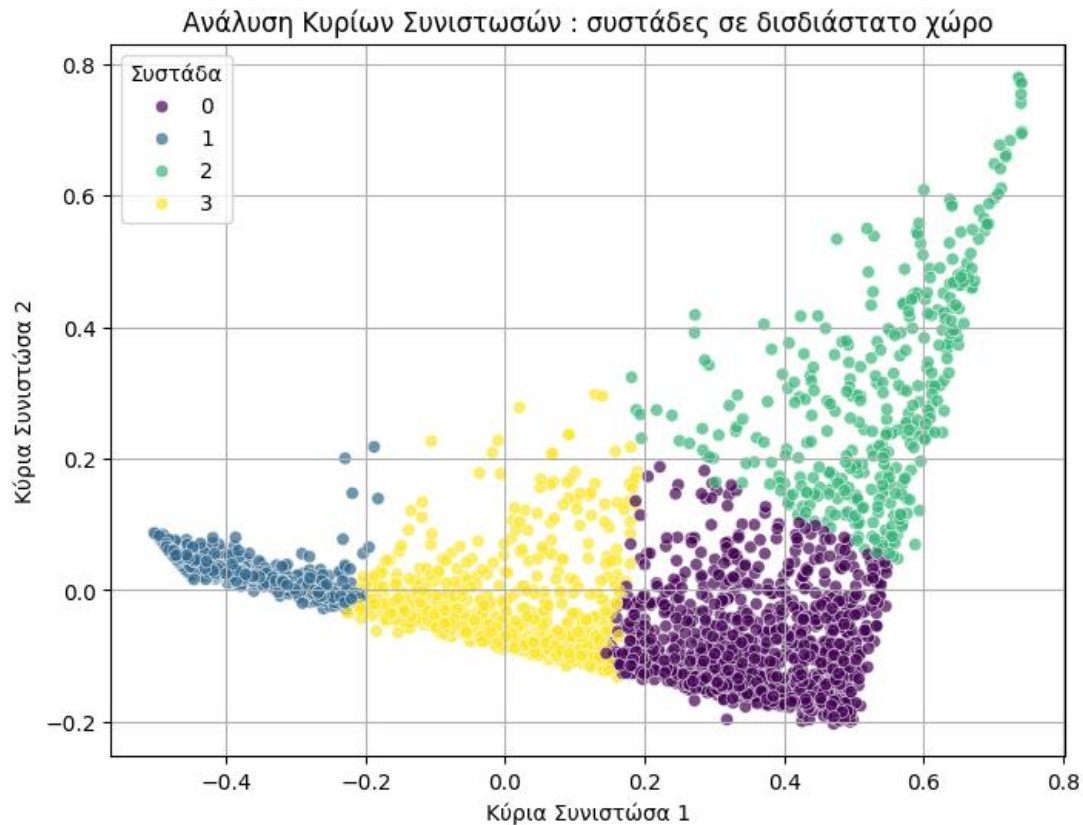


Σχήμα 5.10: Κατανομή οικιακών καταναλωτών ανά συστάδα

Συγκεκριμένα, παρατηρείται η εξής κατανομή:

- **Συστάδα 0:** 1238 καταναλωτές
- **Συστάδα 1:** 1501 καταναλωτές
- **Συστάδα 2:** 362 καταναλωτές
- **Συστάδα 3:** 1146 καταναλωτές

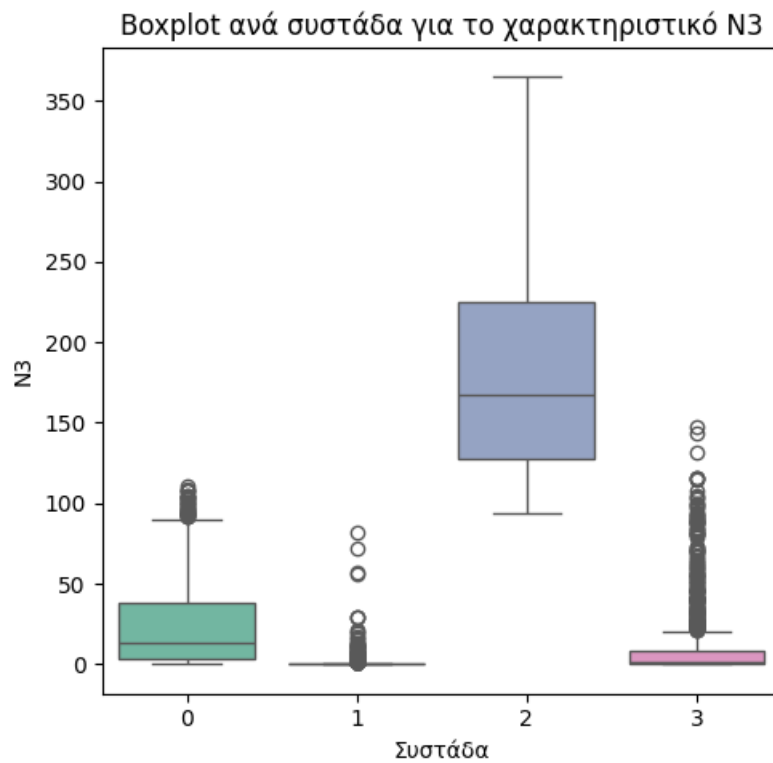
Ακολούθως, πραγματοποιείται οπτικοποίηση των αποτελεσμάτων ομαδοποίησης μέσω της μεθόδου Ανάλυσης Κυρίων Συνιστωσών (Principal Component Analysis – PCA). Η τεχνική αυτή επιτρέπει την προβολή των ομάδων σε δύο διαστάσεις μέσω ανάλυσης των δύο κυρίων συνιστωσών με κανονικοποιημένες τιμές. Με αυτόν τον τρόπο διευκολύνεται η γραφική αναπαράσταση της κατανομής των παρατηρήσεων στο νέο δισδιάστατο χώρο. Το συγκεκριμένο διάγραμμα απεικονίζεται στο Σχήμα 5.11.



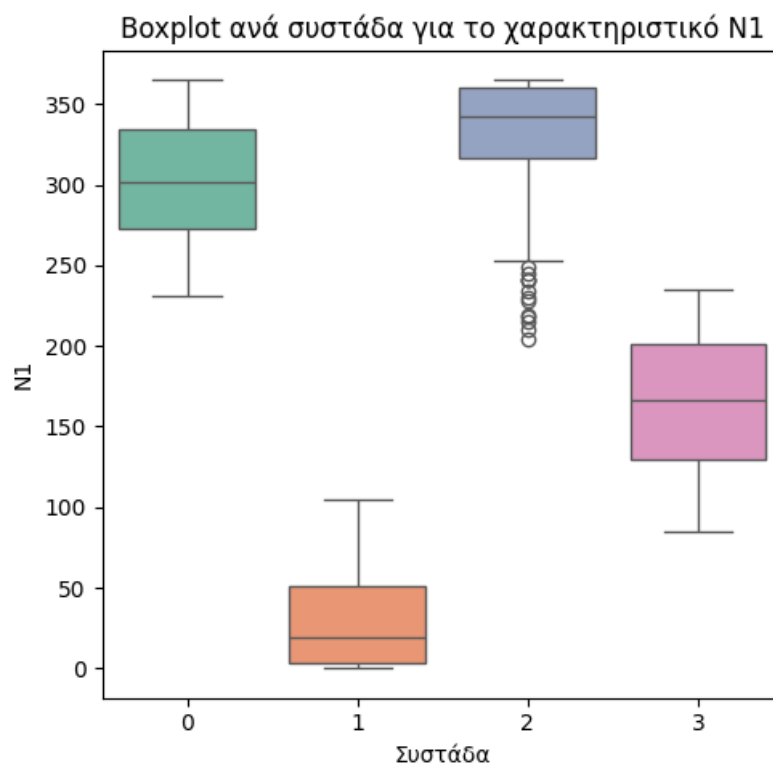
Σχήμα 5.11: Ανάλυση Κυρίων Συνιστωσών

Παρατηρείται ότι οι ομάδες εμφανίζουν σαφή διαχωρισμό μεταξύ τους, χωρίς εμφανείς επικαλύψεις, γεγονός που επιβεβαιώνει την αποτελεσματικότητα της ομαδοποίησης και ενισχύει την ερμηνευσιμότητα των αποτελεσμάτων.

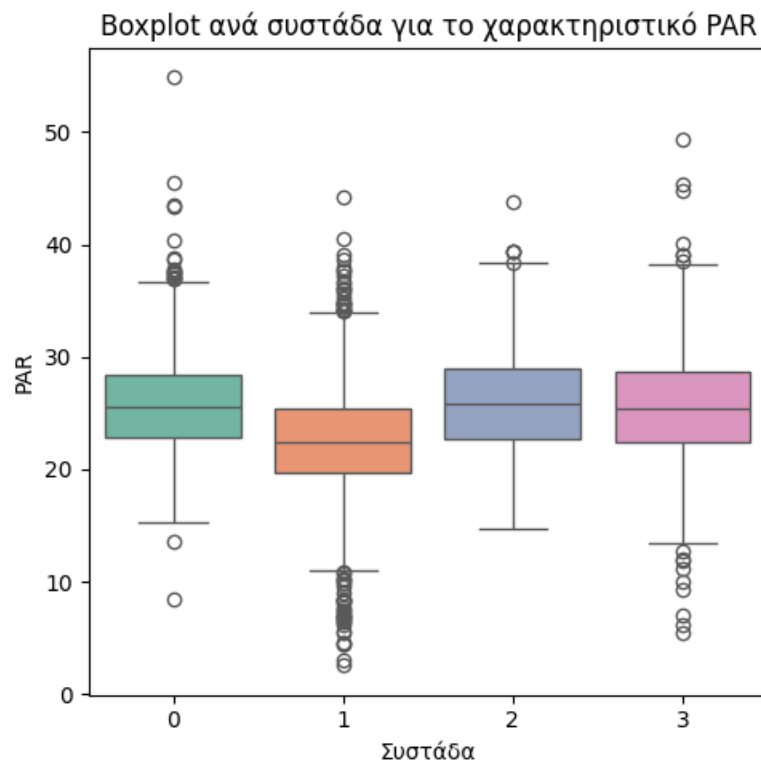
Στη συνέχεια, πραγματοποιείται ανάλυση της καταναλωτικής συμπεριφοράς ανά συστάδα για καθένα από τα τελικά χαρακτηριστικά εισόδου. Για τον σκοπό αυτό χρησιμοποιούνται διαγράμματα τύπου Boxplot, τα οποία παρουσιάζονται στα Σχήματα 5.12, 5.13 και 5.14. Μέσω των γραφημάτων αυτών αποτυπώνεται η κατανομή των τιμών εντός κάθε συστάδας και επιτρέπεται η σύγκριση μεταξύ των ομάδων.



Σχήμα 5.12: Boxplot ανά συστάδα για το χαρακτηριστικό που αφορά την υψηλή κατανάλωση



Σχήμα 5.13: Boxplot ανά συστάδα για το χαρακτηριστικό που αφορά τη συχνότητα κατανάλωσης



Σχήμα 5.14: Boxplot ανά συστάδα για το χαρακτηριστικό που αφορά το ποσοστό κατανάλωσης τις ώρες αιχμής σε σχέση με τη συνολική ημερήσια κατανάλωση

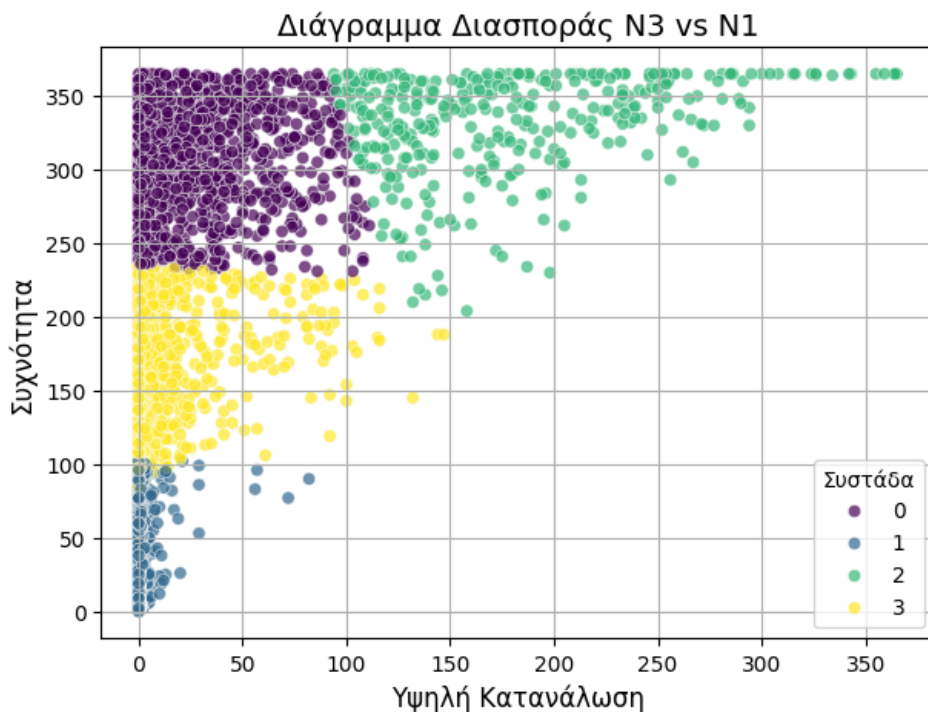
Από την παρατήρηση των Σχημάτων 5.12 έως 5.14 προκύπτουν τα εξής βασικά συμπεράσματα:

- Η συστάδα 0 παρουσιάζει σχετικά μέτριες προς υψηλές τιμές κατανάλωσης, κυρίως στα χαρακτηριστικά που αφορούν την υψηλή κατανάλωση και συχνότητα, με υψηλές τιμές στη συχνότητα. Παρατηρείται επίσης αυξημένη διασπορά, με παρουσία ακραίων τιμών.
- Η συστάδα 1 χαρακτηρίζεται από χαμηλές τιμές κυρίως στο χαρακτηριστικό που αφορά τη συχνότητα κατανάλωσης, γεγονός που δείχνει πιο ομοιόμορφη ημερήσια κατανομή κατανάλωσης, χωρίς έντονη κατανάλωση στις αιχμές. Όσον αφορά το χαρακτηριστικό που αφορά τις ημέρες με υψηλή κατανάλωση, παρατηρείται ύπαρξη ακραίων τιμών. Οι τιμές αυτές υποδηλώνουν συγκεκριμένες, σπάνιες περιπτώσεις που παρουσιάζουν σημαντικά υψηλότερη κατανάλωση κατά τις ώρες αιχμής από το αναμενόμενο για τη συγκεκριμένη συστάδα. Αυτή η διακύμανση μπορεί να οφείλεται σε εποχιακούς καταναλωτές.
- Η συστάδα 2 εμφανίζει σταθερά υψηλές τιμές και στα τρία χαρακτηριστικά, υποδηλώνοντας έντονη και συχνή κατανάλωση κατά τις ώρες αιχμής. Οι τιμές εκτείνονται σε μεγαλύτερο εύρος, ενώ καταγράφονται και μεμονωμένες ακραίες παρατηρήσεις.
- Η συστάδα 3 συγκεντρώνει καταναλωτές με πολύ χαμηλές τιμές στο χαρακτηριστικό που αφορά την υψηλή κατανάλωση, κάτι που υποδηλώνει περιορισμένη συχνότητα έντονης κατανάλωσης, αλλά παρουσιάζει ποικιλία στα υπόλοιπα δύο

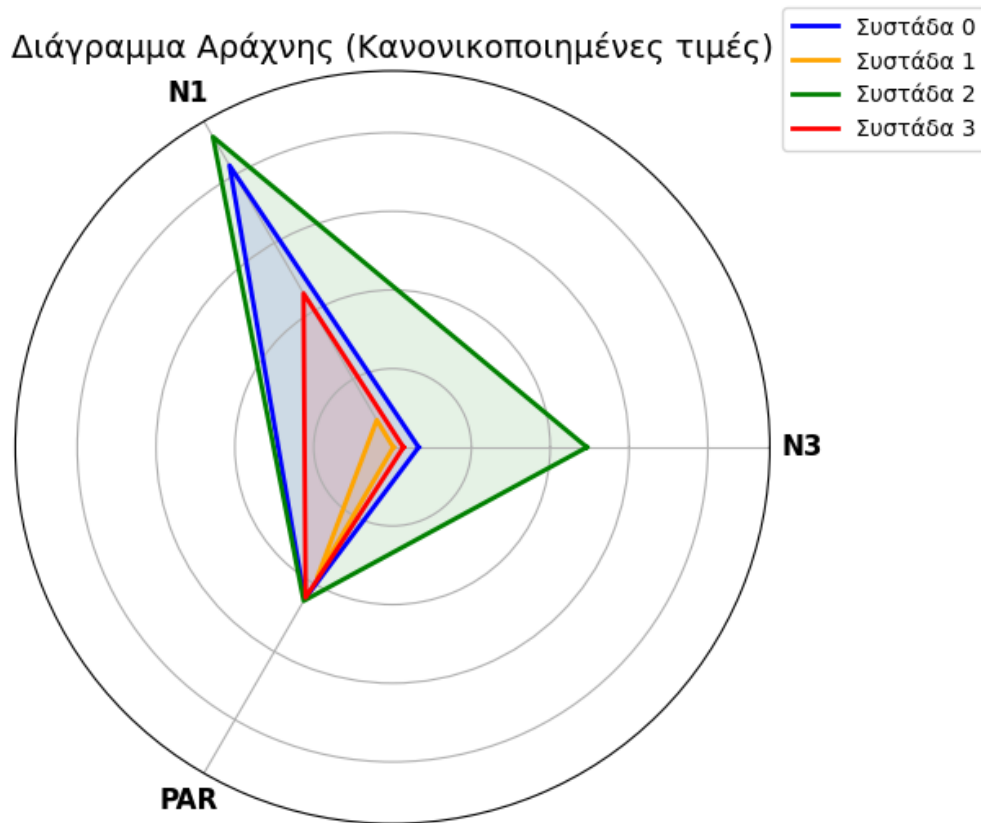
χαρακτηριστικά. Η διασπορά των τιμών είναι αισθητή, με εντοπισμένες περιπτώσεις αυξημένης απόκλισης. Η παρουσία των ακραίων τιμών παρουσιάζει παρόμοια συμπεριφορά και πιθανές ερμηνείες με αυτή που παρατηρήθηκε στις ακραίες τιμές της συστάδας 1.

Τα παραπάνω αποτελέσματα ενισχύουν τη διαφοροποίηση των ομάδων και τεκμηριώνουν την επιτυχία της ομαδοποίησης ως προς την αποτύπωση διαφορετικών προτύπων καταναλωτικής συμπεριφοράς. Βέβαια, κάποιοι καταναλωτές δεν έχουν απόλυτα διακριτή ομάδα. Επιπλέον, παρατηρείται ότι το χαρακτηριστικό που αφορά το ποσοστό κατανάλωσης τις ώρες αιχμής σε σχέση με τη συνολική ημερήσια, παρουσιάζει συγκρίσιμες τιμές μεταξύ όλων των ομάδων, γεγονός που επιβεβαιώνει την κοινή συμπεριφορά των καταναλωτών ως προς τις ώρες αιχμής. Το αποτέλεσμα αυτό υποδηλώνει ότι, ανεξαρτήτως συστάδας, καταγράφεται τάση αύξησης της κατανάλωσης κατά τις συγκεκριμένες ώρες, επιβεβαιώνοντας τον εντοπισμό της χρονικής περιόδου αιχμής.

Προκειμένου να αποτυπωθεί ολοκληρωμένα η καταναλωτική συμπεριφορά των ομάδων ακολουθεί η ανάλυση της σχέσης μεταξύ των δύο βασικότερων χαρακτηριστικών εισόδου (N_3 και N_1) μέσω Διαγράμματος Διασποράς (Scatter Plot), καθώς και η αναπαράσταση της συνολικής καταναλωτικής συμπεριφοράς κάθε συστάδας με τη βοήθεια Διαγράμματος Αράχνης (Radar Plot). Στο Σχήμα 5.15, απεικονίζεται το Διάγραμμα Διασποράς, όπου κάθε σημείο αντιστοιχεί σε έναν οικιακό καταναλωτή. Το Σχήμα 5.16 απεικονίζει τη μέση καταναλωτική συμπεριφορά κάθε συστάδας ως προς τα κανονικοποιημένα χαρακτηριστικά.



Σχήμα 5.15: Διάγραμμα Διασποράς μεταξύ N_3 και N_1 ανά συστάδα



Σχήμα 5.16: Διάγραμμα Αράχνης των κανονικοποιημένων χαρακτηριστικών ανά συστάδα

Στο Σχήμα 5.15 παρατηρείται ότι οι καταναλωτές που βρίσκονται στο επάνω δεξί τμήμα του γραφήματος συγκεντρώνουν υψηλές τιμές και στα δύο χαρακτηριστικά και συνεπώς παρουσιάζουν τη μεγαλύτερη καταλληλότητα για παροχή ευελιξίας, αποτελώντας ιδιαίτερο ενδιαφέρον για τον ΦοΣΕ Φορτίου, με κυρίαρχη τη συστάδα 2. Αμέσως επόμενη ως προς την ένταση και τη συχνότητα της κατανάλωσης αναδεικνύεται η συστάδα 0, η οποία παρουσιάζει επίσης αυξημένες τιμές. Στο Σχήμα 5.16 διαπιστώνεται ξανά ότι η συστάδα 2 εμφανίζει τις υψηλότερες τιμές σε όλα τα χαρακτηριστικά, επιβεβαιώνοντας την κυρίαρχη παρουσία της κατά τις ώρες αιχμής. Συνεπώς, οι καταναλωτές των ομάδων 0 και 2 μπορούν να θεωρηθούν οι πλέον κατάλληλοι υποψήφιοι για ενσωμάτωση σε προγράμματα ευελιξίας.

Προς πληρέστερη κατανόηση της κατανομής των δεδομένων, παρατίθεται η γραφική απεικόνιση της κατανομής πιθανότητας των τριών επιλεγμένων χαρακτηριστικών εισόδου. Το Σχήμα 5.17 παρουσιάζει τη συνολική κατανομή των κανονικοποιημένων τιμών για τα χαρακτηριστικά στο σύνολο των καταναλωτών, μέσω Διαγραμμάτων Πυκνότητας.



Σχήμα 5.17: Κατανομή πιθανότητας των κανονικοποιημένων χαρακτηριστικών N_3 , N_1 , και PAR

Από την παρατήρηση του Σχήματος 5.17 προκύπτουν τα εξής συμπεράσματα:

- Το χαρακτηριστικό N_3 εμφανίζει έντονη ασυμμετρία, με την πλειονότητα των τιμών συγκεντρωμένη κοντά στο μηδέν, γεγονός που φανερώνει ότι οι περισσότεροι καταναλωτές έχουν περιορισμένο αριθμό ημερών με αυξημένη κατανάλωση.
- Το χαρακτηριστικό N_1 παρουσιάζει διπολική μορφή κατανομής, με συσσωρεύσεις τόσο στις χαμηλές όσο και στις υψηλές τιμές, γεγονός που αποκαλύπτει την ύπαρξη διακριτών ομάδων χρηστών ως προς τη συχνότητα κατανάλωσης.
- Το χαρακτηριστικό PAR ακολουθεί κανονική κατανομή, με το μεγαλύτερο μέρος των τιμών συγκεντρωμένο γύρω από το μέσο, υποδεικνύοντας σταθερότητα στη σχετική κατανάλωση κατά τις ώρες αιχμής για το σύνολο των χρηστών.

Η παραπάνω ανάλυση προσφέρει πολύτιμη οπτική πληροφόρηση για τη στατιστική φύση των μεταβλητών και ενισχύει την κατανόηση της συμπεριφοράς των καταναλωτών πριν και μετά την εφαρμογή της ομαδοποίησης.

Η υλοποίηση των μεθόδων Elbow, Silhouette Score, Calinski–Harabasz, του αλγορίθμου K-Means, καθώς και της Ανάλυσης Κυρίων Συνιστωσών (PCA) γίνεται αξιοποιώντας τη βιβλιοθήκη Scikit-learn [26] της γλώσσας προγραμματισμού Python, ενώ η γραφική απεικόνιση όλων των διαγραμμάτων γίνεται με χρήση της βιβλιοθήκης Matplotlib [18]. Στα Σχήματα 5.18 και 5.19, παρουσιάζεται ο κώδικας που χρησιμοποιήθηκε σε αυτό το στάδιο της μεθοδολογίας.

```

# K-means Clustering
from sklearn.cluster import KMeans
from sklearn.metrics import silhouette_score, calinski_harabasz_score

# Define parameters for K-Means:
# init: Use 'k-means++' for smarter centroid initialization (improves convergence)
# n_init: Run the algorithm 10 times with different initializations and choose the best
# max_iter: Allow up to 450 iterations per run for convergence
# random_state: Set seed for reproducibility of results

kmeans_kwargs = {
    "init": "k-means++",
    "n_init": 10,
    "max_iter": 450,
    "random_state": 42
}

# Lists to store evaluation metrics
inertia = []
silhouette_scores = []
calinski_scores = []

# Test k values from 2 to 10
k_values = range(2, 11)
for k in k_values:
    kmeans = KMeans(n_clusters=k, **kmeans_kwargs)
    cluster_labels = kmeans.fit_predict(normalized_features)

    # Store metrics
    inertia.append(kmeans.inertia_)
    silhouette_score_value = silhouette_score(normalized_features, cluster_labels)
    silhouette_scores.append(silhouette_score_value)
    calinski_scores.append(calinski_harabasz_score(normalized_features, cluster_labels))

    # Print silhouette score for each k
    print(f"Silhouette Score for k={k}: {silhouette_score_value:.4f}")

# Elbow Method - Plot Inertia (WCSS)
plt.figure(figsize=(8, 5))
plt.plot(k_values, inertia, marker='o', label="Inertia")
plt.xlabel('Number of Clusters (k)')
plt.ylabel('Inertia (WCSS)')
plt.title('Elbow Method for Optimal Clusters')
plt.grid(True)
plt.show()

# Silhouette Score Plot
plt.figure(figsize=(8, 5))
plt.plot(k_values, silhouette_scores, marker='o', linestyle='--', color='g', label="Silhouette Score")
plt.xlabel('Number of Clusters (k)')
plt.ylabel('Silhouette Score')
plt.title('Silhouette Score for Optimal Clusters')
plt.grid(True)
plt.show()

# Calinski-Harabasz Score Plot
plt.figure(figsize=(8, 5))
plt.plot(k_values, calinski_scores, marker='o', linestyle='--', color='r', label="Calinski-Harabasz Score")
plt.xlabel('Number of Clusters (k)')
plt.ylabel('Calinski-Harabasz Score')
plt.title('Calinski-Harabasz Score for Optimal Clusters')
plt.grid(True)
plt.show()

```

Σχήμα 5.18: Υλοποίηση σε Python των μεθόδων Elbow, Silhouette, Calinski–Harabasz

```
# PCA Visualization
from sklearn.decomposition import PCA

pca = PCA(n_components=2)
pca_result = pca.fit_transform(normalized_features)
pca_df = pd.DataFrame(pca_result, columns=["PC1", "PC2"], index=features.index)
pca_df["Cluster"] = features["Cluster"]

plt.figure(figsize=(8, 6))
sns.scatterplot(x="PC1", y="PC2", hue="Cluster", palette="viridis", data=pca_df, alpha=0.7)
plt.title("PCA: Clusters in 2D Space")
plt.xlabel("Principal Component 1")
plt.ylabel("Principal Component 2")
plt.grid(True)
plt.show()
```

Σχήμα 5.19: Υλοποίηση σε Python της μεθόδου Ανάλυσης Κύριων Συνιστωσών

Μετά την ολοκληρωμένη ανάλυση των χαρακτηριστικών και την εφαρμογή του αλγορίθμου ομαδοποίησης, κάθε συστάδα αξιολογείται ως προς την καταλληλότητά της για συμμετοχή σε προγράμματα Απόκρισης Ζήτησης. Η αξιολόγηση βασίζεται στη συχνότητα και την ένταση της κατανάλωσης κατά τις ώρες αιχμής, καθώς και στην ομοιογένεια της συμπεριφοράς εντός κάθε ομάδας. Συγκεκριμένα:

- **Συστάδα 0:** Αν και η κατανάλωση κυμαίνεται σε μέτρια επίπεδα, παρατηρείται έντονη συχνότητα αιχμής. Το γεγονός αυτό καθιστά τους καταναλωτές της ομάδας αυτής εν δυνάμει υποψήφιους για συμμετοχή σε προγράμματα Απόκρισης Ζήτησης, καθώς η συχνότητα αποτελεί κρίσιμο παράγοντα για ένα ΦοΣΕ Φορτίου.
- **Συστάδα 1:** Η ομάδα αυτή χαρακτηρίζεται από χαμηλή συμμετοχή στις ώρες αιχμής και συνολικά περιορισμένη κατανάλωση. Συνεπώς, δεν θεωρείται κατάλληλη για ένταξη σε δράσεις ευελιξίας και δεν παρουσιάζει ιδιαίτερο ενδιαφέρον για ένα ΦοΣΕ Φορτίου.
- **Συστάδα 2:** Πρόκειται για την ομάδα με την υψηλότερη καταναλωτική δραστηριότητα κατά τις ώρες αιχμής. Οι καταναλωτές αυτής της συστάδας παρουσιάζουν υψηλές τιμές τόσο σε επίπεδο κατανάλωσης, όσο και στη συχνότητα εμφάνισης αυξημένης ζήτησης τις ώρες αιχμής. Ως εκ τούτου, θεωρούνται οι πλέον κατάλληλοι υποψήφιοι για ένταξη σε σχήματα ευέλικτης διαχείρισης ζήτησης.
- **Συστάδα 3:** Οι καταναλωτές αυτής της ομάδας εμφανίζουν μέτρια προς χαμηλή κατανάλωση, με ελαφρώς αυξημένη συχνότητα τις ώρες αιχμής. Ωστόσο, η ανομοιογένεια της συμπεριφοράς εντός της συστάδας μειώνει την αξιοπιστία της ομάδας για επιχειρησιακή αξιοποίηση. Μόνο ένας μικρός αριθμός καταναλωτών μπορεί να θεωρηθεί κατάλληλος για συμμετοχή.

Συνοψίζοντας, η παραπάνω κατηγοριοποίηση προσφέρει στοχευμένη πληροφόρηση για την αξιοποίηση των καταναλωτών από ένα ΦοΣΕ Φορτίου στο πλαίσιο της διαχείρισης ζήτησης. Με βάση την ένταση και τη συχνότητα της κατανάλωσης κατά τις ώρες αιχμής, καθώς και τη συνοχή της συμπεριφοράς εντός κάθε ομάδας, οι συστάδες 0 και 2 αναδεικνύονται ως οι πλέον κατάλληλες για συμμετοχή σε προγράμματα Απόκρισης

Ζήτησης. Συγκεκριμένα, η συστάδα 2 περιλαμβάνει τους καταναλωτές με τη μεγαλύτερη δυναμική, ενώ η συστάδα 0 εμφανίζει επίσης συνεπή καταναλωτική συμπεριφορά στις ώρες αιχμής, καθιστώντας τη σημαντική επόμενη επιλογή.

5.3 ΠΡΑΚΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

Στο παρόν κεφάλαιο, εξετάζεται η ένταξη των οικιακών καταναλωτών σε πραγματικά προγράμματα ευελιξίας. Προς τον σκοπό αυτό, αξιοποιούνται πραγματικά δεδομένα αιτημάτων καθοδικής ευελιξίας για την περιοχή του Λονδίνου, όπως δημοσιεύονται από τον διαχειριστή δικτύου διανομής ηλεκτρικής ενέργειας UK Power Networks [28]. Επιλέγονται δέκα αντιπροσωπευτικά αιτήματα, βάσει των οποίων πραγματοποιείται φιλτράρισμα του αρχικού συνόλου δεδομένων. Τα αιτήματα αυτά προέρχονται από το έτος 2024 και αφορούν τοπικές αγορές ευελιξίας. Για τη διατήρηση της χρονικής συμβατότητας με το υπόλοιπο σύνολο δεδομένων κατανάλωσης, επιλέγονται αιτήματα χαμηλής ισχύος και προσαρμόζονται σε αντίστοιχες ημερομηνίες του έτους 2013, χωρίς βλάβη της γενικότητας της ανάλυσης. Τα αιτήματα παρουσιάζονται στον Πίνακα 5.1.

Πίνακας 5.1: Αιτήματα ευελιξίας από τον διαχειριστή UK Power Networks για επιλεγμένες ώρες που αντιστοιχούν στην αιχμή καταναλωτών της παρούσας διπλωματικής εργασίας

Α/Α αιτήματος ευελιξίας	Ημερομηνία	Χρονική διάρκεια (h)	Ωριαία Ισχύς (kW)	Ζητούμενη Ενέργεια (kWh)
1	5/2/2013	[17–19]	1268	3804
2	4/2/2013	[17–19]	1268	3804
3	4/2/2013	[17–19]	934	2802
4	5/2/2013	[17–19]	934	2802
5	4/2/2013	[17–19]	651	1953
6	5/2/2013	[17–19]	651	1953
7	4/2/2013	[17–19]	630	1890
8	5/2/2013	[17–19]	630	1890
9	4/2/2013	[17–19]	599	1797
10	5/2/2013	[17–19]	599	1797

Με βάση τα χρονικά διαστήματα των αιτημάτων αυτών, διαμορφώνονται δύο σενάρια πρακτικών εφαρμογών. Στο πρώτο σενάριο, εξετάζεται η κατάλληλη επιλογή καταναλωτών χωρίς ομαδοποίηση, ενώ στο δεύτερο υλοποιείται στοχευμένη επιλογή καταναλωτών για παροχή ευελιξίας με βάση τα αποτελέσματα της μοντελοποίησης της ομαδοποίησης. Μέσω της σύγκρισης των δύο προσεγγίσεων, αναδεικνύεται η αξία της ομαδοποίησης μέσω της κατάλληλης επιλογής συμμετεχόντων με την καλύτερη δυνατότητα παροχής ευελιξίας.

Στο Σενάριο Α, η επιλογή των καταναλωτών για την κάλυψη κάθε αιτήματος ευελιξίας βασίζεται αποκλειστικά στη στιγμιαία κατανάλωσή τους ιστορικά κατά την ακριβή ώρα αυτού. Πιο συγκεκριμένα, οι καταναλωτές ταξινομούνται σε φθίνουσα σειρά με βάση τη μετρούμενη κατανάλωση και επιλέγονται διαδοχικά, ξεκινώντας από τον περισσότερο ενεργοβόρο, μέχρι να επιτευχθεί ο απαιτούμενος στόχος ισχύος. Στο Σχήμα 5.20 παρουσιάζονται τα αποτελέσματα της ανάλυσης, όπου καταγράφονται η χρονική περίοδος του αιτήματος, ο απαιτούμενος στόχος σε MW, ο αριθμός των καταναλωτών που επιλέγονται για να το καλύψουν, καθώς και ο αριθμός αυτών που ανήκουν στις συστάδες 0 και 2, δηλαδή στις

ομάδες που έχουν αναδειχθεί ως κατάλληλες για συμμετοχή σε δράσεις ευελιξίας. Η τελευταία αυτή τιμή επιτρέπει τη σύγκριση με το Σενάριο Β, στο οποίο αξιοποιείται η πληροφορία της ομαδοποίησης.

Σενάριο Α: χωρίς ομαδοποίηση							
A/A	Έτος	Μήνας	Ημέρα	Χρονική διάρκεια (h)	Στόχος kWh	Καταναλωτές	Συστάδες 0,2
1	2013	2	5	[17-19]	3804	314	271
2	2013	2	4	[17-19]	3804	346	292
3	2013	2	4	[17-19]	2802	224	198
4	2013	2	5	[17-19]	2802	205	178
5	2013	2	4	[17-19]	1953	137	122
6	2013	2	5	[17-19]	1953	125	111
7	2013	2	4	[17-19]	1890	132	119
8	2013	2	5	[17-19]	1890	120	107
9	2013	2	4	[17-19]	1797	123	112
10	2013	2	5	[17-19]	1797	112	99

Σχήμα 5.20: Αιτήματα ευελιξίας για το Σενάριο Α, όπου δεν έχει γίνει ομαδοποίηση

Για παράδειγμα, στο αίτημα 1 του Σεναρίου Α, απαιτείται ενέργεια ίση με 3804 kWh, η οποία καλύπτεται από 314 καταναλωτές, εκ των οποίων οι 271 ανήκουν στις συστάδες 0 ή 2. Ενδεικτικά, ο καταναλωτής με τη μεγαλύτερη δυνατότητα παροχής ευελιξίας τις ώρες 17:00 έως 20:00 είναι ο καταναλωτής με αναγνωριστικό MAC000105 και διαθέσιμη ενέργεια 35,69 kWh. Η πλήρης κάλυψη του αιτήματος απαιτεί και τη συμμετοχή καταναλωτών με χαμηλότερο δυναμικό, όπως ο καταναλωτής με αναγνωριστικό MAC001445, του οποίου η συνεισφορά ανέρχεται σε 8,58 kWh.

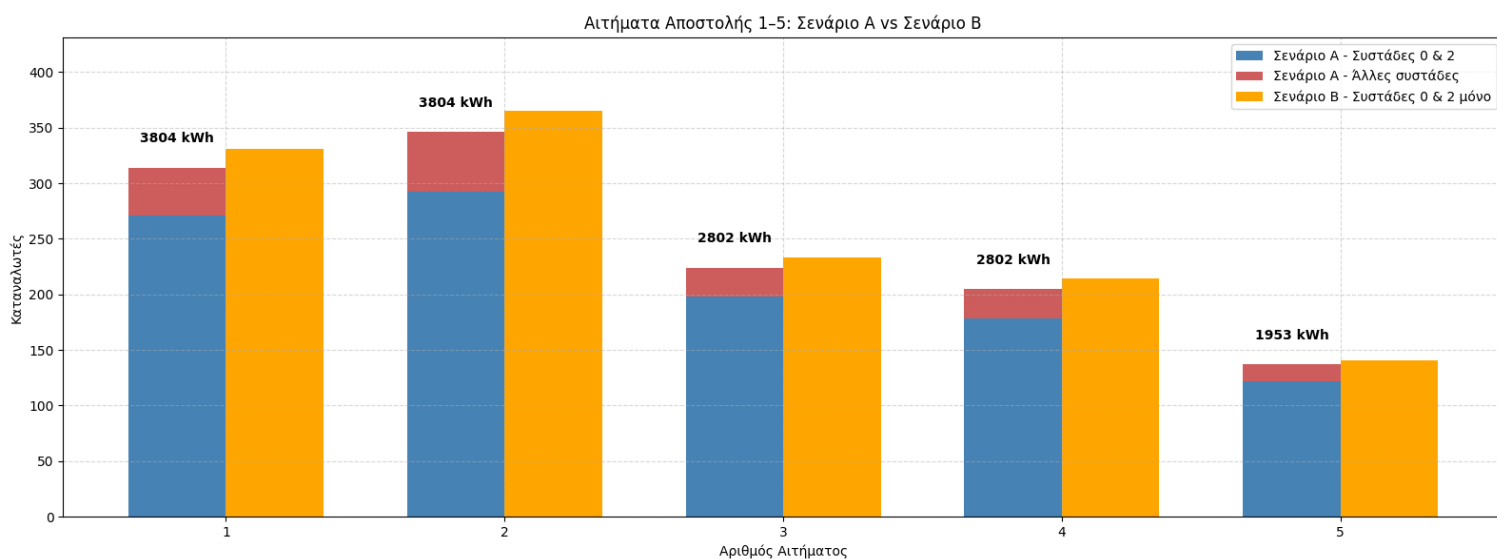
Στο Σενάριο Β, η διαδικασία επιλογής των καταναλωτών ακολουθεί την ίδια μεθοδολογία με το Σενάριο Α, δηλαδή οι καταναλωτές ταξινομούνται σε φθίνουσα σειρά βάσει της στιγμιαίας τους κατανάλωσης τη χρονική στιγμή του αιτήματος και επιλέγονται διαδοχικά έως την κάλυψη του απαιτούμενου στόχου ισχύος. Ωστόσο, σε αντίθεση με την προηγούμενη προσέγγιση, εδώ επιτρέπεται η επιλογή καταναλωτών μόνο από τις συστάδες 0 και 2, οι οποίες, σύμφωνα με τα αποτελέσματα της ομαδοποίησης που παρουσιάστηκαν στην Ενότητα 5.2, έχουν χαρακτηριστεί ως οι πλέον κατάλληλες για συμμετοχή σε δράσεις ευελιξίας. Με τον τρόπο αυτό, αξιοποιείται η πληροφορία της ομαδοποίησης προκειμένου να διαμορφωθεί μια πιο στοχευμένη και αποδοτική ενεργοποίηση, περιορίζοντας τη συμμετοχή σε καταναλωτές με αυξημένη πιθανότητα θετικής ανταπόκρισης σε ενεργοποίηση από ένα ΦοΣΕ Φορτίου σε κάποιο αίτημα. Στο Σχήμα 5.21 παρουσιάζονται τα αντίστοιχα αποτελέσματα της εν λόγω ανάλυσης.

Σενάριο Β: με ομαδοποίηση						
A/A	Έτος	Μήνας	Ημέρα	Χρονική διάρκεια (h)	Στόχος kWh	Καταναλωτές (Συστάδες 0 & 2 μόνο)
1	2013	2	5	[17-19]	3804	331
2	2013	2	4	[17-19]	3804	365
3	2013	2	4	[17-19]	2802	233
4	2013	2	5	[17-19]	2802	214
5	2013	2	4	[17-19]	1953	141
6	2013	2	5	[17-19]	1953	129
7	2013	2	4	[17-19]	1890	135
8	2013	2	5	[17-19]	1890	124
9	2013	2	4	[17-19]	1797	126
10	2013	2	5	[17-19]	1797	116

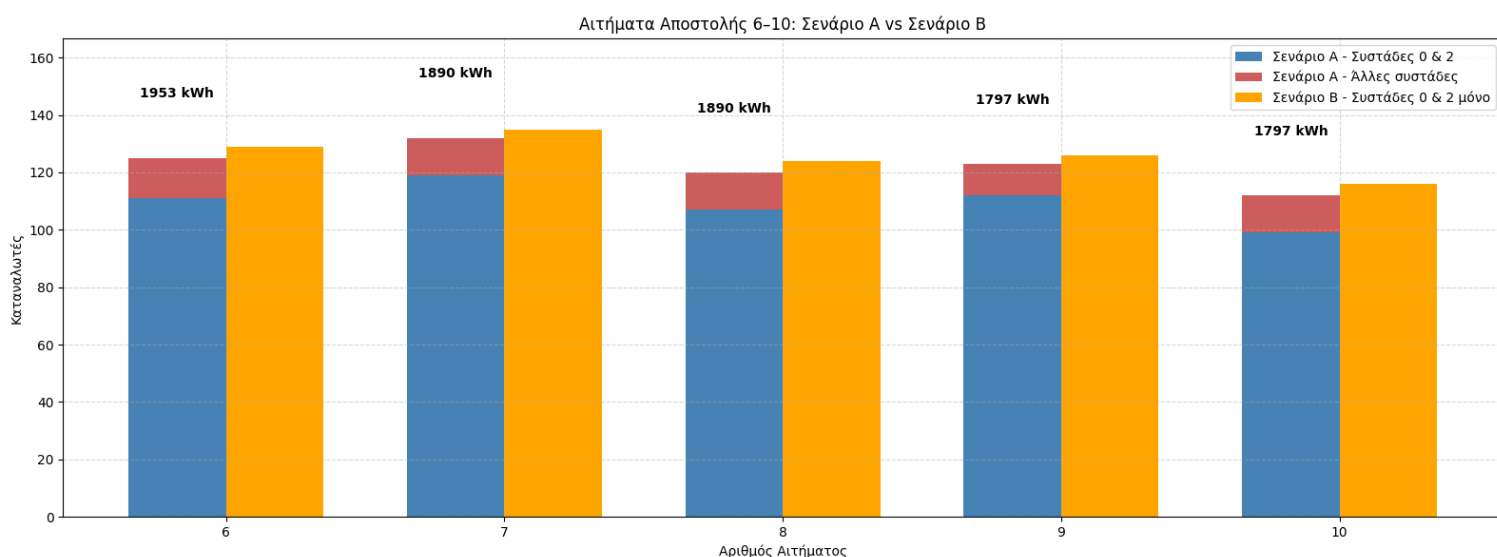
Σχήμα 5.21: Αιτήματα ευελιξίας για το Σενάριο Β, όπου έχει γίνει ομαδοποίηση

Σύμφωνα με το Σχήμα 5.21, στο αίτημα 1 του Σεναρίου Β, απαιτείται ενέργεια ίση με 3804 kWh, η οποία καλύπτεται από 331 καταναλωτές, εκ των οποίων όλοι ανήκουν στις συστάδες 0 ή 2. Ενδεικτικά, ο καταναλωτής με τη μεγαλύτερη δυνατότητα παροχής ευελιξίας τις ώρες 17:00 έως 20:00 είναι ο καταναλωτής με αναγνωριστικό MAC000105 και διαθέσιμη ενέργεια 35,69 kWh. Η πλήρης κάλυψη του αιτήματος απαιτεί και τη συμμετοχή καταναλωτών με χαμηλότερο δυναμικό, όπως ο καταναλωτής με αναγνωριστικό MAC005315, του οποίου η συνεισφορά ανέρχεται σε 7,96 kWh.

Για την καλύτερη αποτύπωση των διαφορών μεταξύ των δύο σεναρίων, παρατίθενται στα Σχήματα 5.22 και 5.23, συγκριτικό γράφημα των αποτελεσμάτων για κάθε ένα από τα αιτήματα ευελιξίας. Προς διευκόλυνση της ανάγνωσης, το σύνολο των δέκα αιτημάτων έχει χωριστεί σε δύο διαγράμματα των πέντε αιτημάτων έκαστο, διατηρώντας τη δομή και τη σύγκριση των σεναρίων αναλλοίωτη. Στον οριζόντιο άξονα απεικονίζεται ο αύξων αριθμός του αιτήματος, ενώ κάθε στήλη αντιστοιχεί στον αριθμό των καταναλωτών που απαιτούνται για την επίτευξη του αντίστοιχου στόχου σε kWh. Οι στήλες του Σεναρίου Α χωρίζονται επιπλέον σε δύο υποκατηγορίες: οι μπλε μπάρες αναπαριστούν τον αριθμό των καταναλωτών που ανήκουν στις συστάδες 0 και 2, ενώ οι κόκκινες αφορούν καταναλωτές από τις υπόλοιπες συστάδες. Αντίθετα, στο Σενάριο Β, οι πορτοκαλί μπάρες αφορούν αποκλειστικά καταναλωτές από τις συστάδες 0 και 2.



Σχήμα 5.22: Σύγκριση Σεναρίων A και B για τα αιτήματα ευελιξίας 1 έως 5



Σχήμα 5.23: Σύγκριση Σεναρίων A και B για τα αιτήματα ευελιξίας 6 έως 10

Το γράφημα καταδεικνύει τη συγκριτική απόδοση των δύο προσεγγίσεων. Παρατηρείται ότι, για ίσους στόχους παροχής ευελιξίας, ο αριθμός καταναλωτών που απαιτείται στο Σενάριο B είναι πάντα ίσος ή μεγαλύτερος σε σχέση με το συνολικό πλήθος του Σεναρίου A. Αυτό είναι αναμενόμενο καθώς στο σενάριο A επιλέγονται οι καταναλωτές με τη μεγαλύτερη κατανάλωση άρα και δυνατότητα παροχής ευελιξίας, με την υπόθεση ότι είναι διατεθειμένοι να μεταφέρουν πλήρως την κατανάλωσή τους σε άλλες ώρες της ημέρας. Αντίθετα, στο Σενάριο B επιλέγονται εκείνοι οι καταναλωτές του Σεναρίου A, οι οποίοι ανήκουν στις συστάδες 0 και 2, με την μεγαλύτερη δυνατότητα παροχής ευελιξίας. Συμπληρωματικά, επιλέγονται οι αμέσως επόμενοι καταναλωτές που ανήκουν στις συστάδες αυτές ώστε να καλυφθεί το ύψος του αιτήματος ευελιξίας. Ωστόσο, το σύνολο των καταναλωτών στο Σενάριο B προέρχεται αποκλειστικά από ομάδες υψηλής καταλληλότητας,

γεγονός που ενισχύει τη λειτουργική αξιοπιστία της ενεργοποίησης και καθιστά την προσέγγιση με ομαδοποίηση πιο αποτελεσματική.

Στο τελευταίο στάδιο της παρούσας ανάλυσης, επιχειρείται η ποσοτική αποτίμηση της ενεργειακής συνεισφοράς των καταναλωτών, που χαρακτηρίστηκαν ως μη κατάλληλοι για συμμετοχή σε δράσεις ευελιξίας (καταναλωτές εκτός των συστάδων 0 και 2). Η αξιολόγηση αφορά αποκλειστικά το Σενάριο Α, όπου η επιλογή καταναλωτών πραγματοποιείται ανεξαρτήτως συστάδας και εστιάζει στους καταναλωτές που συμμετείχαν στα αιτήματα, αλλά δεν ανήκουν στις προτεινόμενες συστάδες. Σκοπός της ανάλυσης είναι να εκτιμηθεί ο βαθμός συνέπειας των καταναλωτών που δεν ανήκουν στις αξιόπιστες συστάδες 0 και 2, και επομένως η πιθανότητα αξιόπιστης ενεργοποίησης των καταναλωτών αυτών από ένα ΦοΣΕ Φορτίου. Αν και αυτοί οι καταναλωτές είχαν ιστορικά παρουσιάσει υψηλά επίπεδα κατανάλωσης ενέργειας, προκύπτει ότι τέτοια γεγονότα δεν αποτελούν συστηματική ή προβλέψιμη συμπεριφορά με βάση το ιστορικό τους προφίλ. Επομένως, η επιλογή τέτοιων καταναλωτών από τους ΦοΣΕ Φορτίου μπορεί να θέσει σε κίνδυνο την ενεργειακή κάλυψη του εκάστοτε αιτήματος ευελιξίας. Για τον σκοπό της ανάλυσης αυτής, υπολογίζεται μια μετρική για τους εν λόγω «κόκκινους», με βάση τα Σχήματα 5.22 και 5.23, καταναλωτές, στηριζόμενη στα εξής βήματα:

1. Επιλέγεται η χρονική περίοδος κατά την οποία υποβλήθηκαν τα δέκα αιτήματα ευελιξίας.
2. Καταγράφονται, για κάθε αίτημα, η απαιτούμενη ζήτηση σε kWh, το πλήθος των συνολικών καταναλωτών που συμμετέχουν, καθώς και ο αριθμός των «κόκκινων» καταναλωτών (εκτός συστάδων 0 και 2).
3. Υπολογίζεται η ποσότητα κατανάλωσης που καλύπτει η υποομάδα των «κόκκινων» καταναλωτών P_{red} .
4. Για κάθε αίτημα, υπολογίζεται το πλήθος των ημερών εντός του 2013, στις οποίες η κατανάλωση των συγκεκριμένων «κόκκινων» καταναλωτών υπερβαίνει την τιμή ισχύος P_{red} που δίνουν στο αίτημα, σύμφωνα με την παρακάτω μαθηματική εξίσωση (5.1):

$$P_{historical}(avg. consumption from 17 - 20) > P_{red} \quad (5.1)$$

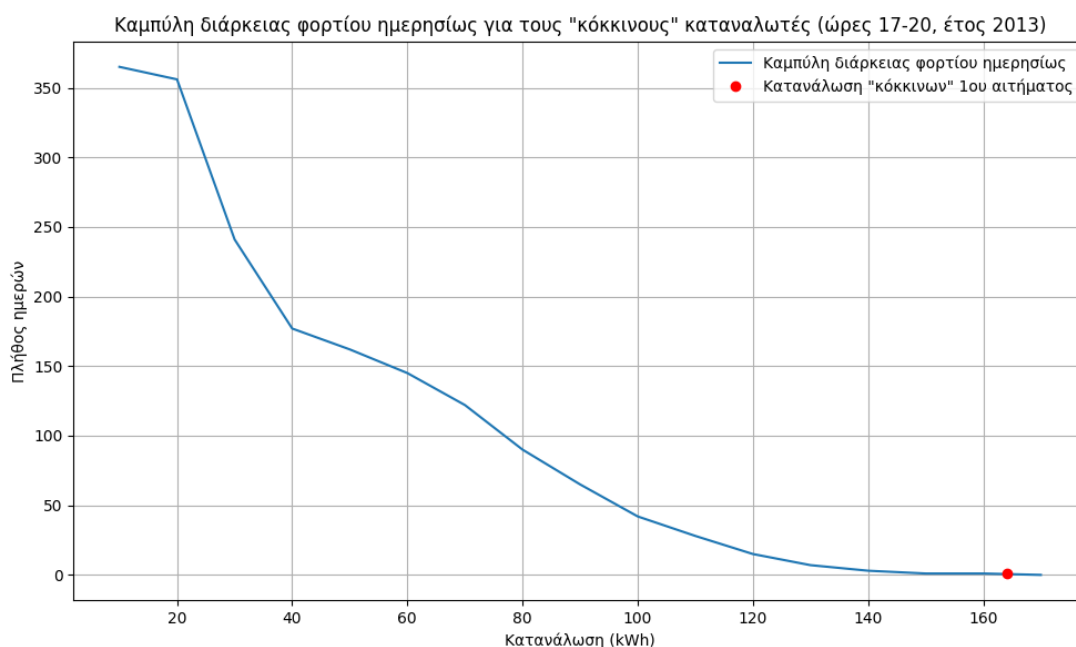
5. Τέλος υπολογίζεται ο δείκτης συνέπειας με βάση την ιστορικότητα των καταναλωτών αυτών, δηλαδή το ποσοστό πιθανότητας εμφάνισης κατανάλωσης μεγαλύτερης από P_{red} σε όλο το έτος 2013.

Τα αποτελέσματα της ανάλυσης απεικονίζονται στον Πίνακα 5.2.

Πίνακας 5.1: Δείκτης συνέπειας με βάση την ιστορικότητα των «κόκκινων» καταναλωτών

Α/Α αιτήματος ευελιξίας	P_{red} (kWh)	Δείκτης συνέπειας (%)
1	164,18	0,27
2	177,04	0,27
3	100,7	0,27
4	115,22	0,27
5	65,39	0,27
6	69,03	0,27
7	58,26	0,27
8	65,11	0,27
9	50,92	0,55
10	65,11	0,27

Προκειμένου να οπτικοποιηθεί η εκτίμηση της αξιοπιστίας των καταναλωτών που δεν ανήκουν στις συστάδες 0 και 2, στο Σχήμα 5.24 απεικονίζεται, ενδεικτικά, η καμπύλη διάρκειας αθροιστικού φορτίου για τους «κόκκινους» καταναλωτές του 1^{ου} αιτήματος κατά τις ώρες αιχμής 17-20.

**Σχήμα 5.24:** Κατανομή ημερών με υπέρβαση κατανάλωσης για τους «κόκκινους» καταναλωτές του 1^{ου} αιτήματος (ώρες 17–20, έτος 2013)

Από τα αποτελέσματα της ανάλυσης διαπιστώνεται ότι η αξιοπιστία ενεργοποίησης των καταναλωτών που δεν ανήκουν στις συστάδες 0 και 2 είναι ιδιαίτερα περιορισμένη. Όπως προκύπτει από τον Πίνακα 5.2, ο δείκτης συνέπειας για όλα τα αιτήματα είναι πολύ

χαμηλός, γεγονός που υποδηλώνει ότι σε ελάχιστες ημέρες του έτους 2013 οι εν λόγω καταναλωτές παρουσιάζουν επαρκή επίπεδα κατανάλωσης, ώστε να ανταποκριθούν στο εκάστοτε αίτημα. Η εικόνα αυτή αποτυπώνεται και στο Σχήμα 5.24, όπου παρατηρείται ότι η συνολική ενέργεια που απαιτείται από τους συγκεκριμένους καταναλωτές στο πρώτο αίτημα υπερβαίνεται ιστορικά σε μόνο μία ημέρα εντός του έτους. Τα ευρήματα ενισχύουν τη σημασία της στοχευμένης επιλογής καταναλωτών βάσει ομαδοποίησης, καθώς η ενεργοποίηση συμμετεχόντων με περιορισμένη ιστορική συνέπεια ενδέχεται να μειώσει την αξιοπιστία εκπλήρωσης των απαιτήσεων ευελιξίας.

ΣΥΜΠΕΡΑΣΜΑΤΑ - ΕΠΕΚΤΑΣΕΙΣ

6.1 ΣΥΝΟΨΗ ΤΗΣ ΕΡΓΑΣΙΑΣ ΚΑΙ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΤΗΣ

Στην παρούσα διπλωματική εργασία μελετήθηκε η ομαδοποίηση οικιακών καταναλωτών ηλεκτρικής ενέργειας με σκοπό την αξιολόγηση της ικανότητάς τους να παρέχουν ευελιξία στο σύστημα διανομής ηλεκτρικής ενέργειας, στο πλαίσιο της ενεργειακής μετάβασης και των αυξανόμενων απαιτήσεων των δικτύων διανομής. Με τη χρήση πραγματικών δεδομένων κατανάλωσης από έξυπνους μετρητές, αναπτύχθηκε μεθοδολογία ταξινόμησης των καταναλωτών σε ομάδες με παρόμοια πρότυπα συμπεριφοράς, επιτρέποντας την ταυτοποίηση των καταλληλότερων για συμμετοχή σε δράσεις Απόκρισης Ζήτησης.

Η ανάλυση βασίστηκε σε σύνολο δεδομένων για 5.567 οικιακούς καταναλωτές στο Λονδίνο, με χρονική ανάλυση μισής ώρας για τη χρονική περίοδο 2011–2014. Μετά από κατάλληλη επεξεργασία, η τελική ανάλυση ομαδοποίησης πραγματοποιήθηκε για 4.247 καταναλωτές για το έτος 2013. Η διαδικασία ομαδοποίησης υλοποιήθηκε με τον αλγόριθμο K-Means, χρησιμοποιώντας τη γλώσσα προγραμματισμού Python. Για την εκτίμηση του βέλτιστου αριθμού συστάδων εφαρμόστηκαν μέθοδοι, όπως η Elbow, καθώς και δείκτες αξιολόγησης, όπως οι Silhouette και Calinski-Harabasz.

Αρχικά υπολογίστηκαν οκτώ διαφορετικά χαρακτηριστικά που αποτυπώνουν τη συμπεριφορά κατά τις περιόδους αιχμών, ωστόσο έπειτα από ανάλυση ευαισθησίας επιλέχθηκαν τελικά τρία χαρακτηριστικά, τα οποία παρουσίασαν τη μεγαλύτερη διακριτική ικανότητα. Ο βέλτιστος αριθμός συστάδων προέκυψε ίσος με τέσσερα, εξασφαλίζοντας ικανοποιητικό διαχωρισμό και ερμηνευσιμότητα των προφίλ κατανάλωσης φορτίου.

Αναλύοντας τα προφίλ που προέκυψαν από την ομαδοποίηση, διαπιστώθηκε ότι οι καταναλωτές της Συστάδας 2 (362 νοικοκυριά) διαθέτουν έντονη και συχνή δραστηριότητα κατά τις ώρες αιχμής, γεγονός που τους καθιστά την καταλληλότερη ομάδα για στοχευμένη συμμετοχή σε προγράμματα ευελιξίας. Παράλληλα, οι 1.238 καταναλωτές της Συστάδας 0, που επίσης παρουσιάζουν σημαντική παρουσία κατά τις περιόδους αιχμής, αποτελούν έναν συμπληρωματικό πυρήνα. Συνολικά, διαμορφώνεται ένα δυναμικό σύνολο περίπου 1.600 καταναλωτών, το οποίο μπορεί να αξιοποιηθεί επιχειρησιακά από έναν ΦοΣΕ Φορτίου για τη διαχείριση αιχμών. Αντίθετα, η Συστάδα 1 χαρακτηρίζεται από χαμηλή κατανάλωση και απουσία ενδιαφέροντος, ενώ η Συστάδα 3 ενδέχεται να χρησιμοποιηθεί συμπληρωματικά σε στοχευμένα, βραχυχρόνια σχήματα Απόκρισης Ζήτησης.

Στη συνέχεια, για την αξιολόγηση της πρακτικής ωφελιμότητας της ομαδοποίησης, εξετάστηκε η εφαρμογή των παραπάνω αποτελεσμάτων σε πραγματικά αιτήματα ευελιξίας του έτους 2024, όπως δημοσιεύονται από τον διαχειριστή UK Power Networks. Επιλέχθηκαν αιτήματα με χαρακτηριστικά ισχύος συμβατά με τα υπό μελέτη προφίλ κατανάλωσης και αντιστοιχίστηκαν σε ημερομηνίες του 2013, ώστε να διατηρηθεί η χρονική συνοχή με τα

διαθέσιμα δεδομένα κατανάλωσης, χωρίς να αλλοιώνεται η γενικότητα της ανάλυσης. Διαμορφώθηκαν δύο εναλλακτικά σενάρια:

- Σενάριο Α (χωρίς ομαδοποίηση): η επιλογή καταναλωτών έγινε αποκλειστικά βάσει της στιγμιαίας ιστορικής κατανάλωσής τους την ώρα του αιτήματος με φθίνουσα σειρά, θεωρώντας πλήρη συμμόρφωση μετακίνησης φορτίου για παροχή ευελιξίας.
- Σενάριο Β (με ομαδοποίηση): επιλέχθηκαν μόνο καταναλωτές από τις συστάδες 0 και 2, εφαρμόζοντας την ίδια κατάταξη βάσει στιγμιαίας κατανάλωσης με φθίνουσα σειρά.

Τα αποτελέσματα ανέδειξαν ότι η επιλογή των συστάδων 0 και 2 βασίστηκε στη διαπιστωμένη παρουσία στιγμιαίας υψηλής κατανάλωσης κατά τις ώρες αιχμής, γεγονός που συνδέεται άμεσα με τη δυνατότητα συμμετοχής σε προγράμματα ευελιξίας. Η ανάλυση επιβεβαίωσε πως οι καταναλωτές εκτός των δύο αυτών συστάδων εμφανίζουν χαμηλή συχνότητα κατανάλωσης υψηλής ενέργειας σε κρίσιμες χρονικές περιόδους, γεγονός που τεκμηριώνει τον περιορισμένο ρόλο τους στην παροχή εκτιμώμενης ευελιξίας. Η διαφοροποίηση αυτή δεν είναι μόνο στατιστικά εμφανής, αλλά και λειτουργικά κρίσιμη, καθώς ενισχύει την αξιοπιστία επιλογής των κατάλληλων καταναλωτών για εφαρμογές ΦοΣΕ Φορτίου και δράσεις Απόκρισης Ζήτησης.

Συνοψίζοντας, η διπλωματική εργασία ανέδειξε ότι η ανάλυση καταναλωτικών προτύπων μέσω ομαδοποίησης και η στοχευμένη ενεργοποίηση καταναλωτών ενισχύουν τη διαχειρισσιμότητα των προγραμμάτων ευελιξίας, προσφέροντας πολύτιμη υποστήριξη στις λειτουργίες ενός ΦοΣΕ Φορτίου στο πλαίσιο της μετάβασης σε ένα πιο ευέλικτο και αποδοτικό σύστημα διανομής ηλεκτρικής ενέργειας.

6.2 ΠΙΘΑΝΕΣ ΕΠΕΚΤΑΣΕΙΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Η παρούσα διπλωματική εργασία μπορεί μελλοντικά να επεκταθεί ως ακολούθως:

- Διερεύνηση εναλλακτικών τεχνικών ομαδοποίησης, όπως ιεραρχικές (hierarchical), η μέθοδος DBSCAN και οι αλγόριθμοι fuzzy clustering. Η εφαρμογή και σύγκριση των τεχνικών αυτών σε διαφορετικά προφίλ κατανάλωσης ενδέχεται να προσφέρει μεγαλύτερη ερμηνευσιμότητα και βελτιωμένη απόδοση, επιτρέποντας την αναγνώριση καταναλωτικών ομάδων με μεγαλύτερη ακρίβεια και σταθερότητα.
- Συνδυασμός της μεθοδολογίας ομαδοποίησης με προβλεπτικά μοντέλα ευελιξίας σε πραγματικό χρόνο, όπως είναι τα μοντέλα μηχανικής μάθησης ή οι μέθοδοι χρονικών σειρών. Με αυτόν τον τρόπο, θα είναι δυνατή η εκ των προτέρων εκτίμηση της πιθανής συμμετοχής κάθε ομάδας καταναλωτών σε αίτημα ευελιξίας, γεγονός που θα ενισχύσει τη λειτουργική αξιοπιστία των ΦοΣΕ Φορτίου και την αποδοτικότητα των ενεργειακών παρεμβάσεων.
- Ενσωμάτωση κοινωνικοοικονομικών χαρακτηριστικών, όπως εισόδημα, τύπος κατοικίας ή ενεργειακή ευαισθητοποίηση, με στόχο την κατανόηση των παραγόντων που επηρεάζουν τη διάθεση συμμετοχής σε προγράμματα ευελιξίας. Η συσχέτιση των τεχνικών δεδομένων κατανάλωσης με κοινωνικά και οικονομικά κίνητρα μπορεί να οδηγήσει σε αποτελεσματικότερα και πιο δίκαια προγράμματα Απόκρισης Ζήτησης.

- Διερεύνηση της εφαρμογής της προτεινόμενης μεθοδολογίας σε καταναλωτές από διαφορετικά γεωγραφικά και αγοραστικά περιβάλλοντα, με σκοπό την αποτίμηση της προσαρμοστικότητάς της σε ετερογενή δεδομένα και την αξιολόγηση των αποτελεσμάτων υπό διαφορετικές συνθήκες. Η επέκταση αυτή δύναται να ενισχύσει τη δυνατότητα γενίκευσης των συμπερασμάτων και να υποστηρίξει την ανάπτυξη ευέλικτων πλαισίων ομαδοποίησης για ευρύτερη εφαρμογή.
- Αναζήτηση και ενσωμάτωση πιο πρόσφατων δεδομένων κατανάλωσης για τη συγκριτική αξιολόγηση των αποτελεσμάτων. Μέσω της χρήσης κατάλληλων συντελεστών προσαρμογής ή της άμεσης απόκτησης επικαιροποιημένων μετρήσεων, μπορεί να εξεταστεί η διαχρονική εγκυρότητα των συστάδων και η προσαρμοστικότητα της μεθοδολογίας σε σύγχρονες τάσεις κατανάλωσης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1]. H. Golmohamadi, “Demand-Side Flexibility in Power Systems: A Survey of Residential, Industrial, Commercial, and Agricultural Sectors,” *Sustainability*, Vol. 14, No. 13, 7916, June 2022.
- [2]. C. Menos-Aikateriniadis, I. Lamprinos and P.S. Georgilakis, “Particle swarm optimization in residential demand-side management: A review on scheduling and control algorithms for demand response provision,” *Energies*, Vol. 15, No. 6, 2211, March 2022.
- [3]. R. Müller, “Grid congestion: what is it and how do we solve it?,” *Tibo Energy*, April 2024, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://tibo.energy/blog/grid-congestion-explained/>
- [4]. N. Χατζηαργυρίου, “Μια προσεχτική ματιά στο Blackout της Ιβηρικής Χερσονήσου,” *Energypress*, May 2025, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://energypress.gr/news/mia-prosehtiki-matia-sto-blackout-tis-ibirikis-hersonisoy>
- [5]. M. Hojjat, M.H. Javidi and D. Yazdanpanah, “Stochastic congestion management considering power system uncertainties: a chance-constrained programming approach,” *Turkish Journal of Electrical Engineering and Computer Sciences*, Vol. 24, No. 1, pp. 61-75, January 2016.
- [6]. A. Corliss, “Introduction to energy system flexibility,” *NationalgridESO*, 2020, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://www.neso.energy/document/189851/download>
- [7]. P. Ponnaganti, R. Sinha, J.R. Pillai and B. Bak-Jensen, “Flexibility provisions through local energy communities: A review,” *Next Energy*, Vol. 1, No. 2, 100022, June 2023.
- [8]. S. Ostovar, M. Moeini-Aghaie and M.B. Hadi, “Flexibility provision of residential energy hubs with demand response applications,” *IET Generation, Transmission & Distribution*, Vol. 16, No. 8, pp. 1668-1679, April 2022.
- [9]. Q. Zhang, and J. Li, “Demand response in electricity markets: A review,” In *Proc. of 2012 International Conference on the European Energy market*, Florence, Italy, 10–12 May 2012.
- [10]. H. Rashid, P. Singh and K. Ramamritham, “Revisiting selection of residential consumers for demand response programs,” In *Proc. of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments*, New York, USA, 8 Nov. 2017.
- [11]. J.C. Holyhead, S.D. Ramchurn and A. Rogers, “Consumer targeting in residential demand response programmes,” In *Proc. of the 2015 ACM sixth International Conference on Future Energy Systems*, New York, USA, 14 July 2015.

- [12]. R. Kaur and D. Gabrijelčič, "Behavior segmentation of electricity consumption patterns: A cluster analytical approach," *Knowledge-based systems*, Vol. 251, 109236, September 2022.
- [13]. UK Power Networks, "SmartMeter Energy Use Data in London Households," *Opinion Research and General Statistics (GLA)*, June 2015, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households>
- [14]. Python Documentation, "General Python FAQ," [Online]. Διαθέσιμο στην ιστοσελίδα: <https://docs.python.org/3/faq/general.html#python-in-the-real-world>
- [15]. Pandas, "Pandas Documentation," [Online]. Διαθέσιμο στην ιστοσελίδα: <https://pandas.pydata.org/docs/>
- [16]. NumPy, "NumPy Documentation," [Online]. Διαθέσιμο στην ιστοσελίδα: <https://numpy.org/doc/>
- [17]. A.J. Pimm, T.T. Cockerill, and P.G. Taylor, "The potential for peak shaving on low voltage distribution networks using electricity storage," *Journal of Energy Storage*, Vol. 16, pp. 231-242, April 2018.
- [18]. Matplotlib, "Matplotlib Documentation," [Online]. Διαθέσιμο στην ιστοσελίδα: <https://matplotlib.org/stable/users/index.html>
- [19]. K.K. Rangi, "How Min-Max Scaler Works," *Medium*, August 2024. [Online]. Διαθέσιμο στην ιστοσελίδα: <https://medium.com/@iamkamleshurangi/how-min-max-scaler-works-9fbeb9347da>
- [20]. G.E. Okereke, M.C. Bali, C.N. Okwueze, E.C. Ukekwe, S.C. Echezona and C.I. Ugwu, "K-means clustering of electricity consumers using time-domain features from smart meter data," *Journal of Electrical Systems and Information Technology*, Vol. 10, No. 1, 2, January 2023.
- [21]. M. Banoula, "K Means Clustering Algorithm in Machine Learning," *Simplilearn*, June 2025. [Online]. Διαθέσιμο στην ιστοσελίδα: <https://www.simplilearn.com/tutorials/machine-learning-tutorial/k-means-clustering-algorithm>
- [22]. A. Tomar, "Elbow Method in K-Means Clustering: Definition, Drawbacks, vs. Silhouette Score," *Builtin*, March 2025, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://builtin.com/data-science/elbow-method>
- [23]. T. Joshi, "Evaluating Clustering Algorithm – Silhouette Score," *Medium*, March 2021, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://tushar-joshi-89.medium.com/silhouette-score-a9f7d8d78f29>
- [24]. Sv. Misha, "Calinski-Harabasz Index for K-Means Clustering Evaluation using Python," *Towards Data Science*, March 2022, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://towardsdatascience.com/calinski-harabasz-index-for-k-means-clustering-evaluation-using-python-4feFeb2988e/>
- [25]. Γ. Τσεκούρας, *Συμβολή Στη Βραχυπρόθεσμη Και Μεσοπρόθεσμη Πρόβλεψη Ζήτησης Φορτίου Και Ενέργειας Συστημάτων Ηλεκτρικής Ενέργειας Με Χρήση Μεθόδων Αναγνώρισης Προτύπων*. Διδακτορική Διατριβή, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Οκτώβριος

2006. Διαθέσιμο στην ιστοσελίδα: <http://artemis.cslab.ece.ntua.gr:8080/jspui/handle/123456789/8561>
- [26]. Scikit-learn, “Scikit-learn, Machine Learning in Python” [Online]. Διαθέσιμο στην ιστοσελίδα: <https://scikit-learn.org/stable/#>
- [27]. Seaborn, “Seaborn: statistical data visualization,” [Online]. Διαθέσιμο στην ιστοσελίδα: <https://seaborn.pydata.org/>
- [28]. UK Power Networks, “Flexibility Dispatches,” *Open Data Portal*, April 2023, [Online]. Διαθέσιμο στην ιστοσελίδα: <https://ukpowernetworks.opendatasoft.com/explore/dataset/ukpn-flexibility-dispatches/information/>
- [29]. Π. Σ. Γεωργιλάκης, Σύγχρονα Συστήματα Μεταφοράς και Διανομής Ηλεκτρικής Ενέργειας. Ηλεκτρονικό Βιβλίο, Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (ΣΕΑΒ), Αθήνα, 2015. Διαθέσιμο στην ιστοσελίδα: <http://repository.kallipos.gr/handle/11419/2013>
- [30]. Π. Σ. Γεωργιλάκης, Οικονομική και Αξιόπιστη Λειτουργία Σύγχρονων Συστημάτων Ηλεκτρικής Ενέργειας. Ηλεκτρονικό Βιβλίο, Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις, Αθήνα, 2023. Διαθέσιμο στην ιστοσελίδα: <https://repository.kallipos.gr/handle/11419/9369>