



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

ΑΝΑΠΤΥΞΗ ΟΛΟΚΛΗΡΩΜΕΝΗΣ ΜΕΘΟΔΟΛΟΓΙΑΣ ΠΡΟΒΛΕΨΗΣ ΖΗΤΗΣΗΣ ΝΕΡΟΥ ΓΙΑ ΤΗΝ ΠΕΡΙΟΧΗ ΤΗΣ ΑΤΤΙΚΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Καραμαζιώτη Παναγιώτη-Ιωάννη



Επιβλέπων: Βασίλειος Ασημακόπουλος
Καθηγητής ΕΜΠ

Υπεύθυνος: Ράπτης Αχιλλέας
Υπ. Διδάκτωρ ΕΜΠ

Αθήνα , Φεβρουάριος 2016

“Water is the driving force of all nature”

Leonardo Da Vinci



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

ΑΝΑΠΤΥΞΗ ΟΛΟΚΛΗΡΩΜΕΝΗΣ ΜΕΘΟΔΟΛΟΓΙΑΣ ΠΡΟΒΛΕΨΗΣ ΖΗΤΗΣΗΣ ΝΕΡΟΥ ΓΙΑ ΤΗΝ ΠΕΡΙΟΧΗ ΤΗΣ ΑΤΤΙΚΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Καραμαζιώτη Παναγιώτη-Ιωάννη

Επιβλέπων: Βασίλειος Ασημακόπουλος
Καθηγητής ΕΜΠ

Υπεύθυνος: Ράπτης Αχιλλέας
Υπ. Διδάκτωρ ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την ...^η Μαρτίου 2016

.....
Β. Ασημακόπουλος,
Καθηγητής Ε.Μ.Π

.....
Ι. Ψαρράς,
Καθηγητής Ε.Μ.Π.

.....
Δ. Ασκούνης
Αναπλ. Καθηγητής Ε.Μ.Π

Αθήνα , Φεβρουάριος 2016

.....
ΚΑΡΑΜΑΖΙΩΤΗ ΠΑΝΑΓΙΩΤΗ ΙΩΑΝΝΗ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π

Copyright © ΠΑΝΑΓΙΩΤΗ ΙΩΑΝΝΗ ΚΑΡΑΜΑΖΙΩΤΗ , 2016

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 με 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 τη συγγραφέα.

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

Περίληψη

Η παρούσα διπλωματική έχει ως σκοπό την ανάπτυξη μιας αξιόπιστης μεθοδολογίας βραχυπρόθεσμης πρόβλεψης ζήτησης νερού για την περιοχή της περιφέρειας της Αττικής. Η σημαντικότητα της επιστήμης των προβλέψεων στην κατανάλωση νερού είναι πολύ σημαντική για διάφορους λόγους. Ειδικότερα, η βραχυπρόθεσμη πρόβλεψη ζήτησης νερού του τελικού καταναλωτή είναι απαραίτητη τόσο για την πρόβλεψη επενδύσεων και εσόδων που θέλει να κάνει ο διαχειριστής του δικτύου όσο και για τον τελικό καθορισμό της τιμής ανά μονάδα προϊόντος. Επιπλέον, εξαιτίας της σημαντικότητας του νερού για την επιβίωση οποιασδήποτε μορφής ζωής πάνω στη γη και με βάση τις νέες οδηγίες της Ε.Ε σχετικά με την διαχείριση των υδάτινων πόρων στην Ευρωπαϊκή περιφέρεια κρίνεται απαραίτητη η ανάπτυξη ενός πανελλαδικού σχεδίου διαχείρισης των υδάτων, στο οποίο θα εντάσσεται και μία αξιόπιστη μεθοδολογία πρόβλεψης ζήτησης νερού. Φιλοδοξία αυτής της διπλωματικής είναι να αποτελέσει μία επιστημονική βάση για την συμμετοχή της επικρατούσας μεθοδολογίας στο παραπάνω σχέδιο διαχείρισης υδάτινων πόρων.

Αρχικά, παρουσιάζεται η Ελληνική πραγματικότητα της διαχείρισης και κατανάλωσης υδάτων δίνοντας ξεχωριστή έμφαση στο νομό Αττικής τον οποίο και αφορά η παρούσα διπλωματική. Έπειτα γίνεται μία αναλυτική επισκόπηση των μεθόδων που χρησιμοποιούνται για τη βραχυπρόθεσμη πρόβλεψη νερού. Ακόμη, παρουσιάζονται και οι εφτά μέθοδοι που χρησιμοποιούνται για την παρούσα διπλωματική προκειμένου να συμβάλλουν με την ακρίβεια τους στις παραγόμενες προβλέψεις. Επιπλέον, αναλύονται τα σφάλματα μέτρησης της ακρίβειας των μεθόδων πρόβλεψης δίνοντας μεγαλύτερη έμφαση στο κριτήριο MAPE το οποίο είναι και το βασικό κριτήριο επιλογής για τη βέλτιστη μεθοδολογία.

Μετά την παρουσίαση του θεωρητικού υποβάθρου της διπλωματικής και έχοντας τα δεδομένα ζήτησης νερού στον τελικό καταναλωτή από την ΕΥΔΑΠ εφαρμόζονται βήμα προς βήμα οι τρεις μεθοδολογίες και συγκρίνονται με βάση τα τελικά αποτελέσματα που δίνουν σε ένα προκαθορισμένο διάστημα (διάστημα Β) στο οποίο είναι ήδη γνωστές οι μηνιαίες παρατηρήσεις των χρονοσειρών. Η συγκεντρωτική παρουσίαση των αποτελεσμάτων φανερώνει ότι η πρώτη μεθοδολογία που περιλαμβάνει την ανάλυση των χρονοσειρών σε Clusters, την παραγωγή επιμέρους προβλέψεων και ύστερα την συνάθροιση τους για την παραγωγή της τελικής πρόβλεψης είναι η προτιμότερη. Τέλος, για τις επιμέρους μεθόδους πρόβλεψης, φαίνεται ότι ο ισοβαρής συνδυασμός των τριών καλύτερων μεθόδων πρόβλεψης είναι ο πιο αποτελεσματικός συγκριτικά με τις μεμονωμένες μεθόδους.

Λέξεις κλειδιά: Τεχνικές προβλέψεων, Πρόβλεψη ζήτησης νερού, Ολοκληρωμένη μεθοδολογία παραγωγής προβλέψεων, Διαχείριση υδάτινων πόρων

Abstract

This thesis aims to develop a reliable methodology for water demand forecasting for the region of Attica, Greece. The importance of water demand forecasting is undeniably huge for several reasons. In particular, mid-term water demand forecasts of the end-consumer is essential both for the provision of investment the water utility operator needs to do for the network, as well as for the final determination of the price per unit of water. Furthermore, because of the importance of water for the survival of any form of life on earth and in the light of new EU directives on the management of water resources in the European territory, the development of a nationwide water management plan that will include a reliable methodology for water demand forecasting is an absolute necessity. The ambition of this project is to provide a solid scientific basis for the inclusion of the best methodology for water demand forecasting, in the above water management plan.

Initially, we present the Greek reality of water consumption and management by giving special emphasis in Attica. Following this, there is a detailed overview of the methods used for the mid-term water demand forecasting. There is also a complete presentation of all seven forecasting methods used in this thesis to contribute with their accuracy in the produced forecasts. Furthermore, we analyze the measurement of accuracy errors of the forecasting methods by putting more emphasis on the criterion MAPE which is the main selection criterion for the optimal methodology.

After the presentation of the theoretical background of the thesis, we develop the three forecasting methodologies, one of which is the most preferable because of a lower final MAPE error. Having the water demand data by EYDAP we implement step by step the three methodologies and compare them against a predetermined period (period B) in which it is already known the real monthly observations of the water demand time series. The aggregate results reveal that the first methodology(No1) is the best methodology and involves the analysis of time series into clusters, producing individual estimates for each cluster and then aggregating them to produce the final prediction. Finally, for individual forecasting methods, it seems that the equal-weight combination of the three best forecasting methods each time, is the most trustworthy from all the other individual methods.

Key words: Forecasting techniques, Water demand forecast, Intergrated methodology, Water management

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή κ. Βασίλειο Ασημακόπουλο, τόσο για την ανάθεση της συγκεκριμένης εργασίας όσο και για τις πολύτιμες συμβουλές και διορθώσεις του. Η γνώση του και η εμπειρία του δεν θα μπορούσαν να παρά να είναι το ιδανικότερο πρότυπο για μένα.

Θα ήθελα επίσης να ευχαριστήσω τον καθηγητή κ. Κωνσταντίνο Νικολόπουλο για τις συμβουλές που μου έδωσε τόσο για την παρούσα διπλωματική όσο και για την μετέπειτα πορεία μου σαν διπλωματούχος.

Αυτή η διπλωματική δεν θα μπορούσε να είχε ολοκληρωθεί αν δεν είχα την αμέριστη και διαρκή συμπαράσταση του υποψήφιου διδάκτορα κ. Αχιλλέα Ράπτη. Οι συμβουλές και οι παρατηρήσεις του υπήρξαν καταλύτης για την εκπόνηση της συγκεκριμένης διπλωματικής.

Τέλος, θα ήταν παράλειψη μου αν δεν ευχαριστούσα την οικογένεια μου για τα εφόδια που έχουν δώσει ώστε να μπορέσω να ολοκληρώσω τις σπουδές μου αλλά και τους φίλους μου για την συνεχή συμπαράσταση και εμπύχωση που μου έδιναν όλα αυτά τα χρόνια.

Περιεχόμενα

1	Ευρεία Περίληψη	15
1.1	Εισαγωγή	15
1.2	Τεχνικές προβλέψεων	16
1.2.1	Μέθοδοι Χρονοσειρών	16
1.2.2	Δείκτες Αξιολόγησης	18
1.3	Μεθοδολογία	20
1.4	Αποτελέσματα και Συμπεράσματα	22
1.4.1	Συνοπτική Παρουσίαση Αποτελεσμάτων	23
1.4.2	Συνοπτική Παρουσίαση Συμπερασμάτων	24
1.5	Προεκτάσεις	24
2	Εισαγωγή	27
2.1	Νερό και Ελληνική Πραγματικότητα	27
2.1.1	Γενικά Στοιχεία Ζήτησης	27
2.1.2	Επάρκεια Υδάτων και τα Εμπόδια Ορθής Αξιοποίησης τους	30
2.2	Σημασία Πρόβλεψης Ζήτησης Νερού	31
2.3	Κατηγορίες Πρόβλεψης Ζήτησης Νερού	33
2.4	Παράγοντες που Επηρεάζουν την Ζήτηση Νερού	34
2.5	Παρουσίαση Έρευνας	36
2.6	Το Αντικείμενο της Διπλωματικής	38
3	Θεωρία Τεχνικών Προβλέψεων	41
3.1	Γενικά περί προβλέψεων	41
3.2	Θεωρία Χρονοσειρών	43
3.2.1	Εισαγωγή	43
3.2.2	Αναπαράσταση Χρονοσειρών	43
3.2.3	Ποιοτικά Χαρακτηριστικά Χρονοσειρών	47
3.3	Μοντέλα Πρόβλεψης	52

3.3.1	Ποσοτικές Μέθοδοι.....	52
3.3.2	Κριτικές Μέθοδοι	54
3.3.3	Τεχνολογικές Μέθοδοι.....	54
3.4	Στατιστική Ανάλυση	55
3.4.1	Βασική Στατιστική Ανάλυση	55
3.4.2	Δείκτες Ακρίβειας Προβλέψεων.....	57
3.5	Μέθοδοι Πρόβλεψης	60
3.5.1	Απλοϊκή μέθοδος (Naive)	61
3.5.2	Μέθοδοι Κινητού Μέσου Όρου	61
3.5.3	Απλή Γραμμική Παλινδρόμηση	63
3.5.4	Μέθοδοι Εκθετικής Εξομάλυνσης.....	64
3.5.5	Μέθοδος Θ (Theta).....	71
3.5.6	Βελτιστοποιημένη Μέθοδος Θ (Optimised Theta)	73
3.5.7	Αυτό-παλινδρομικές Μέθοδοι Κινητού Μέσου Όρου (Auto-Regressive Integrated Moving Average / ARIMA)	75
3.5.8	Αλγόριθμος Πολλαπλής Συνάθροισης (Multi-Aggregation Prediction Algorithm / MAPA).....	78
3.5.9	Νευρωνικά Δίκτυα.....	80
3.5.10	Συνδυασμένες Μέθοδοι Πρόβλεψης.....	82
3.6	Ιεραρχική Δομή Χρονοσειρών.....	83
3.7	Επιλογή Κατάλληλης Μεθόδου Πρόβλεψης	84
4	Μεθοδολογία Διπλωματικής.....	87
4.1	Δεδομένα	87
4.2	Ομαδοποίηση Δεδομένων (Clustering).....	93
4.2.1	Είδη Ιεραρχικής Ομαδοποίησης Δεδομένων	94
4.2.2	Η Συσσωρευτική Τεχνική	95
4.2.3	Η Απόσταση μεταξύ δύο Cluster	95
4.3	Μέθοδος Έρευνας Διπλωματικής	98
4.3.1	Αναλυτική Παρουσίαση των Μεθοδολογιών	103
5	Παρουσίαση Αποτελεσμάτων	105
5.1	Αποτελέσματα Ανάλυσης cluster.....	105

5.2	Αποτελέσματα Μεθοδολογιών	108
5.2.1	Αποτελέσματα Μεθοδολογίας Νο1	108
5.2.2	Αποτελέσματα Μεθοδολογίας Νο2	113
5.2.3	Αποτελέσματα Μεθοδολογίας Νο3	121
6	Συμπεράσματα και Προεκτάσεις	125
6.1	Συμπεράσματα.....	125
6.1.1	Επιλογή Βέλτιστης Μεθοδολογίας.....	125
6.1.2	Επιπλέον Συμπεράσματα	126
6.2	Προεκτάσεις.....	131
	Παράρτημα.....	135
	Βιβλιογραφία	169

Πίνακας Εικόνων

Εικόνα 1.1-Στάδια μεθόδου διπλωματικής	20
Εικόνα 2.1-Χρήση(%) υδάτινων πόρων στην Ελλάδα	27
Εικόνα 2.2-Η διαχρονική εξέλιξη της ζήτησης νερού(μετρούμενη από τα διύλιστήρια) από το 1980 έως το 2013 (Πηγή ΕΥΔΑΠ).....	29
Εικόνα 2.3-Δείκτες εκμετάλλευσης νερού το 2000(αριστερά) και για το προβλεπόμενο 2030(δεξιά).....	33
Εικόνα 2.4-Περίληψη ερευνητικών μεθόδων που έχουν καταγραφη την περίοδο 2000-2010.....	36
Εικόνα 3.1-Μηνιαία κατανάλωση νερού για μία περιοχή του δήμου Αθηνών	44
Εικόνα 3.2-Εποχιακό διάγραμμα: μηνιαία κατανάλωση νερού για μία περιοχή του δήμου Αθηνών	44
Εικόνα 3.3-Αποτύπωμα Άνθρακα – Εξοικονόμηση Καυσίμου.....	45
Εικόνα 3.4-Μήτρα διαγραμμάτων διασποράς	46
Εικόνα 3.5-Μοντέλα τάσης-εποχικότητας εκθετικής εξομάλυνσης	65
Εικόνα 3.6 -Παράδειγμα εφαρμογής μεθόδου.....	72
Εικόνα 3.7-Συνδυασμοί βαρών και τιμών θ	73
Εικόνα 3.8-Στάδια μεθοδολογίας MAPA	78
Εικόνα 3.9-Εξάλειψη εποχικότητας ανάλογα με το επίπεδο συνάνθροισης k	79
Εικόνα 3.10-Παράδειγμα Νευρωνικού δικτύου ενός κρυφού επιπέδου	81
Εικόνα 3.11-Αριθμός ερευνών σχετικά με συνδυαστικές μεθόδους προβλέψεων ανά χρονολογία.....	83
Εικόνα 3.12-Παράδειγμα ιεραρχικής χρονοσειράς	83
Εικόνα 4.1-Κύρια δεδομένα διπλωματικής αποθηκευμένα σε αρχείο .txt	87
Εικόνα 4.2-Boxplot διάγραμμα για τα δεδομένα του Πίνακα 4.1.....	89
Εικόνα 4.3-Ζήτηση νερού(σε m^3) καταγεγραμμένα στον ταχ.κώδικα 15126, Μαρούσι.....	91
Εικόνα 4.4-Ζήτηση νερού(σε m^3) καταγεγραμμένα στον ταχ.κώδικα 15771, Ζωγράφου	92
Εικόνα 4.5-Ζήτηση νερού(σε m^3) καταγεγραμμένα στον ταχ.κώδικα 19003, Μαρκόπουλο	92
Εικόνα 4.6-Συνολική ζήτηση νερού(σε m^3) όπως αποτυπώνεται από τους 274 ταχ.κωδικούς της Αττικής.....	93
Εικόνα 4.7-Στάδια μεθόδου ιεραρχικής συσσωρευτικής ομαδοποίησης δεδομένων	94
Εικόνα 4.8-Ορισμός αποστάσεων μεθόδων <i>single link</i> (κόκκινο βέλος) και <i>complete link</i> (πράσινο βέλος).....	95
Εικόνα 4.9-Παράδειγμα ιεραρχικών cluster με την μέθοδο <i>single link</i>	96
Εικόνα 4.10-Παράδειγμα μεθόδου <i>complete link</i>	96
Εικόνα 4.11-Παράδειγμα μεθόδου <i>group average</i>	97
Εικόνα 4.12-Παράδειγμα εφαρμογής εύρεσης και αντικατάστασης ακραίων τιμών για την περιοχή Άγιος Ιωάννης-Ρέντης	100
Εικόνα 4.13- Χρονοσειρά ζήτησης νερού(αριστερά) και χρονοσειρά ζήτησης νερού ανά μετρητή(δεξιά) για την περιοχή της Αγ.Παρασκευής.....	101
Εικόνα 4.14-Χωρισμός χρονικού εύρους δεδομένων για την διαδικασία παραγωγής προβλέψεων	102
Εικόνα 4.15-Στάδια μεθοδολογίας Νο1.....	103
Εικόνα 4.16-Στάδια μεθοδολογίας Νο2.....	104
Εικόνα 5.1-Μεγέθη cluster μετά τον «καθαρισμό» των δεδομένων από ακραίες και κενές τιμές.....	105
Εικόνα 5.2-Μεγέθη cluster χωρίς τον «καθαρισμό» των δεδομένων από ακραίες και κενές τιμές	105
Εικόνα 5.3-Παραγόμενο δενδρόγραμμα καταμερισμού των cluster.....	107
Εικόνα 5.4-Παρουσίαση γραφικής επίδοσης των μεθόδων <i>Maru</i> , <i>Naive</i> , <i>Συνδυασμένης μεθόδου</i> στο διάστημα A για το cluster 1	108
Εικόνα 5.5-Παρουσίαση γραφικής επίδοσης των μεθόδων <i>Arima</i> , <i>Ets</i> , <i>Theta</i> και <i>Opt.Theta</i> στο διάστημα A για το cluster 1	108
Εικόνα 5.6-Παραγωγή προβλέψεων στο διάστημα B για το πρώτο cluster	109
Εικόνα 5.7-Παραγωγή προβλέψεων για τις υπόλοιπες μεθόδους για το cluster 1.....	109
Εικόνα 5.8-Προσαρμογή των μοντέλων πρόβλεψης στο διάστημα A για το cluster 5	110
Εικόνα 5.9-Προσαρμογή των υπόλοιπων μοντέλων πρόβλεψης στο διάστημα A για το cluster 5	110
Εικόνα 5.10-Παραγωγή προβλέψεων στο διάστημα B για το πέμπτο cluster.....	111

<i>Εικόνα 5.11-Παραγωγή των υπόλοιπων προβλέψεων στο διάστημα B για το πέμπτο cluster.....</i>	<i>111</i>
<i>Εικόνα 5.12-Προσαρμογή των μεθόδων πρόβλεψης για τον ταχ.κώδικα 15771,Ζωγράφου</i>	<i>114</i>
<i>Εικόνα 5.13-Προσαρμογή των υπόλοιπων μεθόδων πρόβλεψης για τον ταχ.κώδικα 15771,Ζωγράφου</i>	<i>114</i>
<i>Εικόνα 5.14-Προσαρμογή των μεθόδων για το διάστημα B.....</i>	<i>115</i>
<i>Εικόνα 5.15-Προσαρμογή των μεθόδων για το διάστημα B.....</i>	<i>115</i>
<i>Εικόνα 5.16-Προσαρμογή των μεθόδων στο διάστημα A</i>	<i>116</i>
<i>Εικόνα 5.17-Προσαρμογή των υπόλοιπων μεθόδων στο διάστημα A</i>	<i>116</i>
<i>Εικόνα 5.18-Προβλέψεις των μεθόδων στο διάστημα B.....</i>	<i>117</i>
<i>Εικόνα 5.19-Προβλέψεις των υπόλοιπων μεθόδων στο διάστημα B.....</i>	<i>117</i>
<i>Εικόνα 5.20-Σωρευτική απεικόνιση σφαλμάτων MAPE όλων των μεθόδων για όλες τις περιοχές της Αττικής.....</i>	<i>119</i>
<i>Εικόνα 5.21-Προσαρμογή μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα A</i>	<i>122</i>
<i>Εικόνα 5.22-Προσαρμογή υπόλοιπων μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα A</i>	<i>122</i>
<i>Εικόνα 5.23-Προβλέψεις μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα</i>	<i>123</i>
<i>Εικόνα 5.24-Προβλέψεις υπόλοιπων μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα B.....</i>	<i>123</i>
<i>Εικόνα 6.1-Γραφήματα Συνολικής ζήτησης και ζήτησης ανά περιοχή εισοδήματος.....</i>	<i>128</i>
<i>Εικόνα 6.2-Συγκεντρωτική παρουσίαση επίδοσης μεθόδων για τις τρεις μεθοδολογίες.....</i>	<i>130</i>

1 Ευρεία Περίληψη

1.1 Εισαγωγή

Είναι γεγονός, ακόμα και αν δεν το συνειδητοποιούμε άμεσα, ότι οι προβλέψεις αποτελούν αναπόσπαστο κομμάτι της καθημερινότητας του σύγχρονου ανθρώπου. Από την απλή απόφαση του πρωινού ξυπνήματος για να είμαστε έγκαιρα παρών σε μία πρωινή εκκρεμότητα, μέχρι την απόφαση για επένδυση εκατομμυρίων ή την πρόγνωση του καιρού προκειμένου να πάρουμε μαζί μας ομπρέλα, μας οδηγούν στην λήψη κάποιας απόφασης σχετικά με τις τοπικές καιρικές συνθήκες ή την κατάσταση των οικονομιών. Η αβεβαιότητα της εποχής που διανύουμε έχει καταστήσει την επιστήμη των προβλέψεων σαν απαραίτητο εργαλείο οικονομικού προγραμματισμού των επιχειρήσεων θέτοντας ως μεγάλη πρόκληση την σωστή και ακριβής θεμελίωση μεθοδολογιών για τον καθορισμό των προβλέψεων.

Η πρόβλεψη της ζήτησης νερού είναι ένας έντονα ενεργός κλάδος της επιστήμης των προβλέψεων που μετρά μερικές δεκαετίες επιστημονικές εξέλιξης και εφαρμογής. Η σημαντικότητα του κλάδου αυτού είναι τεράστια καθώς δεν υπάρχει πολυτιμότερο αγαθό για κάθε μορφή γήινης ζωής από το νερό. Γνωρίζουμε ότι ο άνθρωπος χρειάζεται περίπου 2 λίτρα τη μέρα για την σωστή λειτουργία του οργανισμού του, οπότε με πρόχειρους υπολογισμούς, καθημερινά απαιτούνται περίπου 10δισ λίτρα νερού (10hm^3) για όλον τον πληθυσμό της γης. Αντιλαμβανόμαστε ότι τα νούμερα είναι τεράστια για ένα αγαθό τόσο πολύτιμο όσο το νερό και η ανάγκη για την σωστή διαχείριση του είναι κάτι παραπάνω από επιτακτική.

Εκτός όμως από την σημαντικότητα πρόβλεψης του νερού για την εξυπηρέτηση της ανθρώπινης ζωής, η πρόβλεψη ζήτησης του νερού έχει αντίκτυπο και σε οικονομικά μεγέθη. Για την ακρίβεια, η διαχειρίστρια εταιρία νερού είναι ο άμεσα ενδιαφερόμενος καθώς η γνώση της μελλοντικής κατανάλωσης νερού είναι απαραίτητη για τον οικονομικό προγραμματισμό της μονάδας, αλλά και για την χρηματοδότηση μεγάλων επενδύσεων στην επέκταση και συντήρηση του δικτύου μεταφοράς του νερού. Ακόμη, η πρόβλεψη νερού χρησιμεύει και στην έγκαιρη ειδοποίηση του εποπτεύοντος προσωπικού σχετικά με αναμενόμενες ακραίες διακυμάνσεις κατανάλωσης ώστε να προβούν σε ανάλογες ενέργειες για την αποφυγή διαταραχών και ανεπιθύμητων καταστάσεων. Η αντιμετώπιση τέτοιων καταστάσεων θα οδηγήσει σε αποφυγή έκτακτων ενεργειών οι οποίες θα ανεβάσουν τα κόστη παραγωγής με αποτέλεσμα ένα

μέρος αυτών να περάσει στο τελικό καταναλωτή με τις ανάλογες συνέπειες στην υπόλοιπη οικονομική δραστηριότητα της τοπικής οικονομίας.

1.2 Τεχνικές προβλέψεων

Για την εξαγωγή προβλέψεων αλλά και γενικότερα για την στατιστική ανάλυση και τη μελέτη μιας μεταβλητής, το πρώτο σημαντικό βήμα είναι η συλλογή και η οργάνωση των ιστορικών στοιχείων της μεταβλητής αυτής. Τα δεδομένα που θα συλλεχθούν θα πρέπει να είναι όσο το δυνατό πιο έγκυρα και επικαιροποιημένα έτσι, ώστε η μετέπειτα επεξεργασία τους με κάποια μέθοδο πρόβλεψης να επιτρέψει την επίτευξη της κατά το δυνατόν καλύτερης ακρίβειας. Μετά τη συγκέντρωση των δεδομένων, εφαρμόζονται η κατάλληλοι μέθοδοι πρόβλεψης και προεκτείνονται στο μέλλον. Τέλος, η προέκταση των χρονοσειρών ακολουθείται από την αξιολόγηση των τιμών που παρήχθησαν με την χρήση των κατάλληλων δεικτών στατιστικού λάθους.

1.2.1 Μέθοδοι Χρονοσειρών

Οι τρεις μεγάλες κατηγορίες στις οποίες εντάσσονται οι κατηγορίες προβλέψεων που έχουν αναπτυχθεί ως σήμερα είναι :

- Ποσοτικές (μέθοδοι χρονοσειρών, αιτιοκρατικές μέθοδοι)
- Κριτικές
- Τεχνολογικές

Για την παρούσα διπλωματική χρησιμοποιήθηκαν οι ποσοτικοί μέθοδοι και συγκεκριμένα οι μέθοδοι των χρονοσειρών καθώς είναι ευκολότερα υλοποιήσιμοι σε σχέση με τις αιτιοκρατικές μεθόδους. Στην παρούσα διπλωματική υλοποιούνται επτά μέθοδοι πρόβλεψης:

1. Απλοϊκή μέθοδος (Naive)
2. Μέθοδοι Εκθετικής Εξομάλυνσης (Exponential Smoothing)
3. Αυτοπαλινδρομικές μέθοδοι Κινητού Μέσου Όρου (Arima)
4. Μέθοδος Θ (Theta)
5. Βελτιστοποιημένη μέθοδος Θ (Optimal Theta)
6. Αλγόριθμος Πολλαπλής Συνάθροισης (MAPA)
7. Συνδυασμός τριών βέλτιστων μεθόδων (Combined method)

Η απλοϊκή μέθοδος (Naive) αποτελεί την πιο απλή στατιστική μέθοδο, της οποίας η πρόβλεψη προκύπτει ως η πραγματική παρατήρηση της ακριβώς προηγούμενης

χρονικής περιόδου. Έχει καθιερωθεί η χρήση των αποτελεσμάτων της ως σημείο αναφοράς για τις υπόλοιπες μεθόδους.

Οι μέθοδοι εκθετικής εξομάλυνσης είναι ιδιαίτερα δημοφιλείς από την αρχή κιόλας της εμφάνισής τους. Δίνουν μεγαλύτερη βαρύτητα στα πιο πρόσφατα δεδομένα της χρονοσειράς, καθώς θεωρούν ότι αυτά παρέχουν περισσότερη πληροφορία συγκριτικά με τα παλαιότερα δεδομένα της. Το μοντέλο σταθερού επιπέδου (SES) υποθέτει την απουσία τάσης από τα δεδομένα και παράγει προβλέψεις με οριζόντια προέκταση των δεδομένων. Το μοντέλο γραμμικής τάσης (Holt) αποτελεί προέκταση του μοντέλου SES, στο οποίο πλέον γίνεται διαχείριση της συνιστώσας τάσης που παρατηρείται στα δεδομένα πολλών χρονοσειρών. Η παραγωγή των προβλέψεων γίνεται με απλή προέκταση μιας ευθείας γραμμής. Τέλος, το μοντέλο φθίνουσας τάσης (Damped) μειώνει σταδιακά την επίδραση της τάσης στα αποτελέσματα πρόβλεψης και ελέγχει το ρυθμό αύξησης των τιμών τους.

Η μέθοδος *Theta* αποτελεί μια μονοδιάστατη μέθοδο βασιζόμενη στη μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς μέσω της παραμέτρου θ , που εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των δεδομένων. Οι παραγόμενες γραμμές, γραμμές *Theta*, προεκτείνονται ξεχωριστά και στη συνέχεια γίνεται συνδυασμός τους με βάρη (ίσα για την κλασική μέθοδο) για την παραγωγή της τελικής πρόβλεψης. Ανάλογα την τιμή της παραμέτρου θ , δίνεται μεγαλύτερη έμφαση στη μακροπρόθεσμη (μικρό θ) είτε στη βραχυπρόθεσμη (μεγάλο θ) συμπεριφορά των δεδομένων.

Υπάρχουν αρκετές εργασίες που προτείνουν μεθόδους βελτιστοποίησης της κλασικής μεθόδου θ . Σε αυτήν την διπλωματική χρησιμοποιούμε το μοντέλο βελτιστοποίησης όπως περιγράφεται από την εργασία : *The Optimised Theta method (José Augusto Fioruci et al, 2015)*. Σε αυτήν την εργασία, επεκτείνεται η κλασική μέθοδος θ , επιλέγοντας την γραμμή θ που περιγράφει καλύτερα την βραχυπρόθεσμη συμπεριφορά της χρονοσειράς, διατηρώντας ως έχει την μακροπρόθεσμη συνιστώσα της (για $\theta_1 = 0$). Για την επιλογή αυτής της γραμμής θ , ελαχιστοποιείται μία συνάρτηση σφαλμάτων που βασίζεται σε προβλέψεις σφαλμάτων πάνω σε ένα εύρος τιμών (validation sample). Ο συνδυασμός των προβλέψεων που προέρχονται από τις δύο αυτές γραμμές θ πραγματοποιείται μέσω κατάλληλων βαρών (όχι απαραίτητα [0.5,0.5] όπως στην κλασική μέθοδο θ) που διασφαλίζουν την ανασύνθεση της αρχικής χρονοσειράς.

Οι αυτοπαλινδρομικές μέθοδοι κινητού μέσου όρου (*Arima*) είναι μία ευρέως διαδεδομένη μέθοδος πρόβλεψης η οποία μπορεί να χρησιμοποιηθεί για μονοδιάστατες (ένα ενιαίο σύνολο δεδομένων) ή πολυπαραγοντικές (συγκρίνοντας δύο ή περισσότερα σύνολα δεδομένων) αναλύσεις. Η τεχνική που χρησιμοποιείται από τη μέθοδο ARIMA περιλαμβάνει τους κινητούς μέσους όρους, αφαίρεση της τάσης, και τις μεθόδους παλινδρόμησης για τον εντοπισμό και την απομάκρυνση της αυτοσυσχέτισης στα δεδομένα. Ένα μοντέλο *Arima* εφαρμόζεται σε στάσιμες χρονοσειρές και χρησιμοποιεί

ένα γραμμικό συνδυασμό από προηγούμενες τιμές της χρονοσειράς με μία σειρά από προβλέψεις παρελθόντων τιμών των σφαλμάτων πρόβλεψης.

Ο αλγόριθμος πολλαπλής συνάθροισης (MAPA) είναι μία μεθοδολογία πρόβλεψης όπου κάθε χρονοσειρά αναλύεται σε επιμέρους χρονοσειρές μέσω χρονικής συνάθροισης (για παράδειγμα, η μετατροπή μηνιαίων χρονοσειρών σε διμηνιαίες, τετραμηνιαίες κ.λ.π), σε κάθε τέτοια χρονοσειρά εφαρμόζεται ένα μοντέλο εκθετικής εξομάλυνσης (ETS) και γίνονται προβλέψεις για τα αντίστοιχα χαρακτηριστικά της χρονοσειράς. Έπειτα, τα χαρακτηριστικά αυτά από κάθε επίπεδο συνάθροισης συνδυάζονται για την δημιουργία της τελικής πρόβλεψης. Η καινοτομία του αλγορίθμου αυτού έγκειται στο γεγονός ότι προτείνει τον συνδυασμό των επιμέρους χαρακτηριστικών των προβλέψεων για κάθε επίπεδο και για κάθε είδος χαρακτηριστικών και όχι των προβλέψεων αυτών καθ'αυτών. Δηλαδή, συνδυάζονται ξεχωριστά οι συνιστώσες της τάσης, του επιπέδου και της εποχικότητας και μετά συναθροίζονται για να εκφράσουν τις τελικές προβλέψεις.

Η μέθοδος συνδυασμού των τριών βέλτιστων μεθόδων (Combined method) είναι μία μέθοδος που αναπτύχθηκε εξολοκλήρου για τις ανάγκες αυτής της διπλωματικής. Με βάση τον συγκεκριμένο αλγόριθμο, επιλέγονται οι τρεις μέθοδοι που έχουν τις καλύτερες επιδόσεις προσαρμογής στα δεδομένα των χρονοσειρών (in sample model fitting) και συνδυάζονται ισοβαρώς για τον καθορισμό της τελικής πρόβλεψης. Η υλοποίηση της μεθόδου είναι αρκετά εύκολη και τα αποτελέσματα που έδωσε ήταν άκρως ικανοποιητικά όπως θα δείξουμε και πιο κάτω.

Επιπλέον, υπάρχουν οι μέθοδοι γραμμικής παλινδρόμησης, αλλά και αρκετές ακόμη μέθοδοι πιο σύνθετες από αυτές που αναφέρθηκαν προηγουμένως και που χρησιμοποιούνται στη βραχυπρόθεσμη πρόβλεψη ζήτησης νερού. Μερικά από αυτά τα μοντέλα είναι τα τεχνητά νευρωνικά δίκτυα και διάφορες προσαρμογές των μεθόδων παλινδρόμησης.

1.2.2 Δείκτες Αξιολόγησης

Η στατιστική ανάλυση αφορά ουσιαστικά την εύρεση βασικών στατιστικών δεικτών και συμβάλλει στον εντοπισμό των χαρακτηριστικών της χρονοσειράς για τη μετέπειτα ορθότερη αντιμετώπιση της. Σε συνδυασμό με τη γραφική παράσταση, επιτρέπει στους αναλυτές να έχουν μια γρήγορη, δομημένη και ταυτόχρονα συνολική εικόνα για το σύνολο της χρονοσειράς. Στη διπλωματική μας έχουν υπολογιστεί βασικοί στατιστικοί δείκτες ακρίβειας προβλέψεων και για τον υπολογισμό τους, εκτός από την πραγματική χρονοσειρά, απαιτείται και η χρονοσειρά που προκύπτει από την εφαρμογή των μεθόδων πρόβλεψης. Παρακάτω θα παρουσιάσουμε μερικούς βασικούς στατιστικούς δείκτες για τον υπολογισμό των σφαλμάτων:

- Μέσο σφάλμα (Mean Error):

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

- Μέσο απόλυτο σφάλμα (Mean Absolute Error):

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

- Μέσο τετραγωνικό σφάλμα (Mean Squared Error):

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

- Ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error):

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

- Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error):

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

- Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (symmetric Mean Absolute Percentage Error):

$$sMAPE = \frac{1}{n} \sum_{i=1}^n \frac{2 \cdot |Y_i - F_i|}{|Y_i + F_i|} \cdot 100\%$$

- Μέσο απόλυτο κανονικοποιημένο σφάλμα (Mean Absolute scaled Error):

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

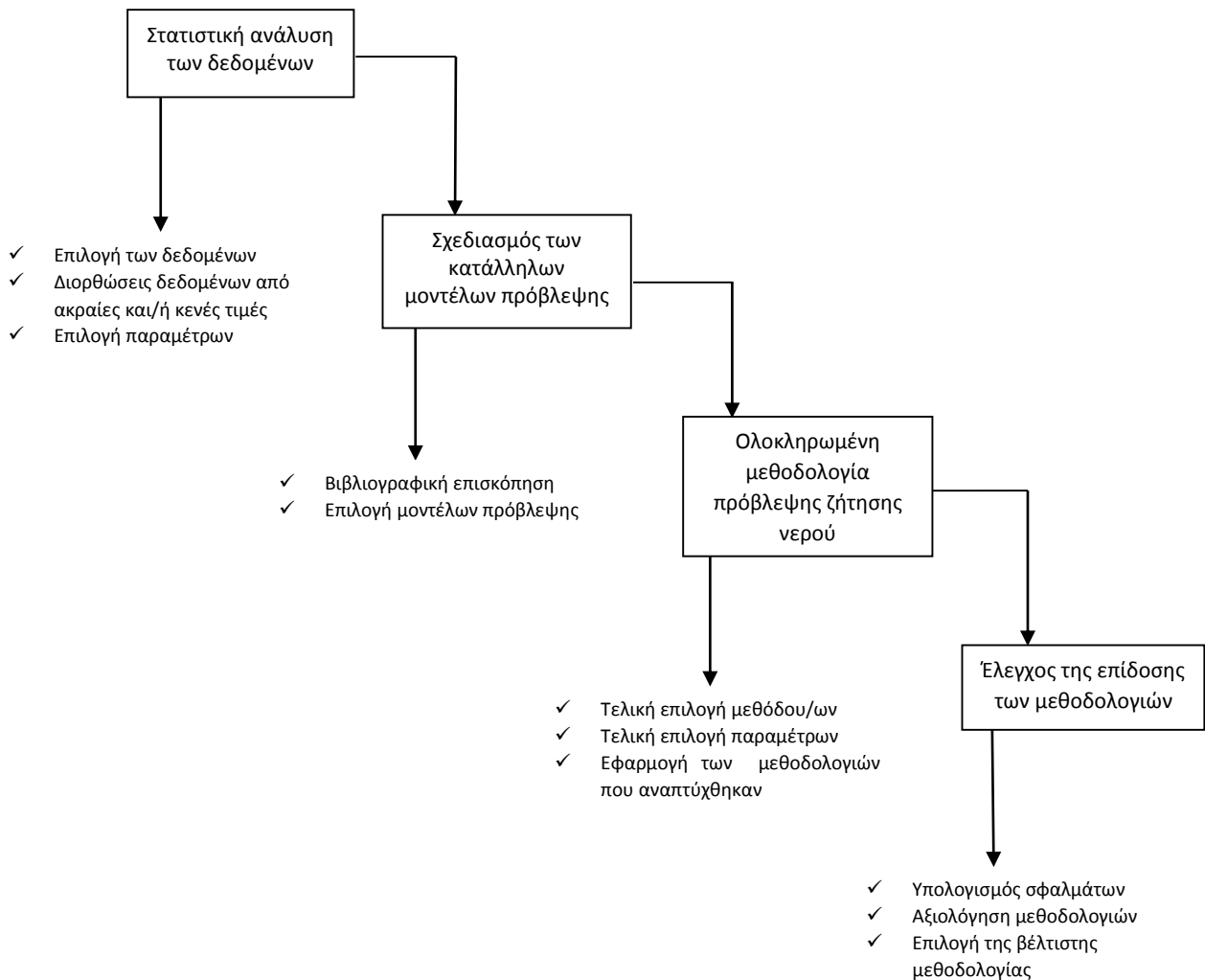
Τα σφάλματα Mean Error (ME), Mean Absolute Error (MAE), Mean Squared Error (MSE) και Root Mean Squared Error (RMSE) βασίζονται κατά κύριο λόγο στις διαφορές των τιμών πρόβλεψης από τις τιμές των πραγματικών τιμών. Τα υπόλοιπα σφάλματα εκφράζουν τα σφάλματα σε ποσοστιαία μορφή και είναι εξαιρετικά χρήσιμη μία τέτοια προσέγγιση καθώς καθιστά δυνατή τη σύγκριση σφαλμάτων από χρονοσειρές διαφορετικού επιπέδου μέσης τιμής. Το τελευταίο σφάλμα Mean Absolute Scaled Error (MASE) είναι ένα σχετικό μέτρο (relative measure) που χρησιμοποιεί το σφάλμα μίας

μεθόδου, στην συγκεκριμένη περίπτωση τη μέθοδο Naive ως benchmark– ορόσημο για να το συγκρίνει με το σφάλμα της υπό εξέτασης μεθόδου πρόβλεψης.

1.3 Μεθοδολογία

Η υλοποίηση μιας μεθοδολογίας η οποία βασίζεται σε ανάλυση τόσο μεγάλου όγκου δεδομένων και κατασκευή μεθοδολογιών ανάλογα με τη φύση των δεδομένων, δεν μπορεί να γίνει με παράλληλη διενέργεια όλων των απαραίτητων διεργασιών. Για την ομαλή διεξαγωγή και την ολοκλήρωσή της, απαιτείται ο διαχωρισμός σε επιμέρους βήματα το καθένα από τα οποία θα συμπληρώνει μια επιμέρους ενότητα της μεθοδολογίας και θα είναι άρρηκτα με τα επόμενα στάδια.

Τα προτεινόμενα στάδια υλοποίησης της μεθοδολογίας είναι τα παρακάτω:



Εικόνα 1.1-Στάδια μεθόδου διπλωματικής

Βασικό βήμα για την παραγωγή των προβλέψεων είναι η επιλογή και η ανάλυση των δεδομένων του μεγέθους που μελετάται. Στην διπλωματική μας, είχαμε τις μηνιαίες ενδείξεις κατανάλωσης για κάθε μετρητή και για κάθε έναν από τους 274 διαθέσιμους ταχυδρομικούς κωδικούς της Αττικής. Τα δεδομένα αυτά ήταν οι επίσημες μετρήσεις της ΕΥΔΑΠ που διαχειρίζεται το δίκτυο ύδρευσης και άρδευσης της Αττικής. Το χρονικό διάστημα για το οποίο έχει γίνει πλήρη καταγραφή των δεδομένων είναι από 1/2007 έως 12/2013. Υπήρχαν και κάποιες μετρήσεις από το 2005 και 2006 αλλά ήταν ελλιπής και ως εκ τούτου κρίθηκαν αναξιόπιστες να χρησιμοποιηθούν στη διπλωματική.

Με βάση ένα αρχικό αλγόριθμο που αναπτύχθηκε, δημιουργούμε 274 χρονοσειρές μηνιαίας ζήτησης 84 παρατηρήσεων η καθεμιά (84 μήνες από 1/2007 έως 12/2013). Απεικονίζοντας κάποιους από αυτούς τους ταχυδρομικούς κωδικούς, αντιλαμβανόμαστε ότι οι περισσότερες χρονοσειρές έχουν εμφανή εποχικότητα. Ξεκινώντας από τον Ιανουάριο κάθε έτους, η κατανάλωση πέφτει το Φεβρουάριο και αυξάνεται μέχρι τους καλοκαιρινούς μήνες. Ύστερα η κατανάλωση δείχνει να μειώνεται ξανά προοδευτικά μέχρι το τέλος του έτους. Κάποιες άλλες χρονοσειρές παρουσιάζουν τυχαίες διακυμάνσεις επίσης.

Αφού έχουμε στο μυαλό μας αυτές τις παρατηρήσεις, προχωράμε στο πρώτο βήμα υλοποίησης της μεθόδου που είναι ο «καθαρισμός» των τιμών από κενές, μηδενικές και ακραίες τιμές. Στα δεδομένα που είχαμε για το σκοπό της διπλωματικής αυτής, υπήρχαν αρκετές χρονοσειρές που είχαν μεμονωμένες τιμές ως μηδενικές ή άγνωστες. Για την περίπτωση των μηδενικών τιμών, εξαιτίας της μικρής συχνότητας εμφάνισης τους, θεωρήθηκε ότι προήλθαν από λάθος καταγραφή του πληροφοριακού συστήματος και έτσι δεν θεωρήθηκαν σαν χρονοσειρές διακοπτόμενης ζήτησης. Για τον εντοπισμό και αντικατάσταση των ακραίων τιμών, χρησιμοποιήθηκε το πακέτο `tsoutliers` για την γλώσσα R.

Μετά από το καθαρισμό των δεδομένων, προχωράμε στην κατασκευή των κατάλληλων μοντέλων πρόβλεψης με βάση την υπάρχουσα βιβλιογραφία. Εδώ χρησιμοποιήσαμε τις μεθόδους εκθετικής εξομάλυνσης, τις αυτοπαλινδρομικές μεθόδους κινητού μέσου όρου, τη μέθοδο Θ και την βελτιστοποιημένη μέθοδο της, τη μέθοδο πολλαπλής συνάθροισης, την μέθοδο συνδυασμού τριών βέλτιστων μεθόδων και την απλοϊκή μέθοδο. Η προσαρμογή των παραπάνω μεθόδων επιλέχθηκε να γίνει στις πρώτες 67 μηνιαίες παρατηρήσεις (διάστημα A), ενώ ο έλεγχος των προβλέψεων να γίνεται στις υπόλοιπες 17 παρατηρήσεις (διάστημα B). Κριτήριο για το αν μία μέθοδος πρόβλεψης κάνει καλή προσαρμογή σε δεδομένα ή αλλιώς αν δίνει ικανοποιητική ακρίβεια προβλέψεων, είναι το σφάλμα MAPE.

Το επόμενο στάδιο είναι ο καθορισμός των μεθοδολογιών που θα συγκριθούν ως προς τα αποτελέσματα που δίνουν. Οι μεθοδολογίες που αναπτύχθηκαν περιληπτικά είναι οι εξής:

Μεθοδολογία Νο1: Από τις 274 τελικές χρονοσειρές για κάθε ένα ταχυδρομικό κώδικα, δημιουργούμε $n=10$ cluster (ο αριθμός των cluster δεν ήταν προκαθορισμένος αλλά προέκυψε από την ανάλυση cluster που κάναμε πάνω στα δεδομένα των χρονοσειρών) και σε κάθε ένα από αυτά κάνουμε προβλέψεις με βάση τις μεθόδους χρονοσειρών που παρουσιάσαμε παραπάνω. Οι μέθοδοι που δίνουν το μικρότερο σφάλμα MAPE στο διάστημα A, συναθροίζονται (bottom-up aggregation) για τον σχηματισμό της συνολικής πρόβλεψης της μεθοδολογίας Νο1 στο διάστημα B.

Μεθοδολογία Νο2: Από τις 274 τελικές χρονοσειρές για κάθε ένα ταχυδρομικό κώδικα, παράγουμε προβλέψεις με βάση τις επτά μεθόδους που παρουσιάσαμε πιο πάνω. Οι μέθοδοι που δίνουν το μικρότερο σφάλμα προσαρμογής στο διάστημα A, συναθροίζονται (bottom-up aggregation) πρώτα στα 10 cluster της μεθοδολογίας Νο1 και ύστερα σε μία τελική χρονοσειρά για τον σχηματισμό της συνολικής πρόβλεψης της μεθοδολογίας Νο2 στο διάστημα B.

Μεθοδολογία Νο3: Από τις 274 τελικές χρονοσειρές για κάθε ένα ταχυδρομικό κώδικα, τις αθροίζουμε ανά μήνα και προκύπτει μία συνολική χρονοσειρά που αντιπροσωπεύει τη συνολική μηνιαία ζήτηση όλων των περιοχών της Αττικής για τις οποίες είχαμε διαθέσιμα δεδομένα. Έπειτα παράγουμε προβλέψεις με βάση τα μοντέλα πρόβλεψης του κεφαλαίου 3 που έχουμε στην διάθεση μας. Η μέθοδος που δίνει το μικρότερο σφάλμα προσαρμογής στο διάστημα A θα χρησιμοποιηθεί για την παραγωγή προβλέψεων στο διάστημα B οι οποίες αποτελούν τις τελικές προβλέψεις της μεθοδολογίας Νο3.

Έχοντας τις τελικές προβλέψεις όλων των μεθοδολογιών για το διάστημα B, συγκρίνουμε κάθε μια με τα πραγματικά δεδομένα της συνολικής χρονοσειράς στο διάστημα B. Η μεθοδολογία που έχει το *μικρότερο* σφάλμα MAPE εκεί, είναι και η προτιμότερη.

1.4 Αποτελέσματα και Συμπεράσματα

Έχοντας ακολουθήσει πιστά τα βήματα υλοποίησης της μεθοδολογίας όπως περιγράφονται σχηματικά στην εικόνα 1.1, σε αυτήν την ενότητα θα παρουσιάσουμε συνοπτικά τα αποτελέσματα των αλγορίθμων που δημιουργήσαμε για τους σκοπούς αυτής της διπλωματικής αλλά θα κάνουμε και συνοπτική παρουσίαση των τελικών συμπερασμάτων μας από την πειραματική διαδικασία της.

1.4.1 Συνοπτική Παρουσίαση Αποτελεσμάτων

Στον πίνακα 6.1 του κεφαλαίου 6 παρουσιάζονται οι τελικές επιδόσεις των τριών μεθοδολογιών που αναπτύχθηκαν. Παρατηρούμε ότι η μεθοδολογία Νο1 είναι η προτιμότερη καθώς το σφάλμα MAPE που δίνει στο διάστημα Β είναι κατά 16% μικρότερο από τη μεθοδολογία Νο2 και κατά 41% βελτιωμένο από τη μεθοδολογία Νο3. Επομένως, αν κάποιος θα ήθελε να κάνει προβλέψεις με βάση τα δεδομένα αυτής της διπλωματικής, θα ήταν προτιμότερο να χρησιμοποιήσει τα βήματα της εικόνας 1.1 και επιπλέον να χρησιμοποιήσει την μεθοδολογία Νο1 αντί των υπολοίπων.

Αρχικά, παρουσιάζοντας τα δεδομένα μας, ο πίνακας 4.1 και ο πλήρης αντίστοιχος πίνακας του παραρτήματος Π-3 απεικονίζουν κάποιους δείκτες (μέση τιμή, τυπική απόκλιση, μέσος όρος κ.α.) για την φύση των δεδομένων ζήτησης νερού. Βασικός πίνακας είναι και ο πίνακας 4. με τον πλήρη πίνακα Π-4 του παραρτήματος και απεικονίζει την μέση τιμή του αριθμού των μετρητών για κάθε περιοχή διαθέσιμη στα δεδομένα μας. Τα δεδομένα αυτά είναι πολύ χρήσιμα για τις προβλέψεις ζήτησης νερού στη μεθοδολογία Νο1.

Επιπλέον, ο πίνακας 5.2 του κεφαλαίου 5.2 απεικονίζει τα σφάλματα MAPE (διάστημα Α) των μεθόδων πρόβλεψης για όλα τα διαθέσιμα clusters που δημιουργήθηκαν από την μεθοδολογία Νο1. Στο παράρτημα, υπάρχει και ο αντίστοιχος πίνακας Π-5 ο οποίος αφορά τα αφιλτράριστα αρχικά δεδομένα ζήτησης. Ακόμα, ο πίνακας 5.1 και η εικόνα 5.1 δείχνουν τον καταμερισμό των clusters στους διαθέσιμους ταχυδρομικούς κώδικες. Η εικόνα 5.2 απεικονίζει τον καταμερισμό των clusters εάν στη μέθοδο μας δεν είχε προηγηθεί ο «καθαρισμός» των δεδομένων όπως έχει περιγραφεί. Επιπλέον, ο πίνακας Π-8 του παραρτήματος απεικονίζει τις βέλτιστες προβλέψεις για το διάστημα ελέγχου των τελικών προβλέψεων κάθε μεθοδολογίας(διάστημα Β)

Περνώντας στη μεθοδολογία Νο2, ο πίνακας 5.3 και αντίστοιχος πλήρης πίνακας Π-6 του παραρτήματος απεικονίζει το σφάλμα MAPE και των επτά μεθόδων πρόβλεψης για τους διαθέσιμους ταχυδρομικούς κώδικες. Συμπληρωματικά δίνεται και η εικόνα 5.20 όπου απεικονίζει το σωρευτικό σφάλμα MAPE όλων των μεθόδων. Με διακεκομμένη γραμμή σημειώνεται το όριο κάτω από το οποίο οι προσαρμογές των μεθόδων πρόβλεψης για τις αντίστοιχες περιοχές είναι ικανοποιητική. Τέλος, στον πίνακα Π-7 του παραρτήματος δίνονται οι βέλτιστες προβλέψεις για κάθε περιοχή και είναι αυτές οι οποίες θα συνδυαστούν ώστε να δώσουν την μοναδική τελική πρόβλεψη της μεθοδολογίας Νο2 μέσω συνάθροισης bottom-up.

Για τη μεθοδολογία Νο3, ο πίνακας 5.5 απεικονίζει το σφάλμα MAPE της προσαρμογής των μοντέλων πρόβλεψης στη χρονοσειρά συνολικής κατανάλωσης.

1.4.2 Συνοπτική Παρουσίαση Συμπερασμάτων

Περνώντας στην καταγραφή των τελικών συμπερασμάτων, μπορούμε να πούμε τα εξής:

1. Σύμφωνα με τους πίνακες 6.2 αλλά και τις εικόνες 5.1 και 5.2 είναι προφανές ότι ο «καθαρισμός» των πρωτόλειων δεδομένων από κενές, μηδενικές και ακραίες τιμές είναι επιτακτικός καθώς οδηγεί σε βελτίωση της αξιοπιστίας των προβλέψεων και για τις τρεις μεθοδολογίες που εφαρμόστηκε.
2. Η ζήτηση νερού στην Αττική έχει κατά κύριο λόγο εποχικό χαρακτήρα. Η κατανάλωση σε γενικές γραμμές αυξάνεται όσο πλησιάζουμε προς τους καλοκαιρινούς μήνες και η μειώνεται όσο απομακρυνόμαστε από αυτούς. Το μοτίβο αυτό ακολουθείται για όλα τα έτη που κάλυπταν τα δεδομένα μας.
3. Θέλοντας να κάνουμε μία ομαδοποίηση μεταξύ περιοχών με χαμηλό κατά κεφαλήν εισόδημα και περιοχών με μεσαίο/υψηλό κατά κεφαλήν εισόδημα, είδαμε ότι η κατανάλωση και στις δύο κοινωνικές ομάδες φθίνει μόνο που στην πρώτη ομάδα η πτώση της κατανάλωσης είναι κατά 38% ταχύτερη.
4. Σύμφωνα με τους πίνακες 6.3-6.6 και την εικόνα 6.2 η μέθοδος με την καλύτερη συνολική επίδοση είναι η μέθοδος συνδυασμού των τριών βέλτιστων μεθόδων πρόβλεψης. Εξαιτίας της αρκετά καλύτερης επίδοσης της στις μεθοδολογίες No1 και No2 κερδίζει την πρωτιά παρόλο που στην τρίτη μεθοδολογία η επίδοση της είναι η τέταρτη καλύτερη. Στην δεύτερη θέση βρίσκεται η οικογένεια των μεθόδων εκθετικής εξομάλυνσης με σταθερά καλή παρουσία και στις τρεις μεθοδολογίες, ενώ στην τρίτη θέση βρίσκεται η μέθοδος ARIMA η οποία τείνει να προσαρμόζεται καλύτερα σε χρονοσειρές που έχουν εμφανή στασιμότητα από ότι σε χρονοσειρές με εποχικά χαρακτηριστικά.
5. Η κλασική μέθοδος Θ και η βελτιστοποιημένη μέθοδος Θ έχουν παρόμοια επίδοση με την κλασική μέθοδο να υπερέχει στα σημεία και να καταλαμβάνει την τέταρτη συνολικά θέση επίδοσης ακρίβειας.

1.5 Προεκτάσεις

Οι μεθοδολογίες που εξετάστηκαν στα πλαίσια αυτής της διπλωματικής και οι μέθοδοι πρόβλεψης που χρησιμοποιήθηκαν μπορεί να έδωσαν ικανοποιητικά ποσοστά ακρίβειας σφάλματος αλλά αυτό δεν σημαίνει ότι σε μία μελλοντική προσέγγιση δεν θα

μπορέσουν να βελτιωθούν ακόμα περισσότερο. Συγκεκριμένα, μία μελλοντική έρευνα θα μπορούσε να επικεντρωθεί στα εξής σημεία:

1. Ενσωμάτωση των μετεωρολογικών παραγόντων και κυρίως του κλίματος και της θερμοκρασίας στα μοντέλα πρόβλεψης. Εξέταση της ενσωμάτωσης της με διάφορους συνδυασμούς (ακραίες τιμές, συνδυασμός παρελθοντικών τιμών θερμοκρασίας κλπ).
2. Ενσωμάτωση οικονομικών παραγόντων στα μοντέλα πρόβλεψης (μεταβολή ΑΕΠ, εξέλιξη ποσοστών ανεργίας και απασχόλησης κλπ). Η διαδικασία αυτή είναι εξαιρετικά πολύπλοκη και μπορεί να επιτευχθεί ενδεχομένως με μοντέλα πολλών εισόδων.
3. Χρησιμοποίηση νευρωνικών δικτύων για την παραγωγή προβλέψεων. Η χρησιμοποίηση αυτή της τεχνικής κερδίζει έδαφος τα τελευταία χρόνια και μπορεί να δώσει ικανοποιητικότερη ακρίβεια για τα δεδομένα αυτής της εργασίας.
4. Τα συμπεράσματα που βγήκαν από τα την παρούσα διπλωματική θα μπορούσαν να χρησιμοποιηθούν σαν βάση για έναν ευρύτερο σχεδιασμό διαχείρισης υδάτων τόσο στην Αττική, όσο και στον ευρύτερο Ελλαδικό χώρο.

2 Εισαγωγή

2.1 Νερό και Ελληνική Πραγματικότητα

Στις επόμενες υποενότητες 2.1.1-2.1.2 θα παρουσιάσουμε κάποια στατιστικά στοιχεία που αφορούν το νερό αποκλειστικά για τον Ελλαδικό χώρο. Η ύπαρξη αυτού του κεφαλαίου έχει σαν στόχο να προσανατολίσει τον αναγνώστη σχετικά με τα δεδομένα που αφορούν την γενικότερη κατάσταση του νερού στην Ελληνική επικράτεια κάνοντας ιδιαίτερη αναφορά στη περίπτωση του νομού Αττικής.

2.1.1 Γενικά Στοιχεία Ζήτησης

Από την συνολική ζήτηση νερού στην Ελλάδα, η αγροτική χρήση κατέχει την μερίδα του λέοντος με ποσοστό 86% , από το οποίο το 96% χρησιμοποιείται για άρδευση και απ' αυτό το 80% χάνεται σε απώλειες από τον υδρολογικό κύκλο της κάθε περιοχής.



Εικόνα 2.1-Χρήση(%) υδάτινων πόρων στην Ελλάδα

Παρακάτω δίνεται ένας πίνακας που αφορά τις ζητήσεις νερού για εργασίες ύδρευσης και άρδευσης αντίστοιχα σε όλα τα γεωγραφικά διαμερίσματα της Ελλάδας για το έτος 2011. Οι αριθμοί αυτοί δίνουν μία καλή εικόνα του επιμερισμού της ζήτησης νερού παρόλο που δεν αφορούν την χρονολογία συγγραφής αυτής της διπλωματικής.

Υδατικά διαμερίσματα	Άρδευση(hm^3)	Ύδρευση(hm^3)
Δυτικής Πελοποννήσου	201,0	23,0
Βόρειας Πελοποννήσου	401,5	41,7
Ανατολικής Πελοποννήσου	324,9	22,1
Δυτικής Στερεάς Ελλάδας	366,5	22,4
Ηπείρου	153,5	33,9
Αττικής	99,0	420,0
Ανατ. Στερεάς Ελλάδας	773,7	41,6
Θεσσαλίας	550,0	69,0
Δυτικής Μακεδονίας	609,4	43,7
Κεντρικής Μακεδονίας	527,6	99,8
Ανατολικής Μακεδονίας	627,0	32,0
Θράκης	825,2	27,9
Κρήτης	320,0	42,3
Νήσων Αιγαίου	80,2	37,2

Πίνακας 2.1-Επιμερισμός ζήτησης ανά υδατικό διαμέρισμα και ανά είδος χρησιμοποίησης

2.1.1.1 Η εικόνα της Αττικής

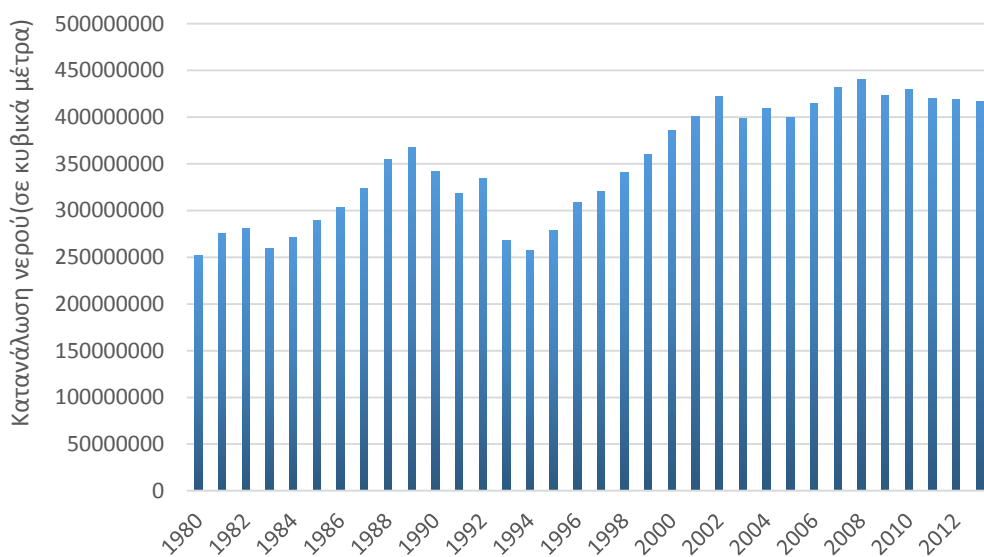
Το υδατικό διαμέρισμα(ΥΔ) της Αττικής περιλαμβάνει όλη την έκταση του νομού Αττικής καθώς και κάποιες εκτάσεις από του νομούς Βοιωτίας, Κορινθίας και Κυκλάδων. Αποτελεί το 2,5% της συνολικής εκτάσεως της χώρας, αλλά συγκεντρώνει άνω του 35% του πληθυσμού της. Συνεπώς, εδώ παρατηρείται η μεγαλύτερη χρήση νερού για ύδρευση, ίση με το 44% του συνόλου της χώρας. Η ζήτηση νερού στην Αττική ανέρχεται περίπου στα 520 hm^3 /έτος, ενώ οι ίδιοι υδατικοί πόροι της Αττικής δεν ξεπερνούν τα 490 hm^3 /έτος, με πιο σημαντική πηγή αντλήσεως νερού στην Αττική τον ταμιευτήρα του Μαραθώνα με ωφέλιμη χωρητικότητα 32,2 hm^3 /έτος. Πραγματοποιείται, έτσι, μεταφορά σημαντικών ποσοτήτων νερού από τα ΥΔ Δυτικής και Ανατολικής Στερεάς Ελλάδος και συγκεκριμένα από τη λίμνη Υλίκη (από το 1955), από τον ταμιευτήρα του Μόρνου (1980) και αυτόν του Αγίου Δημητρίου στον Εύηνο ποταμό (από το 1996). Η μέση ετήσια κατανάλωση της Αθήνας σε νερό από τους ταμιευτήρες Μόρνου και Υλίκης κατά την περίοδο 1982-2007 ήταν περίπου 320 hm^3 , ενώ την ίδια περίοδο η μέση ετήσια απόληψη νερού από τους ίδιους ταμιευτήρες ήταν 350 hm^3 . Η δε συνολική μέση ετήσια κατανάλωση της Αθήνας σε νερό στην ίδια περίοδο ήταν 334,3 hm^3 (Πηγή: ΕΥΔΑΠ). Συνεπώς, η μεταφορά νερού από την Υλίκη και τον Μόρνο συνεπάγεται απώλειες κατά την μεταφορά, ύψους 40 hm^3 /έτος περίπου.

Βάσει των σημερινών συνθηκών, τον Ιούλιο η προσφορά υδάτινων πόρων στο ΥΔ της Αττικής (με τις μεταφορές που προαναφέρθηκαν) μόλις φθάνει για την ικανοποίηση της ζήτησεως. Παραδείγματος χάριν, την 25.6.2008 το απολήψιμο απόθεμα νερού στους

ταμειυτήρες του Μαραθώνα, της Υλίκης, του Μόρνου και του Ευήνου ανερχόταν στα 666,5 hm³, έναντι 1.020 hm³ την 25.6.2007.

Η εξαιρετικά χαμηλή συμβολή των υδάτινων πόρων της Αττικής για κάλυψη της ζήτησης σχετίζεται και με τη συστηματική υποβάθμιση τόσο των επιφανειακών όσο και των υπόγειων υδάτων αυτού του ΥΔ. Η υποβάθμιση αυτή οφείλεται αφενός στην συγκέντρωση του πληθυσμού, των βιομηχανικών και άλλων δραστηριοτήτων στην περιοχή αυτή και αφετέρου, στην έλλειψη ή στην αδυναμία εφαρμογής ενός ικανού ρυθμιστικού πλαισίου που θα εξασφαλίζει την αποφυγή της ανωτέρω υποβαθμίσεως.

Γενικότερα, στο ΥΔ της Αττικής η κρίση του νερού είναι ήδη αισθητή, κυρίως κατά τους καλοκαιρινούς μήνες. Τα τελευταία έτη τον Ιούλιο η ζήτηση νερού στο ΥΔ της Αττικής μόλις που καλύπτεται από την προσφορά. Η κατάσταση αναμένεται να χειροτερεύσει μεσοπρόθεσμα (10-25 έτη) σύμφωνα με τα σενάρια που διατυπώνονται στο Εθνικό Πρόγραμμα Διαχείρισης Υδάτινων Πόρων του ΕΜΠ, λόγω της αυξήσεως της ζήτησης τον Ιούλιο στα 81 hm³ (κατά 44,64%), που θα επιφέρει έλλειμμα στο υδατικό ισοζύγιο ισοδύναμο με το 30% αυτής της ζήτησεως.



Εικόνα 2.2-Η διαχρονική εξέλιξη της ζήτησης νερού(μετρούμενη από τα διύλιστήρια) από το 1980 έως το 2013 (Πηγή ΕΥΔΑΠ)

Τα τελευταία χρόνια παρατηρείται μία σταθεροποίηση -με ελαφρώς πτωτική τάση- της τάξης του 0.8% για την κατανάλωση νερού στην Αττική, όπως μπορεί να παρατηρηθεί στην εικόνα 2.1.3. Παρατηρείται επίσης ένα «βύθισμα» της κατανάλωσης στις αρχές τις δεκαετίας του '90 το οποίο οφείλεται στην εκστρατεία εξοικονόμησης νερού που έλαβε μέρος τότε (σε συνδυασμό με την αύξηση της τιμής του νερού) προκειμένου να αντιμετωπιστούν τα μειωμένα αποθέματα νερού της Αττικής. Η κατανάλωση νερού έφτασε ξανά στα επίπεδα του 1991 το 1997, τέσσερα χρόνια μετά το τέλος της εκστρατείας ενημέρωσης του κοινού για την εξοικονόμηση νερού.

2.1.2 Επάρκεια Υδάτων και τα Εμπόδια Ορθής Αξιοποίησης Τους

Σύμφωνα με έρευνες, στην Ελλάδα δεν παρουσιάζεται πρόβλημα έλλειψης υδάτων σε καμία από τις 13 υδατικές περιφέρειες της χώρας. Ωστόσο υπάρχουν ενδείξεις ότι το υδατικό δυναμικό της χώρας μειώνεται. Σύμφωνα με τα αποτελέσματα της έρευνας του προγράμματος ENSEMBLES (<http://www.ensembles-eu.org>), στην κεντρική και βόρειο Ελλάδα παρατηρείται μείωση των κατακρημνισμάτων τις πέντε τελευταίες δεκαετίες. Η μείωση ξεκινά από 30 mm και τοπικά φτάνει στα 150 mm ανά δεκαετία. Παράλληλα η σύγκριση της απορροής των ποταμών το χρονικό διάστημα 1900-1970 με το χρονικό διάστημα 1971-1998 δείχνει μειωμένη ροή σε πανελλαδικό επίπεδο σε ποσοστό 5 ως 10%, εξαιρουμένης της Ηπείρου όπου η πτώση περιορίζεται στο 2 με 5% (Milly et al 2005). Σύμφωνα με τις προβλέψεις για την Ελλάδα αναμένεται μείωση του υδατικού δυναμικού τουλάχιστον κατά 25% το έτος 2030 ενώ ήδη από το 2005 η διαθεσιμότητα νερού, με εξαίρεση την οροσειρά της Πίνδου, είναι στη χαμηλότερη κλίμακα σε σχέση με χώρες της ΕΕ.

Γενικά, η Ελλάδα διαθέτει σχετικά επαρκείς ποσότητες υδατικών πόρων αλλά ταυτόχρονα αντιμετωπίζει σημαντικά προβλήματα όσον αφορά στην αξιοποίηση και βέλτιστη διαχείριση τους (ΕΕ 2007). Σύμφωνα με τη μελέτη αυτή, οι κυριότεροι λόγοι που προκαλούν προβλήματα στην αξιοποίηση των υδατικών πόρων της χώρας είναι:

- Η άνιση κατανομή των υδατικών πόρων στο χώρο. Η δυτική Ελλάδα δέχεται πολύ μεγαλύτερα ύψη βροχών από την ανατολική. Ειδικότερα, 24% της συνολικής έκτασης της χώρας δέχεται το 36% των συνολικών ατμοσφαιρικών κατακρημνισμάτων, ενώ ακόμη μεγαλύτερη είναι η διαφοροποίηση στα ποσοστά της επιφανειακής απορροής.
- Η ανομοιόμορφη κατανομή των υδατικών πόρων στο χρόνο, με μεγάλη συγκέντρωση βροχοπτώσεων κατά τη χειμερινή περίοδο. Στη νότια Ελλάδα το 80,9% των ετήσιων βροχοπτώσεων συγκεντρώνεται σ' αυτή την περίοδο, ενώ το θερινό ύψος της βροχής αυξάνει προς βορρά, και στα βορειότερα τμήματα παίρνει τη μεγαλύτερη τιμή του, δηλαδή 20% του ετήσιου όγκου.
- Η άνιση κατανομή της ζήτησης στο χώρο, αναντίστοιχη με την κατανομή της προσφοράς. Ο άξονας Θεσσαλονίκη – Αθήνα - Πάτρα που παρουσιάζει τη μεγαλύτερη συγκέντρωση πληθυσμού και δραστηριοτήτων, δεν διαθέτει σημαντικούς υδατικούς πόρους.

- Η ανομοιόμορφη κατανομή της ζήτησης στο χρόνο, αναντίστοιχη με την κατανομή της προσφοράς. Ο μεγαλύτερος καταναλωτής του χρησιμοποιούμενου νερού, η γεωργία (87%), το καταναλώνει την ξηρή περίοδο. Την ίδια περίοδο και ειδικότερα τους μήνες Ιούλιο-Αύγουστο, διπλασιάζεται λόγω τουρισμού και η κατανάλωση νερού ύδρευσης.
- Η γεωμορφολογία της χώρας. Ο έντονος οριζόντιος και κατακόρυφος διαμελισμός, καθώς και η δομή και διάταξη των πετρωμάτων, έχουν ως αποτέλεσμα τη δημιουργία πολλών μικρών υδατορεμάτων με χειμαρρική κυρίως δίαιτα, επιφανειακή απορροή μικρής διάρκειας, αυξημένη κατείδυση και συχνά πλημμυρικά φαινόμενα.
- Η εξάρτηση της βόρειας Ελλάδας από τις επιφανειακές απορροές ποταμών που προέρχονται από γειτονικά κράτη (περίπου 13 hm³ ετησίως). Το 23% του συνολικού υδατικού δυναμικού της χώρας προέρχεται από γειτονικές χώρες, καθιστώντας ιδιαίτερα σημαντική την ανάγκη για προώθηση διασυνοριακών δράσεων διαχείρισης υδατικών πόρων.
- Το μεγάλο ανάπτυγμα ακτών (15.021 km) σε συνδυασμό με τη λιθολογική σύσταση των πετρωμάτων που συντείνει, λόγω της εντατικής εκμετάλλευσης παράκτιων υδροφορέων, στην υφαλμύρινσή τους.

2.2 Σημασία Πρόβλεψης Ζήτησης Νερού

Η πρόβλεψη ζήτησης νερού είναι ένας κλάδος της επιστήμης των προβλέψεων και αποσκοπεί στην ανάπτυξη των κατάλληλων μεθοδολογιών που θα δίνουν αξιόπιστες προβλέψεις ζήτησης νερού, βασιζόμενες στην διαθεσιμότητα ενός όγκου ιστορικών δεδομένων.

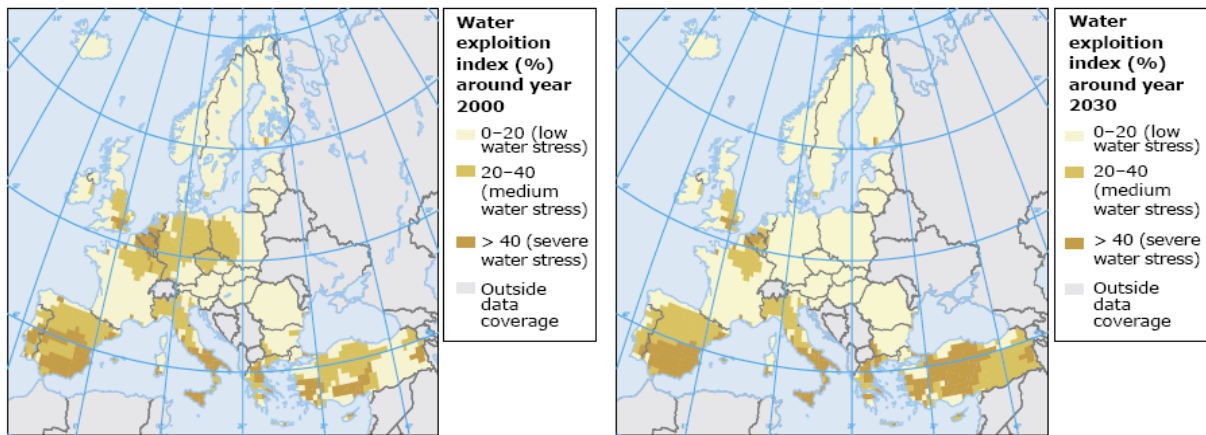
Η μοντελοποίηση της ζήτησης και η δημιουργία μεθοδολογιών για την πρόβλεψη της είναι σημαντικής ωφέλειας για πολλούς λόγους:

- Η επένδυση κεφαλαίων σχετιζόμενων με την επέκταση-συντήρηση του υπάρχοντος δικτύου ύδρευσης είναι αρκετά κοστοβόρα και μπορεί να αγγίξει μερικά δεκάδες ή και εκατοντάδες εκατομμύρια ευρώ. Είναι επομένως επιτακτική η ανάπτυξη και μοντελοποίηση αξιόπιστων μεθόδων πρόβλεψης της ζήτησης νερού για αποφυγή περιττών επενδύσεων που δεν θα αποφέρουν τα προβλεπόμενα στο κοινωνικό σύνολο.

- Παρόλο που το νερό είναι ένα σχετικά φτηνό αγαθό, είναι απαραίτητο για την ζωή του ανθρώπου και η τοποθεσία πόσιμου νερού μπορεί να είναι αντικείμενο διενέξεων και γενικότερης αναστάτωσης. Η ακριβής πρόβλεψη ζήτησης νερού μπορεί να εξομαλύνει αυτές τις αντιθέσεις και να εδραιώσει μία αξιόπιστη βάση διαχείρισης των υδάτων από εταιρίες διαχείρισης νερού.
- Πολλές περιοχές της Ελλάδας(Θεσσαλικός κάμπος, Κυκλάδες, Αττική κ.α.) κατά τους καλοκαιρινούς-άνυδρους μήνες αντιμετωπίζουν προβλήματα στα αποθεματικά του πόσιμου νερού που έχουν στην διάθεση τους. Επομένως η δημιουργία ενός πλάνου βιώσιμης συντήρησης του νερού για την δύσκολη περίοδο του καλοκαιριού μπορεί να γίνει υποχρεωτική στο άμεσο μέλλον. Μία σωστή ανάλυση των ωφελειών και της βελτιστοποίησης του κόστους από μία τέτοια ενέργεια δεν μπορεί να γίνει αν δεν υπάρχει πρώτα μία σαφής εικόνα της αναμενόμενης ζήτησης νερού.

Παράλληλα, σε ευρωπαϊκό επίπεδο θα προωθηθούν πολιτικές και κανόνες για την διαχείριση των υδάτινων πόρων που συμπεριλαμβάνονται στις στρατηγικές ανάπτυξης της Ε.Ε με την ονομασία *Europe 2020*. Πιο συγκεκριμένα, όσον αφορά την διαχείριση υδάτων αναφέρει, μεταξύ άλλων, ότι θα πρέπει να γίνεται τακτικός έλεγχος της κατάστασης των υδάτινων πόρων μέσω αξιοποίησης της υπάρχουσας πληροφορίας σε καθορισμένο χώρο και χρόνο αλλά και να γίνεται πρόβλεψη για τυχόν έκτακτα γεγονότα (ξηρασίες κλπ.) που θα επηρεάσουν την ισορροπία του υδάτινου οικοσυστήματος. Γίνεται επομένως ευθεία αναφορά στην ανάπτυξη σύγχρονων συστημάτων παρακολούθησης της χρήσης νερού και μεθοδολογιών για την πρόβλεψη της μέσω ενός αλληλοσυνδεόμενου πανευρωπαϊκού συστήματος δεδομένων. Επιπλέον, στην αξιολόγηση "*Water scarcity and droughts policy,2012*" αναφέρεται πως το 10% της έκτασης της Ε.Ε. και το 15% του συνόλου του Ευρωπαϊκού πληθυσμού έχει επηρεαστεί από καταστάσεις έλλειψης νερού. Όπως γίνεται αντιληπτό και από την Εικόνα 2.3 ,είναι πολύ σημαντικό να αναπτυχθούν ικανές μεθοδολογίες πρόβλεψης ζήτησης νερού, καθώς η εκμετάλλευση του νερού αναμένεται να αυξηθεί στον Ελλαδικό χώρο για το μακροπρόθεσμο μέλλον(2030) κάτι που, αν δεν ληφθούν τα κατάλληλα μέτρα, θα οδηγήσει σε συνθήκες μειωμένης διαθεσιμότητας του πόσιμου νερού για τον πληθυσμό αλλά και τις βιομηχανίες.

Γενικότερα, η πρόβλεψη ζήτησης νερού βοηθάει στην βελτίωση του συστήματος ύδρευσης για το οποίο εφαρμόζεται. Βραχυπρόθεσμο, μακροπρόθεσμο ή μεσοπρόθεσμο μοντέλα προβλέψεων ,αναπτυγμένα με προσοχή στην ακρίβεια και την αποδοτικότητα τους , βοηθούν τις εταιρίες διαχείρισης να προσφέρουν το νερό με το λιγότερο δυνατό κόστος για τον τελικό καταναλωτή.



Εικόνα 2.3-Δείκτες εκμετάλλευσης νερού το 2000(αριστερά) και για το προβλεπόμενο 2030(δεξιά)

Η παρούσα διπλωματική χρησιμοποιεί τα επίσημα μηνιαία ιστορικά δεδομένα κατανάλωσης νερού στο Νομό Αττικής για την περίοδο 2007-2013 από την Εταιρεία Ύδρευσης και Αποχέτευσης Πρωτεύουσας (ΕΥΔΑΠ) και συγκρίνει τρεις ευρύτερες μεθοδολογίες-προσεγγίσεις ως προς την ακρίβεια τους σχετικά με τα αποτελέσματα που δίνουν σε ήδη γνωστές τιμές κατανάλωσης νερού οικιακών ή βιομηχανικών καταναλωτών. Έτσι δεν παράγουμε τιμές πρόβλεψης για το άγνωστο μέλλον(μετά το 2013), αλλά χρησιμοποιούμε τα ήδη υπάρχοντα δεδομένα (τέλη 2012 έως 2013) για να επαληθεύσουμε σε αυτό το χρονικό περιθώριο ποια μέθοδος είναι η πιο αξιόπιστη. Κάθε μεθοδολογία χρησιμοποιεί τις μεθόδους που θα περιγραφούν διεξοδικότερα στο κεφάλαιο 3.

2.3 Κατηγορίες Πρόβλεψης Ζήτησης Νερού

Η βιβλιογραφία που αναφέρεται στις προβλέψεις, κάνει συχνά αναφορά στον διαχωρισμό του χρονικού εύρους πρόβλεψης σε βραχυπρόθεσμο, μεσοπρόθεσμο και μακροπρόθεσμο. Τυπικά, κάθε χρονικός ορίζοντας πρόβλεψης χρησιμοποιείται για διαφορετικούς σκοπούς και παράγουν διαφορετική ακρίβεια πρόβλεψης.

- **Μακροπρόθεσμη πρόβλεψη**

Το χρονικό εύρος εδώ είναι από μία δεκαετία και άνω και σε ορισμένες περιπτώσεις αγγίζει τα 30 χρόνια. Αφορούν προβλέψεις που έχουν σχέση με την επέκταση των δικτύων ύδρευσης που απαιτούν μεγάλο χρηματικό ποσό επένδυσης και η προέγκρισή τους θα πάρει αρκετό χρόνο. Όσο ο χρονικός ορίζοντας πρόβλεψης αυξάνεται, τόσο η πρόβλεψη είναι ευάλωτη σε λάθος υπολογισμούς. Επομένως σε τέτοιες περιπτώσεις, οι διαχειριστές των υδάτων προσπαθούν να αποφύγουν το σενάριο μίας πολύ μεγάλης επέκτασης, μεγαλύτερης της αναγκαίας, που θα οδηγούσε σε αυξημένα κόστη συντήρησης που στην τελική θα τα

επωμίζονταν οι τελικοί καταναλωτές, αλλά και μίας επέκτασης πολύ μικρότερης της απαιτούμενης, καθώς θα οδηγούσε σε έλλειψη πόσιμου νερού.

- **Μεσοπρόθεσμη πρόβλεψη**

Αυτές οι προβλέψεις (από ένα έως δέκα χρόνια) είναι κατάλληλες για βελτιώσεις στο σύστημα διανομής του νερού και για καθορισμό των τιμών πώλησης των μονάδων νερού και επικεντρώνονται στην μεταβλητότητα της κατανάλωσης νερού σε ένα σταθερό ή ελαφρά μεταβαλλόμενο αριθμό καταναλωτών.

- **Βραχυπρόθεσμη πρόβλεψη**

Αυτές οι προβλέψεις λειτουργούν υποστηρικτικά στις λειτουργίες των εταιριών διαχείρισης αλλά και στις λογιστικές και οικονομικές διαδικασίες.

- **Πολύ-βραχυπρόθεσμη πρόβλεψη**

Έχουν εύρος μερικές ώρες, μέρες ή εβδομάδες. Αυτές οι προβλέψεις χρησιμοποιούν προσομοιώσεις του καιρού με στόχο, μέσω της παραγωγής προβλέψεων, να βελτιστοποιήσει τις καθημερινές εργασίες της μονάδας διαχείρισης των υδάτων και στην διαφύλαξη της ασφάλειας του συστήματος.

Τύπος πρόβλεψης	Ορίζοντας πρόβλεψης	Εφαρμογές
Μακροπρόθεσμος	Δεκαετίες, 10-40 χρόνια	Κατασκευή και επέκταση δικτύου διανομής
Μεσοπρόθεσμος	Έως 10 χρόνια	Συντήρηση υπάρχοντος δικτύου και προσαρμογή τιμών
Βραχυπρόθεσμος	1-2 χρόνια	Λογιστικός σχεδιασμός, πρόβλεψη επενδύσεων και εσόδων, έλεγχος λειτουργιών
Πολύ βραχυπρόθεσμος	Ώρες, μέρες, έως δύο εβδομάδες	Βελτιστοποίηση καθημερινών εργασιών

Πίνακας 2.2-Πίνακας τύπου προβλέψεων και αντίστοιχων εφαρμογών του

2.4 Παράγοντες που Επηρεάζουν την Ζήτηση Νερού

Ένας αριθμός παραγόντων επηρεάζουν την ζήτηση νερού μεταξύ αυτών ο πληθυσμός, η ανεργία, ή διείσδυση της τεχνολογίας, ο καιρός κ.α. Παρακάτω αναλύονται οι κυριότεροι παράγοντες.

- **Πληθυσμός, Εργασία, Τεχνολογία.** Η μεταβολή του πληθυσμού είναι από τους μεγαλύτερους παράγοντες που μεταβάλλουν την τάση της κατανάλωσης νερού. Είναι χαρακτηριστικό ότι περιοχές της Ελλάδος με αυξανόμενο πληθυσμό συνεισφέρουν προσθετικά στην ζήτηση νερού.

Ακόμα η κατανάλωση επηρεάζεται και από τον κύκλο εργασιών των επιχειρήσεων και των βιομηχανιών αλλά και από τις αλλαγές στο προσωπικό εισόδημα. Έχει παρατηρηθεί ότι η κατανάλωση νερού αυξάνεται όταν το εισόδημα αυξάνεται δεδομένου ότι όλοι οι άλλοι παράγοντες που επηρεάζουν την κατανάλωση παραμένουν σταθεροί. Η κατανάλωση νερού είναι πιθανότερο να πέσει όταν για παράδειγμα, κάποιο μέλος της οικογένειας χάσει την δουλειά του.

Ακόμα η έλευση και διεξόδυση της τεχνολογίας είναι πιθανότερο να επηρεάσει την ζήτηση νερού με παρόμοιο τρόπο. Για παράδειγμα η χρησιμοποίηση αυτόματων πλυντηρίων πιάτων στις οικίες των καταναλωτών μπορεί να συμβάλλει στην εξοικονόμηση νερού.
- **Καιρός και Κλίμα.** Ίσως οι πιο προφανείς παράγοντες που επηρεάζουν την ζήτηση νερού. Τα εποχιακά χαρακτηριστικά στις μηνιαίες χρονοσειρές πολλών περιοχών που είχαμε στα δεδομένα μας, οφείλονται σε μεγάλο βαθμό στα συγκεκριμένα χαρακτηριστικά του κλίματος της κάθε εποχής. Για παράδειγμα η κατανάλωση νερού αυξάνεται το καλοκαίρι καθώς πολύ χρησιμοποιούν το νερό για πότισμα σε κήπους ή βεράντες κάτι που τους χειμερινούς μήνες υφίσταται σε μικρότερο βαθμό. Ακόμη, για την εξάλειψη της πιθανότητας αφυδάτωσης, οι άνθρωποι τείνουν να καταναλώνουν περισσότερες ποσότητες νερού το καλοκαίρι απ'ότι το χειμώνα.
- **Τιμή.** Η τιμή πώλησης κάθε μονάδας νερού παίζει σημαντικό ρόλο στις προβλέψεις μακροπρόθεσμου, μεσοπρόθεσμου και βραχυπρόθεσμου ορίζοντα. Τόσο η κατανάλωση νερού όσο και τα έσοδα της μονάδας διαχείρισης επηρεάζονται με την μεταβολή της τιμής πώλησης κάθε μονάδας νερού. Μία αύξηση τιμών θα μπορούσε στο άμεσο μέλλον να αλλάξει την στάση των καταναλωτών σχετικά με την κατανάλωση νερού. Αυτές οι αλλαγές θα μπορούσε να είναι η ορθότερη χρησιμοποίηση του νερού κατά την διάρκεια προσωπικής περιποίησης ή κατά τον καθαρισμό του αυτοκινήτου. Σε επόμενο στάδιο, αν οι αυξήσεις δεν απορροφηθούν από τον πληθωρισμό, θα μπορούσε οι καταναλωτές να χρησιμοποιούν συσκευές σωστότερης διαχείρισης νερού ώστε να αντισταθμίσουν σε μεταγενέστερο χρόνο τις επιπτώσεις της αύξησης της τιμής του.
- **Άλλοι παράγοντες.** Σε ορισμένες περιπτώσεις οι εταιρίες διαχείρισης δίνουν κίνητρα για την εξοικονόμηση νερού όπως για παράδειγμα η χρήση συσκευών ελεγχόμενης ροής νερού σε βρύσες ή καζανάκια. Αυτά τα προγράμματα επειδή

οδηγούν συνήθως σε μείωση της πραγματικής κατανάλωσης και άρα και των εσόδων, είναι προσωρινού χαρακτήρα και θέτονται σε ισχύ σε ακραίες περιπτώσεις(εκτεταμένη ξηρασία κ.α.).

Ακόμα, η υποβάθμιση της ποιότητας του συστήματος διανομής, απόρροια κακής συντήρησης, οδηγεί σε συνήθως ακαθόριστη ποσότητα υδάτινων διαρροών από χαλασμένους αγωγούς μεταφοράς. Στις περιπτώσεις που αυτές οι ποσότητες καταγράφονται στους τελικούς καταναλωτές, εκτός από το αίσθημα μπερδέματος και δυσαρέσκειας που δημιουργούνται(αφού στην ουσία ο καταναλωτής έχει να πληρώσει κάτι το οποίο δεν έχει χρησιμοποιήσει και αποτελεί παράβλεψη της μονάδας διαχείρισης) οδηγεί συχνά στην ελάττωση της πραγματικής ζήτησης.

2.5 Παρουσίαση Έρευνας

Περιοδικότητα	Μονοπαραγοντικές Χρονοσειρές (Univariate time series)	Παλινδρόμηση χρονοσειρών	Scenario based	Συστήματα υποστήριξης αποφάσεων	Νευρωνικά δίκτυα με συμβατικές μεθόδους	Composite/ Hybrid	Σύνολο
Ωριαία	-	-	-	-	Herrera et al.(2010)	Zhou et al.(2002)	2
Ημερήσια	-	-	Goodchild (2003)	Froukh(2001), Jain and Ormsbee (2001) and Levin et al.(2006)	Cutore et al.(2008) and Jain and Ormsbee(2002)	Caido(2010), Gato et al.(2007) and Zhou et al.(2000)	10
Εβδομαδιαία	-	-	-	-	Adomowski(2010), Bougadis et al.(2005) and Jain et al.(2001)	-	3
Μηνιαία	-	Brekke et al.(2002) and Polebitski and Palmer(2010)	-	Mohammed and Al-Mualla(2010)	Firat et al. (2009)	Altunkaynak et al.(2005)	6
Ετήσια	Alhumoud(2008)	Lee et al. (2010)	Burney et al.(2001), Wei et al.(2010), Williamson et al. (2002)	Feng et al.(2007)	-	Wang et al.(2009) and Wu et al.(2010)	8
Πολ/πλες περιοδικότητες	-	-	-	-	Ghiassi et al. (2008) and Jentgen et al. (2007)	Alvisi et al.(2007), Aly and Wanakule (2004)	4
Σύνολο	1	3	4	6	9	10	33

Εικόνα 2.4-Περίληψη ερευνητικών μεθόδων που έχουν καταγραφεί την περίοδο 2000-2010

Σε αυτό το κεφάλαιο θα παρουσιάσουμε περιληπτικά τις δημοφιλέστερες μεθόδους που έχουν δημοσιευτεί σε ερευνητικά περιοδικά στο θέμα των προβλέψεων ζήτησης νερού σε αστικά κέντρα κατά την περίοδο 2000 έως 2010 [Emmanuel Donkor et al. 2014]. Τα αποτελέσματα της επισκόπησης δείχνουν ότι παρόλο που μία ποικιλία μεθόδων και μοντέλων έχει προσελκύσει την προσοχή, οι εφαρμογές αυτών διαφέρουν, καθώς εξαρτώνται από την μεταβλητή πρόβλεψης (ζήτηση κατά κεφαλήν, ημέρα/ώρα αιχμής κ.α.), την περιοδικότητα της (ωριαία, μηνιαία κ.λ.π.) αλλά και τον ορίζοντα πρόβλεψης. Ενώ μέθοδοι με τεχνητά νευρωνικά δίκτυα είναι πιθανότερο να χρησιμοποιηθούν για βραχυπρόθεσμες προβλέψεις, τα οικονομετρικά μοντέλα συνδυασμένα με προσομοιώσεις ή βασιζόμενα-σε-σενάρια προβλέψεις (scenario based forecasting) τείνουν να χρησιμοποιούνται περισσότερο σε μακροπρόθεσμες προβλέψεις.

Από την παραπάνω εικόνα μπορούν να εξαχθούν κάποια συμπεράσματα:

- Η χρησιμοποίηση των νευρωνικών δικτύων κερδίζει έδαφος σε σχέση με πιο απλές μεθόδους, όπως οι μονοπαραγοντικές χρονοσειρές.
- Οι μέθοδοι παλινδρόμησης έχουν χρησιμοποιηθεί εκτενώς. Στις περισσότερες περιπτώσεις χρησιμοποιούνται σε συνδυασμό με νευρωνικά δίκτυα ή με μονοπαραγοντικές χρονοσειρές. Το γενικό συμπέρασμα είναι ότι τα νευρωνικά δίκτυα αποδίδουν καλύτερα σε βραχυπρόθεσμες προβλέψεις, εξαιτίας της ικανότητας που έχουν να αντιλαμβάνονται τις μη-γραμμικότητες των δεδομένων. Σε αντιδιαστολή, για μέσο/μάκρο-πρόθεσμες προβλέψεις δεν έχουν γίνει πολλές έρευνες και η αποτελεσματικότητά τους εκεί δεν έχει αποδειχθεί.
- Σχετικά με την περιοδικότητα της μεταβλητής ζήτησης, βλέπουμε ότι οι έρευνες επικεντρώνονται περισσότερο σε παραλλαγές μεταβλητών ημερήσιας ή ετήσιας περιοδικότητας. Έτσι φαίνεται ότι οι έρευνες αυτές εξυπηρετούν κυρίως την αντιμετώπιση προβλημάτων σε καθημερινές εργασίες και σε στρατηγικές ανάπτυξης όπως βελτιστοποίηση συστήματος διανομής και σχεδιασμού χωρητικότητας για τις εταιρίες διαχείρισης υδάτων.
- Για ετήσια περιοδικότητα της μεταβλητής ζήτησης, οι έρευνες έχουν επικεντρωθεί αρκετά στην χρησιμοποίηση μοντέλων πρόβλεψης που βασίζονται σε σενάρια (scenario based) και σε συστήματα υποστήριξης αποφάσεων.
- Η χρησιμοποίηση παλινδρόμησης για τη μοντελοποίηση μηνιαίας ζήτησης, ακολουθεί την γενική άποψη ότι βραχυπρόθεσμες-μεσοπρόθεσμες καταναλώσεις επηρεάζονται από μεταβλητές σχετιζόμενες με τον καιρό ενώ οι μακροπρόθεσμες καταναλώσεις από κοινωνικό-οικονομικούς παράγοντες (Adomowski, 2010).

Στο παράρτημα στο τέλος της διπλωματικής υπάρχουν οι πίνακες Π.1 και Π.2 στους οποίους γίνεται σχολιασμός κάποιων επιπλέον ερευνών που αφορούν την ζήτηση νερού σε αστικό περιβάλλον.

2.6 Το Αντικείμενο της Διπλωματικής

Συνοψίζοντας όσα ειπώθηκαν στις προηγούμενες ενότητες αυτού του κεφαλαίου, εξαιτίας της σημαντικότητας του νερού σαν αγαθό απαραίτητο για κάθε μορφή ζωής, αλλά και εξαιτίας των νέων πολιτικών που προωθεί η Ευρωπαϊκή Ένωση για την βιωσιμότητα του οικοσυστήματος αλλά και για την αποφυγή έκτακτων καταστάσεων έλλειψης νερού, είναι πολύ σημαντικό να υιοθετηθούν στην Ελλάδα σύγχρονες πρακτικές πρόβλεψης της καταναλωτικής ζήτησης νερού σε μελλοντικό χρόνο. Όπως φάνηκε και προηγουμένως, στην Ελλάδα, παρόλο που τα αποθέματα νερού φαίνεται να επαρκούν για τις ανάγκες του πληθυσμού της για το άμεσο μέλλον, είναι πολύ πιθανό σε 10-15 χρόνια κάτι τέτοιο να πάψει να ισχύει. Όλα αυτά σε συνδυασμό με την μη ύπαρξη μίας ενιαίας πανελλαδικής στρατηγικής διαχείρισης υδάτων κάνουν επιτακτική την άμεση προώθηση επιστημονικών προσεγγίσεων-μεθοδολογιών σχετικών με την πρόβλεψη της ζήτησης νερού. Η διπλωματική αυτή έχει σαν στόχο να αναπτύξει μία βέλτιστη μεθοδολογία ως προς την ακρίβεια της στις προβλέψεις ζήτησης νερού που δίνει σε ένα σταθερό εύρος χρόνου. Αναπτύσσοντας τα στάδια, καταλήξαμε σε τρεις βασικές μεθοδολογίες τις οποίες και συγκρίναμε στα αποτελέσματα που δίνουν. Επομένως θα μπορούσε, κατά την γνώμη μας, τα αποτελέσματα αυτής της έρευνας να αποτελέσουν βάση για μία λεπτομερέστερη στρατηγική παραγωγής προβλέψεων, που θα ενταχθεί σε μία ευρύτερη στρατηγική διαχείρισης υδάτινων πόρων στον ελλαδικό χώρο.

Η εργασία αυτή αφορά 274 ταχυδρομικούς κωδικούς (ΤΚ) του νομού Αττικής και έχουν μακράν το μεγαλύτερο μερίδιο κατανάλωσης στην Ελλάδα (πάνω από το 30% του συνολικής ζήτησης νερού) για την περίοδο από 1/2007 έως και 12/2013 (84 μήνες). Αυτούς τους μήνες τους χωρίζουμε σε δύο μέρη:

1. *διάστημα Α*: το 80% αυτών (67 μήνες) θα χρησιμοποιηθεί για τον καθορισμό των παραμέτρων των μοντέλων πρόβλεψης (model forecast fitting).
2. *διάστημα Β*: οι υπόλοιποι 17 μήνες χρησιμοποιούνται για την παραγωγή προβλέψεων των μοντέλων αυτών.

Οι τρεις μεθοδολογίες που συγκρίνονται, περιληπτικά είναι οι εξής:

Μεθοδολογία Νο1: Από τις 274 τελικές χρονοσειρές για κάθε ένα ΤΚ, δημιουργούμε $n=10$ cluster (ο αριθμός των cluster δεν ήταν προκαθορισμένος

αλλά προέκυψε από την ανάλυση cluster που κάναμε πάνω στα δεδομένα των χρονοσειρών) και σε κάθε ένα από αυτά κάνουμε προβλέψεις με βάση τις μεθόδους που εξηγούμε στο κεφάλαιο 3. Οι μέθοδοι που δίνουν το μικρότερο σφάλμα στο διάστημα A, συναθροίζονται (bottom-up aggregation) για τον σχηματισμό της συνολικής πρόβλεψης.

Μεθοδολογία Νο2: Από τις 274 τελικές χρονοσειρές για κάθε ένα TK, παράγουμε προβλέψεις με βάση τις μεθόδους του κεφαλαίου 3. Οι μέθοδοι που δίνουν το μικρότερο σφάλμα προσαρμογής στο διάστημα A, συναθροίζονται (bottom-up aggregation) πρώτα σε 10 cluster και ύστερα σε μία χρονοσειρά για τον σχηματισμό της συνολικής πρόβλεψης.

Μεθοδολογία Νο3: Από τις 274 τελικές χρονοσειρές για κάθε ένα TK, τις αθροίζουμε ανά μήνα και προκύπτει μία συνολική χρονοσειρά που αντιπροσωπεύει τη συνολική ζήτηση. Έπειτα παράγουμε προβλέψεις με βάση τα έξι μοντέλα πρόβλεψης του κεφαλαίου 3 που έχουμε στην διάθεση μας. Η μέθοδος που δίνει το μικρότερο σφάλμα προσαρμογής στο διάστημα A θα χρησιμοποιηθεί για την παραγωγή προβλέψεων στο διάστημα B.

Από τις τελικές χρονοσειρές που προέκυψαν για κάθε μία μεθοδολογία, συγκρίνεται η ακρίβεια των προβλέψεων που δόθηκαν στο διάστημα B με τα πραγματικά δεδομένα. Η μέθοδος που δίνει το μικρότερο σφάλμα σε εκείνο το διάστημα, είναι και η προτιμότερη μέθοδος.

3 Θεωρία Τεχνικών Προβλέψεων

3.1 Γενικά περί προβλέψεων

Είναι γεγονός ότι καλούμαστε καθημερινά, σε διάφορες περιστάσεις, να πάρουμε αποφάσεις για διάφορα θέματα που μας απασχολούν, βασιζόμενοι σε προβλέψεις που μπορούμε να κάνουμε ανά πάσα στιγμή. Αυτές οι προβλέψεις μπορεί να αφορούν αποφάσεις μελλοντικές σε ατομικό ή συλλογικό επίπεδο. Ένα απλό παράδειγμα θα μπορούσε να είναι η ποσότητα του φαγητού που θα καταναλωθεί σε ένα τραπέζι φίλων που σκεφτόμαστε να κάνουμε για την ονομαστική μας εορτή. Η απόφαση αυτή θα μπορούσε να ληφθεί με βάση των αριθμό των ατόμων που είναι καλεσμένα με σκοπό να αγοράσουμε τόσα υλικά έτσι ώστε όλοι οι παρευρισκόμενοι να απολαύσουν ένα χορταστικό και γευστικό γεύμα.

Φυσικά αυτή η διαδικασία μπορεί να επαναληφθεί και σε υψηλότερο επίπεδο, όταν προβλέψεις καλούνται να κάνουν μικρές ή μεγάλες επιχειρήσεις και οργανισμοί σαν βασικό μέρος της στρατηγικής τους. Για παράδειγμα, μία εταιρία για να μπορέσει να βγάλει το πρόγραμμα παραγωγής των προϊόντων της για το επόμενο εξάμηνο, καλείται να βγάλει μία αξιόπιστη μεθοδολογία πρόβλεψης ζήτησης τέτοια ώστε να μπορεί να ικανοποιεί την εκάστοτε ζήτηση που θα υπάρχει στα προϊόντα της αλλά και να έχει διαθέσιμο απόθεμα ανά πάσα στιγμή ώστε να διαχειρίζεται τυχόν απότομες αυξήσεις στην ζήτηση των προϊόντων της, ενώ ταυτόχρονα το απόθεμα της να μην είναι τέτοιο που να αυξάνει υπερβολικά τα κόστη αποθήκευσης και συντήρησης των αγαθών.

Χωρίς αμφιβολία, το ενδιαφέρον και η σημασία της πρόβλεψης έχει αυξηθεί ραγδαία τα τελευταία 30 χρόνια. Το ενδιαφέρον αυτό προέρχεται τόσο από τον ακαδημαϊκό κόσμο όσο και από πρακτικά ασχολούμενους με την επιστήμη των προβλέψεων. Το κύριο χαρακτηριστικό που τροφοδοτεί αυτή την αύξηση ενδιαφέροντος είναι η *αβεβαιότητα*. Η πρόβλεψη έχει να κάνει με την εκτίμηση συμπεριφοράς συγκεκριμένων μεταβλητών στο μέλλον λαμβάνοντας υπόψιν πληθώρα παραμέτρων για το σκοπό αυτό. Ακριβώς επειδή δεν ξέρουμε τι μπορεί να μας επιφυλάσσει το μέλλον, υπάρχει πάντα ο κίνδυνος «εκτροχιασμού» των προβλέψεων ακόμα και μετά από μία εμπειριστατωμένη έρευνα της εκτιμώμενης συμπεριφοράς του μέλλοντος. Η αβεβαιότητα που αποτελεί και το πιο σημαντικό «εχθρό» της επιστήμης των προβλέψεων έχει κατηγοριοποιηθεί από

τον Μακρυδάκη και τους συνεργάτες του στο βιβλίο «Χορεύοντας με την Τύχη». Αναφέρουν λοιπόν δυο είδη αβεβαιότητας που συναντάμε καθημερινά:

- Η «αβεβαιότητα του μετρό» αναφέρεται σε συνεχείς μικρές τυχαίες διακυμάνσεις τόσο της ιδιωτικής όσο και της επιχειρηματικής καθημερινότητας. Ο όρος προήλθε από την επιπλέον χρονική διάρκεια που θα χρειαστεί ένας συρμός για την διαδρομή μεταξύ δυο απομακρυσμένων σταθμών, λόγω ενός τεχνικού προβλήματος, της πολυκοσμίας ή μιας στάσης εργασίας.
- Η «αβεβαιότητα της καρύδας» αναφέρεται σε ένα εντελώς απρόσμενο και σπάνιο γεγονός που να έχει σημαντικές επιδράσεις και συνέπειες. Το όνομα αυτού του είδους η αβεβαιότητα το πήρε από το απρόσμενο γεγονός που μπορεί να συμβεί καθώς περπατάμε στο δρόμο και να πέσει στο κεφάλι μας μια καρύδα! Η «αβεβαιότητα της καρύδας» αντιπαραβάλλεται με μεγάλες φυσικές ή οικονομικές καταστροφές που είναι δύσκολο να προβλεφθεί το πότε και το που θα συντελεστούν καθώς και πόσο μεγάλες θα είναι οι επιδράσεις τους.

Ο τομέας των προβλέψεων έχει δεχθεί σφοδρές κριτικές και είχε να αντιμετωπίσει τη μεγάλη δυσαρέσκεια σχετικά με την ανικανότητα των μεθόδων να προειδοποιήσουν έγκαιρα για επερχόμενες αλλαγές καθώς και για τα μεγάλα σφάλματα στις προβλέψεις. Συγχρόνως όμως, λανθασμένες ενέργειες οι οποίες προκαλούνται από ασταθή περιβάλλοντα, μη αναμενόμενες εξελίξεις, ασυνέχειες και άλλα αυξάνουν την ανάγκη για πραγματοποίηση προβλέψεων. Όταν δεν υπάρχει αβεβαιότητα στο περιβάλλον και τα πάντα κυλούν ομαλά και ακολουθούν την αναμενόμενη πορεία τους, δεν υπάρχει καμία ουσιαστική ανάγκη για προβλέψεις. Η ειρωνεία έγκειται στο γεγονός ότι σε περιόδους που η ικανότητα πρόβλεψης είναι αισθητά μειωμένη (περίοδοι που χαρακτηρίζονται από συνεχείς και απότομες μεταβολές) η ζήτηση και η ανάγκη για προβλέψεις παρουσιάζει κατακόρυφη αύξηση. Πρακτικά το παραπάνω γεγονός επιβεβαιώνεται σε περιόδους οικονομικών και άλλων κρίσεων κατά την διάρκεια των οποίων η αναζητήσις για συμβούλους προβλέψεων αυξάνεται.

Η βασική αιτία των δυσμενών κριτικών προέρχεται από τις αυξημένες προσδοκίες των ίδιων των χρηστών του. Η πρόβλεψη δεν αποτελεί κάποια μυστική μέθοδος γνώσης του μέλλοντος και άρα η ύπαρξη σφαλμάτων στις προβλέψεις είναι αναπόφευκτη. Αυτό που πραγματικά έχει σημασία είναι να εκτιμώνται όσο πιο ρεαλιστικά γίνεται τα πλεονεκτήματα και οι περιορισμοί των διαφόρων μεθόδων πρόβλεψης και να λαμβάνονται όλα αυτά υπόψη όταν χρησιμοποιούνται οι προβλέψεις σε διαδικασίες σχεδιασμού και λήψης αποφάσεων. Αυτή είναι άλλωστε και η νοοτροπία που ακολουθήθηκε σε αυτήν εδώ την διπλωματική. Όπως θα φανεί και παρακάτω, τέλειες προβλέψεις δεν υπάρχουν, μπορούμε όμως μελετώντας σωστά τα δεδομένα και τις μεθόδους που έχουμε στην διάθεση μας να προσομοιώσουμε όσο το δυνατόν βέλτιστα την αληθινή συμπεριφορά των υπό μελέτη μεταβλητών.

3.2 Θεωρία Χρονοσειρών

3.2.1 Εισαγωγή

Μια από τις διάφορες κατηγοριοποιήσεις δεδομένων η οποία έχει επικρατήσει στον τομέα των προβλέψεων είναι ο διαχωρισμός των δεδομένων σε δυο βασικές κατηγορίες. Τα διαστρωματικά στοιχεία (cross – sectional data) και οι χρονολογικές σειρές (time series), που έχουν επικρατήσει με την ονομασία χρονοσειρές αποτελούν τις 2 αυτές κατηγορίες. Οι χρονοσειρές αποτελούν ένα σύνολο διαδοχικών παρατηρήσεων της τιμής κάποιου φυσικού ή άλλου μεγέθους ανηγμένες στο χρόνο. Υπάρχουν δύο είδη χρονοσειρών:

- *Ντετερμινιστικές* χρονοσειρές όπου οι μελλοντικές παρατηρήσεις μπορούν να προσδιοριστούν επ' ακριβώς από τις εξαρτημένες διαδοχικές παρελθοντικές τιμές.
- *Στοχαστικές* χρονοσειρές, όπου σε αντίθεση με την παραπάνω κατηγορία, το μέλλον περιγράφεται μερικώς από το παρελθόν εξαιτίας της επίδρασης του «τυχαίου παράγοντα» στα δεδομένα. Αυτή η μορφή των χρονοσειρών είναι και αυτή που συναντάται συνήθως, έτσι, θεωρείται γενικά ότι οι χρονοσειρές αντιπροσωπεύουν κυρίως στοχαστικές διαδικασίες και όχι ντετερμινιστικά συστήματα.

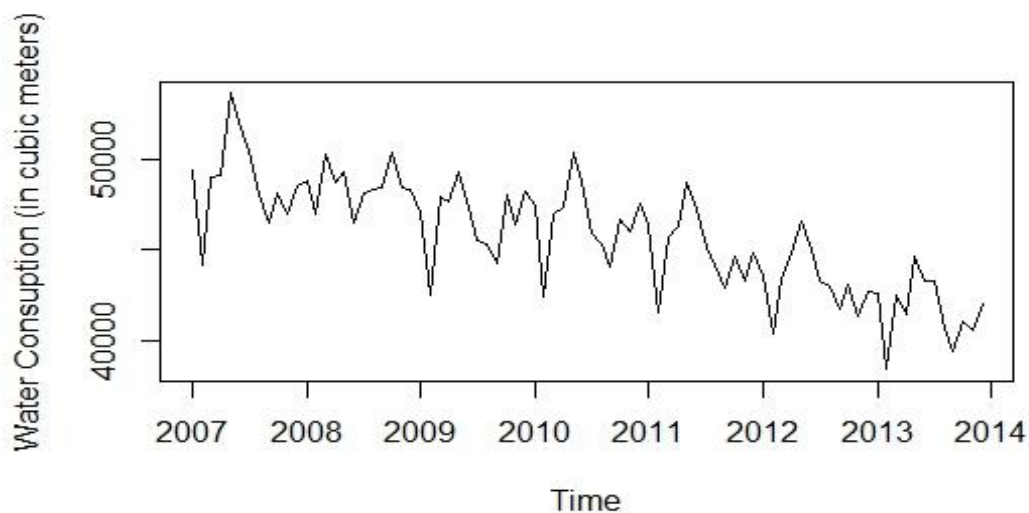
Ο διαχωρισμός των χρονοσειρών που αναφέρθηκε δεν είναι πάντα τόσο προφανής ώστε άμεσα να μπορούμε να χαρακτηρίσουμε μια χρονοσειρά. Παρόλα αυτά, η κατηγοριοποίηση είναι απαραίτητη για την αναγνώριση και την κατανόηση των παραμέτρων που επηρεάζουν την εξέλιξη μιας χρονοσειράς ανεξάρτητα με το είδος των δεδομένων τους.

3.2.2 Αναπαράσταση Χρονοσειρών

Σε περιπτώσεις όπου είναι αναγκαία η οπτική αναπαράσταση δεδομένων μεγάλου όγκου, διάφορες μορφές δισδιάστατης αναπαράστασης προέκυψαν. Οι κυριότερες είναι:

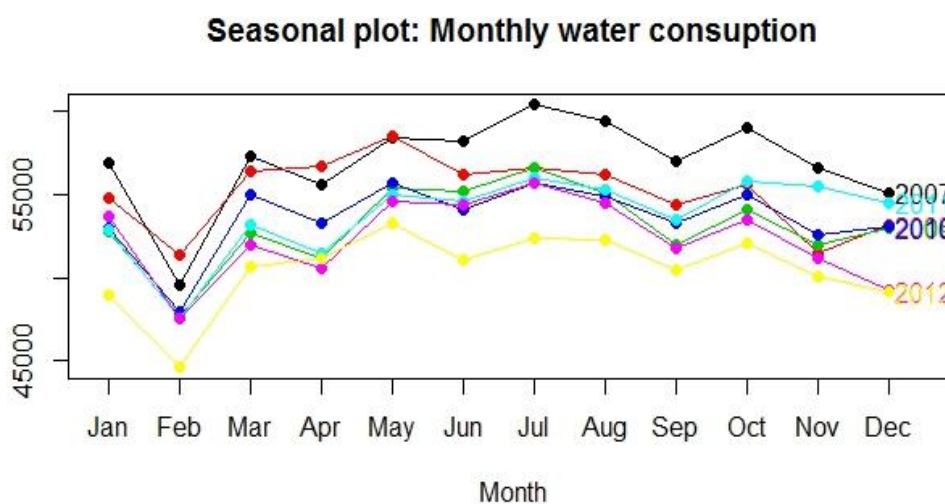
1. *Διαγράμματα Χρόνου* (time plots): Είναι το πλέον προφανές και χρησιμοποιούμενο διάγραμμα και αναπαριστά τα διαθέσιμα δεδομένα στην πάροδο του χρόνου. Μέσω των διαγραμμάτων αυτών γίνονται άμεσα αντιληπτά τα βασικά χαρακτηριστικά των χρονοσειρών όπως η τάση και η

εποχιακότητα. Η χρήση των γραφημάτων αυτών είναι ευρεία σε σχεδόν όλους τους τομείς των επιχειρήσεων αλλά και της καθημερινότητας (καταγραφή καταναλώσεων, πωλήσεις κ.α.)



Εικόνα 3.1-Μηνιαία κατανάλωση νερού για μία περιοχή του δήμου Αθηνών

2. **Εποχιακά Διαγράμματα** (seasonal plots): Αφορούν χρονοσειρές με έντονη εποχικότητα όπως αυτή της εικόνας 3.2.2.1. Αυτό που στην ουσία κάνουν είναι να κατακερματίσουν την δοσμένη χρονοσειρά σε επιμέρους τμήματα με σκοπό να διαφανούν κάποια από τα επιμέρους χαρακτηριστικά της.

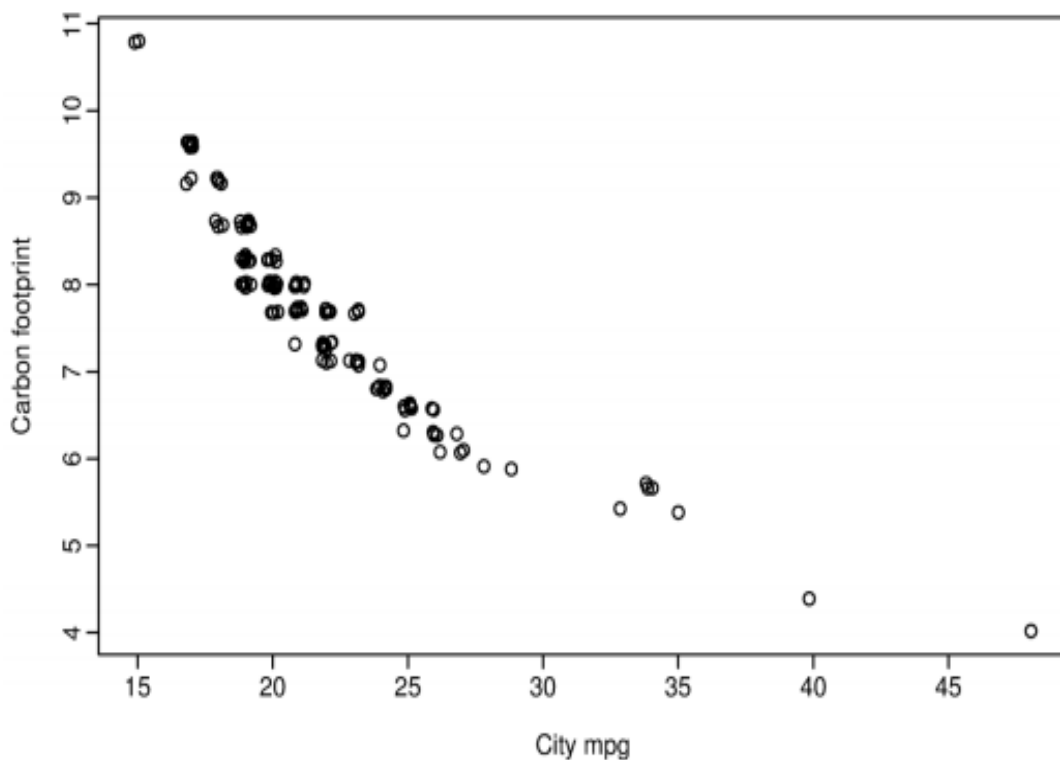


Εικόνα 3.2-Εποχιακό διάγραμμα: μηνιαία κατανάλωση νερού για μία περιοχή του δήμου Αθηνών

Πρακτικά, αυτό που γίνεται εμφανές στο παραπάνω διάγραμμα είναι ότι παρατηρείται μία απότομη πτώση της κατανάλωσης σε όλα τα έτη κατά το μήνα Φεβρουάριο ενώ και στους επόμενους μήνες(με εξαίρεση τον Δεκέμβρη) οι

αυξήσεις και οι μειώσεις στις καταναλώσεις συμφωνούν σε σχέση με την χρονολογία αλλά σε διαφορετική αναλογία κάθε φορά(π.χ. όλες οι χρονοσειρές παρουσιάζουν πτώση το μήνα Νοέμβρη από τον προηγούμενο μήνα, αλλά με διαφορετική ένταση κάθε χρονιά).

3. *Διαγράμματα διασποράς (scatter plots)* : Η χρήση αυτού του τύπου γραφημάτων γίνεται κυρίως για σύγκριση μεταξύ διαφορετικών προϊόντων, υπηρεσιών ή οποιαδήποτε άλλη δυνατή σύγκριση. Το γράφημα παρουσιάζει τα διαφορετικά αυτά επιλεγμένα δεδομένα και αποτυπώνει την σύγκριση σε σχέση μεταξύ δύο διαφορετικών μεγεθών ή χαρακτηριστικών που αφορούν αυτά τα δεδομένα.

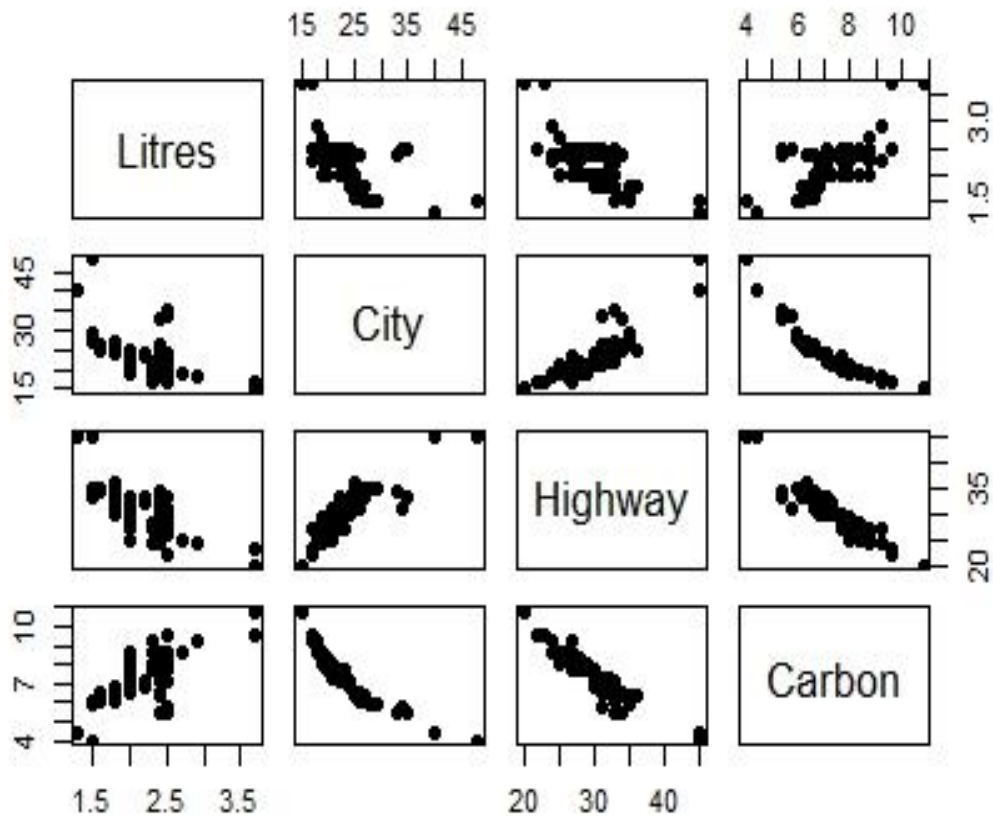


Εικόνα 3.3-Αποτύπωμα Άνθρακα – Εξοικονόμηση Καυσίμου

Από την γραφική αναπαράσταση και εν γένει την οπτικοποίηση των ιστορικών δεδομένων καθίσταται ευκολότερη η διαδικασία αναγνώρισης των βασικών χαρακτηριστικών της χρονοσειράς καθώς και η εύρεση ακραίων και ιδιαίτερων τιμών, των οποίων η διόρθωση ή η αντιμετώπιση είναι απόφαση του αναλυτή. Το παραπάνω σχήμα δείχνει τη σχέση μεταξύ του αποτυπώματος άνθρακα και την οικονομία καυσίμου για τα μικρά αυτοκίνητα από ένα data-set που επεξεργαστήκαμε μέσω του εργαλείου Rstudio χρησιμοποιώντας το πακέτο 'fpp'. Κάθε σημείο στο γράφημα δείχνει έναν τύπο οχήματος. Υπάρχει μια ισχυρή σχέση μεταξύ του μη-γραμμικού μεγέθους του αποτυπώματος άνθρακα του αυτοκινήτου και πόλης με βάση την οικονομία καυσίμου. Οχήματα με καλύτερη οικονομία καυσίμου έχουν μικρότερο αποτύπωμα άνθρακα από τα οχήματα που έκαναν

χρήση πολλών καυσίμων. Ωστόσο, η σχέση δεν είναι γραμμική όπως γίνεται αντιληπτό.

4. Τέλος, όταν υπάρχουν διάφορες δυνητικές μεταβλητές πρόβλεψης, είναι χρήσιμο να σχεδιάσει κανείς την κάθε μεταβλητή σε σχέση με μία άλλη μεταβλητή. Αυτά τα οικόπεδα μπορούν να οργανωθούν σε μια συνολική μήτρα που λέγεται *Μήτρα διαγραμμάτων διασποράς* (scatter-plot matrix)



Εικόνα 3.4-Μήτρα διαγραμμάτων διασποράς

3.2.3 Ποιοτικά Χαρακτηριστικά Χρονοσειρών

Οι παραδοσιακές μέθοδοι ανάλυσης των χρονοσειρών ασχολούνται με την ανάλυση της διακύμανσης της χρονοσειράς σε τέσσερα βασικά συστατικά: την *τάση*, την *κυκλικότητα*, την *εποχικότητα* και τις *μη κανονικές διακυμάνσεις*. Παρακάτω θα αναφερθούμε ξεχωριστά σε κάθε ένα από τα στοιχεία που προσδιορίζουν μία χρονοσειρά:

- Τάση (trend, T): Αποτελεί το πρώτο συστατικό μιας χρονοσειράς και ορίζεται ως μια «μακροπρόθεσμη» μεταβολή του μέσου επιπέδου τιμών μιας χρονοσειράς. Ο ορισμός της τάσης, αν και κοινά αποδεκτός, δημιουργεί ένα πρόβλημα σχετικό με ποια μεταβολή θεωρείται μακροπρόθεσμη έτσι ώστε να μπορεί να εξεταστεί η αντίστοιχη αύξηση ή μείωση στο μέσο επίπεδο. Η απάντηση στο ερώτημα αυτό ποικίλει ανάλογα με την φύση των εξεταζόμενων δεδομένων. Απαραίτητη σε κάθε περίπτωση είναι η ύπαρξη ικανοποιητικού όγκου ιστορικών δεδομένων έτσι ώστε να μπορεί με ασφάλεια να εξαχθεί κάποιο συμπέρασμα σχετικά με την τάση. Η τάση στην γενική της εικόνα μπορεί να είναι ανοδική, πτωτική ή σταθερή και μπορεί να εκτιμηθεί κατά προσέγγιση με μια ευθεία γραμμή ή μια εκθετική καμπύλη ή άλλη οποιαδήποτε άλλη οικογένεια καμπυλών.
- Εποχικότητα (seasonality, S): Ένα εποχιακό μοτίβο υπάρχει όταν μία χρονοσειρά επηρεάζεται από εποχιακά μοτίβα (π.χ. 3-μηνιαία δεδομένα, μηνιαία κ.α.). Η εποχικότητα είναι πάντα συγκεκριμένης και γνωστής περιόδου. Η εποχικότητα είναι μετά την τάση το πιο εύκολα αναγνωρίσιμο χαρακτηριστικό μιας χρονοσειράς από την γραφική της αναπαράσταση. Επίσης εύκολος είναι και ο τρόπος αντιμετώπισης της επίδρασης της εποχικότητας στα δεδομένα μιας χρονοσειράς καθώς οι αλλαγές που προκαλεί στα δεδομένα επαναλαμβάνονται κατά την πάροδο του χρόνου, στα ίδια χρονικά διαστήματα και με την ίδια ποσοστιαία αλλαγή. Παραδείγματα εποχιακών χρονοσειρών είναι οι πωλήσεις παγωτών που εκτοξεύονται το καλοκαίρι ή ο αριθμός επιβατών πλοίων και αεροπλάνων που πολλαπλασιάζονται κατά τις περιόδους των γιορτών.
- Κυκλικότητα (cyclic, C): Υπάρχει κυκλικό μοτίβο όταν εμφανίζονται δεδομένα που παρουσιάζουν διακυμάνσεις ως προς τα πλάτη τους (ανεβαίνουν και πέφτουν) που δεν είναι της σταθερής περιόδου. Η διάρκεια των διακυμάνσεων είναι τουλάχιστον 2 χρόνων συνήθως. Η κυκλικότητα οφείλεται κατά κύριο λόγο σε εξωγενείς παράγοντες. Πολλοί συγχέουν την κυκλική συμπεριφορά με την εποχιακή συμπεριφορά, αλλά είναι πραγματικά αρκετά διαφορετική. Εάν οι διακυμάνσεις δεν έχουν σταθερό χρονικό διάστημα, τότε υπάρχει κυκλικότητα. Εάν η περίοδος είναι αμετάβλητη και συνδέονται με κάποια πτυχή του ημερολογίου, τότε το

μοτίβο είναι εποχιακό. Σε γενικές γραμμές, το μέσο μήκος των κύκλων είναι μεγαλύτερο από το μήκος του εποχιακού μοτίβου, και το μέγεθος των κύκλων τείνει να είναι πιο μεταβλητό από το μέγεθος του εποχιακού μοτίβου.

- Μη κανονικές διακυμάνσεις (abnormal fluctuations): Είναι εκείνες οι παρατηρήσεις που εμφανίζονται στην γραφική απεικόνιση της χρονοσειράς ως απότομες αλλαγές στο πρότυπο συμπεριφοράς της. Τέτοιες αλλαγές που έχουν είτε παροδικό είτε ακόμα και μόνιμο χαρακτήρα δε θα μπορούσαν να προβλεφθούν με χρήση αποκλειστικά των ιστορικών δεδομένων. Οι αλλαγές με παροδική διάρκεια που η επίδρασή τους διαρκεί για σύντομο χρονικό διάστημα ονομάζονται *outliers* ή *special events*. Η αναγνώριση και το φιλτράρισμα αυτών τιμών από τα πραγματικά δεδομένα είναι απαραίτητο να γίνει έτσι να οδηγηθούμε σε σωστά συμπεράσματα κατά την επεξεργασία των δεδομένων αυτών. Στις μη κανονικές διακυμάνσεις είναι και οι αλλαγές με μόνιμο χαρακτήρα στην χρονοσειρά που οι αλλαγές τους θα συνεχιστούν και στο μέλλον ονομάζονται *level shifts*. Ένα παράδειγμα αλλαγής level shift είναι η πτώση του επιπέδου των πωλήσεων μιας εταιρείας λόγω εισαγωγής στην αγορά μιας ανταγωνίστριας εταιρείας στον ίδιο κλάδο.
- Τυχειότητα (randomness, R) : Αναφέρεται σε τυχαίες και μη προβλέψιμες διακυμάνσεις των δεδομένων μιας χρονοσειράς. Είναι η διακύμανση που παραμένει στη χρονοσειρά έπειτα από την απομόνωση της τάσης, της κυκλικότητας και της εποχικότητας . Οι διακυμάνσεις τυχειότητας ενδέχεται να είναι ασυνέχειες με παροδικό ή μόνιμο χαρακτήρα (ασυνήθιστες τιμές, outliers, ή αλλαγή επιπέδου, level shift, αντίστοιχα).

Οι τέσσερις βασικές συνιστώσες των χρονοσειρών μπορούν να απομονωθούν από αυτές με μεθόδους αποσύνθεσης. Με απομόνωση της εποχικότητας και της τυχειότητας προκύπτει η σειρά τάσης-κύκλου κάθε αρχικής χρονοσειράς, στην οποία εφαρμόζονται εν συνεχεία οι επιλεγμένες μέθοδοι πρόβλεψης και οι τιμές που προκύπτουν από την εκάστοτε πρόβλεψη εποχικοποιούνται ξανά, ώστε να ληφθεί το τελικό αποτέλεσμα. Με τη διαδικασία αυτή επιτυγχάνεται αύξηση της αξιοπιστίας των αποτελεσμάτων της πρόβλεψης και δίνεται στο χρήστη η δυνατότητα να εισάγει και τη δική του κρίση, βάσει της εμπειρίας και της διαίσθησής του. Ο υπολογισμός της σειράς τάσης-κύκλου γίνεται με διάφορες μεθόδους (εξομάλυνση ή αποσύνθεση), με την απλούστερη αυτών να είναι η εξομάλυνση με χρήση κινητών μέσων όρων. Αμέσως μετά αναλύεται η εξομάλυνση με χρήση των κινητών μέσων όρων, ενώ στην επόμενη παράγραφο (παράγραφος 3.2.3.2) αναλύεται η μέθοδος της κλασικής αποσύνθεσης η οποία είναι δημοφιλέστερη και η πιο αποτελεσματική μέθοδος γενικά.

3.2.3.1 Μέθοδοι Εξομάλυνσης με Κινητούς Μέσους Όρους

Η μέθοδος των κινητών μέσων όρων δεν παρέχει πληροφορίες σχετικά με την εποχικότητα και την τυχαιότητα της χρονοσειράς τις οποίες απομακρύνει από αυτή για την παραγωγή της σειράς τάσης-κύκλου. Πιο λεπτομερώς, υπάρχουν τέσσερα είδη κινητών μέσων όρων για την παραγωγή της εξομαλυμένης σειράς. Σημαντική είναι η επιλογή του μήκους n πριν τον υπολογισμό των κινητών μέσων όρων. Με την αύξηση του μήκους n επιτυγχάνεται και μεγαλύτερη εξομάλυνση των δεδομένων, ωστόσο πολύ μεγάλη τιμή του μήκους οδηγεί στον υπολογισμό της μέσης τιμής των δεδομένων της χρονοσειράς και δεν διακρίνονται η τάση και ο κύκλος που μας ενδιαφέρουν, ενώ επιπλέον αυξάνεται πολύ ο αριθμός των κενών τιμών στην αρχή και στο τέλος της παραγόμενης σειράς. Να σημειωθεί ότι στην περίπτωση έντονης εμφάνισης εποχικότητας προτιμάται η χρήση μήκους ίσου (ή μεγαλύτερου) με το μήκος της περιόδου της εποχικότητας.

- Ο απλός κινητός μέσος όρος (ΚΜΟ) πρόκειται για τον απλό μέσο όρο n τιμών (όπου n : περιττός για τη διατήρηση συμμετρίας στους υπολογισμούς) της αρχικής χρονοσειράς γύρω από την εκάστοτε παρατήρηση. Υπολογίζοντας τον απλό κινητό μέσο όρο για όλες τις παρατηρήσεις της χρονοσειράς προκύπτει η επιθυμητή σειρά τάσης-κύκλου.
- Ο σταθμισμένος κινητός μέσος όρος (ΣΚΜΟ) αποτελεί παραλλαγή του απλού ΚΜΟ με συμμετοχή των γειτονικών παρατηρήσεων στον υπολογισμό της σειράς τάσης-κύκλου με άνισα βάρη. Το άθροισμα από τα βάρη πρέπει να ισούται με τη μονάδα, δίνεται προσοχή ώστε να είναι συμμετρικά ως προς την εκάστοτε (τρέχουσα) παρατήρηση, ενώ η επιλογή τους γίνεται με στόχο να δοθεί μεγαλύτερη βαρύτητα σε παρατηρήσεις πλησίον της εκάστοτε παρατήρησης και μικρότερη στις πιο μακρινές παρατηρήσεις. Η χρήση του σταθμισμένου κινητού μέσου όρου οδηγεί σε περισσότερο εξομαλυμένη σειρά τάσης-κύκλου.
- Στο διπλό κινητό μέσο όρο (ΔΚΜΟ) με διπλή εφαρμογή του απλού ΚΜΟ, με ίσα ή και άνισα μήκη, οδηγούμαστε σε διπλή εξομάλυνση, αλλά και στην ύπαρξη περισσότερων κενών τιμών συγκριτικά πάντα με τον απλό κινητό μέσο όρο (ΚΜΟ). Ακόμη, η εφαρμογή του ΔΚΜΟ ισοδυναμεί με την εφαρμογή του ΣΚΜΟ με ορισμένα βάρη (εξαρτώμενα από την κάθε περίπτωση).
- Ο κεντρικός κινητός μέσος όρος (ΚΚΜΟ) αποτελεί συνδυασμό του απλού και του διπλού κινητού μέσου όρου και χρησιμεύει για τον υπολογισμό κινητού μέσου όρου άρτιου μήκους. Η εφαρμογή του κεντρικού κινητού μέσου όρου ισοδυναμεί με ένα διπλό κινητό μέσο όρο $2 \cdot n$ (όπου n το μήκος του απλού ΚΜΟ), καθώς και με ένα σταθμισμένο κινητό μέσο όρο με ορισμένα βάρη. Η χρήση του ΚΚΜΟ είναι

πλέον συνήθης επί χρονοσειρών με παρατηρήσεις μηνιαίες, τριμηνιαίες, εξαμηνιαίες, κ.λ.π για την εξομάλυνση της όποια εποχικότητας ενδέχεται να παρουσιάζουν (με δώδεκα, τέσσερις και δύο παρατηρήσεις ανά έτος αντίστοιχα).

3.2.3.2 Κλασική Αποσύνθεση

Όσον αφορά στη μέθοδο της αποσύνθεσης, η πιο κοινή είναι η κλασική μέθοδος αποσύνθεσης, η οποία εφαρμόζεται τόσο σε χρονοσειρά με πολλαπλασιαστική όσο και με προσθετική διατύπωση της αποσύνθεσης (με την πολλαπλασιαστική να αποτελεί τον πιο συνηθισμένο τρόπο διατύπωσης με καλύτερα αποτελέσματα).

Το πολλαπλασιαστικό μοντέλο διατυπώνεται ως εξής:

$$Y_t = S_t \times T_t \times C_t \times R_t \quad (3.2.3.2.1)$$

ενώ το προσθετικό είναι :

$$Y_t = S_t + T_t + C_t + R_t$$

όπου S,T,C,R πρόκειται για τις συνιστώσες εποχικότητας, τάσης, κύκλου και τυχαιότητας αντίστοιχα για τη χρονική περίοδο t.

Υπάρχει η δυνατότητα παραγωγής προβλέψεων μέσω της αποσύνθεσης, όπου έπειτα από ανάλυση της αρχικής χρονοσειράς στις συνιστώσες της ακολουθεί η σύνθεσή τους ξανά για τη λήψη της τελικής τιμής πρόβλεψης. Η βασική συνιστώσα που αφαιρείται από τα δεδομένα είναι η εποχικότητα, η οποία έπειτα από την παραγωγή της πρόβλεψης ενσωματώνεται και πάλι στην τελική χρονοσειρά. Στη συνέχεια αναλύονται τα βήματα που ακολουθούνται κατά την κλασική αποσύνθεση σε χρονοσειρά πολλαπλασιαστικής σχέσης των συνιστωσών της [σχέση (3.2.3.2.1)]. Η μεθοδολογία της κλασικής αποσύνθεσης αποτελείται από πέντε βασικά βήματα:

- 1) Αρχικά, επιλέγεται ένα από τα τέσσερα είδη κινητού μέσου όρου που σημειώθηκαν στην παράγραφο 3.1 με βάση το μήκος της εποχικότητας των δεδομένων. Ο υπολογιζόμενος κινητός μέσος όρος είναι απαλλαγμένος από εποχικότητα, ενώ ενδέχεται να περιέχει μόνο μικρή τυχαιότητα, δίνεται έτσι μια καλή εικόνα της πορείας των δεδομένων της χρονοσειράς σχετικά με την τάση και την κυκλικότητα. Μπορεί να θεωρηθεί επομένως ότι ισχύει το εξής:

$$KMO(n)_t = T_t \times C_t$$

Στην παραπάνω σχέση συμβολίζονται με n το μήκος του κινητού μέσου όρου, με T η συνιστώσα της τάσης και με C η συνιστώσα του κύκλου της χρονοσειράς. Σε περίπτωση που τα δεδομένα έχουν μήκος εποχικότητας άρτιο αριθμό, τότε προτιμάται η χρήση ΚΚΜΟ αντίστοιχου μήκους.

- 2) Έπειτα διαιρούνται τα πραγματικά δεδομένα με τους κεντρικούς μέσους όρους που υπολογίστηκαν στο 1^ο βήμα, όπως φαίνεται στην επόμενη σχέση. Λαμβάνονται με τον τρόπο αυτό οι λόγοι εποχικότητας, που χαρακτηρίζονται ακόμη από την εποχικότητα και την τυχαιότητα της χρονοσειράς.

$$\frac{Y_t}{\text{ΚΚΜΟ}(n)_t} = \frac{T_t \times C_t \times S_t \times R_t}{T_t \times C_t} = S_t \times R_t$$

- 3) Στο βήμα αυτό, στόχος είναι η απαλοιφή της τυχαιότητας από τους λόγους εποχικότητας που μόλις υπολογίστηκαν. Για να συμβεί αυτό, υπολογίζονται οι μέσοι όροι των αντίστοιχων λόγων εποχικότητας, οι οποίοι αποτελούν τους δείκτες εποχικότητας της χρονοσειράς (της αρχικής χρονοσειράς). Ενδέχεται βέβαια να χρειάζεται ακόμη μια επεξεργασία αυτών των μέσων όρων ώστε το άθροισμά τους να είναι ίσο με το μήκος της εποχικότητας. Η διαδικασία αυτή ονομάζεται κανονικοποίηση των δεικτών εποχικότητας της χρονοσειράς και γίνεται με υπολογισμό των αντίστοιχων μέσων όρων των λόγων εποχικότητας, χωρίς τις ακραίες τιμές (τη μέγιστη και την ελάχιστη τιμή τους). Συγκεκριμένα, έπειτα από τον υπολογισμό των λόγων εποχικότητας, αγνοείται η μέγιστη και η ελάχιστη τιμή τους και υπολογίζεται ο μέσος όρος των υπόλοιπων. Οι νέοι μέσοι όροι ονομάζονται μη κανονικοποιημένοι δείκτες εποχικότητας και το πλήθος τους ισούται με το μήκος της εποχικότητας. Μετά, υπολογίζεται ο συντελεστής κανονικοποίησης, ως ο λόγος του μήκους της εποχικότητας προς το άθροισμα όλων των μη κανονικοποιημένων δεικτών εποχικότητας. Με πολλαπλασιασμό των μη κανονικοποιημένων δεικτών επί τον συντελεστή κανονικοποίησης προκύπτουν εν τέλει οι κανονικοποιημένοι δείκτες εποχικότητας.

- 4) Στο βήμα αυτό, διαιρούνται τα δεδομένα της πραγματικής χρονοσειράς με τους αντίστοιχους κανονικοποιημένους δείκτες εποχικότητας που υπολογίστηκαν στο 3^ο βήμα. Παράγεται έτσι η αποεποχικοποιημένη σειρά, η οποία περιέχει πλέον τις υπόλοιπες τρεις συνιστώσες της, τάση, κύκλο και τυχαιότητα, έχοντας απαλλαγεί από την εποχικότητα.

$$\frac{Y_t}{S_t} = \frac{T_t \times C_t \times S_t \times R_t}{S_t} = T_t \times C_t \times R_t$$

5) Ενδιαφέρον πλέον έχει η απαλοιφή και της τυχαιότητας από την ήδη αποεποχικοποιημένη σειρά. Υπολογίζεται, για το σκοπό αυτό, ο κινητός μέσος όρος μήκους $n = 3$ ή $n = 6$ παρατηρήσεων την αποεποχικοποιημένης σειράς, ώστε να παραχθεί μια ικανοποιητική και ακριβής σειρά τάσης-κύκλου. Για βέλτιστα αποτελέσματα συνίσταται η χρήση διπλού κινητού μέσου όρου, ΔΚΜΟ(3x3). Στην περίπτωση που ζητείται η συνιστώσα της τυχαιότητας, υπολογίζεται με βάση τη σχέση:

$$\frac{T_t \times C_t \times R_t}{\text{ΚΚΜΟ}(3 \times 3)_t} = \frac{T_t \times C_t \times R_t}{T_t \times C_t} = R_t$$

3.3 Μοντέλα Πρόβλεψης

Υπάρχουν τρεις μεγάλες κατηγορίες μεθόδων πρόβλεψης:

- Ποσοτικές
- Κριτικές
- Τεχνολογικές

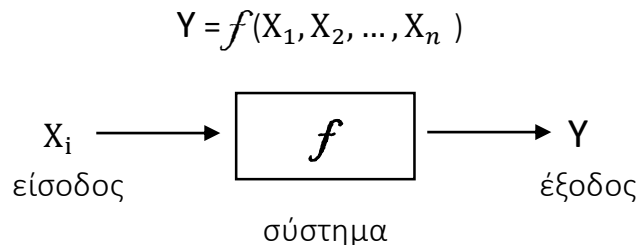
Παρακάτω αναλύονται αυτές οι κατηγορίες τεχνικών προβλέψεων

3.3.1 Ποσοτικές Μέθοδοι

Οι ποσοτικές μέθοδοι διακρίνονται με την σειρά τους σε δύο μεγάλες κατηγορίες: το *μοντέλο χρονοσειρών* (time series model) και το *αιτιοκρατικό ή επεξηγματικό μοντέλο* (explanatory model). Μελετώντας τα χαρακτηριστικά και τις ιδιότητες αυτών των μοντέλων γίνεται ευκολότερη η κατανόηση των βασικών υποθέσεων πάνω στις οποίες στηρίζεται κάθε ποσοτική μέθοδος, ενώ ταυτόχρονα εντοπίζονται τα πλεονεκτήματα και μειονεκτήματα της εφαρμογής σε συγκεκριμένες καταστάσεις. Να σημειωθεί εδώ ότι σε αυτήν την διπλωματική χρησιμοποιήθηκαν αποκλειστικά μέθοδοι χρονοσειρών. Παρακάτω γίνεται η παρουσίαση τους.

- Μοντέλο χρονοσειρών
Είναι ξεκάθαρα το πιο διαδεδομένο είδος ποσοτικής μεθόδου. Για την χρησιμοποίησή του θα πρέπει να υπάρχει ικανός όγκος δεδομένων του υπό πρόβλεψη μεγέθους σε προηγούμενες και σταθερές χρονικές περιόδους (ιστορικά δεδομένα). Όσο περισσότερα δεδομένα έχουμε, τόσο ικανότερο θα είναι το

παραγόμενο μοντέλο πρόβλεψης. Οι προβλέψεις πραγματοποιούνται αφού πρώτα αναγνωρισθεί ένα συγκεκριμένο πρότυπο στα δεδομένα το οποίο επαναλαμβάνεται με το χρόνο και παραμένει σταθερό. Έπειτα της αναγνώρισης του προτύπου γίνεται η προέκταση του σε τόσες περιόδους όσοι και το επιθυμητό χρονικό εύρος πρόβλεψης. Η μαθηματική σχέση που περιγράφει το μοντέλο είναι:



όπου X_1, X_2, \dots, X_n είναι τα ιστορικά δεδομένα έως και την χρονική περίοδο n . Η έξοδος Y είναι η παραγόμενη πρόβλεψη και f είναι το μοντέλο πρόβλεψης.

Σε αυτές τις μεθόδους συγκαταλέγονται γνωστοί μέθοδοι προβλέψεων όπως η εξομάλυνση και οι αυτοπαλινδρομικές μέθοδοι κινητού μέσου όρου (AR.I.M.A.) οι οποίες θα επεξηγηθούν διεξοδικότερα σε επόμενη ενότητα. Γίνεται επομένως αντιληπτό ότι μία από τις αδυναμίες του συγκεκριμένου μοντέλου είναι ότι δεν μπορεί να παράξει αξιόπιστες προβλέψεις σε περιπτώσεις που υπάρχουν παράγοντες εξωγενείς που επηρεάζουν το υπό πρόβλεψη μέγεθος και δεν μπορούν να προσδιοριστούν με βάση την φύση των ιστορικών στοιχείων.

- Αιτιοκρατικό μοντέλο

Το μοντέλο αυτό στηρίζεται στην βασική υπόθεση ότι υπάρχει μία σταθερή σχέση μεταξύ του υπό πρόβλεψη μεγέθους και ορισμένων παραμέτρων που το επηρεάζουν. Το μαθηματικό μοντέλο περιγράφεται με παρόμοιο τρόπο όπως και στο μοντέλο των χρονοσειρών:

$$Y = f(X_1, X_2, \dots, X_n),$$

όπου X_i με $i = 1, 2, \dots, n$ είναι οι ανεξάρτητες μεταβλητές που έχει αποδειχθεί ότι επηρεάζουν το υπό πρόβλεψη μέγεθος.

Υπάρχει όμως μία σοβαρή διαφορά μεταξύ των δύο μοντέλων. Ενώ στα μοντέλα χρονοσειρών η συνάρτηση που περιγράφεται το σύστημα είναι προκαθορισμένη και προσδιορίζεται από το πρότυπο που υποθέτουμε ότι ακολουθούν τα ιστορικά δεδομένα, στα αιτιοκρατικά μοντέλα αποτελεί προϊόν υπολογιστικών πράξεων μεταξύ διαφόρων εισόδων της ανεξάρτητης μεταβλητής X_i και του ακριβή τύπου πρόβλεψης f . Το βασικό μειονέκτημα της μεθόδου αυτής είναι ότι χρειάζεται μεγάλο όγκο τιμών ακόμα και για τις ανεξάρτητες μεταβλητές που έχει αναγνωρισθεί ότι επηρεάζουν τις μελλοντικές τιμές, προκειμένου να προσδιορίσει

το κατάλληλο μοντέλο πρόβλεψης και να παράξει αξιόπιστα δεδομένα. Στις επεξηγηματικές μεθόδους ανήκουν οι μέθοδοι παλινδρόμησης όπως και οι οικονομετρικές μέθοδοι.

3.3.2 Κριτικές Μέθοδοι

Οι κριτικές μέθοδοι πρόβλεψης (Judgemental forecasting) δεν έχουν τις ίδιες απαιτήσεις σε δεδομένα όπως οι ποσοτικές μέθοδοι. Τα δεδομένα των κριτικών μεθόδων αποτελούν προϊόν γνώσης, διαίσθησης και κριτικής σκέψης από πλευράς έμπειρων στελεχών και ειδικών σε θέματα προβλέψεων. Οι κριτικές μέθοδοι πρόβλεψης μπορούν να συνυπολογίσουν ειδικά γεγονότα και ενέργειες, ενώ ταυτόχρονα μπορούν να αντισταθμίσουν ανεπάρκειες και ελλείψεις σε ιστορικά δεδομένα. Παράλληλα μπορούν να συνδυαστούν και με κατάλληλες στατιστικές μεθόδους σε περιπτώσεις όπου τα πολύπλοκα μοντέλα που χρησιμοποιούνται δεν βγάζουν τα αναμενόμενα αποτελέσματα. Το μεγάλο πρόβλημα των μεθόδων αυτών είναι ότι σε πολλές περιπτώσεις δεν είναι αμερόληπτες (biased forecasting) καθώς πολλές φορές οι προσδοκίες ή η αισιοδοξία και η απαισιοδοξία μπορούν να επηρεάσουν αρνητικά την ποιότητα της κριτικής πρόβλεψης.

3.3.3 Τεχνολογικές Μέθοδοι

Οι τεχνολογικές μέθοδοι πρόβλεψης χρησιμοποιούνται για μακροπρόθεσμες προβλέψεις σχετικά με τεχνολογικά, οικονομικά, κοινωνικά και πολιτικά θέματα. Διακρίνονται σε δύο κατηγορίες: στις *διερευνητικές* (exploratory) και στις *κανονιστικές* (normative). Οι διερευνητικές μέθοδοι ξεκινούν από το παρελθόν ή το παρόν και εξετάζοντας όλες τις πιθανές περιπτώσεις οδηγούνται στο μέλλον. Στον αντίποδα, υπάρχουν οι κανονιστικές μέθοδοι που πρώτα καθορίζουν όλους τους μελλοντικούς στόχους και έπειτα εξετάζουν τη δυνατότητα επίτευξης τους λαμβάνοντας υπόψιν τους περιορισμούς, τους διαθέσιμους πόρους αλλά και τις τεχνολογίες.

3.4 Στατιστική Ανάλυση

Ο υπολογισμός βασικών στατιστικών δεικτών μίας χρονοσειράς αποτελεί σημαντικό κομμάτι ανάλυσης της που βοηθάει στην αποσαφήνιση των μεθόδων που θα χρησιμοποιηθούν για την παραγωγή των προβλέψεων. Επιτρέπει στους αναλυτές να έχουν μία σαφή και ολοκληρωμένη εικόνα για τα ποιοτικά χαρακτηριστικά της χρονοσειράς και για αυτό η στατιστική ανάλυση τόσο των χρονοσειρών, όσο και των προβλέψεων και κατ' επέκταση και των μοντέλων πρόβλεψης έχει αποτελέσει αντικείμενο πολλών ερευνητών του τομέα των προβλέψεων, Makridakis (1993) και Armstrong (2001). Η στατιστική ανάλυση χωρίζεται σε δύο βασικές κατηγορίες που παρουσιάζονται στις επόμενες ενότητες.

3.4.1 Βασική Στατιστική Ανάλυση

Οι βασικοί στατιστικοί δείκτες που υπολογίζονται στον n αριθμό παρατηρήσεων της μελετώμενης χρονοσειράς είναι:

- Μέση τιμή, $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$
- Μέγιστη και Ελάχιστη τιμή (Maximum and Minimum) της χρονοσειράς

- Τυπική απόκλιση, $\sigma = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}$

Ο δείκτης αυτός εκφράζει πόσο μεγάλη είναι η διασπορά των τιμών των δεδομένων γύρω από τη μέση τιμή,

- Διακύμανση, σ^2
Δείχνει το βαθμό συγκέντρωσης των παρατηρήσεων γύρω από τη μέση τιμή (όμοια με την τυπική απόκλιση, αλλά με τα μεγέθη υψωμένα στο τετράγωνο).
- Συνδιακύμανση, $Cov(X, Y) = \frac{1}{n} \sum_{i=1}^n [(Y_i - \bar{Y}) \cdot (X_i - \bar{X})]$

Ο δείκτης αυτός εμφανίζει τη σχέση ανάμεσα σε δύο μεταβλητές, συγκρίνοντας τη φορά μεταβολής τους. Δείχνει δηλαδή αν μεταβάλλονται προς την ίδια (θετική συνδιακύμανση) ή προς την αντίθετη κατεύθυνση (αρνητική συνδιακύμανση)

και τη συσχέτιση ή μη των συνόλων των δύο μεταβλητών (όταν δεν υπάρχει συσχέτιση η συνδιακύμανση ισούται πάντα με το μηδέν).

- Συντελεστής γραμμικής συσχέτισης, $r_{xy} = \frac{\sum_{i=1}^n [(Y_i - \bar{Y}) \cdot (X_i - \bar{X})]}{\sum_{i=1}^n (Y_i - \bar{Y})^2 \cdot \sum_{i=1}^n (X_i - \bar{X})^2}$

Είναι το μέτρο της συγκέντρωσης των σημείων ενός διαγράμματος διασποράς γύρω από την ευθεία παλινδρόμησης, που δηλώνει τη συσχέτιση μεταξύ δύο συνόλων δεδομένων (κατεύθυνση και βαθμό συσχέτισης).

- Για $0 < r < 1$ οι δύο μεταβλητές X και Y είναι θετικά γραμμικά συσχετισμένες.
- Για $-1 < r < 0$ οι δύο μεταβλητές είναι αρνητικά γραμμικά συσχετισμένες.
- Για $r = \pm 1$ υπάρχει τέλεια θετική γραμμική συσχέτιση, οπότε όλα τα σημεία βρίσκονται επί της ευθείας παλινδρόμησης.
- Για $r = 0$ δεν υπάρχει καθόλου συσχέτιση μεταξύ των δύο μεταβλητών.

Να σημειώσουμε ότι όσο η τιμή του δείκτη r τείνει στις ακρινές τιμές του διαστήματος $[-1, 1]$ στο οποίο κινείται, τόσο πιο ισχυρή είναι η συσχέτιση των δύο μεταβλητών μεταξύ τους.

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

Ο δείκτης αυτός εκφράζει τη συσχέτιση μεταξύ παρατηρήσεων της ίδιας πλέον μεταβλητής, οι οποίες έχουν χρονική διαφορά k περιόδους. Η τιμή του κινείται στο διάστημα $[0,1]$ και όταν είναι κοντά στο άνω όριο της φανερώνει μεγάλη συσχέτιση, ενώ το αντίθετο συμβαίνει για τιμή κοντά στο μηδέν. Η σημασία του δείκτη αυτού φαίνεται στην περίπτωση που μας ενδιαφέρει ο εντοπισμός της εποχικής συμπεριφοράς μια χρονοσειράς (με αλγεβρικό τρόπο).

- Συντελεστής μεταβλητότητας, $Cv = \frac{\sigma}{\bar{Y}}$

Πρόκειται για κανονικοποιημένο μέτρο της διασποράς των παρατηρήσεων δείγματος ή πληθυσμού και πλεονεκτεί έναντι της απλής τυπικής απόκλισης (σ) λόγω της μη συμπερίληψης της επίδρασης του επιπέδου των παρατηρήσεων. Δε γίνεται όμως να υπολογιστεί σε περίπτωση μηδενικής μέσης τιμής.

3.4.2 Δείκτες Ακρίβειας Προβλέψεων

Η σημασία της συγκεκριμένης κατηγορίας της στατιστικής ανάλυσης στον κλάδο των προβλέψεων είναι καίριας σημασίας καθώς αποτελεί βασικό εργαλείο για την αξιολόγηση μεθόδων αλλά και για τον χαρακτηρισμό τους σχετικά με τον τρόπο προσέγγισης την μεθοδολογία πρόβλεψης και την πραγματική χρονοσειράς. Κύρια έννοια για να οριστούν οι μετέπειτα δείκτες της στατιστικής ακρίβειας προβλέψεων αποτελεί το σφάλμα, δηλαδή η διαφορά μεταξύ της πραγματικής τιμής και της πρόβλεψης για μία περίοδο, το οποίο ορίζεται ως εξής:

$$e_i = Y_i - F_i$$

Δεδομένα	Πρόβλεψη
Y_1	F_1
Y_2	F_2
Y_3	F_3
...	...
Y_{n-2}	F_{n-2}
Y_{n-1}	F_{n-1}
Y_n	F_n
	F_{n+1}
	...
	F_{n+h}

forecast model fitting
(in-sample forecasts)

Όπως φαίνεται και από το παραπάνω πίνακα, κάθε φορά που θέλουμε να παράξουμε μία πρόβλεψη για h περιόδους στο μέλλον, το μοντέλο παράγει τιμές από την χρονική στιγμή t_1 έως και t_{n+h} . Οι τιμές F_1 έως F_n είναι οι προσαρμοσμένες τιμές του μοντέλου πρόβλεψης (fitted values) πάνω στα πραγματικά δεδομένα και συνήθως χρησιμεύουν στο προσδιορισμό των παραμέτρων του μοντέλου για πιο αξιόπιστες προβλέψεις, ενώ οι τιμές F_{n+1} έως F_{n+h} είναι οι προβλέψεις για το μέλλον που οριστικοποιούνται αφού το μοντέλο έχει καταλήξει στον πλήρη καθορισμό όλων των παραμέτρων του.

Είναι προφανές πως η τιμή του σφάλματος δεν μπορεί να υπολογιστεί αν δεν υπάρχουν για την ίδια χρονική περίοδο τόσο οι πραγματικές τιμές της χρονοσειράς όσο επίσης και οι τιμές πρόβλεψης. Οπότε μπορεί να γίνει διαχωρισμός των σφαλμάτων, σε σφάλμα του μοντέλου πρόβλεψης (in-sample error) το οποίο προκύπτει από τις διαφορές των πραγματικών τιμών της χρονοσειράς που είναι ήδη διαθέσιμες και των τιμών του μοντέλου πρόβλεψης για αυτές τις χρονικές περιόδους (F_1 έως F_n) και στο πραγματικό σφάλμα (out-of-sample error) που προκύπτει από τη διαφορά της πραγματικής μελλοντικής τιμής της χρονοσειράς που θα γίνει γνωστή μετά από το αντίστοιχο χρονικό διάστημα και τις πρόβλεψης που έχει παραχθεί από το αντίστοιχο μοντέλο για εκείνη τη χρονική περίοδο.

Οι δείκτες που ακολουθούν είναι ορισμένοι έτσι ώστε να εκφράζουν το σφάλμα της μεθόδου πρόβλεψης για n περιόδους, αλλά με κατάλληλες αλλαγές στις περιόδους που αναφέρονται μπορεί να εκφράσουν και το πραγματικό σφάλμα της πρόβλεψης. Παρακάτω λοιπόν, γίνεται παράθεση των βασικότερων δεικτών σφάλματος:

- Μέσο σφάλμα (Mean Error):

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

Το μέσο σφάλμα (Mean Error) εκφράζει ένα μέτρο της συστηματικότητας του σφάλματος. Συγκεκριμένα, όσο μικρότερη τιμή λαμβάνει (κοντά στο μηδέν), τόσο περισσότερο χαρακτηρίζονται τα σφάλματα από τυχαιότητα. Και όσο απομακρύνεται η τιμή του σφάλματος από το μηδέν, τότε αναπτύσσεται μια τάση για απαισιοδοξία των προβλέψεων για θετικές τιμές του σφάλματος και μια τάση για αισιοδοξία για αρνητικές τιμές αντίστοιχα.

- Μέσο απόλυτο σφάλμα (Mean Absolut Error):

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

Το μέσο απόλυτο σφάλμα (Mean Absolute Error) δείχνει την ακρίβεια της μεθόδου, χωρίς να δείχνει την κατεύθυνση της πρόβλεψης. Όσο μεγαλύτερη η τιμή του δείκτη αυτού, τόσο μικρότερη η ακρίβεια του μοντέλου πρόβλεψης, όπως φαίνεται και από τον τύπο υπολογισμού του. Δεν είναι πάντα κατάλληλο για χρονοσειρές διαφορετικού ύψους και ίδιας διασποράς, καθώς δεν είναι δυνατή η λήψη συγκεκριμένων πληροφοριών.

- Μέσο τετραγωνικό σφάλμα (Mean Squared Error):

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

Το μέσο τετραγωνικό σφάλμα (Mean Squared Error) είναι ένα μέτρο σύγκρισης της ακρίβειας πρόβλεψης των διαφόρων μοντέλων που χρησιμοποιούνται και χρησιμοποιείται για τον υπολογισμό των βέλτιστων παραμέτρων εξομάλυνσης. Ιδιαίτερη βαρύτητα δίνεται στα μεγάλα σφάλματα, λόγω του τετραγωνισμού των σφαλμάτων, σε αντίθεση με τα μικρά σφάλματα που δεν επηρεάζουν τόσο έντονα το σφάλμα.

- *Ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error):*

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

Επειδή το μέσο τετραγωνικό σφάλμα εκφράζεται στις μονάδες τις αρχικής χρονοσειράς υψωμένες στο τετράγωνο, υπολογίζεται η τετραγωνική ρίζα αυτού, ώστε να είναι το σφάλμα εκφρασμένο στις μονάδες της αρχικής χρονοσειράς. Οι ιδιότητες της ρίζας του μέσου τετραγωνικού σφάλματος (Root Mean Squared Error) παραμένουν ως και του μέσου τετραγωνικού σφάλματος.

- *Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error):*

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

Το μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error) λαμβάνει μη αρνητικές τιμές όπως φαίνεται και από τον παρακάτω τύπο. Είναι από τα δύο κύρια χρησιμοποιούμενα σφάλματα για την εκτίμηση της ακρίβειας των μοντέλων πρόβλεψης. Σε περιπτώσεις μεγάλων θετικών τιμών των δεδομένων της χρονοσειράς προτιμάται, λόγω της απλότητάς του. Ακόμη, στην αξιολόγηση της ακρίβειας μιας μεθόδου πρόβλεψης εφαρμοζόμενης σε περισσότερες από μία χρονοσειρές και μάλιστα διαφορετικού επιπέδου τιμών, χρησιμοποιείται το σφάλμα αυτό. Παρουσιάζει ωστόσο αδυναμία εφαρμογής σε χρονοσειρές διακοπτόμενης ζήτησης, αφού καταλήγει σε απροσδιοριστία.

Στην διπλωματική αυτή χρησιμοποιήθηκε αυτός ο τύπος σφάλματος για τον προσδιορισμό των καλύτερων προβλέψεων, καθώς δίνει εύληπτα και σαφή αποτελέσματα. Αρχικά υπήρχε η σκέψη να χρησιμοποιηθεί ο δείκτης MAE όμως απορρίφθηκε καθώς κάποιος θα έπρεπε να ξέρει το μέσο όρο των χρονοσειρών ώστε να μπορεί να κρίνει αν ένα μοντέλο αποδίδει καλύτερα από κάποιο άλλο.

- *Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (symmetric Mean Absolute Percentage Error):*

$$sMAPE = \frac{1}{n} \sum_{i=1}^n \frac{2 \cdot |Y_i - F_i|}{|Y_i + F_i|} \cdot 100\%$$

Το συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (Symmetric Mean Absolute Percentage Error) χρησιμοποιείται μαζί με το μέσο απόλυτο ποσοστιαίο σφάλμα κατά κύριο λόγο στην ανάλυση της ακρίβειας προβλέψεων. Σε αντίθεση όμως

με το MAPE, το sMAPE δεν παρουσιάζει απροσδιοριστία κατά την εφαρμογή του σε χρονοσειρές διακοπτόμενης ζήτησης. Αποφεύγει τα μεγάλα σφάλματα σε περίπτωση έντονων διαφορών ανάμεσα στην απόλυτη ποσοστιαία τιμή σφάλματος όταν η πραγματική τιμή είναι μεγαλύτερη της πρόβλεψης και το ανάποδο. Το μειονέκτημα αυτού του δείκτη είναι ότι δεν αντιμετωπίζονται οι αισιόδοξες προβλέψεις με αντίστοιχο τρόπο όπως οι απαισιόδοξες και επομένως δεν είναι τόσο συμμετρικός όσο παρουσιάζεται.

- Μέσο απόλυτο κανονικοποιημένο σφάλμα (Mean Absolute scaled Error):

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

Το σφάλμα αυτό προτάθηκε αρχικά από Hyndman και Koehler(2006) και είναι ένα σχετικό μέτρο (relative measure) που χρησιμοποιεί το σφάλμα μίας μεθόδου, στην συγκεκριμένη περίπτωση τη μέθοδο Naive ως benchmark - ορόσημο για να το συγκρίνει με το σφάλμα της υπό εξέταση μεθόδου πρόβλεψης. Στα θετικά της μεθόδου είναι ότι πολύ σπάνια θα δώσει απροσδιοριστία (σε αντίθεση με το MAPE) και δίνει ίσα βάρη στα μικρά και στα μεγάλα σφάλματα(σε αντίθεση με τα MSE και RMSE).

3.5 Μέθοδοι Πρόβλεψης

Στην παρούσα διπλωματική υλοποιούνται επτά μέθοδοι πρόβλεψης:

1. Απλοϊκή μέθοδος (Naive)
2. Μέθοδοι Εκθετικής Εξομάλυνσης (Exponential Smoothing)
3. Αυτοπαλινδρομικές μέθοδοι Κινητού Μέσου Όρου (Arima)
4. Μέθοδος Θ (Theta)
5. Βελτιστοποιημένη μέθοδος Θ (Optimal Theta)
6. Αλγόριθμος Πολλαπλής Συνάθροισης (MAPA)
7. Συνδυασμός τριών βέλτιστων μεθόδων (Combined method)

Ακολούθως θα παρουσιαστούν αναλυτικότερα μία προς μία οι προαναφερόμενοι μέθοδοι, ο τρόπος εφαρμογής τους όπως επίσης και η σημασία τους στην επιστήμη των προβλέψεων, καθώς έχουν χρησιμοποιηθεί στα πλαίσια της παρούσας διπλωματικής εργασίας. Παράλληλα θα παρουσιαστούν περιληπτικά και οι μέθοδοι νευρωνικών δικτύων, απλής παλινδρόμησης και κινητών μέσων όρων για λόγους πληρότητας του

κεφαλαίου αλλά και γιατί αποτελούν βασικές μέθοδοι πρόβλεψης σε πολλές ερευνητικές εργασίες.

3.5.1 Απλοϊκή μέθοδος (Naive)

Η μέθοδος Naive αποτελεί την πιο απλή μέθοδο πρόβλεψης. Η συγκεκριμένη μέθοδος δίνει ως πρόβλεψη για την επόμενη χρονική περίοδο την ίδια τιμή με την παρατήρηση που είχε σημειωθεί την προηγούμενη ακριβώς χρονική περίοδο. Έχει καλή απόδοση για προβλέψεις μίας περιόδου μπροστά σε αποεποχικοποιημένες χρονοσειρές καθώς η αναμενόμενη τιμή της πρόβλεψης δεν διαφέρει σημαντικά από την τελευταία παρατήρηση που είναι διαθέσιμη. Συνήθως όμως, δεν δίνει ακριβείς προβλέψεις με αποτέλεσμα να μην χρησιμοποιείται τόσο ως μέθοδος πρόβλεψης όσο ως σημείο αναφοράς (benchmark) για άλλες, πιο πολύπλοκες μεθόδους. Η μαθηματική σχέση που περιγράφει αυτή τη μέθοδο πρόβλεψης είναι:

$$F_t = Y_{t-1}$$

3.5.2 Μέθοδοι Κινητού Μέσου Όρου

Οι μέθοδοι μέσων όρων εκτός από την χρησιμότητά τους για την εξομάλυνση των ιστορικών δεδομένων και κατά συνέπεια την ομαλοποίηση των χρονοσειρών μπορούν να χρησιμοποιηθούν και ως μέθοδοι πρόβλεψης. Κάποιοι από τους μέσους όρους που χρησιμοποιούμε για την παραγωγή προβλέψεων είναι ο απλός μέσος όρος και κινητός μέσος όρος.

3.5.2.1 Απλός Κινητός Μέσος Όρος

Η μέθοδος του απλού μέσου όρου στηρίζεται στην εύρεση του μέσου όρου όλων των παρατηρήσεων και στη χρήση αυτής της τιμής για πρόβλεψη. Συνεπώς η πρόβλεψη δίνεται βάσει της παρακάτω σχέσης:

$$F_{t+1} = \frac{1}{t} \sum_{i=1}^t Y_i$$

Η χρήση αυτής της μεθόδου ενδείκνυται για περιπτώσεις που οι παρατηρήσεις δεν παρουσιάζουν τάση ή αξιοπρόσεκτη εποχικότητα.

3.5.2.2 Κινητός Μέσος Όρος

Ένας τρόπος να διαχειριστεί η επιρροή των παρελθουσών παρατηρήσεων στην πρόβλεψη όταν έχει επιλεγθεί ως μέθοδος πρόβλεψης η μέθοδος των μέσων όρων είναι να καθοριστεί το μήκος του μέσου όρου των παρατηρήσεων που θα ληφθούν υπ' όψη στην εξαγωγή της πρόβλεψης. Έτσι το μοντέλο καθώς μια νέα παρατήρηση γίνεται διαθέσιμη θα ανανεώνεται με αποτέλεσμα να γίνεται πιο ακριβές αφού λαμβάνει υπόψη του δεδομένα κοντά στο παρόν. Η επιλογή του μήκους του μέσου όρου ο οποίος θα υπολογιστεί και η τιμή του θα οριστεί ως πρόβλεψη μετατρέπει τον απλό μέσο όρο σε κινητό μέσο όρο. Ο όρος κινητός μέσος όρος χρησιμοποιείται για να περιγράψει τη διαδικασία καθώς όταν μία νέα παρατήρηση γίνεται διαθέσιμη, τότε υπολογίζεται ο νέος μέσος όρος των τελευταίων παρατηρήσεων του συγκεκριμένου μήκους που έχει επιλεγθεί. Αυτός ο νέος μέσος όρος θα είναι η τιμή της πρόβλεψης που παράγεται από αυτήν την μέθοδο για την επόμενη χρονική περίοδο. Είναι σημαντικό να τονιστεί ότι το πλήθος των παρατηρήσεων που χρησιμοποιούνται για την εξαγωγή του μέσου όρου παραμένει σταθερό καθ' όλη τη διαδικασία πρόβλεψης και περιλαμβάνει πάντα τις πιο πρόσφατες παρατηρήσεις. Η σχέση που χρησιμοποιείται για την εφαρμογή της μεθόδου του κινητού μέσου όρου, οποίος συμβολίζεται: ΚΜΟ(k) είναι:

$$F_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^t Y_i$$

Μία ενδεχόμενη σύγκριση του κινητού μέσου όρου και του απλού μέσου όρου θα παρουσίαζε ενδιαφέρον για να γίνει αντιληπτή η διαφορά στην χρήση τους. Το πλεονέκτημα του κινητού μέσου όρου είναι η σημασία που δίνεται πάντα στις τελευταίες ίσου πλήθους παρατηρήσεις ενώ τα μειονεκτήματα που παρουσιάζει είναι ότι απαιτεί περισσότερο χώρο αποθήκευσης δεδομένων, διότι πρέπει να αποθηκευτούν όλες οι παρατηρήσεις από τις οποίες θα εξάγεται κάθε φορά ο μέσος όρος και όχι απλά η τιμή του μέσου όρου. Το μειονέκτημα των μεθόδων μέσων όρων είναι ότι καμία από τις δύο μεθόδους δεν μπορεί να διαχειριστεί με επιτυχία κύρια χαρακτηριστικά των χρονοσειρών όπως είναι η τάση και η εποχικότητα κατά την εφαρμογή τους για την παραγωγή προβλέψεων. Για αυτό το λόγο τα μοντέλα μέσων όρων χρησιμοποιούνται κυρίως για την εξάλειψη της εποχικότητας και της τυχαιότητας από τις χρονοσειρές, ώστε να προκύψει μια εκτίμηση της γραμμής τάσης-κύκλου. Έτσι λοιπόν τα μοντέλα αυτά θα μπορούσαμε να πούμε ότι είναι κατά κύριο λόγο, εργαλεία αποσύνθεσης και όχι εργαλεία πρόβλεψης.

3.5.3 Απλή Γραμμική Παλινδρόμηση

Η μέθοδος της απλής γραμμικής παλινδρόμησης βασίζεται στην υπόθεση ύπαρξης σχέσης ανάμεσα στη μεταβλητή πρόβλεψης (εξαρτημένη μεταβλητή) και σε μια άλλη μεταβλητή (ανεξάρτητη μεταβλητή). Εκτός από την υπόθεση ότι υπάρχει μια τέτοια σχέση, υποθέτουμε ότι η σχέση αυτή είναι και γραμμική. Σκοπός λοιπόν της απλής γραμμικής παλινδρόμησης είναι η έκφραση της σχέσης της μεταβλητής Y από μία ανεξάρτητη μεταβλητή X με την εξίσωση μιας ευθείας γραμμής:

$$Y = \alpha + bX + e$$

Όπου α είναι το αρχικό σημείο (για $b=0$) και b είναι η κλίση της ευθείας, ενώ ο όρος e , δηλώνει το σφάλμα, δηλαδή την απόκλιση της παρατήρησης από της ευθεία που παριστάνεται από την παραπάνω σχέση.

Στόχος της απλής γραμμικής παλινδρόμησης είναι η εκτίμηση των παραμέτρων α και β έτσι ώστε η ευθεία:

$$Y = \alpha + bX$$

να αποτελεί τη "βέλτιστη", δηλαδή να προσαρμόζεται όσο το δυνατόν καλύτερα στα δεδομένα. Το σφάλμα προσαρμογής μπορεί να θεωρηθεί σαν την κατακόρυφη απόκλιση της παρατήρησης από την ευθεία προσαρμογής και δίδεται ως εξής:

$$e_t = Y_i - \hat{Y}_t$$

Όπου η τιμή αντιπροσωπεύει την εκτιμώμενη τιμή (από της ευθεία παλινδρόμησης) και η τιμή αντιστοιχεί στην πραγματική παρατήρηση. Σαν βέλτιστη ευθεία προσαρμογής, επιλέγεται αυτή για την οποία το άθροισμα των τετραγώνων των σφαλμάτων γίνεται ελάχιστο. Η μέθοδος είναι γνωστή σαν μέθοδο Ελαχίστων Τετραγώνων, επιλέγονται δηλαδή οι συντελεστές που ελαχιστοποιούν το άθροισμα των τετραγώνων των διαφορών των πραγματικών τιμών από τις προβλεπόμενες σε κάθε χρονική περίοδο, όπως φαίνεται και στην επόμενη σχέση:

$$(a, b) \mid \min \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right]$$

Οι μαθηματικές σχέσεις για τα b και α είναι:

$$b = \frac{\sum_{i=1}^n [(Y_i - \bar{Y}) \cdot (X_i - \bar{X})]}{\sum_{i=1}^n (X_i - \bar{X})^2}$$
$$\alpha = \bar{Y} - b\bar{X}$$

Όπου \bar{X} και \bar{Y} οι μέσες τιμές των διανυσμάτων X και Y και n ο αριθμός των παρατηρήσεων με βάση τις οποίες υπολογίζεται η ευθεία παλινδρόμησης.

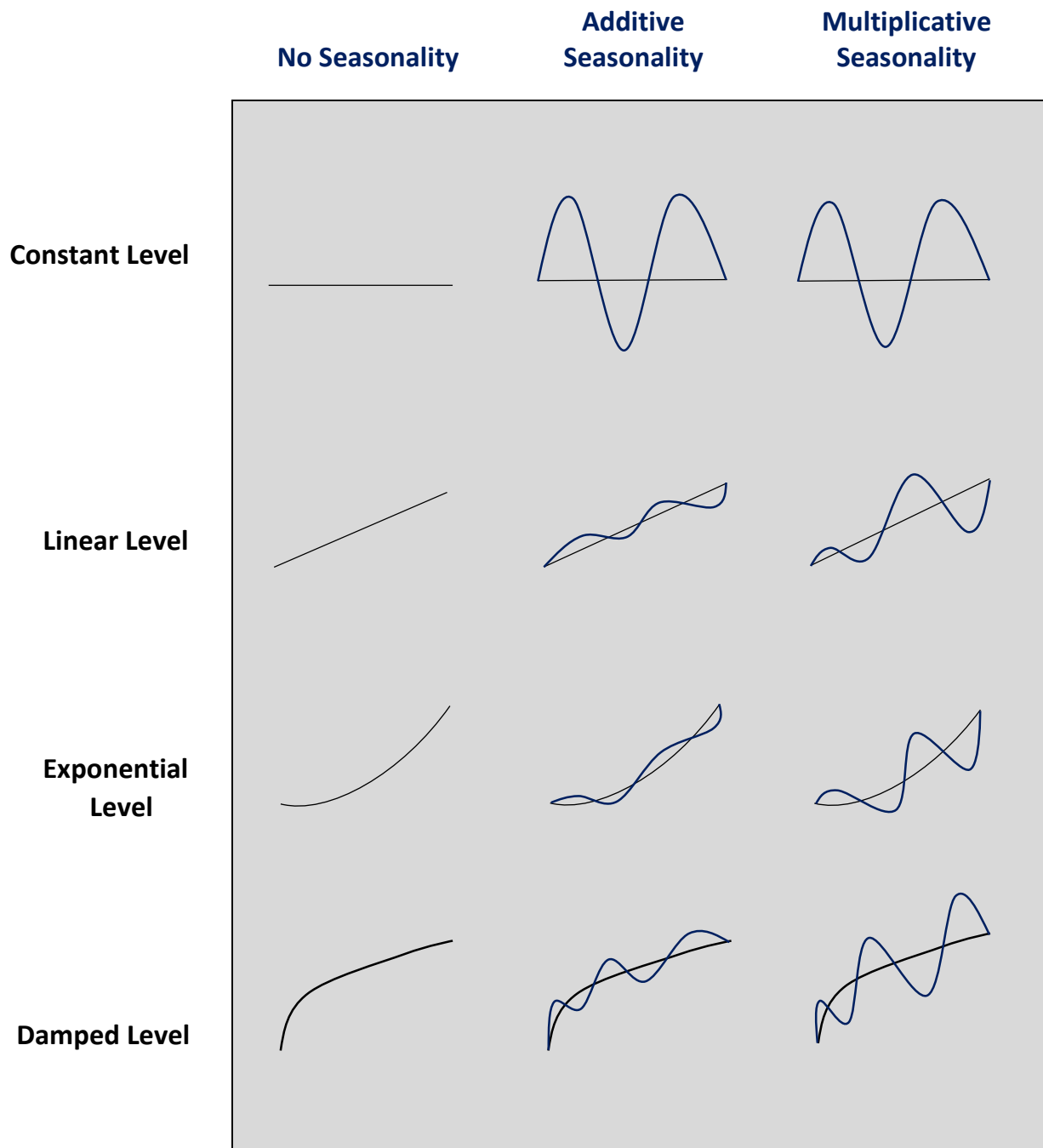
3.5.4 Μέθοδοι Εκθετικής Εξομάλυνσης

Μία επέκταση των μεθόδων μέσου όρου, είναι οι μέθοδοι πρόβλεψης με σταθμισμένο μέσο όρο. Δηλαδή όλες οι παρατηρήσεις να μην έχουν τη ίδια βαρύτητα για την εξαγωγή των προβλέψεων. Είναι συχνό φαινόμενο, οι πιο πρόσφατες παρατηρήσεις να είναι καλύτερος οδηγός για την πρόβλεψη της μελλοντικής τιμής. Γι' αυτόν τον λόγο, δημιουργήθηκε η ανάγκη για μοντέλα πρόβλεψης που θα χρησιμοποιούν τις παλαιότερες παρατηρήσεις με μειωμένη βαρύτητα συγκριτικά με τις πιο πρόσφατες. Σε αυτήν την παράγραφο λοιπόν θα περιγραφούν μέθοδοι οι οποίοι εφαρμόζουν εκθετική μείωση του συντελεστή βαρύτητας όσο πιο παλαιά είναι μια παρατήρηση. Αυτός είναι και ο λόγος που αυτές οι μέθοδοι καλούνται μέθοδοι εκθετικής εξομάλυνσης.

Γενικότερα, η εκθετική εξομάλυνση είναι μέθοδος πρόβλεψης, η οποία προεκτείνει στοιχεία του προτύπου των ιστορικών δεδομένων στο μέλλον. Το μοντέλο της αντίστοιχης πρόβλεψης εφαρμόζεται στη δοθείσα χρονοσειρά αφού πρώτα τα αντίστοιχα δεδομένα έχουν εξομαλυνθεί έτσι ώστε να απομονωθούν τα πραγματικά πρότυπα από τις καθαρά τυχαίες διακυμάνσεις.

Οι μέθοδοι εκθετικής εξομάλυνσης παρουσιάστηκαν για πρώτη φορά γύρω στο 1940 και η άνθισή τους ήρθε το 1960 μαζί με την άνθιση της επιστήμης της πληροφορικής. Οι μέθοδοι εκθετικής εξομάλυνσης είναι ιδιαίτερα δημοφιλείς στο πεδίο των προβλέψεων λόγω της απλότητας τους, των περιορισμένων απαιτήσεων τους για αποθήκευση δεδομένων και του μειωμένου υπολογιστικού φόρτου που απαιτούν. Επίσης παρά την απλότητα που τις διακρίνει, σύμφωνα με αποτελέσματα πρακτικών μελετών, παρουσιάζουν ικανοποιητικά ποσοστά ακρίβειας σε σχέση με πιο πολύπλοκες μεθόδους, διότι δεν επηρεάζονται τις ιδιομορφίες των προτύπων των δεδομένων ούτε από τυχαία εμφανιζόμενες ακραίες τιμές. Βέβαια αντίθετα με την ακρίβεια των προβλέψεων, οι εμπειρικές μελέτες έδειξαν ότι τα μοντέλα γραμμικής και εκθετικής τάσης εμφανίζουν μια υπεραισιοδοξία με αποτέλεσμα να οδηγούν σε ιδιαίτερα υψηλές τιμές προβλέψεων και κατ'επέκταση σε αυξημένες τιμές του στατιστικού δείκτη Mean Error.

Τα μοντέλα εκθετικής εξομάλυνσης χωρίζονται σε κατηγορίες ανάλογα με τη γενική μορφή της γραφικής παράστασης της χρονοσειράς. Σύμφωνα λοιπόν με την κατηγοριοποίηση αυτή προκύπτουν τέσσερα μοντέλα τάσης: τα μοντέλα σταθερού επιπέδου(constant level), γραμμικής τάσης (linear level), εκθετικής τάσης (exponential level) και φθίνουσας τάσης (damped level). Παρακάτω παρουσιάζονται γραφικά τα τέσσερα αυτά μοντέλα τάσης με τα τρία εποχιακά μοντέλα:



Εικόνα 3.5-Μοντέλα τάσης-εποχικότητας εκθετικής εξομάλυνσης

Αρχικά λοιπόν, το μοντέλα σταθερού επιπέδου υποθέτει την απουσία τάσης από τα δεδομένα. Εν γένει χρησιμοποιείται για τις προβλέψεις ενός βήματος, διότι η πρόβλεψη για οποιαδήποτε μελλοντική χρονική στιγμή γίνεται με την προέκταση μιας οριζόντιας ευθείας γραμμής. Αντιθέτως, το μοντέλο γραμμικής τάσης είναι πρακτικά πιο διαδεδομένο διότι η πρόβλεψη γίνεται με προέκταση μίας ευθείας γραμμής συμπεριλαμβανομένης την ύπαρξης της τάσης με αυτόν τον τρόπο. Χαρακτηριστικά παραδείγματα που ενδείκνυται η εφαρμογή μοντέλου εκθετικής τάσης είναι το ποσοστό αύξησης των πωλήσεων στην αρχή του κύκλου ζωής ενός προϊόντος. Λόγω του προβλήματος υπεραισιοδοξίας των μοντέλων γραμμικής και εκθετικής τάσης που

αναφέραμε, προέκυψε η ανάγκη εφαρμογής ενός άλλου μοντέλου, όπως του μοντέλου φθίνουσας τάσης, το οποίο αποτελεί την καλύτερη προσέγγιση σχετικά με τις μακροχρόνιες προβλέψεις, διότι μειώνεται σταδιακά το μέγεθος κατά το οποίο αυξάνονται οι τιμές της χρονοσειράς κάθε χρονική περίοδο. Πιο αναλυτικά ακολουθεί μελέτη και παρουσίαση κάθε μοντέλου για την καλύτερη κατανόηση του.

3.5.4.1 Μοντέλο Σταθερού Επιπέδου – Απλή Εκθετική Εξομάλυνση (Simple Exponential Smoothing)

Το μοντέλο περιγράφεται από τις εξής εξισώσεις:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + a \cdot e_t$$

$$F_{t+1} = S_t$$

Όπου e_t είναι το σφάλμα της πρόβλεψης το οποίο προκύπτει από τη διαφορά της πραγματικής τιμής της χρονοσειράς και της πρόβλεψης για την ίδια χρονική περίοδο t . Ο δείκτης t λοιπόν, αντιπροσωπεύει την χρονική περίοδο. Το S_t είναι το επίπεδο της χρονοσειράς στο τέλος της χρονικής περιόδου t και είναι το επίπεδο της προηγούμενης χρονικής περιόδου και ενός ποσοστού του σφάλματος. Ιδιαίτερο ενδιαφέρον παρουσιάζει η τιμή του ποσοστού αυτού διότι αντιπροσωπεύει την τιμή του συντελεστή εξομάλυνσης και λαμβάνει τιμές από 0 έως 1. Τέλος η τιμή $F_t(m)$ είναι η πρόβλεψη που πραγματοποιείται στο τέλος κάθε περιόδου t και αναφέρεται σε m περιόδους μπροστά. Χαρακτηριστικό της μεθόδου σταθερού επιπέδου είναι ότι η πρόβλεψη για κάθε χρονική περίοδο είναι ίση με το επίπεδο S_t .

Αν θέλαμε να περιγράψουμε με λίγες λέξεις την λειτουργία του μοντέλου σταθερού επιπέδου θα είχαμε την παρακάτω διαδικασία. Σε κάθε χρονική στιγμή, υπολογίζεται το σφάλμα με σκοπό να κρατήσει την τιμή της πρόβλεψης αρκετά κοντά στο επίπεδο της πραγματικής χρονοσειράς. Για την παραγωγή της πρόβλεψης κάθε χρονική στιγμή, πρέπει να έχει υπολογιστεί η τιμή του επιπέδου από την πραγματική χρονοσειρά για την προηγούμενη χρονική στιγμή.

Η ερώτηση που προκύπτει από την λειτουργία του μοντέλου είναι, τι θα γίνει με την πρόβλεψη για την πρώτη χρονική περίοδο για την οποία δεν υπάρχουν ιστορικά δεδομένα. Η απάντηση δίνεται από το αρχικό επίπεδο, σαν πρώτη πρόβλεψη στο συγκεκριμένο μοντέλο χρησιμοποιείται το αρχικό επίπεδο. Είναι λοιπόν δεδομένη η σημασία της σωστής επιλογής του αρχικού επιπέδου του μοντέλου πρόβλεψης για την παραγωγή προβλέψεων με ακρίβεια.

Αρχικό Επίπεδο

Συνήθεις μεθοδολογίες για την πρώτη τιμή του επιπέδου της απλής εκθετικής εξομάλυνσης σταθερού επιπέδου είναι:

- Ο μέσος όρος όλων των παρατηρήσεων
- Ο μέσος όρος των n πρώτων παρατηρήσεων
- Η πρώτη παρατήρηση
- Το σταθερό επίπεδο από το μοντέλο της απλής γραμμικής παλινδρόμησης

Η τιμή της αρχικοποίησης του επιπέδου για την εφαρμογή του μοντέλου, αφήνεται στο ερευνητή και εξαρτάται και από τα χαρακτηριστικά της εκάστοτε χρονοσειράς.

Συντελεστής Εξομάλυνσης

Όπως αναφέραμε, η παράμετρος α αποτελεί το συντελεστή εξομάλυνσης της μεθόδου. Οι τιμές που λαμβάνει ανήκουν στο διάστημα 0 έως 1 και εν γένει το κριτήριο που χρησιμοποιείται για τον προσδιορισμό της είναι η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE). Το κριτήριο αυτό είναι το επικρατέστερο στην βιβλιογραφία. Βέβαια η τιμή της παραμέτρου δεν είναι μοναδική καθώς μπορεί να επιλεχθούν άλλα κριτήρια για την επιλογή της αντίστοιχης τιμής όπως είναι η ελαχιστοποίησης κάποιου άλλου σφάλματος.

Η εύρεση της τιμής της σταθεράς εξομάλυνσης βρίσκεται με αλγοριθμικό τρόπο είτε με γραμμική αναζήτηση ψάχνοντας το ελάχιστο σφάλμα είτε με κάποια άλλη βελτιστοποίηση. Η γραμμική αναζήτηση εν γένει έχει το πρόβλημα χρόνου ιδιαίτερα όταν οι χρονοσειρές που αναφερόμαστε έχουν πολλά δεδομένα.

Ο βέλτιστος συντελεστής εξομάλυνσης καθορίζεται από δύο κύριους παράγοντες οι οποίοι αλληλοεξαρτώνται. Ο πρώτος είναι το ποσοστό θορύβου που υπάρχει στην χρονοσειρά. Όσο περισσότερος θόρυβος υπάρχει στα δεδομένα της χρονοσειράς τόσο μικρότερη πρέπει να είναι η τιμή του συντελεστή εξομάλυνσης για να αποφευχθεί η υπερβολική αντίδραση στον θόρυβο. Ο άλλος παράγοντας είναι η σταθερότητα του μέσου όρου της χρονοσειράς. Αν ο μέσος όρος της χρονοσειράς μεταβάλλεται, ο συντελεστής εξομάλυνσης θα πρέπει να είναι μεγάλος έτσι ώστε οι προβλέψεις να παρακολουθούν τις αντίστοιχες μεταβολές των δεδομένων. Αντίθετα αν ο μέσος όρος είναι σχετικά σταθερός, τότε η τιμή του συντελεστή εξομάλυνσης θα είναι μικρή.

Οι ακραίες τιμές του συντελεστή εξομάλυνσης έχουν καθοριστική σημασία για την παραγωγή της πρόβλεψης. Μηδενική τιμή του συντελεστή εξομάλυνσης σημαίνει ότι η πρόβλεψη θα μείνει ίδια για όλες τις χρονικές περιόδους και ίση με το αρχικό επίπεδο. Γι'αυτό εν γένει χρησιμοποιείται ένα κατώτατο όριο για την τιμή της παραμέτρου έτσι ώστε να αποφεύγεται η μηδενική τιμή. Αντίθετα, η μέγιστη τιμή του συντελεστή εξομάλυνσης δηλαδή να είναι ίσος με τη μονάδα, οδηγεί την τιμή της πρόβλεψης

να ταυτίζεται κάθε φορά με την τιμή της προηγούμενης χρονικής περιόδου, δηλαδή να είναι ίση με την τελευταία τιμή της χρονοσειράς και έτσι το μοντέλο μετατρέπεται στην απλοϊκή μέθοδο (Naive).

3.5.4.2 Μοντέλο Γραμμικής Τάσης (Holt Exponential Smoothing)

Το μοντέλο εξομάλυνσης γραμμικής τάσης επέκταση του μοντέλου απλής εκθετικής εξομάλυνσης η οποία μπορεί επιπρόσθετα να διαχειριστεί την συνιστώσα της τάσης που συχνά συναντάμε στις πραγματικές επιχειρησιακές χρονοσειρές. Παρουσιάζει ομοιότητες με το μοντέλο παλινδρόμησης, όμως σταδιακά αποδίδεται μεγαλύτερη βαρύτητα στα πιο πρόσφατα δεδομένα και το αρχικό σημείο και η κλίση επαναυπολογίζονται σε κάθε χρονική περίοδο. Πρακτικές μελέτες έχουν δείξει ότι οι εξομαλυμένες τιμές του αρχικού σημείου και της κλίσης είναι πολύ πιο ακριβείς από τις αντίστοιχες τιμές που υπολογίζονται αν στα δεδομένα εφαρμοστεί απλή γραμμική παλινδρόμηση. Το μοντέλο της εξομάλυνσης γραμμικής τάσης (Holt Exponential Smoothing, λόγω της εισαγωγής του στην επιστήμη των προβλέψεων το 1957 από τον Holt) μαθηματικά περιγράφεται από τις παρακάτω εξισώσεις:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + T_{t-1} + a \cdot e_t$$

$$T_t = T_{t-1} + b \cdot e_t$$

$$F_{t+m} = S_t + m \cdot T_t$$

Όπου e_t είναι το σφάλμα της πρόβλεψης το οποίο προκύπτει από τη διαφορά της πραγματικής τιμής της χρονοσειράς και της πρόβλεψης για την ίδια χρονική περίοδο t . Το S_t , είναι το επίπεδο της χρονοσειράς στο τέλος της χρονικής περιόδου t και είναι ίσο με το άθροισμα του επιπέδου της χρονικής περιόδου $t-1$, της τάσης την χρονική περίοδο $t-1$ και ενός ποσοστού το σφάλματος πρόβλεψης. Το ποσοστό αυτό καθορίζεται από τον συντελεστή a ο οποίος ορίζεται ως ο συντελεστής εξομάλυνσης του επιπέδου και το πεδίο τιμών του είναι από το 0 έως το 1. Η τάση T_t αντιπροσωπεύει την τάση που υπάρχει στην χρονοσειρά για την περίοδο t και είναι ίση με το άθροισμα της τάσης της χρονικής περιόδου $t-1$ και ενός ποσοστού του σφάλματος της πρόβλεψης. Το ποσοστό αυτό συμβολίζεται με τον συντελεστή b ο οποίος καλείται συντελεστής εξομάλυνσης της τάσης και το πεδίο τιμών του είναι επίσης από το 0 έως το 1. Η ποσότητα F_{t+m} που υπάρχει στην τελευταία από τις σχέσεις που περιγράφουν το μοντέλο είναι η πρόβλεψη που πραγματοποιείται στο τέλος της περιόδου t και αναφέρεται σε m περιόδους μπροστά. Η πρόβλεψη είναι ίση με το άθροισμα του επιπέδου S_t και της τάσης T_t πολλαπλασιασμένη με τον αριθμό m περιόδων του ορίζοντα πρόβλεψης. Από την ανάλυση των σχέσεων που προηγήθηκε,

γίνονται εμφανείς και οι διαφορές μεταξύ του μοντέλου γραμμικής τάσης και μοντέλου γραμμικής παλινδρόμησης. Το επίπεδο στο μοντέλο γραμμικής τάσης είναι το αρχικό σημείο μίας γραμμής τάσης η οποία αντιστοιχεί στη συγκεκριμένη μόνο χρονική περίοδο και μεταβάλλεται ανάλογα με τα δεδομένα κάθε χρονικής περιόδου.

Αρχικό Επίπεδο και Αρχική Τάση

Η αρχικοποίηση τόσο του επιπέδου όσο και της τάσης είναι εξαιρετικά σημαντική και στο μοντέλο γραμμικής εξομάλυνσης. Το αρχικό επίπεδο υπολογίζεται όπως και στην απλή εκθετική εξομάλυνση. Ως αρχική τάση συνήθως χρησιμοποιείται:

- η διαφορά της δεύτερης και πρώτης παρατήρησης ($Y_2 - Y_1$)
- η διαφορά την n -οστής παρατήρησης και πρώτης διαιρεμένη με $n-1$
- η σταθερά κλίσης της κλίσης από το μοντέλο της απλής γραμμικής παλινδρόμησης

Η τελική επιλογή γίνεται πάντα σύμφωνα με τα χαρακτηριστικά και το είδος της χρονοσειράς την οποία θέλουμε να μελετήσουμε.

Συντελεστές Εξομάλυνσης a και b

Σχετικά με τις τιμές των συντελεστών εξομάλυνσης, όπως και στο μοντέλο εκθετικής εξομάλυνσης σταθερού επιπέδου, ποικίλουν οι τιμές τους στο εύρος από το 0 έως του 1, ανάλογα με το κριτήριο επιλογής που θα χρησιμοποιηθεί, όπως η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος που συνηθίζεται ή η ελαχιστοποίηση του μέσου απόλυτου σφάλματος. Από την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος προκύπτουν τιμές και για τα δύο παραμέτρους της μεθόδου. Ουσιαστικά προκύπτει ο καλύτερος συνδυασμός των παραμέτρων σχετικά με την ακρίβεια. Είναι προφανές πως λόγω ύπαρξης δύο παραμέτρων εξομάλυνσης, η εύρεση της καλύτερης τιμής τους γίνεται πιο πολύπλοκη διαδικασία που απαιτεί και περισσότερο χρόνο. Συνεπώς η χρήση αποτελεσματικού αλγορίθμου προς εύρεση αυτών των παραμέτρων κρίνεται σημαντική.

3.5.4.3 Μοντέλο μη Γραμμικής Τάσης (Damped Exponential Smoothing)

Για τις περιπτώσεις στις οποίες το μοντέλο εξομάλυνσης για γραμμική τάση αποτυγχάνει να αποτυπώσει σωστά τη μελλοντική πορεία των χρονοσειρών (εξαιτίας της υπερεκτίμησης που το χαρακτηρίζει), αναπτύχθηκε το μοντέλο μη γραμμικής τάσης από τους Gardner και McKenzie το 1985, το οποίο προκαλεί βαθμιαία μείωση του μεγέθους αύξησης των τιμών της χρονοσειράς σε κάθε χρονική περίοδο.

Οι εξισώσεις που περιγράφουν το μοντέλο εκθετικής εξομάλυνσης μη γραμμικής τάσης είναι οι εξής:

$$S_t = \alpha \cdot X_t + (1 - \alpha) \cdot (S_{t-1} + \phi \cdot T_{t-1})$$

$$T_t = \gamma \cdot (S_t - S_{t-1}) + (1 - \gamma) \cdot \phi \cdot T_{t-1}$$

$$F_{t+m} = S_t + \sum_{i=1}^m \phi^i \cdot T_t$$

Οι περισσότεροι συμβολισμοί είναι γνωστοί ήδη από τις προηγούμενες μεθόδους εκθετικής εξομάλυνσης. Συγκεκριμένα, με e συμβολίζεται η απόκλιση της πραγματικής τιμής από την προβλεπόμενη, με X_t η πραγματική τιμή της χρονοσειράς την τρέχουσα περίοδο t , με S το επίπεδο της χρονοσειράς, με T η τάση της χρονοσειράς, με F η προβλεπόμενη τιμή, με α ο συντελεστής εξομάλυνσης του επιπέδου και με γ ο συντελεστής εξομάλυνσης της τάσης της χρονοσειράς. Και, όπως ήδη σημειώθηκε, με t συμβολίζεται η χρονική περίοδος, ενώ με m ο χρονικός ορίζοντας της πρόβλεψης.

Υπεισέρχεται όμως πλέον μία παράμετρος που ελέγχει το ρυθμό αύξησης των τιμών για την αφαίρεση της θετικής προκατάληψης (που παρουσιάζει το μοντέλο γραμμικής τάσης). Η παράμετρος αυτή ονομάζεται παράμετρος διόρθωσης της τάσης, ϕ , η οποία μπορεί να λάβει τιμές θετικές (ή και τη μηδενική τιμή) χωρίς άνω περιορισμό, σε αντίθεση με τους συντελεστές α και γ . Είναι, ωστόσο, σημαντική η κατηγοριοποίηση της μορφής του μοντέλου, ανάλογα με τα διαστήματα τιμών της παραμέτρου ϕ :

- Για $\phi = 0$ προκύπτει το μοντέλο απλής εκθετικής εξομάλυνσης, καθώς η τάση T δεν έχει κανένα ρόλο στον υπολογισμό της πρόβλεψης.
- Για $0 < \phi < 1$ προκύπτει το μοντέλο φθίνουσας τάσης (Damped Exponential Smoothing), το οποίο μειώνει σταδιακά την επίδραση της τάσης στο αποτέλεσμα της πρόβλεψης και ελέγχει με τον τρόπο αυτό το ρυθμό αύξησης των τιμών τους.
- Για $\phi = 1$ προκύπτει το μοντέλο γραμμικής τάσης όπως φαίνεται από την τελευταία εξίσωση.
- Για $\phi > 1$ προκύπτει το μοντέλο εκθετικής τάσης, το οποίο είναι υπεραισιόδοξο και οδηγεί σε υψηλές τιμές πρόβλεψης. Έτσι συνίσταται η χρήση του μόνο για πρόβλεψη ζήτησης ενός προϊόντος στα αρχικά στάδια εμφάνισής του στην αγορά.

Για την αποφυγή της παραγωγής προκατειλημμένων προβλέψεων λόγω λανθασμένης βελτιστοποίησης του συντελεστή ϕ και για επίτευξη καλύτερων αποτελεσμάτων, περιορίζεται το διάστημα τιμών της παραμέτρου ϕ στο διάστημα $[0,1]$. Η επιλογή των αρχικών τιμών του επιπέδου και της τάσης της χρονοσειράς συνίσταται να γίνεται με εφαρμογή της γραμμικής παλινδρόμησης και ο υπολογισμός του βέλτιστου συνδυασμού των συντελεστών α , γ και ϕ με τη διαδικασία ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος, όπως και στις προηγούμενες μεθόδους.

Έχει παρατηρηθεί ότι το μοντέλο μη γραμμικής τάσης παράγει πιο ακριβείς προβλέψεις σε σχέση με το μοντέλο γραμμικής τάσης, που είναι όμως πιο συχνά χρησιμοποιούμενο.

Μπορεί να πει κανείς ότι το μοντέλο μη γραμμικής τάσης δίνει ικανοποιητικά αποτελέσματα σε περιπτώσεις μεγάλου πλήθους χρονοσειρών, οπότε και απαιτείται πλήρως αυτοματοποιημένο σύστημα στην παραγωγή προβλέψεων και μάλιστα όσο μεγαλύτερος ο χρονικός ορίζοντας πρόβλεψης, τόσο καλύτερο είναι το μοντέλο αυτό (συγκριτικά με τις υπόλοιπες μεθόδους εκθετικής εξομάλυνσης). Τέλος, ικανοποιητικά αποτελέσματα δίνει και σε περιπτώσεις αδυναμίας επιλογής ορισμένου μοντέλου πρόβλεψης.

3.5.5 Μέθοδος θ (Theta)

Η μέθοδος πρόβλεψης Theta [Assimakopoulos and Nikolopoulos, 2000] αποτελεί μια μονοδιάστατη μέθοδο, που βασίζεται στη μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς μέσω της παραμέτρου θ , η οποία και εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των δεδομένων, όπως φαίνεται στην παρακάτω σχέση:

$$Y_t^\theta = \theta \cdot Y_t'' , Y_t'' = Y_t - 2 \cdot Y_{t-1} + Y_{t-2}$$

Η ιδιαίτερη προσέγγιση της μεθόδου έγκειται στο γεγονός ότι διαχωρίζει τα αποεποχικοποιημένα δεδομένα σε γραμμές ξεχωριστές, με έμφαση στη βραχυπρόθεσμη και στη μακροπρόθεσμη τάση τους, ανάλογα με την τιμή της παραμέτρου θ . Συγκεκριμένα, η μείωση της τιμής της παραμέτρου θ επιτυγχάνει αύξηση του βαθμού μείωσης των καμπυλοτήτων, μέχρι που για μηδενική τιμή της παραμέτρου η μέθοδος παράγει γραμμή που ισοδυναμεί με την ευθεία των ελαχίστων τετραγώνων (LRL). Αντίθετα, για τιμές του θ μεγαλύτερες της μονάδας ενισχύονται οι τοπικές καμπυλότητες της χρονοσειράς, αυξάνεται ο βαθμός διόγκωσής της και έτσι προκαλείται τόνωση και της βραχυπρόθεσμης συμπεριφοράς της χρονοσειράς. Η παράμετρος θ λαμβάνει και αρνητικές τιμές, όπως για παράδειγμα $\theta = -1$, η οποία αντιστοιχεί στη συμμετρική χρονοσειρά της αρχικής ως προς την ευθεία γραμμικής παλινδρόμησης.

Έπειτα από την αποσύνθεση της αρχικής χρονοσειράς σε δύο ή περισσότερες γραμμές Theta, προεκτείνεται η καθεμία ξεχωριστά στο μέλλον με οποιαδήποτε στατιστική μέθοδο πρόβλεψης έχει ήδη αναφερθεί και με συνδυασμό των διαφορετικών παραγόμενων προβλέψεων προκύπτει η τελική πρόβλεψη. Να αναφερθεί ότι κάθε γραμμή Theta διατηρεί τη μέση τιμή, αλλά και την κλίση της παλινδρόμησης της αρχικής χρονοσειράς, ενώ δέχονται μεταβολή οι καμπυλότητες και η διακύμανσή της.

Η απλοποιημένη εξίσωση που περιγράφει την κάθε γραμμή Theta είναι η εξής:

$$Y_t^\theta = \theta \cdot Y_t + (1 - \theta) \cdot LRL_t$$

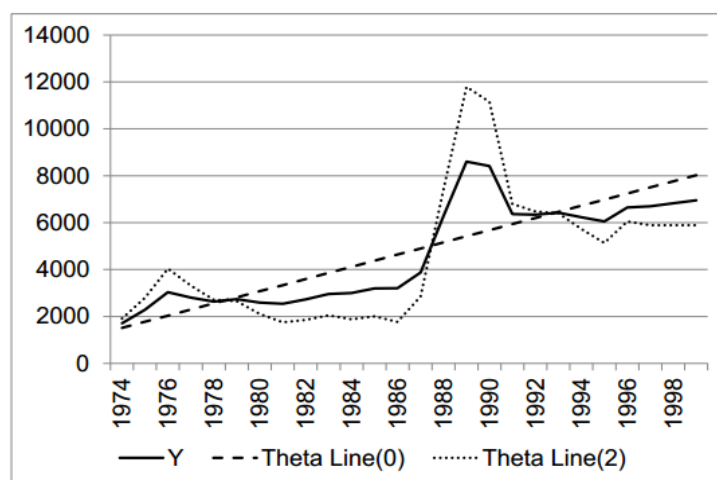
Ενώ το κλασικό μοντέλο Theta, που θα χρησιμοποιηθεί και στην παρούσα εργασία, συνδυάζει τις γραμμές Theta με $\theta=0$ και $\theta=2$ με ίσα βάρη, ώστε να παραχθεί η ζητούμενη πρόβλεψη. Αυτή είναι και η απλούστερη περίπτωση της συγκεκριμένης μεθόδου, στην οποία η γραμμή με $\theta=0$ παρέχει πληροφορία για τη μακροπρόθεσμη τάση της χρονοσειράς και η γραμμή με $\theta=2$ παρέχει πληροφορία για τη βραχυπρόθεσμη τάση της αντίστοιχα. Και η εξίσωση από την οποία προκύπτει η τελική πρόβλεψη της μεθόδου είναι η ακόλουθη:

$$Y_t = \frac{1}{2} \cdot (Y_t^{\theta=1+\alpha} + Y_t^{\theta=1-\alpha}) \quad (3.5.5.1)$$

Συνοπτικά τα βήματα που ακολουθούνται στην κλασική μέθοδο Θ είναι:

1. Έλεγχος Εποχικότητας
2. Αποεποχικοποίηση μέσω της κλασικής μεθόδου πολλαπλασιαστικής αποσύνθεσης
3. Αποσύνθεση της αποεποχικοποιημένης σειράς στις δύο γραμμές theta ($\theta=0$ και $\theta=2$)
4. Πρόβλεψη για κάθε γραμμή ξεχωριστά
5. Συνδυασμός των επιμέρους προβλέψεων
6. Εποχικοποίηση χρησιμοποιώντας του δείκτες εποχικότητας που υπολογίστηκαν σε προηγούμενο βήμα.

Η μέθοδος αυτή στόχο είχε την αύξηση του βαθμού αξιοποίησης της χρήσιμης πληροφορίας που βρίσκεται στα δεδομένα, πριν εφαρμοστεί σε αυτά κάποιο μοντέλο επέκτασής τους στο μέλλον. Μέσω των διαδικασιών της μεθόδου αυτής, γίνεται πιο αποδοτικός ο γραμμικός συνδυασμός των προβλέψεων των συνιστωσών. Καταλήγοντας, συμπεραίνεται ότι η μέθοδος Θ οδηγεί σε καλύτερα αποτελέσματα όσον αφορά μακροπρόθεσμες προβλέψεις της εξέλιξης των δεδομένων ή ενίσχυση της βραχυπρόθεσμης συμπεριφοράς της χρονοσειράς ανάλογα με την τιμή της παραμέτρου θ , όπως προαναφέρθηκε.



Εικόνα 3.6 -Παράδειγμα εφαρμογής μεθόδου

3.5.6 Βελτιστοποιημένη Μέθοδος Θ (Optimised Theta)

Μία σημαντική ιδιότητα της μεθόδου Θ είναι ότι, εξ'ορισμού, η μέθοδος είναι δυναμική. Μπορεί κάποιος να επιλέξει διαφορετικές γραμμές θ και να συνδυάσει τις παραγόμενες προβλέψεις με ίσα ή άνισα βάρη. Η κλασική μέθοδος Θ όπως παρουσιάστηκε πιο πάνω περιορίζει αυτή τη σημαντική ιδιότητα προκαθορίζοντας τις παραμέτρους αυτές με συγκεκριμένες τιμές. Επομένως, αν η επιλογή αυτών των τιμών προέχονταν μέσω πράξεων βελτιστοποίησης, τότε η μέθοδος θα μπορούσε να επικεντρωθεί σε πληροφορίες της χρονοσειράς που έχουν πραγματικά σημασία.

Υπάρχουν αρκετές εργασίες που προτείνουν μεθόδους βελτιστοποίησης της κλασικής μεθόδου Θ. Σε αυτήν την διπλωματική χρησιμοποιούμε το μοντέλο βελτιστοποίησης όπως περιγράφεται από την εργασία: *The Optimised Theta method (José Augusto Fioruci et al, 2015)*. Σε αυτήν την εργασία, επεκτείνεται η κλασική μέθοδος Θ, επιλέγοντας την γραμμή θ που περιγράφει καλύτερα την βραχυπρόθεσμη συμπεριφορά της χρονοσειράς, διατηρώντας ως έχει την μακροπρόθεσμη συνιστώσα της (για $\theta_1 = 0$). Για την επιλογή αυτής της γραμμής θ, ελαχιστοποιείται μία συνάρτηση σφαλμάτων που βασίζεται σε προβλέψεις σφαλμάτων πάνω σε ένα εύρος τιμών (validation sample). Ο συνδυασμός των προβλέψεων που προέρχονται από τις δύο αυτές γραμμές θ πραγματοποιείται μέσω κατάλληλων βαρών (όχι απαραίτητα [0.5,0.5] όπως στην κλασική μέθοδο Θ) που διασφαλίζουν την ανασύνθεση της αρχικής χρονοσειράς.

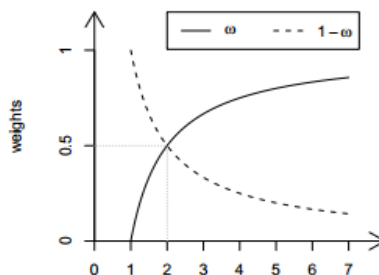
Η σχέση που περιγράφει την μέθοδο αυτή είναι παρόμοια με την (3.5.5.1) και είναι:

$$Y_t = \omega Y_t^{\theta_1} + (1 - \omega) Y_t^{\theta_2}, \quad \omega \in [0,1] \quad (3.5.6.1)$$

όπου ω είναι τα βάρη με τα οποία συνδυάζονται οι δύο γραμμές θ. Υπάρχει μαθηματική σχέση που συνδέει τις τιμές ω , θ_1 και θ_2 και είναι η εξής:

$$\omega := \omega(\theta_1, \theta_2) = \frac{\theta_2 - 1}{\theta_2 - \theta_1}, \quad \theta_1 \leq 1 \leq \theta_2$$

Παρατηρούμε ότι για $\theta_1 = 0$, $\theta_2 = 2$ τότε $\omega = 0.5$ και το μοντέλο συγκλίνει στη κλασική μέθοδο Θ. Για τους υπόλοιπους συνδυασμούς βαρών και τιμών θ, δίνεται η παρακάτω γραφική παράσταση:



Εικόνα 3.7-Συνδυασμοί βαρών και τιμών θ

Με βάση τους τύπους για το ω που δόθηκαν παραπάνω, ο τύπος που δίνει τις προβλέψεις είναι:

$$Y_t = \left(1 - \frac{1}{\theta}\right) (\hat{\alpha} + \hat{\beta}t) + \frac{1}{\theta} Y_t^\theta$$

και οι προβλέψεις για k περιόδους μπροστά είναι:

$$\hat{Y}_{t+k|t} = \left(1 - \frac{1}{\theta}\right) [\hat{\alpha} + \hat{\beta}(t+k)] + \frac{1}{\theta} \hat{Y}_{t+k|t}^\theta \quad (3.5.6.2)$$

όπου $\hat{Y}_{t+k|t}^\theta$ είναι η προέκταση της γραμμής $\theta_2 = \theta$ για k περιόδους μπροστά.

Υπολογισμός παραμέτρου θ_2

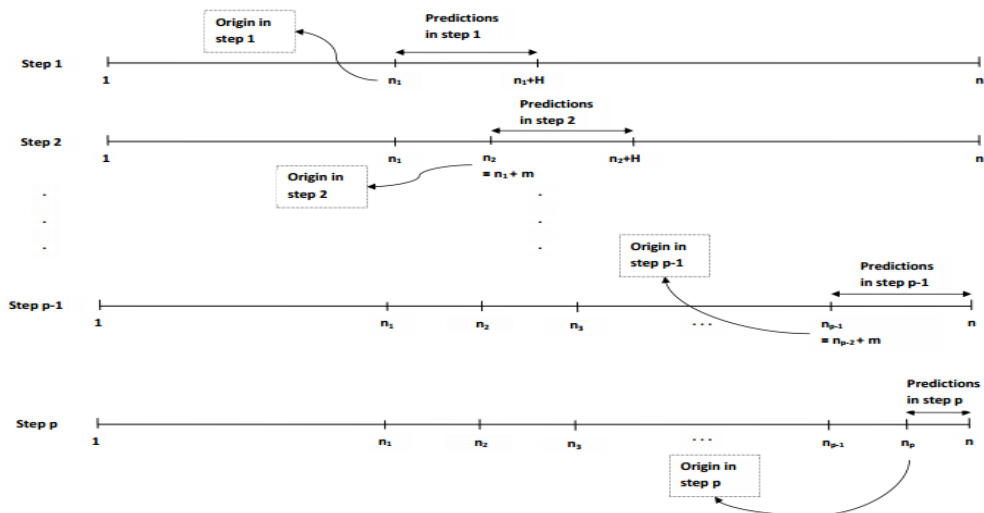
Για τον εντοπισμό των μακροπρόθεσμων συνιστωσών της χρονοσειράς, θεωρούμε ότι $\theta_1 = 0$ και ψάχνουμε να βρούμε την βέλτιστη επιλογή της $\theta_2 = \theta \geq 1$ όπου θα ελαχιστοποιείται η συνάρτηση αποκλίσεων:

$$E(\theta) = \sum_{i=1}^p \sum_{j=1}^{\min(H, n-n_i)} g(y_{n_i+j}, \hat{Y}_{n_i+j|n_i}) \quad (3.5.6.3)$$

Όπου $\hat{Y}_{n_i+j|n_i}$ δίνεται από την εξίσωση (3.5.6.2) και $g(\dots, \dots)$ είναι μία συμμετρική, θετική συνάρτηση σφάλματος όπως το τετραγωνικό σφάλμα $SE(a, b) = (a - b)^2$, το απόλυτο μέσο σφάλμα $AE(a, b) = |a - b|$ κ.α.

Αυτή η αλγοριθμική διαδικασία είναι γνωστή ως Γενικευμένη Εκτίμηση Σφάλματος Κυλιόμενης Αρχής (Generalised Rolling Origin Evaluation/GROE). Μετά το πέρας της, καθορίζεται η βέλτιστη επιλογή της παραμέτρου $\theta_2 = \theta$ και άρα μπορούν να συνδυαστούν οι γραμμές θ για $\theta_1 = 0$ και $\theta_2 = \theta$ με βάρη $1 - \frac{1}{\theta}$ και $\frac{1}{\theta}$ αντίστοιχα για την τελική πρόβλεψη του μοντέλου.

Παρακάτω θα δειχθεί γραφικά ένα γενικό παράδειγμα της μεθόδου GROE:



- n_i : η αρχή των προβλέψεων(τελευταία τιμή του μοντέλου προσαρμογής) στο βήμα $i = 1, 2, \dots, p$
- p : είναι ο αριθμός του επαναπροσδιορισμού της αρχής των προβλέψεων n_i με μέγιστη τιμή: $p^{max} = 1 + \frac{|n-n_1|}{m}$
- m : ο αριθμός της εμπρός μετατόπισης της αρχής(origin) σε κάθε βήμα
- H : ο αριθμός των προβλέψεων (test period) σε κάθε βήμα.
- Γενικά ισχύει ότι $1 < n_1 < n_2 < n_3 < \dots < n_p < n$ όπου $n_{i+1} = n_i + m$ για $i = 1, 2, \dots, p - 1$.

Στο πρώτο βήμα η θέση της αρχής(origin) είναι το σημείο n_1 . Μέχρι το σημείο $n_1 - 1$ γίνεται προσαρμογή του μοντέλου και γίνεται εκτίμηση σφάλματος με βάση την συνάρτηση (3.5.6.3) για τις προβλέψεις στην περιοχή $(n_1, n_1 + H)$.

Στο δεύτερο βήμα η θέση της αρχής επαναπροσαρμόζεται και είναι στο σημείο $n_2 = n_1 + m$. Μέχρι το σημείο $n_2 - 1$ γίνεται προσαρμογή του μοντέλου και γίνεται εκτίμηση σφάλματος με βάση την συνάρτηση (3.5.6.3) για τις προβλέψεις στην περιοχή $(n_2, n_2 + H)$.

Τα βήματα αυτά επαναλαμβάνονται μέχρι το σημείο όπου $n_p + m \geq n$. Μέχρι τότε η βέλτιστη τιμή για την τιμή του θ_2 έχει καθοριστεί και άρα και η συνολική πρόβλεψη με βάση την εξίσωση (3.5.6.1). Σε περίπτωση που τα δεδομένα έχουν εμφανή εποχιακή συνιστώσα, τα παραπάνω βήματα γίνονται για τα αποεποχικοποιημένα δεδομένα, τα οποία μετά την τελική πρόβλεψη με τον συνδυασμό των γραμμών θ_1 και θ_2 επανεποχικοποιούνται.

3.5.7 Αυτό-παλινδρομικές Μέθοδοι Κινητού Μέσου Όρου (Auto-Regressive Integrated Moving Average / ARIMA)

Τα μοντέλα ARIMA (ή αλλιώς μοντέλα Box και Jenkins) παρέχουν μια άλλη προσέγγιση για την πρόβλεψη χρονοσειρών. Η εκθετική εξομάλυνση και τα μοντέλα ARIMA είναι οι δύο πιο ευρέως χρησιμοποιούμενες προσεγγίσεις για την πρόβλεψη χρονοσειρών, και παρέχουν συμπληρωματικές προσεγγίσεις στο πρόβλημα. Ενώ τα μοντέλα της εκθετικής εξομάλυνσης βασίζονται σε μια περιγραφή των τάσεων και της εποχικότητας στα δεδομένα, τα μοντέλα ARIMA έχουν ως στόχο να περιγράψουν τις αυτοσυσχετίσεις στα δεδομένα.

Προτού περιγράψουμε τα μοντέλα ARIMA θα εξηγήσουμε τις έννοιες της στασιμότητας (stationarity) και της τεχνικής της μεθόδου διαφορών (differencing).

Λέμε ότι σε μία χρονοσειρά παρατηρείται *στασιμότητα* όταν τα χαρακτηριστικά της δεν εξαρτώνται από το χρόνο κατά τον οποίο η χρονοσειρά παρατηρείται. Για παράδειγμα μία χρονοσειρά με τάση και εποχικότητα δεν είναι στάσιμη-οι τιμές της χρονοσειράς επηρεάζονται από τυχόν στοιχεία τάσης και εποχικότητας σε διάφορες χρονικές στιγμές. Από την άλλη, μία χρονοσειρά με μεγάλη τυχαιότητα(θόρυβος) είναι στάσιμη-δεν έχει σημασία πότε την παρατηρούμε καθώς θα μοιάζει σε μεγάλο βαθμό σε οποιαδήποτε χρονική περίοδο.

Η τεχνική της *μεθόδου διαφορών* χρησιμοποιείται για να εξαλειφτούν από τις χρονοσειρές τα στοιχεία της τάσης και της εποχικότητας για τον καθορισμό της χρονοσειράς ως στάσιμης. Πρακτικά η τεχνική αυτή υλοποιείται υπολογίζοντας τις διαφορές μεταξύ δύο διαδοχικών τιμών μίας χρονοσειράς:

$$y'_t = y_t - y_{t-1} \quad (\text{διαφορά } 1^{\text{ης}} \text{ τάξης})$$

Πολλές φορές προκειμένου να γίνει μία χρονοσειρά στάσιμη απαιτείται περαιτέρω χρήση της μεθόδου διαφορών:

$$y''_t = y'_t - y'_{t-1} = y_t - 2y_{t-1} + y_{t-2} \quad (\text{διαφορά } 2^{\text{ης}} \text{ τάξης})$$

Μοντέλα Αυτοπαλινδρόμησης

Πρόκειται για ένα μοντέλο στο οποίο περιγράφεται μία μεταβλητή μέσω γραμμικού συνδυασμού παρελθόντων τιμών της μεταβλητής αυτής:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + e_t ,$$

όπου c : σταθερά , e_t : θόρυβος , p : εύρος παρελθόντων τιμών. Το μοντέλο αυτό συμβολίζεται ως AR(p).

Μοντέλα Κινητού Μέσου Όρου

Πρόκειται για ένα μοντέλο, στην ίδια φιλοσοφία του μοντέλου αυτοπαλινδρόμησης, στο οποίο περιγράφεται μία μεταβλητή μέσω γραμμικού συνδυασμού παρελθόντων τιμών των σφαλμάτων πρόβλεψης:

$$y_t = c + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_p e_{t-p} + e_t ,$$

όπου c : σταθερά , e_t : θόρυβος , q : εύρος παρελθόντων τιμών. Το μοντέλο αυτό συμβολίζεται ως MA(q).

Έχοντας εξηγήσει όλα αυτά, έχουμε τον παρακάτω τύπο που περιγράφει τα μοντέλα ARIMA:

$$y'_t = c + \varphi_1 y'_{t-1} + \dots + \varphi_p y'_{t-p} + \theta_1 e_{t-1} + \dots + \theta_p e_{t-p} + e_t , \quad (3.5.7.1)$$

όπου y_t' είναι η χρονοσειρά που έχει προκύψει μετά από εφαρμογή της μεθόδου διαφορών. Αποκαλούμε αυτό το μοντέλο ως ARIMA(p,d,q) , όπου p: εύρος της μεθόδου αυτοπαλινδρόμησης , d: τάξη της πρώτης διαφοράς που εμπλέκεται μέσω της μεθόδου των διαφορών , q: εύρος του μοντέλου κινητού μέσου όρου

Καθορισμός p,d,q

Ο καθορισμός των τιμών p,d,q είναι πολύ βασικός για τον προσδιορισμό ενός επιτυχημένου μοντέλου πρόβλεψης ARIMA. Γενικά δεν υπάρχει προκαθορισμένη μεθοδολογία για την επιλογή τους, και σε περιπτώσεις επεξεργασίας μεγάλου όγκου δεδομένων είναι προτιμότερο να χρησιμοποιούνται αυτοματοποιημένες διαδικασίες. Υπάρχουν όμως κάποιοι άτυποι κανόνες οι οποίοι μπορούν να βοηθήσουν στην επιλογή αυτών των παραμέτρων. Αυτοί οι κανόνες κάνουν χρήση των δεικτών acf (αυτοσυσχέτιση) και pacf (μερική αυτοσυσχέτιση):

1. Τα δεδομένα μπορούν να ακολουθούν το μοντέλο Arima(p,d,0) εάν τα διαγράμματα acf και pacf των διαφορών των τιμών της χρονοσειράς έχουν τα ακόλουθα χαρακτηριστικά:
 - Το διάγραμμα της acf μειώνεται ή είναι ημιτονικής μορφής.
 - Υπάρχει απότομη αύξηση στις τιμές του pacf στο σημείο p.
2. Ομοίως, τα δεδομένα μπορούν να ακολουθούν το μοντέλο Arima(0,d,q) εάν τα διαγράμματα acf και pacf των διαφορών των τιμών της χρονοσειράς έχουν τα ακόλουθα χαρακτηριστικά:
 - Το διάγραμμα της pacf μειώνεται ή είναι ημιτονικής μορφής.
 - Υπάρχει απότομη αύξηση στις τιμές του acf στο σημείο p.

Καθορισμός παραμέτρων παλινδρόμησης

Για τον καθορισμό των παραμέτρων $c, \varphi_1, \theta_1 \dots$ της σχέσης (3.5.7.1) έχουμε δύο μεθόδους:

1. Maximum likelihood estimation (MLE)

Μετά τον καθορισμό των τιμών των p,d,q η γλώσσα R –που χρησιμοποιήσαμε και στην διπλωματική μας- ψάχνει να βρει τις τιμές των παραμέτρων που μεγιστοποιούν την πιθανότητα απόκτησης των δεδομένων που έχουμε παρατηρήσει. Η φιλοσοφία αυτού του αλγορίθμου είναι παρόμοια με αυτή της μεθόδου ελαχίστων τετραγώνων.

2. Akaike's Information criteria (AIC)

Για μοντέλα ARIMA ο δείκτης αυτός παίρνει την μορφή

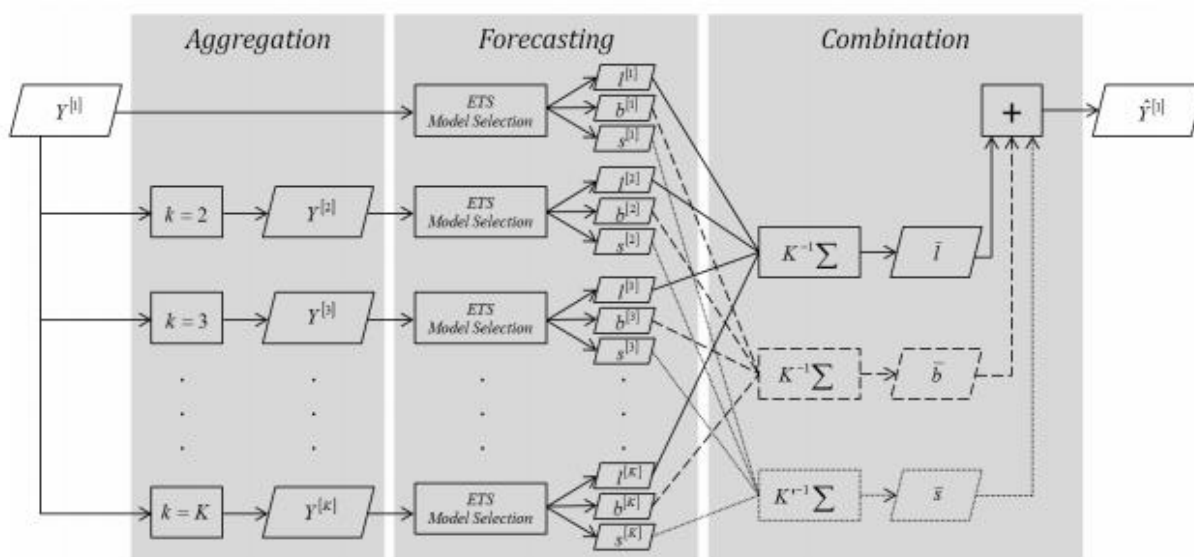
$$AIC_c = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2}$$

όπου $AIC = -2 \log(L) + 2(p + q + k + 1)$, L : πιθανότητα δεδομένων, $k=1$ αν $c \neq 0$ ή $k=0$ αν $c=0$

Στην διπλωματική αυτή, εξαιτίας του μεγάλου όγκου των δεδομένων προς επεξεργασία, ήταν αδύνατον να εκτιμηθεί χειροκίνητα κάθε παράμετρος για το μοντέλο πρόβλεψης ARIMA για κάθε χρονοσειρά ξεχωριστά. Επομένως χρησιμοποιήθηκαν κάποιες αυτοματοποιημένες συναρτήσεις της γλώσσας R οι οποίες, στην πλειοψηφία των περιπτώσεων, δίνουν αξιόπιστες προβλέψεις πολύ κοντά ή και σε ισοδύναμο επίπεδο ακρίβειας με τις βέλτιστες δυνατές.

3.5.8 Αλγόριθμος Πολλαπλής Συνάθροισης (Multi-Aggregation Prediction Algorithm / MAPA)

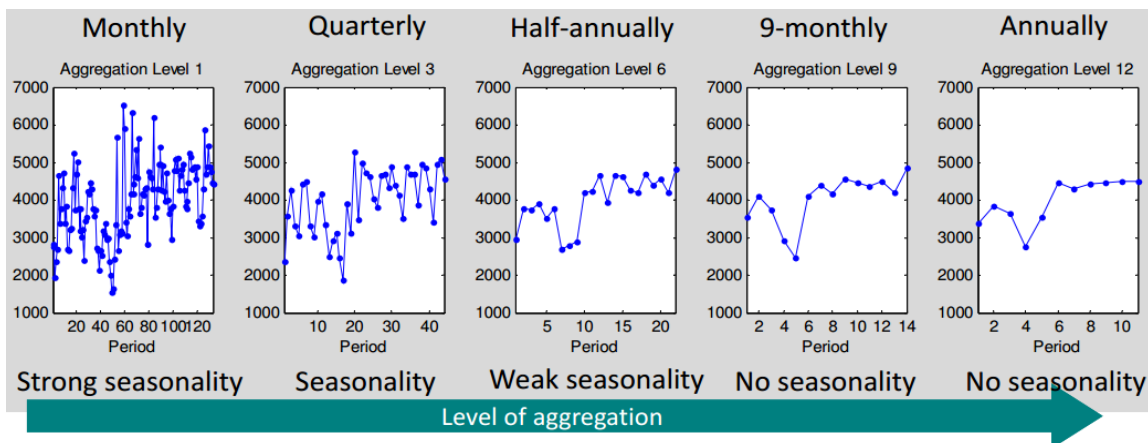
Ο αλγόριθμος πολλαπλής συνάθροισης [Nikolaos Kourentzes et al., 2015] είναι μία πρωτοποριακή μεθοδολογία πρόβλεψης όπου κάθε χρονοσειρά αναλύεται σε επιμέρους χρονοσειρές μέσω χρονικής συνάθροισης, σε κάθε τέτοια χρονοσειρά εφαρμόζεται ένα μοντέλο εκθετικής εξομάλυνσης (ETS) και γίνονται προβλέψεις για τα αντίστοιχα χαρακτηριστικά της χρονοσειράς. Έπειτα, τα χαρακτηριστικά αυτά από κάθε επίπεδο συνάθροισης συνδυάζονται για την δημιουργία της τελικής πρόβλεψης. Παρακάτω φαίνεται γραφικά η μεθοδολογία και στα επόμενα θα εξηγηθούν περισσότερο τα βήματα της:



Εικόνα 3.8-Στάδια μεθοδολογίας MAPA

Βήμα 1^ο: Χρονική Συνάθροιση

Είναι δυνατόν να απομονώσουμε συγκεκριμένα χαρακτηριστικά μίας χρονοσειράς μετατρέποντας τα αρχικά δεδομένα σε εναλλακτικές χρονικές συχνότητες. Έτσι, προτείνεται η συνάθροιση χρονοσειρών σε πολλαπλές χρονοσειρές χαμηλότερης συχνότητας - για παράδειγμα, η μετατροπή μηνιαίων χρονοσειρών σε διμηνιαίες, τετραμηνιαίες κ.λ.π. Σε κάθε επίπεδο συνάθροισης, διαφορετικά χαρακτηριστικά της χρονοσειράς αναδεικνύονται ή απαλούνονται. Σε χαμηλό επίπεδο συνάθροισης(υψηλή συχνότητα δεδομένων) χαρακτηριστικά όπως η εποχικότητα(seasonality) θα ξεχωρίζουν. Όσο το επίπεδο συνάθροισης αυξάνεται, τόσο η εποχικότητα εξομαλύνεται και θα δίνεται περισσότερη σημασία σε χαμηλής συχνότητας χαρακτηριστικά όπως η τάση(trend) και το αρχικό επίπεδο(level).



Εικόνα 3.9-Εξάλειψη εποχικότητας ανάλογα με το επίπεδο συνάθροισης k

Βήμα 2^ο: Προβλέψεις

Μετά την δημιουργία των συναθροισμένων χρονοσειρών $Y^{[k]}$, παράγουμε προβλέψεις για κάθε συναθροισμένη χρονοσειρά. Για το σκοπό αυτό χρησιμοποιείται η οικογένεια μεθόδων της εκθετικής εξομάλυνσης, που παρέχει ένα ολιστικό πλαίσιο μοντελοποίησης όλων των αρχέτυπων χαρακτηριστικών μίας χρονοσειράς όπως η εποχικότητα, η τάση και το επίπεδο. Με λίγα λόγια η χρησιμοποίηση των μοντέλων εκθετικής εξομάλυνσης όπως προτάθηκε από Hyndman et al. (2008) γίνεται γιατί η ανάκτηση των κύριων χαρακτηριστικών της χρονοσειράς είναι ευκολότερη και μεγάλης ακρίβειας. Επομένως, μετά την μετατροπή αυτών των χαρακτηριστικών σε αθροιστική μορφή, οι προβλέψεις σε κάθε επίπεδο αναλύονται στις επιμέρους συνιστώσες:

$$\hat{y}_{i+h}^{[k]} = l_{i+h}^{[k]} + b_{i+h}^{[k]} + s_{i+h}^{[k]},$$

όπου l, b, s είναι οι συνιστώσες επιπέδου, τάσης και εποχικότητας για κάθε πρόβλεψη αντίστοιχα.

Βήμα 3^ο: Συνδυασμός Προβλέψεων

Η παραγωγή προβλέψεων σε διαφορετικά επίπεδα και μεγέθη συχνότητων, οδηγεί στο επόμενο βήμα, στον συνδυασμό τους για την παραγωγή προβλέψεων. Η καινοτομία του αλγορίθμου αυτού έγκειται στο γεγονός ότι προτείνει τον συνδυασμό των επιμέρους χαρακτηριστικών των προβλέψεων για κάθε επίπεδο και για κάθε είδος χαρακτηριστικών και όχι των προβλέψεων αυτών καθ'αυτών. Δηλαδή, συνδυάζονται ξεχωριστά οι συνιστώσες της τάσης, του επιπέδου και της εποχικότητας και μετά συναθροίζονται για να εκφράσουν τις τελικές προβλέψεις. Αυτό γίνεται γιατί αν συνδυάζαμε απλά τις προβλέψεις για κάθε επίπεδο συνάθροισης τότε κάποια χαρακτηριστικά της χρονοσειράς θα επηρέάζονταν (π.χ. ο συνδυασμός χρονοσειράς με εποχικότητα και χρονοσειράς χωρίς εποχικότητα οδηγούσε σε συνδυασμένη χρονοσειρά με κατά το ήμισυ εξομαλυμένη εποχικότητα) και δεν θα έβγαζαν αξιόπιστες προβλέψεις.

3.5.9 Νευρωνικά Δίκτυα

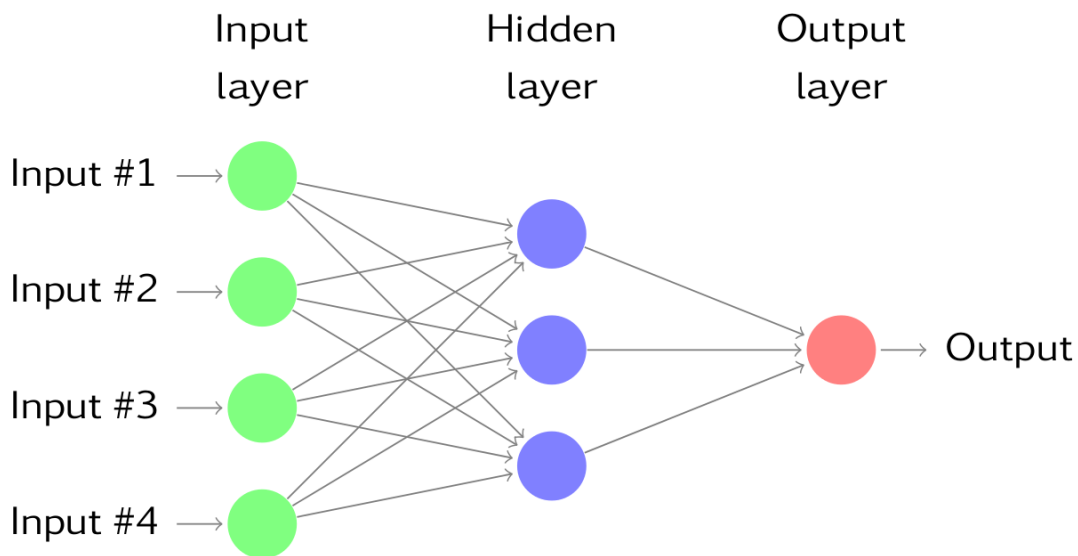
Το νευρωνικό δίκτυο είναι ένα δίκτυο από απλούς υπολογιστικούς κόμβους (νευρώνες, νευρώνια), διασυνδεδεμένους μεταξύ τους. Είναι εμπνευσμένο από το Κεντρικό Νευρικό Σύστημα (ΚΝΣ), το οποίο προσπαθεί να προσομοιώσει.

Οι νευρώνες είναι τα δομικά στοιχεία του δικτύου. Κάθε τέτοιος κόμβος δέχεται ένα σύνολο αριθμητικών εισόδων από διαφορετικές πηγές (είτε από άλλους νευρώνες, είτε από το περιβάλλον), επιτελεί έναν υπολογισμό με βάση αυτές τις εισόδους και παράγει μία έξοδο. Η εν λόγω έξοδος είτε κατευθύνεται στο περιβάλλον, είτε τροφοδοτείται ως είσοδος σε άλλους νευρώνες του δικτύου. Υπάρχουν τρεις τύποι νευρώνων: οι νευρώνες εισόδου, οι νευρώνες εξόδου και οι υπολογιστικοί νευρώνες ή κρυμμένοι νευρώνες. Οι νευρώνες εισόδου δεν επιτελούν κανέναν υπολογισμό, μεσολαβούν απλώς ανάμεσα στις περιβαλλοντικές εισόδους του δικτύου και στους υπολογιστικούς νευρώνες. Οι νευρώνες εξόδου διοχετεύουν στο περιβάλλον τις τελικές αριθμητικές εξόδους του δικτύου. Οι υπολογιστικοί νευρώνες πολλαπλασιάζουν κάθε είσοδό τους με το αντίστοιχο συναπτικό βάρος και υπολογίζουν το ολικό άθροισμα των γινομένων. Το άθροισμα αυτό τροφοδοτείται ως όρισμα στη συνάρτηση ενεργοποίησης, την οποία υλοποιεί εσωτερικά κάθε κόμβος. Η τιμή που λαμβάνει η συνάρτηση για το εν λόγω όρισμα είναι και η έξοδος του νευρώνα για τις τρέχουσες εισόδους και βάρη.

Εάν x_{ki} είναι η i -οστή είσοδος του k νευρώνα, w_{ki} : το i -οστό συναπτικό βάρος του k νευρώνα και $\varphi(\cdot)$ η συνάρτηση ενεργοποίησης του νευρωνικού δικτύου, τότε η έξοδος y_k του k νευρώνα δίνεται από την εξίσωση:

$$y_k = \varphi \left(\sum_{i=0}^N x_{ki} w_{ki} \right)$$

Στον k -οστό νευρώνα υπάρχει ένα συναπτικό βάρος w_{k0} με ιδιαίτερη σημασία, το οποίο καλείται πόλωση ή κατώφλι (bias, threshold). Η τιμή της εισόδου του είναι πάντα η μονάδα, $x_{k0} = 1$. Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός. Η ιδέα προέκυψε από τα βιολογικά νευρικά κύτταρα.



Εικόνα 3.10-Παράδειγμα Νευρωνικού δικτύου ενός κρυφού επιπέδου

Όπως είναι φανερό, οι αριθμοί οι οποίοι συναποτελούν το διάνυσμα εισόδου (κάθε στοιχείο του διανύσματος τροφοδοτείται κατά τη λειτουργία του δικτύου σε έναν νευρώνα εισόδου), αλλά και οι αριθμοί οι οποίοι συναποτελούν το διάνυσμα εξόδου (κάθε στοιχείο του οποίου εμφανίζεται, μετά το πέρας του ολικού υπολογισμού, σε έναν νευρώνα εξόδου), περιγράφουν χαρακτηριστικά του προς επίλυση προβλήματος. Συνήθως αυτό που μας ενδιαφέρει είναι το δίκτυο να απεικονίζει με ορθό τρόπο διανύσματα εισόδου σε κατάλληλα διανύσματα εξόδου, το πρόβλημα δηλαδή είναι η υλοποίηση μίας συνάρτησης πολλαπλών μεταβλητών, κατά κανόνα περίπλοκης και με άγνωστο ακριβή τύπο. Τέτοιες απεικονίσεις έχουν εφαρμογή σε ποικιλία τομέων της επιστήμης και της τεχνολογίας, αφού λειτουργούν ως αριθμητικά μοντέλα για πολλά διαφορετικά ζητήματα. Το ίδιο δίκτυο μπορεί να υλοποιήσει άπειρες διαφορετικές απεικονίσεις, μία για κάθε διαφορετική επιλογή συνόλου συναπτικών βαρών.

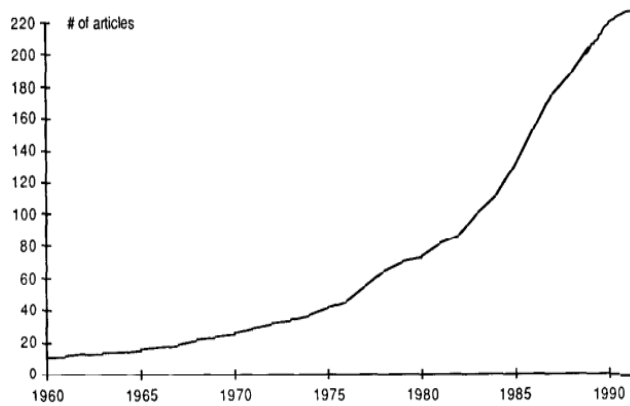
Το κύριο χαρακτηριστικό των νευρωνικών δικτύων είναι η εγγενής ικανότητα μάθησης. Ως μάθηση μπορεί να οριστεί η σταδιακή βελτίωση της ικανότητας του δικτύου να επιλύει κάποιο πρόβλημα (π.χ. η σταδιακή προσέγγιση μίας συνάρτησης). Η μάθηση

επιτυγχάνεται μέσω της εκπαίδευσης, μίας επαναληπτικής διαδικασίας σταδιακής προσαρμογής των παραμέτρων του δικτύου (συνήθως των βαρών και της πόλωσης του) σε τιμές κατάλληλες ώστε να επιλύεται με επαρκή επιτυχία το προς εξέταση πρόβλημα. Αφού ένα δίκτυο εκπαιδευτεί, οι παράμετροί του συνήθως «παγώνουν» στις κατάλληλες τιμές και από εκεί κι έπειτα είναι σε λειτουργική κατάσταση. Το ζητούμενο είναι το λειτουργικό δίκτυο να χαρακτηρίζεται από μία ικανότητα γενίκευσης: αυτό σημαίνει πως δίνει ορθές εξόδους για εισόδους καινοφανείς και διαφορετικές από αυτές με τις οποίες εκπαιδεύτηκε.

3.5.10 Συνδυασμένες Μέθοδοι Πρόβλεψης

Ως γνωστόν, για την επιλογή κατάλληλης μεθόδου πρόβλεψης θα πρέπει τα κύρια χαρακτηριστικά της κάθε χρονοσειράς να αναλυθούν. Σε κάποιες περιπτώσεις, η ανάλυση τους εκτός από επίπονη υπολογιστικά δεν φέρνει ικανοποιητικά αποτελέσματα. Ακόμα και σε περιπτώσεις όπου τα χαρακτηριστικά αυτά καθορίζονται πλήρως, είναι μερικές φορές δύσκολο να βρεθεί μία μέθοδο πρόβλεψης που να εκμεταλλεύεται πλήρως όλη αυτή την πληροφορία. Παρατηρήθηκε ότι ένας τρόπος αύξησης της ακρίβειας των προβλέψεων, είναι ο συνδυασμός διαφορετικών μεθόδων πρόβλεψης. Διαφορετικές μέθοδοι πρόβλεψης, εφαρμοζόμενες στις ίδιες χρονοσειρές, παράγουν διαφορετικά αποτελέσματα καθώς η κάθε πρόβλεψη παρέχει και διαφορετική πληροφορία. Ο συνδυασμός τους παράγει προβλέψεις, αξιοποιώντας περισσότερη πληροφορία στην ίδια τιμή, γεγονός που ενισχύει την ακρίβεια της πρόβλεψης καθώς προσεγγίζεται καλύτερα η πραγματικότητα. Αξιοπρόσεχτο όμως σε αυτό το σημείο είναι να αναφερθεί ο τρόπος με τον οποίο γίνεται να πραγματοποιηθεί ο συνδυασμός των μεθόδων πρόβλεψης. Δύο βασικές τεχνικές επικρατούν για την υλοποίηση του εν λόγω συνδυασμού: α) ο απλός μέσος όρων όλων των προβλέψεων των μεθόδων που θα επιλεγούν και β) ο υπολογισμός του μέσου όρου αλλά με χρήση συντελεστών βαρύτητας. Σε αυτή τη διπλωματική χρησιμοποιήθηκε σαν ξεχωριστή μέθοδος πρόβλεψης ο μέσος όρος των τριών βέλτιστων μεθόδων πρόβλεψης. Καταλήξαμε στο συνδυασμό τριών βέλτιστων μεθόδων και όχι σε κάποιο άλλο νούμερο καθώς ο συγκεκριμένος συνδυασμός έδινε τα καλύτερα αποτελέσματα στο διάστημα ελέγχου επίδοσης των μοντέλων (βέλτιστο in-sample MAPE error). Όπως θα φανεί σε επόμενο κεφάλαιο, η μέθοδος αυτή αν και απλοϊκή σε σύλληψη και ανάπτυξη, δίνει πολύ καλή ακρίβεια στην πλειοψηφία των περιπτώσεων.

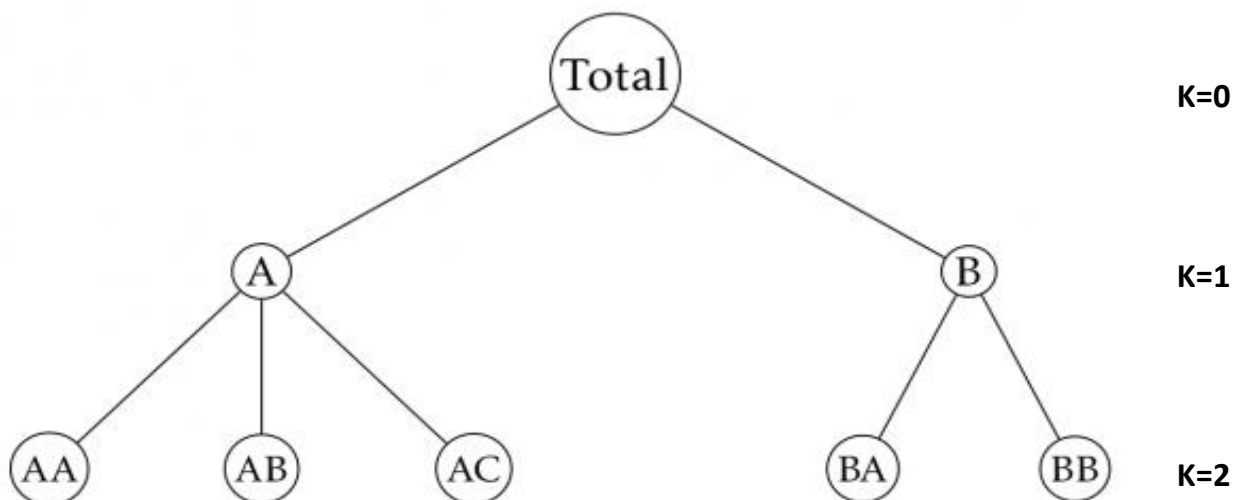
Πάντως γενικότερα η στρατηγική του συνδυασμού των διαφορετικών προβλέψεων, είναι εξαιρετικά εποικοδομητική καθώς από έρευνες και μελέτες όπως αυτές του Clemen (1989) και του Armstrong (2001) που έχουν πραγματοποιηθεί μέχρι σήμερα αποδεικνύεται μείωση του μεγέθους του σφάλματος έως και 6%.



Εικόνα 3.11-Αριθμός ερευνών σχετικά με συνδυαστικές μεθόδους προβλέψεων ανά χρονολογία

3.6 Ιεραρχική Δομή Χρονοσειρών

Οι χρονοσειρές μπορούν συχνά να διαχωρίζονται με φυσικό τρόπο σε μια ιεραρχική δομή, χρησιμοποιώντας χαρακτηριστικά όπως η γεωγραφική θέση, τον τύπο του προϊόντος, κ.λπ. Για παράδειγμα, ο συνολικός αριθμός των ποδηλάτων που πωλούνται από μια αποθήκη ποδηλάτων μπορεί να επιμεριστεί σε μια ιεραρχία των τύπων ποδηλάτων. Μια τέτοια αποθήκη θα πουλήσει ποδήλατα δρόμου, ποδήλατα βουνού, παιδικά ποδήλατα ή υβρίδια. Κάθε ένα από αυτά μπορεί να επιμεριστεί σε περισσότερες κατηγορίες. Παραδείγματος χάριν, τα ποδήλατα για παιδιά μπορούν να χωριστούν σε ποδήλατα ισορροπίας για τα παιδιά κάτω των 4 ετών, ποδήλατα ενιαίας ταχύτητας για παιδιά ηλικίας μεταξύ 4 και 6 και ποδήλατα για παιδιά άνω των 6 ετών. Τέτοιος επιμερισμός επιβάλλει μια ιεραρχική δομή. Αναφερόμαστε σε αυτά ως ιεραρχικές χρονοσειρές.



Εικόνα 3.12-Παράδειγμα ιεραρχικής χρονοσειράς

Στο παράδειγμα της παραπάνω εικόνας, οι χρονοσειρές βάσης AA, AB, AC συναθροίζονται στη χρονοσειρά A. Αντίστοιχα οι χρονοσειρές BA και BB συναθροίζονται στη χρονοσειρά B. Τέλος, οι χρονοσειρές A και B συναθροίζονται στην συνολική χρονοσειρά Total.

Έτσι γίνεται εμφανές ότι οι χρονοσειρές του κατώτερου επιπέδου, για κάθε χρονική στιγμή t , αθροίζονται στις χρονοσειρές του ανώτερου επιπέδου τους. Αυτό μπορεί να παρασταθεί από ένα πίνακα S διαστάσεων $n \times n_k$, όπου n : ο αριθμός των χρονοσειρών που υπάρχουν στην ιεραρχία και n_k : ο αριθμός των χρονοσειρών στο κατώτερο επίπεδο (π.χ. $K=2$). Όλα αυτά, για το παράδειγμα της ιεραρχίας της παραπάνω εικόνας φαίνονται στον παρακάτω τύπο:

$$\begin{bmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \end{bmatrix}$$

ή αλλιώς σε απλοποιημένη μορφή:

$$y_t = S y_{K,t}$$

όπου K ο αριθμός του κατώτερου επιπέδου ιεραρχίας.

Στην παρούσα διπλωματική, φτιάξαμε τα αντίστοιχα ιεραρχικά δέντρα για κάθε μεθοδολογία χρησιμοποιώντας το πακέτο `hts` της γλώσσας R και χρησιμοποιήσαμε την μέθοδο `bottom-up` συνάθροισης για τις βέλτιστες προβλέψεις βάσης που είχαμε βρει, ώστε να συγκρίνουμε την ακρίβεια των `out-of-sample` αποτελεσμάτων του ανώτερου επιπέδου $K=0$.

3.7 Επιλογή Κατάλληλης Μεθόδου Πρόβλεψης

Η εξαγωγή των προβλέψεων λόγω της μεγάλης τεχνολογικής εξέλιξης που υπάρχει, δεν μπορεί να χαρακτηριστεί ως μία δύσκολη διαδικασία. Όμως δε θα μπορούσε κανείς να υποστηρίξει το ίδιο όσον αφορά και την επιλογή της κατάλληλης μεθόδου. Όπως έχει

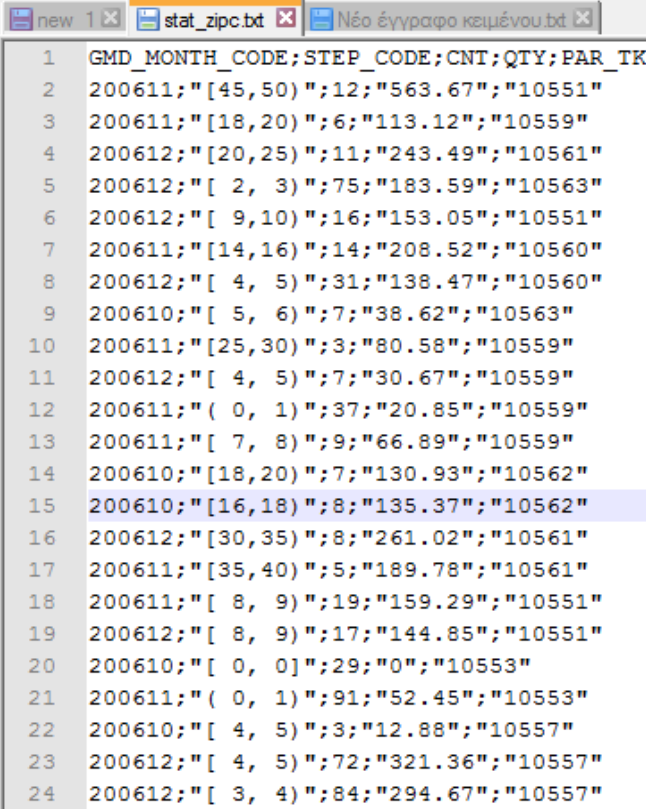
αναφερθεί προηγουμένως, οι μέθοδοι πρόβλεψης κατατάσσονται σε διάφορες κατηγορίες ανάλογα με τις εφαρμογές τους αλλά και τα κύρια χαρακτηριστικά τους, έτσι ώστε να γίνει η διαδικασία της επιλογής τους ανά περίπτωση πιο εύκολη διαδικασία. Κινούμενοι προς αυτήν την κατεύθυνση θα αναφερθούν κάποιοι βασικοί παράγοντες που αντικατοπτρίζουν τις δυνατότητες εφαρμογής των διαθέσιμων μεθόδων. Οι κυριότεροι λοιπόν παράγοντες είναι:

- *Χρονικός ορίζοντας.* Ανάλογα το χρονικό διάστημα στο μέλλον στο οποίο θα αναφέρεται η πρόβλεψη συχνά επιλέγεται και η αντίστοιχη μέθοδος που θα χρησιμοποιηθεί. Οι ποσοτικές μέθοδοι χρησιμοποιούνται περισσότερο για μακροπρόθεσμες προβλέψεις ενώ οι κριτικές μέθοδοι χρησιμοποιούνται περισσότερο για μεσοπρόθεσμες και βραχυπρόθεσμες προβλέψεις. Επίσης σημαντικό στοιχείο είναι και το πλήθος των περιόδων για το οποίο απαιτείται πρόβλεψη.
- *Κόστος.* Αναφερόμενοι σε μία μέθοδο πρόβλεψης, το κόστος της σχετίζεται άμεσα με τον όγκο των δεδομένων που αποτελούν τα ιστορικά στοιχεία και από την πολυπλοκότητα κατά την εφαρμογή της
- *Πρότυπο συμπεριφοράς των δεδομένων.* Δεν είναι δυνατή η εφαρμογή κανενός μοντέλου πρόβλεψης αν πρώτα δεν αναγνωριστεί ένα βασικό πρότυπο συμπεριφοράς των δεδομένων, το οποίο θα αποτελέσει βάση της τεχνικής πρόβλεψης που θα εφαρμοστεί. Τα τέσσερα βασικά πρότυπα που συμπεριφοράς που συχνά εμφανίζονται στις χρονοσειρές και τις περισσότερες φορές συνυπάρχουν είναι το σταθερό πρότυπο, το πρότυπο της τάσης, το εποχιακό και το κυκλικό πρότυπο.
- *Απλότητα και ευκολία στην εφαρμογή της.* Απλές και ευκολονόητες μέθοδοι εν γένει προτιμώνται καθώς είναι και πιο εύκολες στην εφαρμογή τους.
- *Αξιοπιστία.* Η αξιοπιστία σχετικά με τις προβλέψεις, συνδέεται με το επίπεδο λεπτομέρειας που απαιτείται στην αντίστοιχη περίπτωση. Υπάρχουν περιπτώσεις όπου ένα ποσοστό ακρίβεια της πρόβλεψης 10% είναι ικανοποιητικό ενώ άλλες που ακόμα και το μισό ποσοστό από το προαναφερόμενο μπορεί να αποδειχθεί καταστροφικό.

4 Μεθοδολογία Διπλωματικής

4.1 Δεδομένα

Η διαδικασία πρόβλεψης έχει ως πρώτο στάδιο τον καθορισμό του προβλήματος, δηλαδή την αναγνώριση της ανάγκης που υπάρχει και ποιος είναι ο αποδέκτης των αποτελεσμάτων της πρόβλεψης. Στην συγκεκριμένη διπλωματική, αποδέκτης είναι η ΕΥΔΑΠ και στόχος είναι να αναπτύξουμε και να συγκρίνουμε τρεις ευρύτερες μεθοδολογίες ως προς την αξιοπιστία που δίνουν σε πρόβλεψη διαστήματος 17 μηνών, όπου ξέρουμε τις πραγματικές τιμές. Το δεύτερο σημαντικό κομμάτι της διαδικασίας πρόβλεψης είναι η συλλογή και η ανάλυση των δεδομένων. Αυτό το στάδιο βοηθάει το χειριστή ώστε να λάβει καταρχάς υλικό για τη μελέτη του και ακόμη για να αποκτήσει μια σωστή και αντιπροσωπευτική εικόνα των στοιχείων με τα οποία ασχολείται. Έτσι θα μπορεί να έχει μια πιο καλή αντίληψη της κατάστασης των δεδομένων, αλλά θα μπορεί και να αναπτύξει πιο σωστή κριτική σκέψη για την κάθε περίπτωση.



```
new 1 x stat_zipc.txt x Νέο έγγραφο κειμένου.txt x
1 GMD_MONTH_CODE;STEP_CODE;CNT;QTY;PAR_TK
2 200611;"[45,50)";12;"563.67";"10551"
3 200611;"[18,20)";6;"113.12";"10559"
4 200612;"[20,25)";11;"243.49";"10561"
5 200612;"[ 2, 3)";75;"183.59";"10563"
6 200612;"[ 9,10)";16;"153.05";"10551"
7 200611;"[14,16)";14;"208.52";"10560"
8 200612;"[ 4, 5)";31;"138.47";"10560"
9 200610;"[ 5, 6)";7;"38.62";"10563"
10 200611;"[25,30)";3;"80.58";"10559"
11 200612;"[ 4, 5)";7;"30.67";"10559"
12 200611;"( 0, 1)";37;"20.85";"10559"
13 200611;"[ 7, 8)";9;"66.89";"10559"
14 200610;"[18,20)";7;"130.93";"10562"
15 200610;"[16,18)";8;"135.37";"10562"
16 200612;"[30,35)";8;"261.02";"10561"
17 200611;"[35,40)";5;"189.78";"10561"
18 200611;"[ 8, 9)";19;"159.29";"10551"
19 200612;"[ 8, 9)";17;"144.85";"10551"
20 200610;"[ 0, 0]";29;"0";"10553"
21 200611;"( 0, 1)";91;"52.45";"10553"
22 200610;"[ 4, 5)";3;"12.88";"10557"
23 200612;"[ 4, 5)";72;"321.36";"10557"
24 200612;"[ 3, 4)";84;"294.67";"10557"
```

Εικόνα 4.1-Κύρια δεδομένα διπλωματικής αποθηκευμένα σε αρχείο .txt

Στην συγκεκριμένη εργασία, το βασικότερο dataset που χρησιμοποιήθηκε είναι οι άνω του μισού εκατομμυρίου μηνιαίες παρατηρήσεις συνολικής κατανάλωσης στους τελικούς καταναλωτές (οικιακούς, βιομηχανικούς), για κάθε ταχυδρομικό κώδικα (TK) του νομού Αττικής. Ένα πολύ μικρό μέρος των δεδομένων φαίνεται στην εικόνα 4.1, όπου στη στήλη “GMD_MONTH_CODE” η ένδειξη π.χ. 200610 υποδηλώνει το έτος 2006 και τον μήνα Οκτώβριο (10^{ος}), η στήλη “QTY” είναι η ποσότητα ζήτησης σε κυβικά μέτρα τον αντίστοιχο χρόνο και μήνα που καταγράφεται στο αντίστοιχο πλήθος μετρητών από την στήλη “CNT” για τον ταχυδρομικό κώδικα που δηλώνεται από την στήλη “PAR_TK”. Οι υπόλοιπες στήλες δεν θεωρήθηκαν απαραίτητες και δεν χρησιμοποιήθηκαν για τον βασικό κορμό του πρακτικού μέρους της διπλωματικής. Για παράδειγμα, στην σημαδεμένη περίπτωση της εικόνας 4.1 η μέτρηση αφορά τον Οκτώβριο του 2010 για τον TK 10562, και προέρχεται από 8 μετρητές της συγκεκριμένης περιοχής με συνολική μέτρηση 103.5 m³.

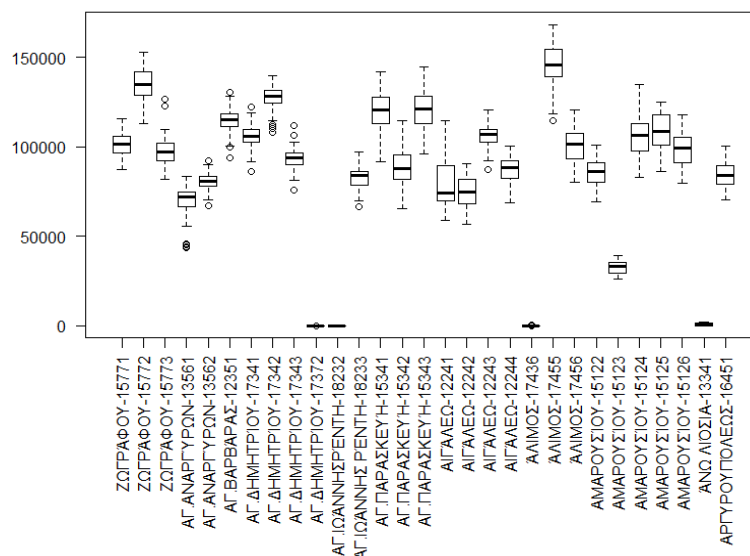
Από αυτά τα δεδομένα, μέσω κώδικα που αναπτύχθηκε στο προγραμματιστικό περιβάλλον RStudio και θα παρουσιαστεί σε παράρτημα στο τέλος της διπλωματικής, φτιάχνουμε 274 χρονοσειρές μηνιαίας ζήτησης νερού που αντιστοιχούν για κάθε διαθέσιμο TK για την χρονική περίοδο από 1/2007 έως 12/2013. Να σημειωθεί ότι στα αρχικά μας δεδομένα υπήρχαν διαθέσιμες μηνιαίες τιμές ζήτησης και για τα έτη 2005 και 2006 αλλά δεν χρησιμοποιήθηκαν στην παρούσα διπλωματική γιατί σε πολλούς TK τα στοιχεία ήταν ελλιπή και επομένως αναξιόπιστα. Έτσι προτιμήθηκε ή έναρξη των παρατηρήσεων να οριστεί το έτος 2007 όπου υπήρχε πλήρης καταγραφή τιμών ζήτησης στην πλειοψηφία των ταχυδρομικών κωδικών.

Περιοχή	Min	Q1(25%)	Median (50%)	Q2(75%)	Max	Mean	Std
ΖΩΓΡΑΦΟΥ-15771	87188.95	96841.678	101581.236	105706.088	115816.06	101024.87	6237.886
ΖΩΓΡΑΦΟΥ-15772	112800.8	128829.388	134551.175	142009.8	152623.14	134583.277	8905.128
ΖΩΓΡΑΦΟΥ-15773	81697.2	92027.323	97079.51	102231.673	126390.94	97430.905	7164.676
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13561	43476.18	66319.53	71766.65	74878.248	83389.93	69735.719	8761.051
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13562	67333.85	77836.488	80616.092	83513.695	92232.09	80371.414	4934.509
ΑΓ.ΒΑΡΒΑΡΑΣ-12351	94082.98	111191.873	115256.73	118122.215	130484.24	114573.471	6555.859
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17341	86211.64	102336.285	106069	109379.703	122207.72	105841.971	6435.051
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17342	108185	124383.145	128121.18	131530.36	139807.24	127352.235	6353.783
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17343	75958.93	90281.02	93909.755	96555.448	111763.78	93491.203	5789.318
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17372	15.47	32.742	48.14	60.888	110.34	50.697	25.931
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18232	88.91	134.745	151.115	169.617	217.34	153.113	29.611
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18233	66321.75	78680.345	84120.36	86379.678	97254.07	82832.827	6322.38
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15341	91884.6	113139.695	120313.595	127537.553	141923.89	119834.244	10557.387

ΑΓ.ΠΑΡΑΣΚΕΥΗ-15342	65206.43	82053.838	87763.645	95421.555	114381.34	88325.223	10405.825
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15343	95753.09	112971.478	121161.02	128384.368	144342.01	120545.265	11583.63
ΑΙΓΆΛΕΩ-12241	58870.77	70063.653	74163.565	89641.638	114492.1	79531.827	13074.096
ΑΙΓΆΛΕΩ-12242	56441.24	68260.503	74845.905	81701.885	90554.5	74930.365	7992.709
ΑΙΓΆΛΕΩ-12243	87150.12	102690.348	106723.85	109866.203	120540.21	106336.536	5934.637
ΑΙΓΆΛΕΩ-12244	68461.25	82396.51	88564.145	91970.19	100554.88	87420.466	6364.231
ΆΛΙΜΟΣ-17436	41.03	64.73	84.995	111.37	209.52	94.798	40.912
ΆΛΙΜΟΣ-17455	114826.2	138995.443	145477.93	154214.018	168306.62	145489.489	12178.283
ΆΛΙΜΟΣ-17456	80099.41	93264.238	101426.475	107182.955	120355.5	100448.872	9564.11
ΑΜΑΡΟΥΣΙΟΥ-15122	69102.98	80420.278	86234.305	90497.468	100713.53	85319.169	7219.297
ΑΜΑΡΟΥΣΙΟΥ-15123	26175.54	29626.825	33141.285	35505.67	39240.22	32758.992	3554.22
ΑΜΑΡΟΥΣΙΟΥ-15124	82814.56	97514.705	106370.465	112606.933	134511.14	105714.709	11167.962
ΑΜΑΡΟΥΣΙΟΥ-15125	86172.96	100741.985	108408.38	117583.073	125187.6	108301.023	10222.277
ΑΜΑΡΟΥΣΙΟΥ-15126	79484.93	91204.835	99020.46	105233.603	117852.43	98189.277	9193.107
ΆΝΘ ΛΙΘΣΙΑ-13341	90.12	112.298	611.745	1583.75	2258.93	826.239	734.07
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16451	70307.33	79357.048	84230.43	89251.64	100377.13	84205.428	6549.9

Πίνακας 4.1-Χαρακτηριστικά ορισμένων χρονοσειρών από τα διαθέσιμα δεδομένα μας

Ο πίνακας 4.1 δείχνει ορισμένα χαρακτηριστικά από τις 29 πρώτες χρονοσειρές των συνολικών διαθέσιμων περιοχών που είχαμε στα δεδομένα μας. Η πρώτη στήλη δείχνει την ελάχιστη μηνιαία τιμή ζήτησης νερού (σε κυβικά μέτρα) από τις συνολικά 84 διαθέσιμες μηνιαίες τιμές για κάθε ΤΚ. Η δεύτερη υπολογίζει το 1^ο τεταρτημόριο, η τρίτη την διάμεσο, η τέταρτη στήλη το 3^ο τεταρτημόριο, την μέγιστη και μέση τιμή για την πέμπτη και έκτη στήλη, και τέλος η τυπική απόκλιση στην έβδομη στήλη. Για γραφική αναπαράσταση των επιμέρους χαρακτηριστικών μπορούμε να σχεδιάσουμε το διάγραμμα boxplot όπως φαίνεται παρακάτω:



Εικόνα 4.2-Boxplot διάγραμμα για τα δεδομένα του Πίνακα 4.1

Η άνω και κάτω ακμή κάθε ορθογώνιας μπάρας αντιστοιχεί στο πρώτο και τρίτο τεταρτημόριο της χρονοσειράς, ενώ τα δύο ακραία στελέχη εκτείνονται μέχρι τη μέγιστη και την ελάχιστη τιμή αντίστοιχα. Μία οριζόντια γραμμή διέρχεται από το ορθογώνιο πλαίσιο και δηλώνει την τιμή της διαμέσου.

Ο πλήρης πίνακας της εικόνας 4.1 δίνεται στο πίνακα Π-3 του παραρτήματος στο τέλος της διπλωματικής εργασίας. Να σημειώσουμε εδώ ότι αυτά τα δεδομένα είναι τα φιλτραρισμένα δεδομένα από τα οποία έχουμε αφαιρέσει τυχόν outliers και έχουμε αντικαταστήσει κενές-άγνωστες τιμές σύμφωνα με την μεθοδολογία που θα περιγράψουμε σε επόμενη ενότητα.

Επιπλέον, ο πίνακας Π-4 του παραρτήματος, παρουσιάζει τον μέσο αριθμό των μετρητών ανά περιοχή. Όπως εξηγείται και στο παράρτημα, προτιμήσαμε να παρουσιάσουμε τον μέσο όρο των τιμών και όχι την εξέλιξη τους σε βάθος χρόνου κατά την διάρκεια των 84 μηνών των διαθέσιμων μετρήσεων που είχαμε στα δεδομένα μας, γιατί πολύ απλά οι διακυμάνσεις γύρω από την μέση τιμή είναι πολύ μικρές και πρακτικά μπορούν να αντιπροσωπευθούν από το μέσο όρο των τιμών τους για την περίοδο 2007-2013. Ο επόμενος πίνακας παρουσιάζει τα δεδομένα για την μέση τιμή του αριθμού των μετρητών σε κάθε ΤΚ που καταγράφεται στον πίνακα 4.2. Στην διπλωματική χρειασθήκαμε τον αριθμό των μετρητών σε μηνιαία βάση ανά περιοχή για να κάνουμε ανάλυση cluster (βλ. ενότητα 4.2).

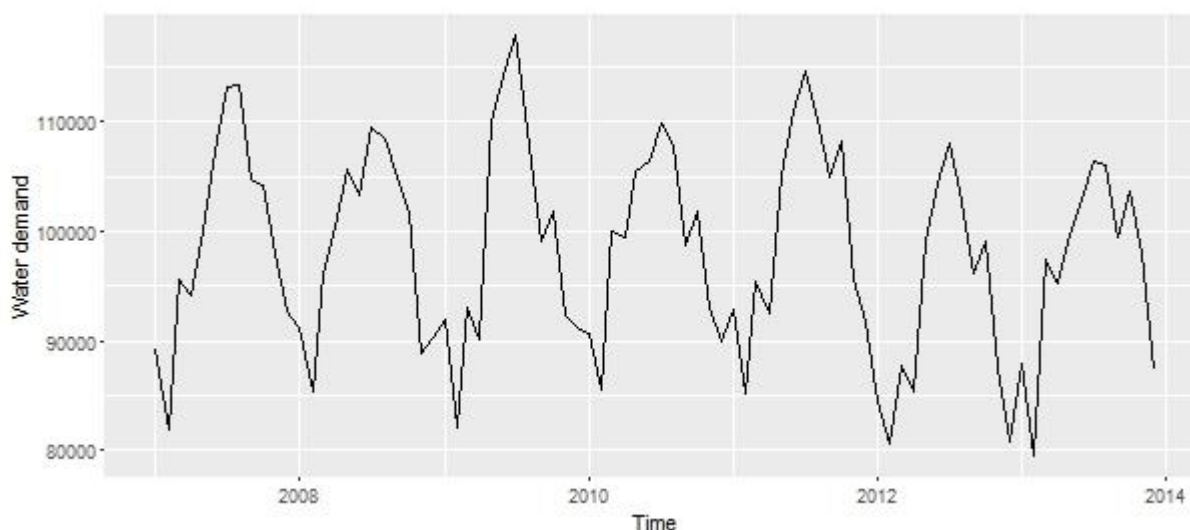
Περιοχή	Μ.Ο. μετρητών
ΖΩΓΡΆΦΟΥ-15771	15778.45238
ΖΩΓΡΆΦΟΥ-15772	19580.29762
ΖΩΓΡΆΦΟΥ-15773	14579.72619
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13561	7522.988095
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13562	8353.52381
ΑΓ.ΒΑΡΒΆΡΑΣ-12351	11327.67857
ΑΓ.ΔΗΜΗΤΡΊΟΥ-17341	10459.70238
ΑΓ.ΔΗΜΗΤΡΊΟΥ-17342	12865.11905
ΑΓ.ΔΗΜΗΤΡΊΟΥ-17343	9636.666667
ΑΓ.ΔΗΜΗΤΡΊΟΥ-17372	3.404761905
ΑΓ.ΙΩΆΝΝΗΣ ΡΈΝΤΗ-18232	10
ΑΓ.ΙΩΆΝΝΗΣ ΡΈΝΤΗ-18233	8570.142857
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15341	11961.17857
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15342	8426.404762
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15343	12178.03571
ΑΙΓΆΛΕΩ-12241	7052.595238
ΑΙΓΆΛΕΩ-12242	7667.404762
ΑΙΓΆΛΕΩ-12243	11217
ΑΙΓΆΛΕΩ-12244	9075.928571
ΆΛΙΜΟΣ-17436	4

ΆΛΙΜΟΣ-17455	14788.88095
ΆΛΙΜΟΣ-17456	8687.285714
ΑΜΑΡΟΥΣΙΟΥ-15122	8788.511905
ΑΜΑΡΟΥΣΙΟΥ-15123	2953.535714
ΑΜΑΡΟΥΣΙΟΥ-15124	9845.5
ΑΜΑΡΟΥΣΙΟΥ-15125	9825.261905
ΑΜΑΡΟΥΣΙΟΥ-15126	9239.416667
ΆΝΩ ΛΙΪΣΙΑ-13341	12.78571429
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16451	8957.77381

Πίνακας 4.2-Μέση τιμή αριθμού μετρητών σε ορισμένους ταχ. κώδικες

Παρακάτω θα παρουσιάσουμε κάποια γραφήματα χρονοσειρών παρμένα από την βάση των δεδομένων μας, με σκοπό να καταδείξουμε κάποιες μορφές χρονοσειρών που είναι παρούσες στα συνολικά μας δεδομένα.

1)



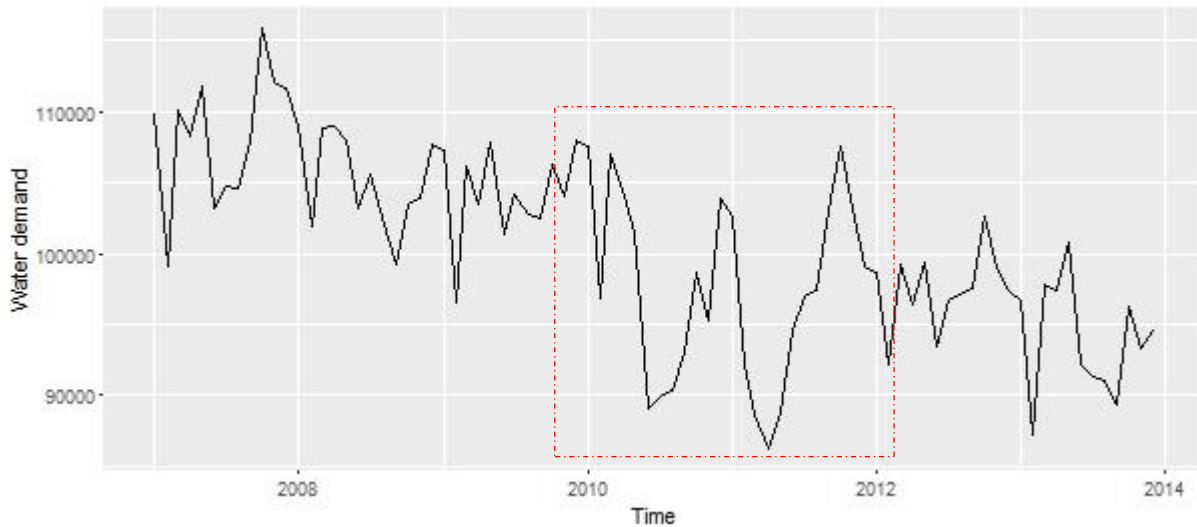
Εικόνα 4.3-Ζήτηση νερού(σε m³) καταγεγραμμένα στον ταχ.κώδικα 15126, Μαρούσι

Από την παραπάνω εικόνα είναι εμφανής η εποχικότητα στα δεδομένα διάρκειας ενός χρόνου, καθώς φαίνεται η κατανάλωση να αυξάνεται μέχρι τους καλοκαιρινούς μήνες όπου καταγράφεται η μέγιστη ζήτηση νερού και κατόπιν, η ζήτηση ελαττώνεται με μικρά скаμπανεβάσματα μέχρι το τέλος του έτους. Επίσης, η χρονοσειρά έχει σχεδόν σταθερή μέση τιμή και δεν έχει εμφανή τάση.

2)

Στην επόμενη χρονοσειρά εικονίζεται η ζήτηση νερού σε περιοχή του δήμου Ζωγράφου.

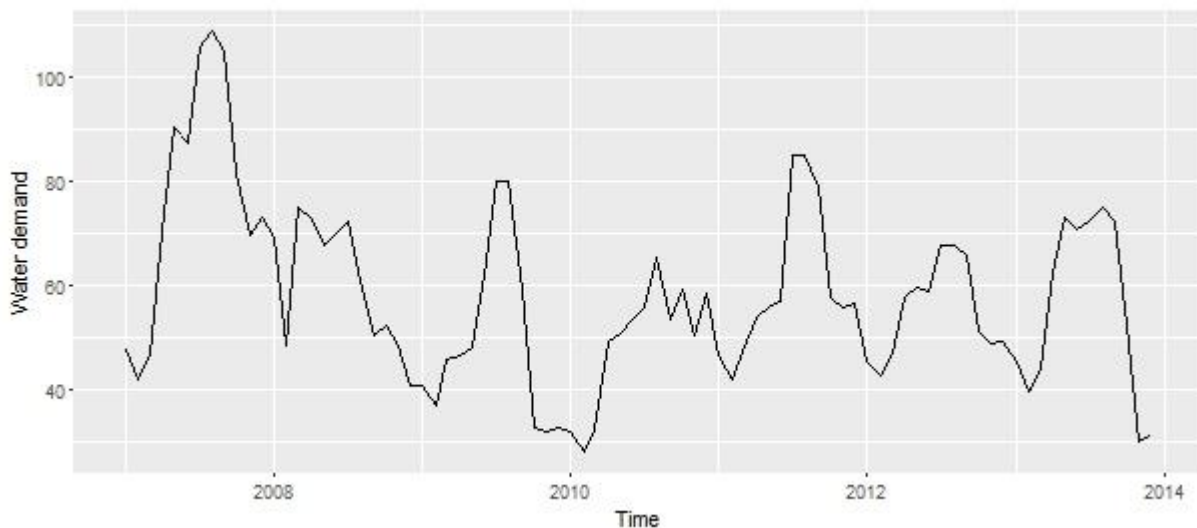
Παρατηρούμε ότι η χρονοσειρά αυτή δεν παρουσιάζει εμφανή εποχικότητα, έχει πτωτική τάση και τυχειότητα ειδικά στις περιοχές εσωτερικά του κόκκινου τετραγώνου.



Εικόνα 4.4-Ζήτηση νερού(σε m^3) καταγεγραμμένα στον ταχ.κώδικα 15771, Ζωγράφου

3)

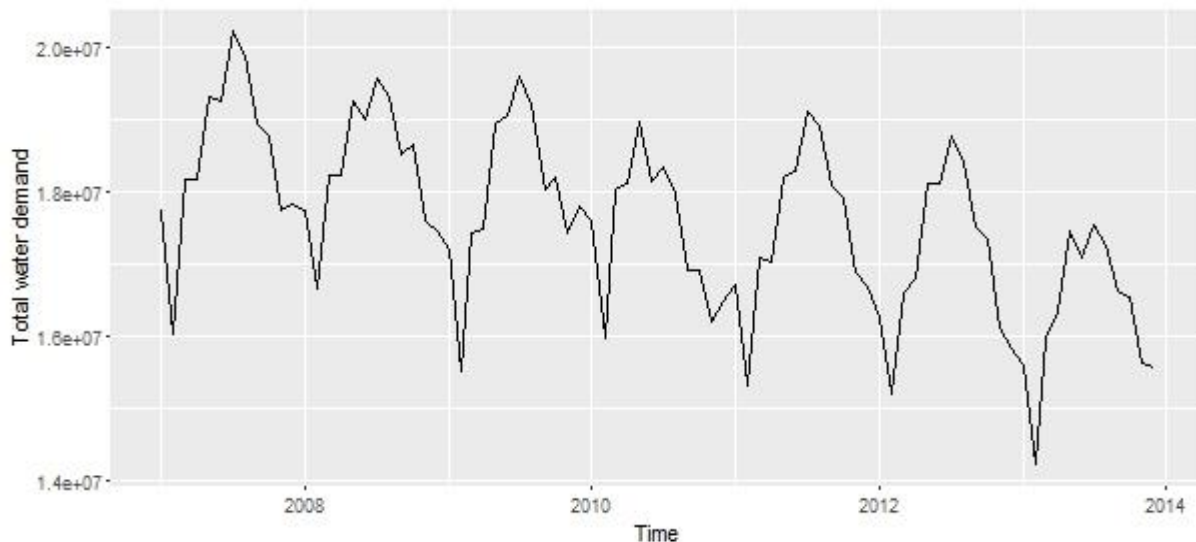
Η παρακάτω εικόνα αφορά την περιοχή Μαρκόπουλου. Όπως μπορούμε να δούμε, τα δεδομένα δεν έχουν εποχικότητα και χαρακτηρίζονται από τυχειότητα και απουσία σαφής τάσης.



Εικόνα 4.5-Ζήτηση νερού(σε m^3) καταγεγραμμένα στον ταχ.κώδικα 19003, Μαρκόπουλο

4)

Στην εικόνα 4.6 παρουσιάζεται η συνολική κατανάλωση νερού όπως αυτή αποτυπώνεται από τα διαθέσιμα δεδομένα. Βλέπουμε ότι έχει σαφή εποχικότητα αλλά και μία ελαφρά πτωτική τάση της τάξης περίπου του 5% κάθε χρόνο.



Εικόνα 4.6-Συνολική ζήτηση νερού(σε m³) όπως αποτυπώνεται από τους 274 ταχ.κωδικούς της Αττικής

4.2 Ομαδοποίηση Δεδομένων (Clustering)

Ένας απλός ορισμός για την ομαδοποίηση ή συσταδοποίηση (clustering) είναι ο ακόλουθος: ομαδοποίηση ονομάζεται η διαδικασία που οργανώνει πρότυπα (παρατηρήσεις, δεδομένα ή διανύσματα χαρακτηριστικών) σε ομάδες (συστάδες-clusters), όπου τα μέλη μιας ομάδας είναι παρόμοια μεταξύ τους σύμφωνα με κάποιο κριτήριο. Πρόκειται για μια μαθηματική τεχνική που έχει σκοπό να προσδιοριστούν οι ομάδες στις οποίες ανήκουν διάφορες ποσότητες δεδομένων, με βάση κάποια κριτήρια ομοιογένειας.

Προτού σχολιάσουμε αναλυτικότερα στις ακόλουθες ενότητες τις τρεις μεθοδολογίες που χρησιμοποιήσαμε σε αυτήν την διπλωματική, θα προσπαθήσουμε να εξηγήσουμε την έννοια της ιεραρχικής ομαδοποίησης δεδομένων (hierarchical data clustering) η οποία τεχνική χρησιμοποιήθηκε από την παρούσα διπλωματική.

Ο στόχος τέτοιων αλγορίθμων είναι να δημιουργήσουν μία αλληλουχία από εμφωλευμένα cluster που κυμαίνονται από μοναχικά cluster μίας τιμής, έως ένα ευρύτερο cluster που περιέχει άλλα μικρότερα στο εσωτερικό του. Αυτή η ιεραρχία των cluster παρουσιάζεται γραφικά από δενδρογράμματα όπως θα δείξουμε και πιο κάτω. Ένα κύριο χαρακτηριστικό αυτής της τεχνικής είναι ότι δεν υποθέτει κάποιο συγκεκριμένο αριθμό cluster, αντιθέτως μπορούν να σχηματιστούν διάφοροι αριθμοί από cluster «κόβοντας» το δενδρογράμμα σε κατάλληλα σημεία.

4.2.1 Είδη Ιεραρχικής Ομαδοποίησης Δεδομένων

Υπάρχουν γενικά δύο τεχνικές ιεραρχικής ομαδοποίησης δεδομένων:

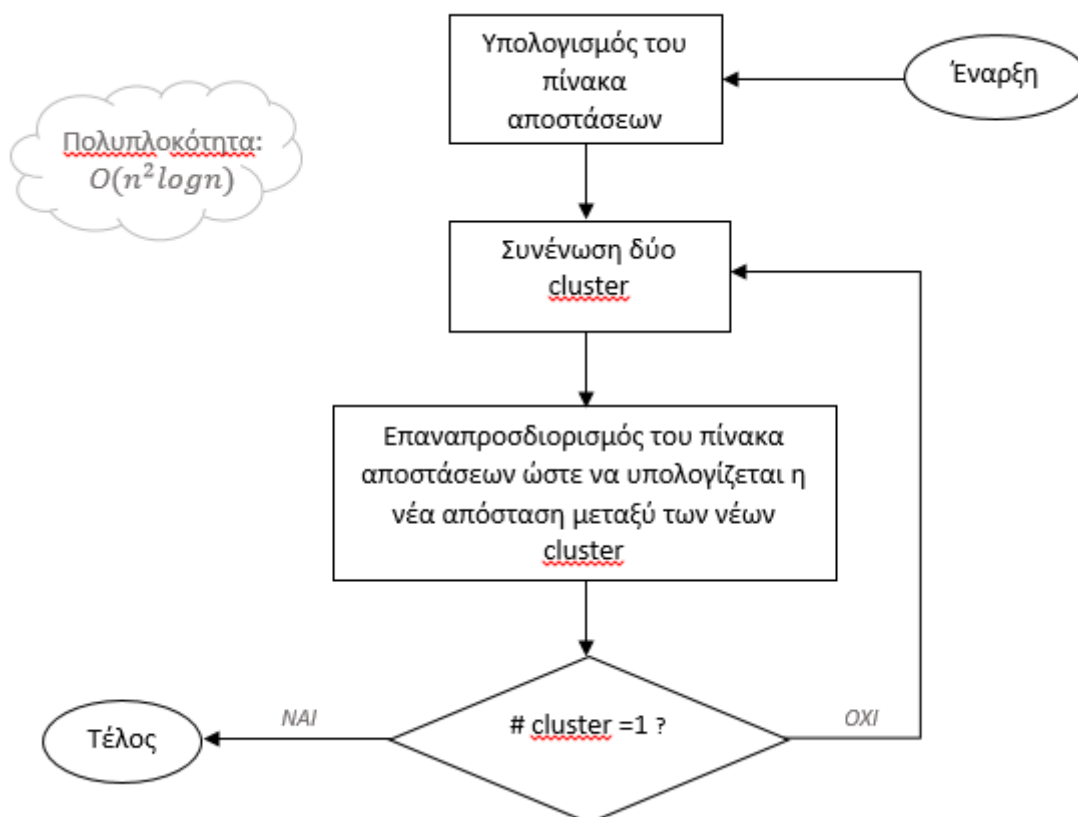
- **Συσσωρευτική τεχνική (Agglomerative)**

Ξεκινά θεωρώντας τα μεμονωμένα δεδομένα ως ξεχωριστές ομάδες (cluster) και σε κάθε βήμα συγχωνεύεται με το πιο κοντινό ζευγάρι cluster. Εδώ θα πρέπει να ξεκαθαριστεί και η έννοια της απόστασης μεταξύ δύο cluster κάτι που θα γίνει στην ενότητα 4.2.3.

- **Διαιρετική τεχνική (Divisive)**

Μία λιγότερο γνωστή τεχνική που ξεκινάει με ένα ευρύτερο cluster που περιλαμβάνει άλλα μικρότερα cluster και καθώς προχωράνε τα βήματα του αλγορίθμου, το αρχικό cluster διασπάται σε μικρότερα μέχρι να απομείνουν cluster μίας τιμής στο κατώτερο επίπεδο.

Στις επόμενες σειρές θα προσπαθήσουμε να εξηγήσουμε την συσσωρευτική τεχνική όπου είναι και αυτή που χρησιμοποιήθηκε στην έρευνα μας.



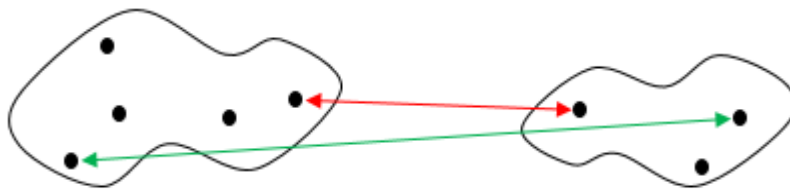
Εικόνα 4.7-Στάδια μεθόδου ιεραρχικής συσσωρευτικής ομαδοποίησης δεδομένων

4.2.2 Η Συσσωρευτική Τεχνική

Η κύρια ιδέα πίσω από αυτήν την τεχνική είναι τα κοντινά δεδομένα να βρεθούν οπωσδήποτε σε κοινά cluster. Ξεκινώντας με τα μοναδικά σημεία σε διδιάστατο χώρο σαν ξεχωριστά cluster, συγχωνεύουμε διαδοχικά δύο cluster μέχρι να μείνει ένα ευρύτερο cluster στο τέλος. Τα αλγοριθμικά βήματα αυτά αποτυπώνονται στο λογικό διάγραμμα της εικόνας 4.7.

4.2.3 Η Απόσταση μεταξύ δύο Cluster

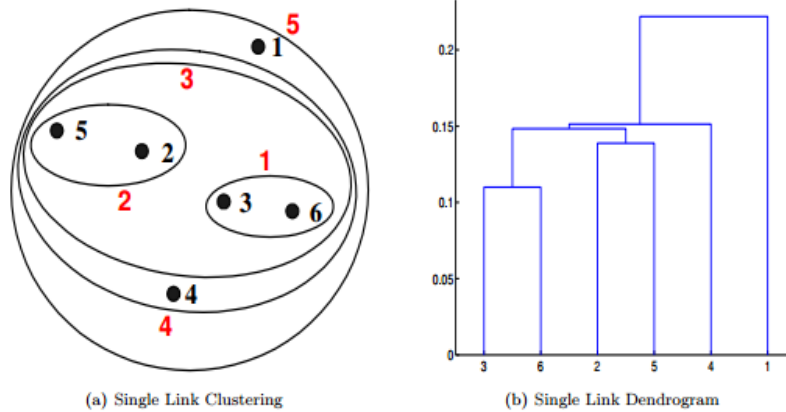
Το κλειδί για τον αλγόριθμο της ιεραρχικής ομαδοποίησης δεδομένων, όπως φάνηκε και από το παραπάνω λογικό διάγραμμα (εικόνα 4.7), είναι ο καθορισμός της απόστασης μεταξύ δύο διαφορετικών cluster. Ανάλογα ποια τεχνική θα επιλέξουμε, τα αποτελέσματα της ανάλυσης cluster θα διαφέρουν αρκετά μεταξύ τους. Παρακάτω θα παρουσιάσουμε τις πιο γνωστές μεθόδους μέτρησης αποστάσεων μεταξύ αντικειμένων δύο διαφορετικών cluster.



Εικόνα 4.8-Ορισμός αποστάσεων μεθόδων *single link*(κόκκινο βέλος) και *complete link*(πράσινο βέλος)

4.2.3.1 Μονή Σύνδεση (Single link)

Με την μέθοδο αυτή, η απόσταση μεταξύ δύο cluster προσδιορίζεται σαν την ελάχιστη απόσταση(κόκκινο βέλος, εικόνα 4.8) μεταξύ δύο οποιονδήποτε τιμών που ανήκουν σε κάθε cluster. Η τεχνική αυτή είναι επιρρεπής σε δεδομένα με θόρυβο και ακραίες τιμές. Ένα παράδειγμα φαίνεται στην εικόνα 4.9 της επόμενης σελίδας.

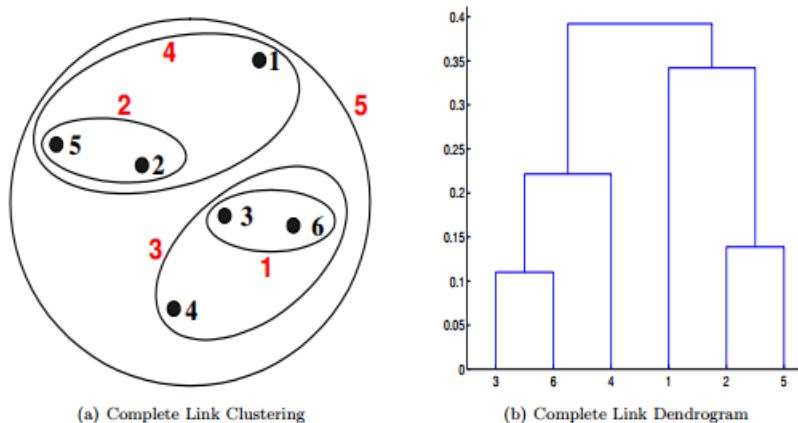


Εικόνα 4.9-Παράδειγμα ιεραρχικών cluster με την μέθοδο single link

Το ύψος στο οποίο κάθε cluster συνενώνεται με κάποιο άλλο στο γράφημα του δενδρογράμματος, καθρεφτίζει το μέτρο της απόστασης που έχουν μεταξύ τους. Στο παραπάνω παράδειγμα, η απόσταση μεταξύ των cluster $\{3,6\}$ και $\{2,5\}$ είναι $\text{dist}(\{3,6\}, \{2,5\}) = \min(\text{dist}(3,2), \text{dist}(6,2), \text{dist}(3,5), \text{dist}(6,5)) = \min(0.1483, 0.2540, 0.2843, 0.3921) = 0.1483$.

4.2.3.2 Πλήρης Σύνδεση (Complete link)

Σε αυτήν την περίπτωση, η απόσταση μεταξύ δύο cluster ορίζεται ως η μέγιστη απόσταση μεταξύ δύο οποιονδήποτε σημείων τους (πράσινη γραμμή, εικόνα 4.8). Η τεχνική αυτή δεν είναι τόσο επιρρεπής στον θόρυβο, αλλά τείνει να ορίζει περισσότερα ξεχωριστά cluster απ'όσα θα έπρεπε.



Εικόνα 4.10-Παράδειγμα μεθόδου complete link

Στο παράδειγμα της εικόνας 4.10, το cluster $\{3,6\}$ συνενώνεται με το cluster $\{4\}$ και όχι με το $\{2,5\}$ γιατί $\text{dist}(\{3,6\}, \{4\}) = \max(\text{dist}(3,4), \text{dist}(6,4)) = \max(0.1513, 0.2216) = 0.2216$ που είναι μικρότερο του $\text{dist}(\{3,6\}, \{2,5\}) = \max(\text{dist}(3,2), \text{dist}(6,2), \text{dist}(3,5), \text{dist}(6,5)) =$

$\max(0.1483, 0.2540, 0.2843, 0.3921) = 0.3921$ και $\text{dist}(\{3,6\}, \{1\}) = \max(\text{dist}(3,1), \text{dist}(6,1)) = \max(0.2218, 0.2347) = 0.2347$.

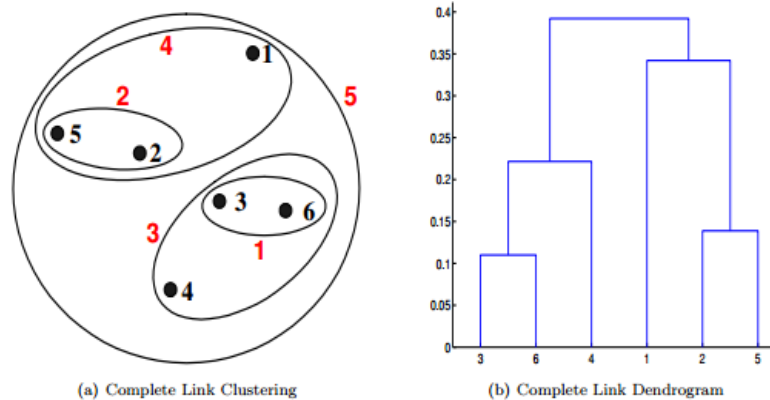
4.2.3.3 Μέση Απόσταση (Group average)

Με αυτήν την μέθοδο, η απόσταση μεταξύ δύο cluster ορίζεται ως η μέση τιμή των αποστάσεων όλων των δυνατών ζευγαριών μεταξύ των δύο cluster. Η επόμενη εξίσωση εκφράζει τον τύπο εύρεσης της απόστασης με αυτή τη μέθοδο:

$$\text{dist}(\text{cluster}_1, \text{cluster}_2) = \sum_{\substack{p \in \text{cluster}_1 \\ q \in \text{cluster}_2}} \frac{\text{dist}(p, q)}{\text{size}(\text{cluster}_1) * \text{size}(\text{cluster}_2)}$$

Για να δείξουμε πως λειτουργεί η μέθοδος έχουμε το παράδειγμα όπου:

$\text{dist}(\{3,6,4\}, \{1\}) = (0.2218 + 0.3688 + 0.2347) / (3 * 1) = 0.2751$. $\text{dist}(\{2, 5\}, \{1\}) = (0.2357 + 0.3421) / (2 * 1) = 0.2889$. $\text{dist}(\{3,6,4\}, \{2,5\}) = (0.1483 + 0.2843 + 0.2540 + 0.3921 + 0.2042 + 0.2932) / (6 * 1) = 0.2637$. Επειδή $\text{dist}(\{3,6,4\}, \{2,5\})$ είναι μικρότερο από $\text{dist}(\{3,6,4\}, \{1\})$ και $\text{dist}(\{2, 5\}, \{1\})$ αυτά τα δύο cluster συγχωνεύονται στο τέταρτο στάδιο.



Εικόνα 4.11-Παράδειγμα μεθόδου group average

4.2.3.4 Μέθοδος του Ward (Ward's method)

Αυτή η μέθοδος είναι μία διαφορετική προσέγγιση σε σχέση με όσες έχουμε δει στις προηγούμενες ενότητες και είναι η μέθοδος που χρησιμοποιήθηκε σε αυτή τη διπλωματική για τον σχηματισμό των cluster. Στόχος του γενικού αλγορίθμου που υλοποιεί αυτή τη μέθοδο είναι σε κάθε βήμα να συγχωνεύονται τα cluster τα οποία

οδηγούν στην μικρότερη αύξηση του αθροίσματος των τετραγώνων των Ευκλείδειων αποστάσεων στο νέο cluster. Η απόσταση στην περίπτωση αυτή ορίζεται ως:

$$\text{dist}(A, B) = \sum_{x \in A} \|x - m_A\|^2 + \sum_{x \in B} \|x - m_B\|^2 + \sum_{x \in A \cup B} \|x - m_{A \cup B}\|^2.$$

Στόχος δηλαδή είναι να επιλεγεί εκείνο το cluster το οποίο θα ελαχιστοποιεί την παρακάτω εξίσωση κόστους:

$$\Delta(A, B) = \frac{n_A n_B}{n_A + n_B} \|m_A - m_B\|^2,$$

όπου n_j είναι ο αριθμός των παρατηρήσεων στο cluster j και m_j είναι το κέντρο του cluster j .

4.3 Μέθοδος Έρευνας Διπλωματικής

Σε αυτήν την ενότητα θα αναλυθούν τα βήματα της μεθόδου που αναπτύξαμε για την σύγκριση των τριών ευρύτερων μεθοδολογιών που παρουσιάστηκαν περιληπτικά στην ενότητα 2.6.

Βήμα 1: Διαχείριση μηδενικών και κενών τιμών

Η συλλογή και διαχείριση των δεδομένων που αποτελούν τις χρονοσειρές, ίσως η βασικότερη πηγή πληροφοριών για την επιστήμη των προβλέψεων, δεν είναι πάντα εύκολη υπόθεση. Υπάρχουν περιπτώσεις ελλειπουσών ή μηδενικών τιμών οι οποίες δημιουργούν προβλήματα στην εφαρμογή των περισσότερων στατιστικών μεθόδων πρόβλεψης και επομένως χρήζουν άμεσης αντιμετώπισης.

Στα δεδομένα που είχαμε για το σκοπό της διπλωματικής αυτής, υπήρχαν αρκετές χρονοσειρές που είχαν μεμονωμένες τιμές ως μηδενικές ή άγνωστες. Η μέθοδος που χρησιμοποιήθηκε για την διαχείριση τους ήταν η εξής:

- Η τιμή προς αντικατάσταση της κενής τιμή ορίζεται ως το ημιάθροισμα της προηγούμενης και της επόμενης παρατήρησης, όταν η χρονοσειρά χαρακτηρίζεται από στασιμότητα και δεν έχει εποχιακή συμπεριφορά.
- Αν η χρονοσειρά παρουσιάζει εποχικότητα, τότε η κενή τιμή ορίζεται ως η μέση τιμή των παρελθοντικών ή/και μελλοντικών τιμών των αντίστοιχων περιόδων.

Διαφορετική αντιμετώπιση χρήζουν οι μηδενικές τιμές οι οποίες αποτελούν μία διαφορετική κατηγορία των «ιδιόρρυθμων» τιμών μιας χρονοσειράς. Αυτές οι τιμές διακρίνονται σε δύο κατηγορίες. Αρχικά, είναι οι μηδενικές τιμές που καταγράφηκαν ως μηδενικές λόγω σφάλματος του πληροφοριακού συστήματος και δεν

ανταποκρίνονται στην πραγματικότητα. Αυτές οι τιμές ουσιαστικά δεν διαφέρουν από τις προαναφερθείσες ελλειπούσες τιμές και είναι όμοιος και ο τρόπος αντιμετώπισης τους. Η δεύτερη κατηγορία των μηδενικών τιμών είναι οι πραγματικά μηδενικές τιμές οι οποίες αναφέρονται σε χρονοσειρές διακοπτόμενης ζήτησης. Αυτές οι τιμές προφανώς δεν μπορούν να συμπληρωθούν με κάποιο από τους προηγούμενους τρόπους καθώς είναι πραγματικές καταγραφείσες μηδενικές τιμές. Οπότε είναι εμφανής η ανάγκη εύρεσης κάποιου διαφορετικού τρόπου αντιμετώπισης τους. Χαρακτηριστικό παράδειγμα εμφάνισης μηδενικών τιμών είναι οι χρονοσειρές διακοπτόμενης ζήτησης, οι οποίες εμφανίζουν ιδιαίτερο ενδιαφέρον καθώς απαιτούν την προσέγγιση βάσει διαφορετικής μεθοδολογίας στην οποία προφανώς και δεν εμπλέκεται η μεταβολή των τιμών τους αλλά η αξιοποίηση της πληροφορίας που λαμβάνεται από τη μηδενική τιμή. Στα δεδομένα της διπλωματικής μας, οι μηδενικές τιμές ήταν μεμονωμένες και εκλήφθησαν σαν σφάλματα του πληροφοριακού συστήματος εισαγωγής δεδομένων.

Βήμα 2 : Διαχείριση ακραίων τιμών (outliers)

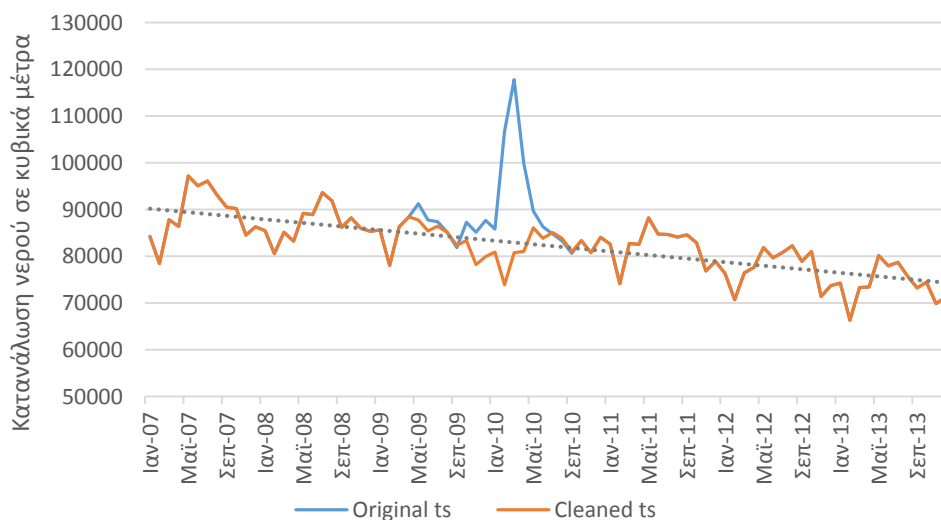
Τα δεδομένα των χρονοσειρών συχνά υφίστανται ξαφνικές αλλαγές που επηρεάζουν την δυναμική των δεδομένων σε παροδική ή μόνιμη βάση. Μερικές φορές, αυτός που διαχειρίζεται τα δεδομένα μπορεί να έχει κάποια πληροφορία για την ύπαρξη αυτών των επιδράσεων(απεργίες, νέοι κανονισμοί, καιρικά φαινόμενα είναι ορισμένα από αυτά) . Σε άλλες όμως περιπτώσεις τέτοιες πληροφορίες δεν είναι διαθέσιμες. Ανεξάρτητα από την τυχόν γνωστή ή άγνωστη πηγή τους, οι αλλαγές αυτές είναι συνήθως μη συστηματικές και δεν μπορούν να συλληφθούν από τα συνήθη μοντέλα χρονοσειρών. Μια προκαταρκτική ανάλυση των δεδομένων μπορεί να αποκαλύψει την ύπαρξη έκτακτων γεγονότων που δεν ανταποκρίνονται στα πρότυπα που μπορούν να συλληφθούν από τα μοντέλα χρονοσειρών, δηλαδή επιδράσεις που είναι εξωγενείς στο μοντέλο. Όταν αυτή η εξωγενής επίδραση δεν σχετίζεται με οποιοδήποτε γνωστό συμβάν τότε μπορούμε να πούμε ότι υπάρχει μια ακραία τιμή(outlier) της χρονολογικής σειράς. Θεωρούμε πέντε είδη των ακραίων τιμών: innovational outlier, additive outlier, level shift, temporary change και seasonal level shift.

Ο εντοπισμός και η διόρθωση της επίδρασης των ακραίων τιμών είναι σημαντική, επειδή έχουν επίπτωση στην επιλογή του μοντέλου, την εκτίμηση των παραμέτρων και, κατά συνέπεια, σχετικά με τις προβλέψεις και τα άλλα αποτελέσματα που επιδιώκονται με την ανάλυση, όπως την εποχική προσαρμογή.

Για την πραγματικότητα αυτής της διπλωματικής, το φιλτράρισμα των αρχικών παρατηρήσεων από ακραίες τιμές είναι πραγματοποιήθηκε χρησιμοποιώντας το πακέτο `tsoutliers` και συγκεκριμένα την συνάρτηση `tso`, της στατιστικής γλώσσας προγραμματισμού R. Τα στάδια υλοποίησης του αλγορίθμου έχουν τα παρακάτω βήματα:

1. Εντοπισμός ακραίων τιμών. Δεδομένου ενός μοντέλου ARIMA, οι ακραίες τιμές εντοπίζονται ελέγχοντας την σημαντικότητα όλων των τύπων των ακραίων τιμών.
2. Απαλοιφή ακραίων τιμών. Δεδομένου ενός συνόλου πιθανόν ακραίων τιμών, ένα μοντέλο ARIMA προσαρμόζεται στα δεδομένα. Η σημαντικότητα των ακραίων τιμών επαναπροσδιορίζεται με βάση το νέο μοντέλο που προσαρμόστηκε, και αν χρειάζεται κάποια αφαιρούνται από το σύνολο των ακραίων τιμών.
3. Επανάληψη των βημάτων 1 και 2 πρώτα για την αρχική χρονοσειρά και έπειτα για την προσαρμοσμένη.

Στην επόμενη εικόνα παρουσιάζεται η επίδραση της εφαρμογής της συνάρτησης tso για την εύρεση και αντικατάσταση των ακραίων τιμών με σκοπό την ομαλοποίηση των δεδομένων. Παρατηρούμε ότι μετά την αφαίρεση των ακραίων τιμών, η χρονοσειρά παρουσιάζει εποχική συμπεριφορά με πτωτική τάση.

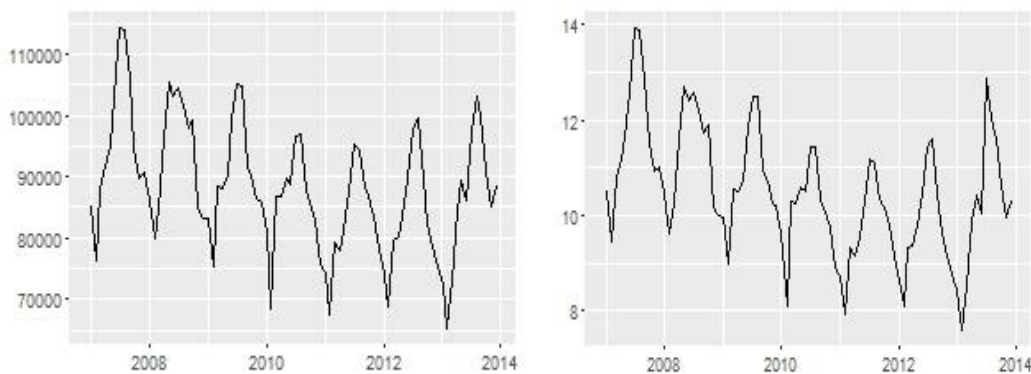


Εικόνα 4.12-Παράδειγμα εφαρμογής εύρεσης και αντικατάστασης ακραίων τιμών για την περιοχή Άγιος Ιωάννης-Ρέντης

Βήμα 3 : Δημιουργία cluster

Θα ξεκινήσουμε την περιγραφή με την εξής διαπίστωση: Δύο χρονοσειρές που παρουσιάζουν την ίδια ακριβώς συμπεριφορά (π.χ. έχουν την ίδια τάση με την ίδια ακριβώς εποχικότητα) αλλά έχουν μεγάλη διαφορά στην μέση τιμή των δεδομένων τους, δεν μπορούν να μπουν σε καμία περίπτωση στο ίδιο cluster. Επομένως εξαιτίας της ποικιλίας των διαθέσιμων δεδομένων ως προς τα μεγέθη που έχουν (ζήτηση μερικών δεκάδων έως μερικών εκατοντάδων χιλιάδων κυβικών μέτρων νερού) ο αριθμός των cluster που θα δημιουργούνται θα ήταν σχεδόν όσος και αριθμός των διαθέσιμων παρατηρήσεων, πράγμα που δεν είναι επιθυμητό. Οπότε, έπρεπε να βρούμε ένα τρόπο ώστε να φέρουμε χρονοσειρές με ίδια ποιοτική συμπεριφορά αλλά διαφορετικές μέσες τιμές στην ίδια κλίμακα. Η λύση μπορεί να βρεθεί αν παρατηρήσουμε ότι όσο περισσότεροι μετρητές υπάρχουν για ένα ταχυδρομικό κώδικα, τόσο μεγαλύτερη είναι η κατανάλωση που καταγράφεται. Υπάρχει δηλαδή μία αναλογικότητα της ζήτησης

νερού με τον αριθμό των μετρητών που δίνουν αυτές τις μετρήσεις. Άρα αν διαιρέσουμε τις τιμές μηνιαίας ζήτησης με τις μηνιαίες τιμές του αριθμού μετρητών, θα μπορέσουμε να μειώσουμε σημαντικά την κλίμακα μίας χρονοσειράς με μεγάλη μέση τιμή χωρίς να επιδράσουμε σημαντικά στα ποιοτικά χαρακτηριστικά της χρονοσειράς και άρα να μπορέσουμε πιο εύκολα να την εντάξουμε σε μία ομάδα με χρονοσειρές παρόμοιας τάξης και χαρακτηριστικών. Για του λόγου το αληθές στις παρακάτω εικόνες παρουσιάζονται στα αριστερά η ζήτηση νερού στην περιοχή της Αγ.Παρασκευής (TK 15342) και στα δεξιά η το πηλίκο της ζήτησης νερού προς τον αριθμό των μετρητών για την ίδια περιοχή. Είναι εμφανές ότι οι δύο χρονοσειρές έχουν ίδια ποιοτικά χαρακτηριστικά αλλά η δεύτερη είναι σε πολύ μικρότερη κλίμακα μεγέθους.



Εικόνα 4.13- Χρονοσειρά ζήτησης νερού(αριστερά) και χρονοσειρά ζήτησης νερού ανά μετρητή(δεξιά) για την περιοχή της Αγ.Παρασκευής

Άρα, έχοντας τις χρονοσειρές ζήτησης νερού ανά μετρητή μπορούμε να ομαδοποιήσουμε τα δεδομένα μας σε cluster χρησιμοποιώντας την συνάρτηση hclust από το πακέτο stats της γλώσσας R. Ο κώδικας αυτός, όπως και όλοι οι υπόλοιποι κώδικες που χρησιμοποιήθηκαν, θα παρουσιαστούν στο παράρτημα στο τέλος της διπλωματικής.

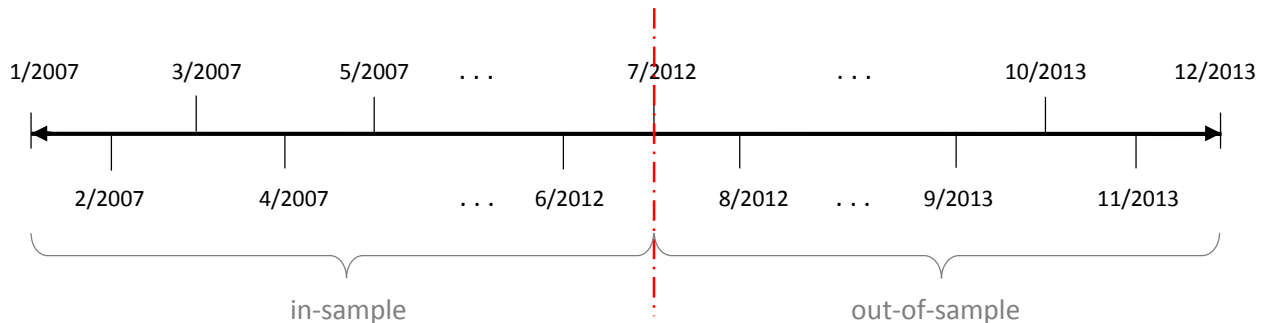
Τα βήματα 1-2 είναι στην ουσία μέθοδοι «καθαρισμού» των δεδομένων από μη επιθυμητές τιμές και έχουν εφαρμογή και στις τρεις μεθοδολογίες που θα συγκρίνουμε και θα παρουσιάσουμε στις επόμενες ενότητες. Αντίθετα, το βήμα 3 για την δημιουργία των cluster αφορά μία μόνο μεθοδολογία από τις τρεις, στην οποία χωρίζουμε τα δεδομένα σε cluster και με βάση αυτά παράγουμε προβλέψεις(βήμα 4) και ελέγχουμε την ακρίβεια τους.

Βήμα 4: Παραγωγή προβλέψεων

Αφού έχουμε καθαρίσει τα δεδομένα από τυχόν ακραίες-κενές τιμές, το επόμενο βήμα είναι η παραγωγή προβλέψεων. Όπως ειπώθηκε και στην ενότητα 2.6 , στόχος της διπλωματικής είναι να αναπτύξει μία βέλτιστη μεθοδολογία για την παραγωγή μεσοπρόθεσμων προβλέψεων. Για αυτό αναπτύξαμε και συγκρίναμε τρεις ευρύτερες

μεθοδολογίες ως προς την ακρίβεια των προβλέψεων που δίνουν. Για να είμαστε σε θέση να απαντήσουμε στο ερώτημα της ακρίβειας, θα πρέπει οι προβλέψεις που θα κάνουμε να αφορούν χρονικές περιόδους για τις οποίες έχουμε τα πραγματικά δεδομένα (θυμίζουμε ότι τα διαθέσιμα δεδομένα μας είναι για την περίοδο από 1/2007 έως 12/2013). Ως εκ τούτου, αποφασίσαμε να χωρίσουμε την χρονική περίοδο των δεδομένων σε δύο διαστήματα (εικόνα 4.14):

- **Διάστημα Α:** Το διάστημα αυτό αποτελεί το 80% των συνολικών δεδομένων και περιλαμβάνει 67 μήνες (από 1/2007 έως 7/2012). Σε αυτό το διάστημα γίνεται η «εκπαίδευση» του κάθε στατιστικού μοντέλου πρόβλεψης (model fitting) όπου προσδιορίζονται οι βέλτιστες παράμετροι ώστε το παραγόμενο μοντέλο να ταιριάζει όσο το δυνατόν καλύτερα σε αυτά τα δεδομένα. Το μοντέλο που θα δώσει το μικρότερο σφάλμα MAPE σε αυτό το διάστημα, επιλέγεται ώστε να επεκταθεί και να δώσει πρόβλεψη στο διάστημα Β.
- **Διάστημα Β:** Αφού προσδιοριστούν μοναδικά οι παράμετροι για το μοντέλο, το χρησιμοποιούμε για την παραγωγή προβλέψεων στο διάστημα Β (out-of-sample forecasting). Θεωρούμε δηλαδή ότι «κρύβουμε» τα δεδομένα μας στο διάστημα αυτό το οποίο αποτελεί το 20% των συνολικών δεδομένων και περιλαμβάνει 17 μήνες (από 8/2012 έως 12/2013).



Εικόνα 4.14-Χωρισμός χρονικού εύρους δεδομένων για την διαδικασία παραγωγής προβλέψεων

Η ακρίβεια των προβλέψεων εξετάζεται στο διάστημα Β με βάση το κριτήριο MAPE.

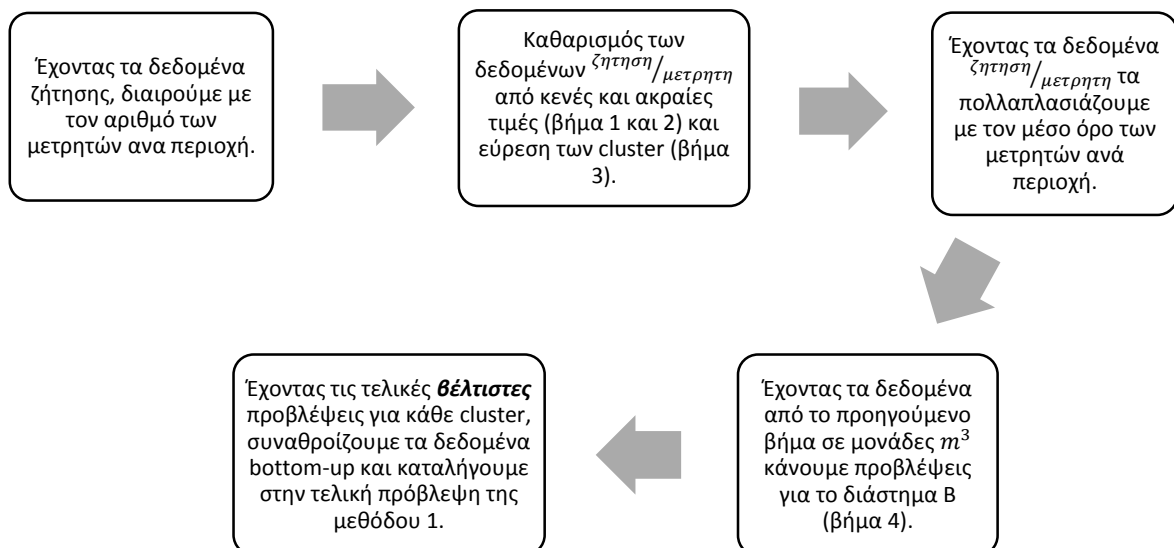
Στην επόμενη ενότητα θα γίνει αναλυτικότερη περιγραφή των τριών μεθοδολογιών που περιγράψαμε περιληπτικά στην ενότητα 2.6.

4.3.1 Αναλυτική Παρουσίαση των Μεθοδολογιών

Μεθοδολογία Νο1

Η μέθοδος αυτή έχει ως εξής: Αφού ανακτήσουμε τα δεδομένα μηνιαίας ζήτησης για κάθε διαθέσιμο ταχυδρομικό κώδικα, τα διαιρούμε με τον αριθμό των μετρητών ανα περιοχή και μήνα. Έχουμε δηλαδή τα δεδομένα ζήτησης ανά μετρητή. Καθαρίζουμε τις τιμές $\frac{\text{ζήτηση}}{\text{μετρητή}}$ από τυχόν κενές και ακραίες τιμές. Σε αυτήν την μεθοδολογία θα χρειαστούμε και το βήμα 3 όπου χωρίζουμε τα δεδομένα μας σε $n=10$ cluster (ο αριθμός των cluster δεν ήταν προκαθορισμένος αλλά προήλθε από ανάλυση τα αποτελέσματα της οποίας θα παρουσιαστούν στο επόμενο κεφάλαιο). Έχοντας λοιπόν βρει τα cluster μας με βάση την ζήτηση ανά μετρητή, πολλαπλασιάζουμε τα δεδομένα αυτά με την μέση τιμή του αριθμού των μετρητών ώστε μετά να προχωρήσουμε στο βήμα 4 και να κάνουμε προβλέψεις σε μονάδες ζήτησης εκφρασμένες σε m^3 με βάση τις επτά μεθόδους χρονοσειρών που αναφέρονται στην αρχή της ενότητας 3.5. Αφού για κάθε cluster εφαρμόσουμε τις επτά προαναφερθείσες μεθόδους πρόβλεψης, κρατάμε αυτήν η οποία δίνει το μικρότερο σφάλμα με βάση το κριτήριο MAPE στο διάστημα A. Έτσι για κάθε cluster έχουμε τις βέλτιστες προβλέψεις στο διάστημα B. Σαν τελικό βήμα είναι η bottom-up συναθροίση των δεδομένων ώστε από τα $n=10$ cluster να καταλήξουμε σε μία χρονοσειρά 17 τιμών που είναι η τελική πρόβλεψη στο διάστημα B που δίνει η μεθοδολογία 1.

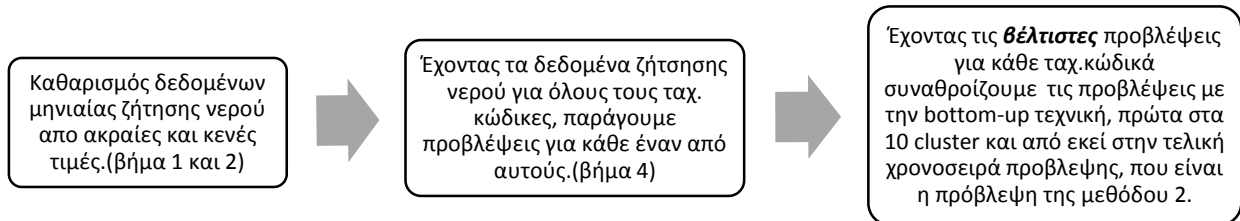
Παρακάτω δίνονται σε διάγραμμα ροής τα στάδια της μεθοδολογίας 1.



Εικόνα 4.15-Στάδια μεθοδολογίας Νο1

Μεθοδολογία Νο2

Για την δεύτερη μεθοδολογία αφού φιλτράρουμε τα δεδομένα ζήτησης για καθένα ταχυδρομικό κώδικα και αφαιρέσουμε τυχόν ακραίες και/ή μηδενικές τιμές (βήμα 1 και 2), παράγουμε τις βέλτιστες προβλέψεις για κάθε ταχυδρομικό κωδικό και ύστερα τις συναθροίζουμε με την τεχνική bottom-up ακολουθώντας την ιεραρχία από τις 274 προβλέψεις βάσης στις συναθροισμένες προβλέψεις για τα δέκα cluster και από εκεί στην τελική χρονοσειρά πρόβλεψης της μεθοδολογίας 2.



Εικόνα 4.16-Στάδια μεθοδολογίας Νο2

Μεθοδολογία Νο3

Εδώ ξεκινάμε πάλι με τα δεδομένα μηνιαίας ζήτησης για κάθε ταχυδρομικό κώδικα. Αφού πρώτα τα καθαρίσουμε (βήμα 1 και 2) από ακραίες και κενές τιμές, τα αθροίζουμε ανά μήνα για όλες τις διαθέσιμες χρονιές(2007-2013) και σχηματίζουμε τη χρονοσειρά συνολικής ζήτησης που αφορά όλες τις διαθέσιμες περιοχές των δεδομένων μας. Σε αυτήν την χρονοσειρά εφαρμόζουμε το βήμα 4 και κάνουμε προβλέψεις με κάθε ένα μοντέλο πρόβλεψης από τα συνολικά έξι που εξετάζουμε. Η βέλτιστη μέθοδος πρόβλεψης θα δώσει κάποιες τιμές για το διάστημα B και αυτές είναι οι τελικές προβλέψεις της μεθοδολογίας 3.

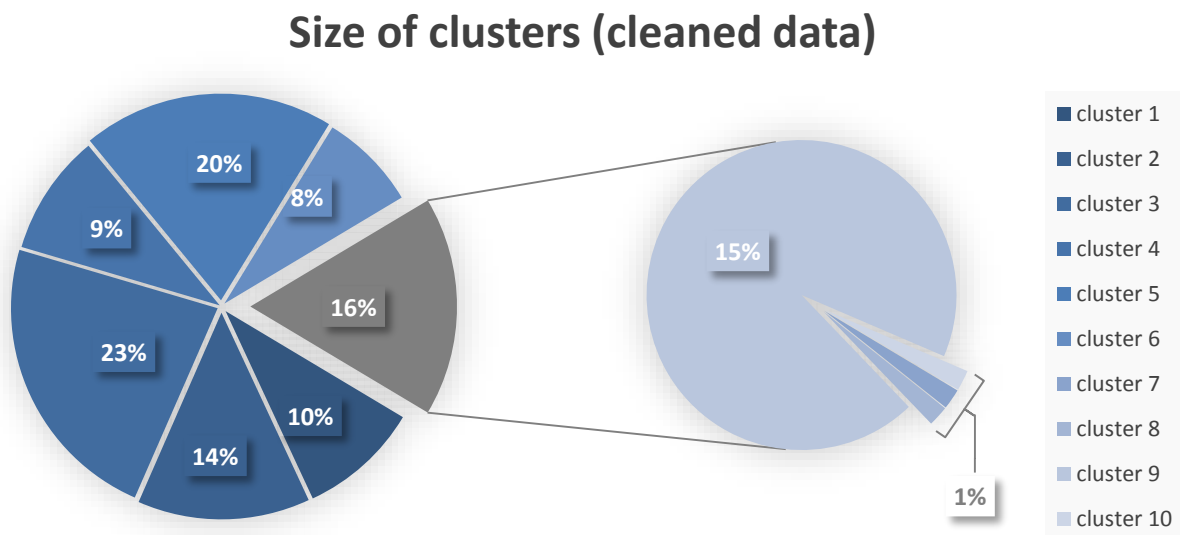
Επιλογή βέλτιστης μεθοδολογίας

Αφού ανακτήσουμε από τις τρεις παραπάνω μεθοδολογίες τις τελικές προβλέψεις τους για το χρονικό διάστημα B, τις συγκρίνουμε ως προς την ακρίβεια τους με τα πραγματικά δεδομένα σε αυτό το διάστημα. Η μεθοδολογία που δίνει το ελάχιστο σφάλμα MAPE είναι και η προτιμότερη.

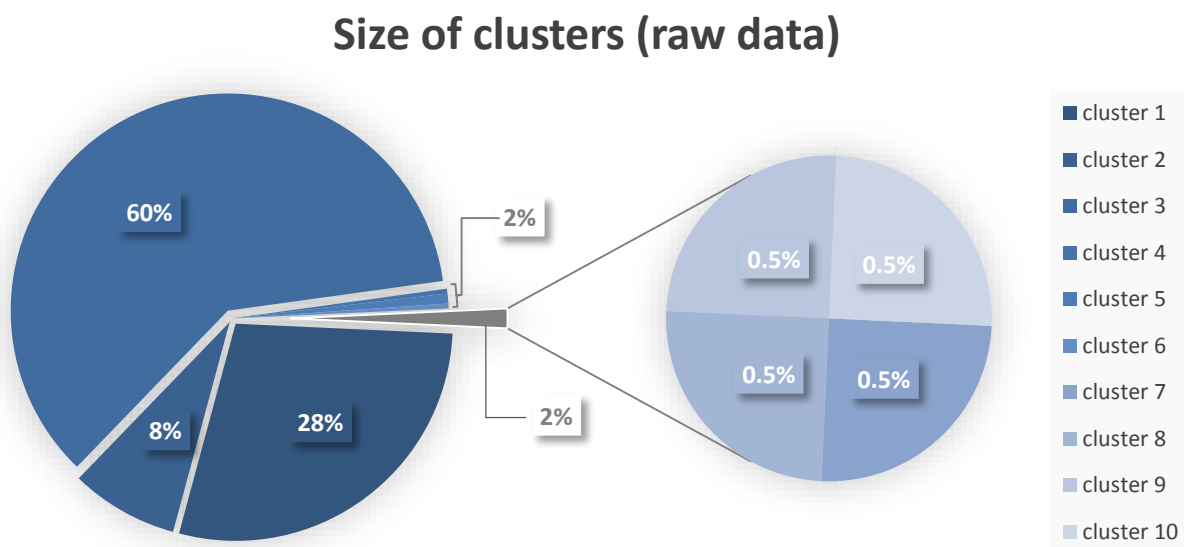
5 Παρουσίαση Αποτελεσμάτων

5.1 Αποτελέσματα Ανάλυσης cluster

Σε αυτήν την ενότητα θα παρουσιάσουμε τα αποτελέσματα της ανάλυσης cluster που περιγράφηκε στο προηγούμενο κεφάλαιο.



Εικόνα 5.1-Μεγέθη cluster μετά τον «καθαρισμό» των δεδομένων από ακραίες και κενές τιμές



Εικόνα 5.2-Μεγέθη cluster χωρίς τον «καθαρισμό» των δεδομένων από ακραίες και κενές τιμές

Παρατηρούμε ότι μετά τον καθαρισμό των δεδομένων *ζήτησης/μετρητή* από τυχόν παρούσες ακραίες και κενές τιμές (βλ. ενότητα 4.3.1), η ανάλυση cluster δίνει μία πιο ομοιόμορφη κατανομή των χρονοσειρών στα δέκα cluster. Σε αντίθεση, η ανάλυση cluster στα αρχικά δεδομένα που δεν έχουν φιλτραριστεί με τις μεθόδους που εξηγήσαμε, δίνουν μία διαφορετική κατανομή των χρονοσειρών στα cluster, κατά πολύ λιγότερο ομοιόμορφη. Όπως θα φανεί και πιο κάτω, αυτό έχει αντίκτυπο και στις προβλέψεις που γίνονται κατά την μεθοδολογία 1, οι οποίες είναι ακριβέστερες στην περίπτωση της κατανομής των cluster όπως στην εικόνα 5.1 σε σχέση με τις αντίστοιχες προβλέψεις όταν η κατανομή των cluster είναι αυτή της εικόνας 5.2.

Ειδικότερα, είναι:

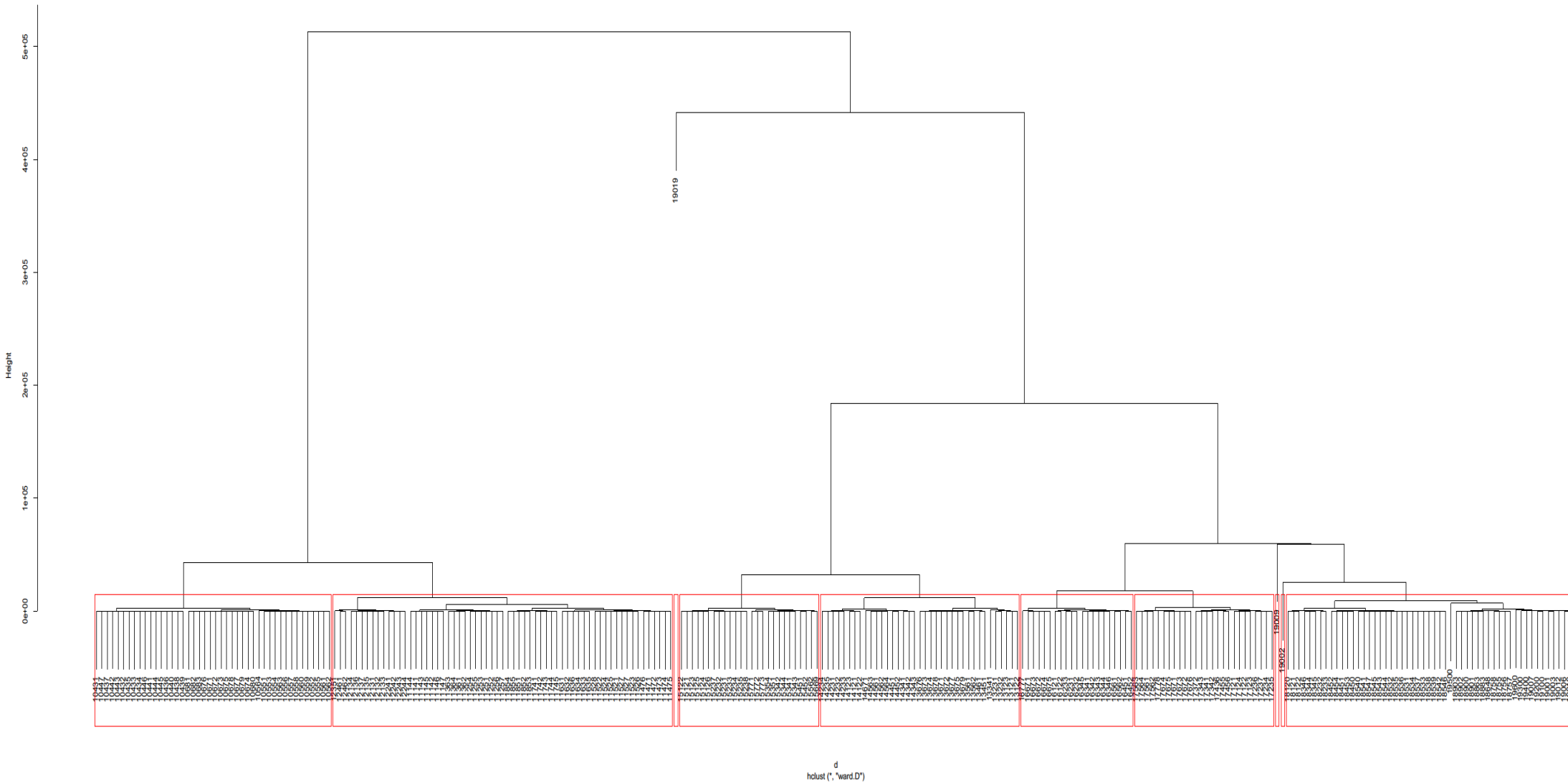
Cluster	Μέγεθος cluster(%) (raw data)	Μέγεθος cluster(%) (cleaned data)
1 ^ο	28%	10%
2 ^ο	8%	14%
3 ^ο	60%	23%
4 ^ο	0.5%	9%
5 ^ο	1%	20%
6 ^ο	0.5%	8%
7 ^ο	0.5%	0.33%
8 ^ο	0.5%	0.33%
9 ^ο	0.5%	15%
10 ^ο	0.5%	0.33%

Πίνακας 5.1-Ποσοστιαία κατανομή των 274 χρονοσειρών ανά cluster

Γίνεται επομένως κατανοητό και με πειραματικό τρόπο η σημαντικότητα του εντοπισμού ακραίων και κενών τιμών από τα αρχικά δεδομένα και η αντικατάστασή τους με μεθόδους που έχει αποδειχθεί ότι δίνουν ικανοποιητικά αποτελέσματα.

Στην επόμενη σελίδα θα παρουσιάσουμε το δενδρόγραμμα που παράχθηκε χρησιμοποιώντας τη συνάρτηση `hclust` της γλώσσας προγραμματισμού R μέσω της μεθόδου του Ward (ενότητα 4.2.3.4). Τα κόκκινα πλαίσια είναι τα διαφορετικά cluster που σχηματίζονται και περικλείουν τους αντίστοιχους ταχυδρομικούς κώδικες που υπάρχουν στην βάση του γραφήματος.

Cluster Dendrogram

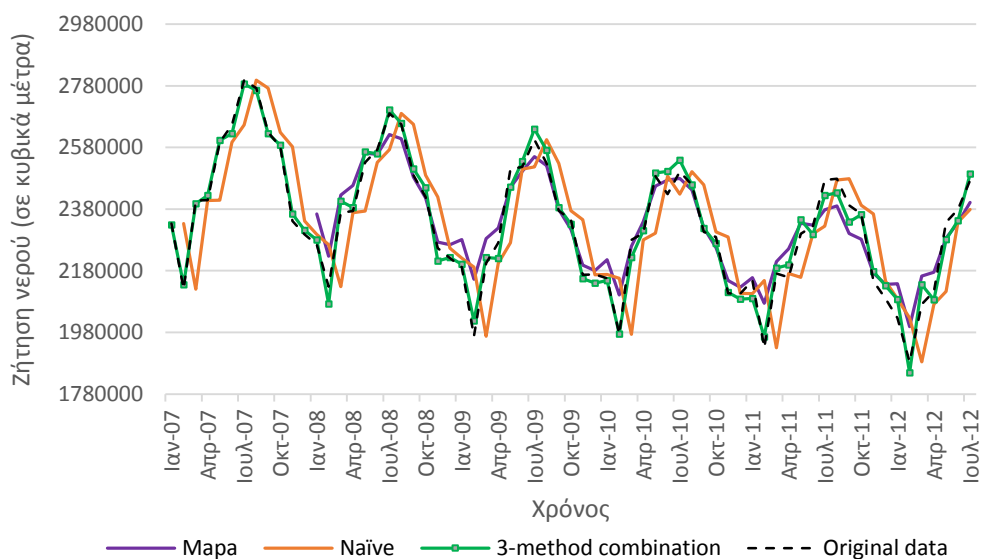


Εικόνα 5.3-Παραγόμενο δενδρόγραμμα καταμερισμού των cluster

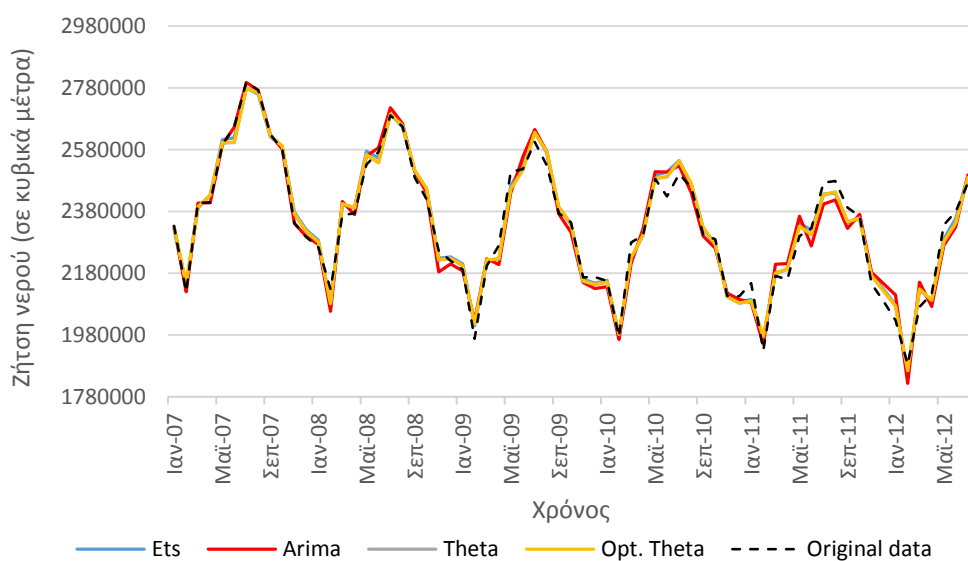
5.2 Αποτελέσματα Μεθοδολογιών

5.2.1 Αποτελέσματα Μεθοδολογίας Νο1

Εδώ θα παρουσιάσουμε τα αποτελέσματα της μεθοδολογίας 1 που περιγράφηκε στην ενότητα 4.3.1.

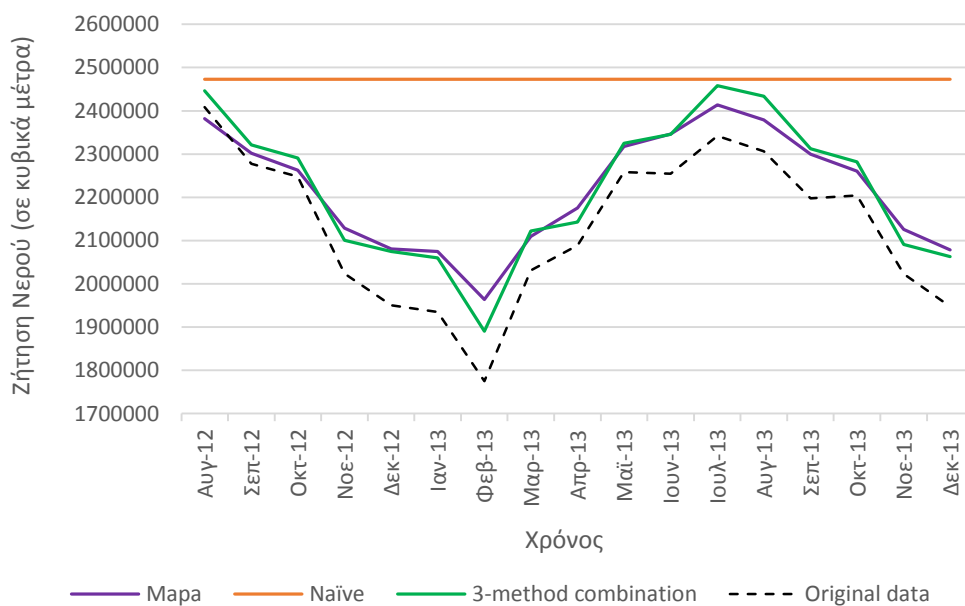


Εικόνα 5.4-Παρουσίαση γραφικής επίδοσης των μεθόδων Mapa,Naive, Συνδυασμένης μεθόδου στο διάστημα A για το cluster 1

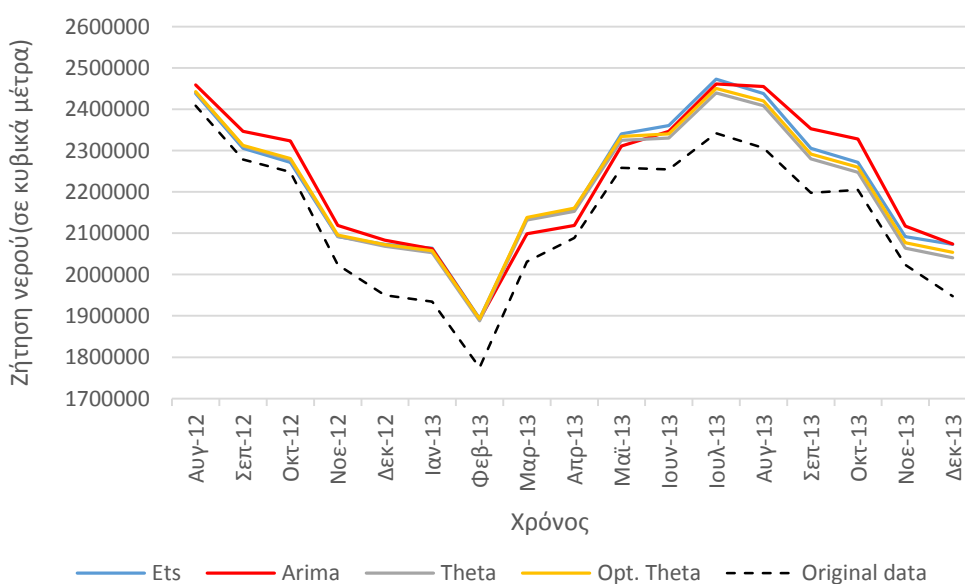


Εικόνα 5.5-Παρουσίαση γραφικής επίδοσης των μεθόδων Arima, Ets, Theta και Opt.Theta στο διάστημα A για το cluster 1

Οι παραπάνω γραφικές παραστάσεις αφορούν την προσαρμογή των μοντέλων Naive, Arima, Ets, MAPA, Theta, optimized Theta και συνδυασμένες μέθοδοι στο διάστημα A (in-sample forecasting) για τα δεδομένα κατανάλωσης του πρώτου cluster όπως φαίνεται από την εικόνα 5.1. Είναι το διάστημα στο οποίο «εκπαιδεύονται» οι μέθοδοι από την πραγματική χρονοσειρά ώστε να δώσουν όσο το δυνατόν καλύτερες προβλέψεις. Από τα γραφήματα καταλαβαίνουμε ότι τα μοντέλα Arima, Ets, Theta, optimized Theta και συνδυασμένες μέθοδοι αποδίδουν πάρα πολύ καλά. Από την άλλη ή μέθοδος MAPA φαίνεται να δυσκολεύεται περισσότερο να προσαρμοστεί στα δεδομένα σε σχέση με τις άλλες μεθόδους.

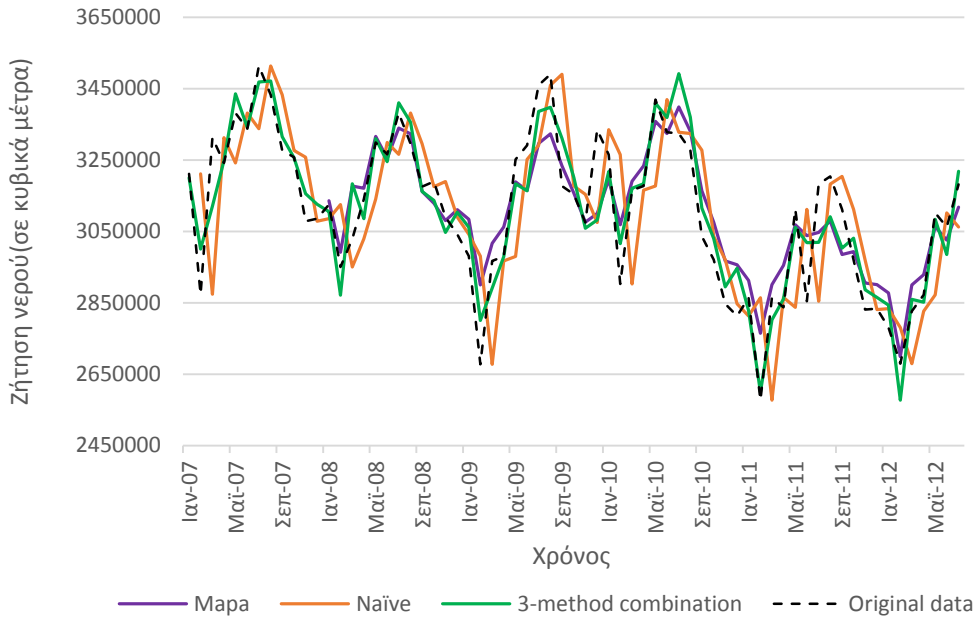


Εικόνα 5.6-Παραγωγή προβλέψεων στο διάστημα B για το πρώτο cluster

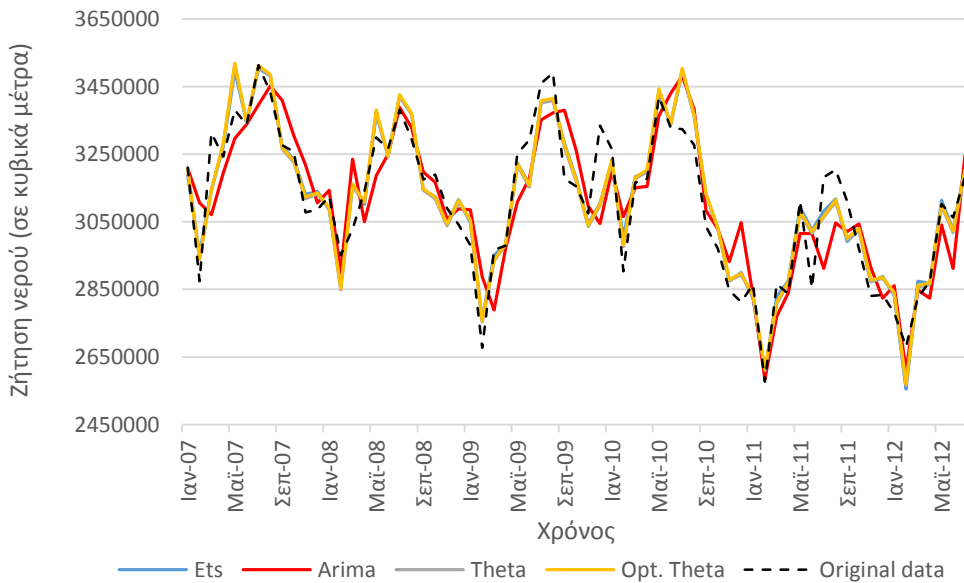


Εικόνα 5.7-Παραγωγή προβλέψεων για τις υπόλοιπες μεθόδους για το cluster 1

Οι γραφικές παραστάσεις των παραπάνω εικόνων (5.6 και 5.7) αφορούν τις προβλέψεις που κάνουμε στο διάστημα B(out-of-sample forecasting). Σύμφωνα με την μεθοδολογία No1 που παρουσιάστηκε, η βέλτιστη μέθοδος πρόβλεψης(σύμφωνα με το κριτήριο MAPE στο διάστημα B) από τις προβλέψεις που υπάρχουν στις δύο παραπάνω εικόνες θα είναι και αυτή που θα χρησιμοποιηθεί σαν την επικρατέστερη πρόβλεψη. Βλέποντας τις προαναφερθείσες εικόνες βλέπουμε ότι οι μέθοδοι Ets, Arima, Theta και συνδυασμένοι μέθοδοι ότι αποδίδουν αρκετά καλά σε σχέση με τις υπόλοιπες.



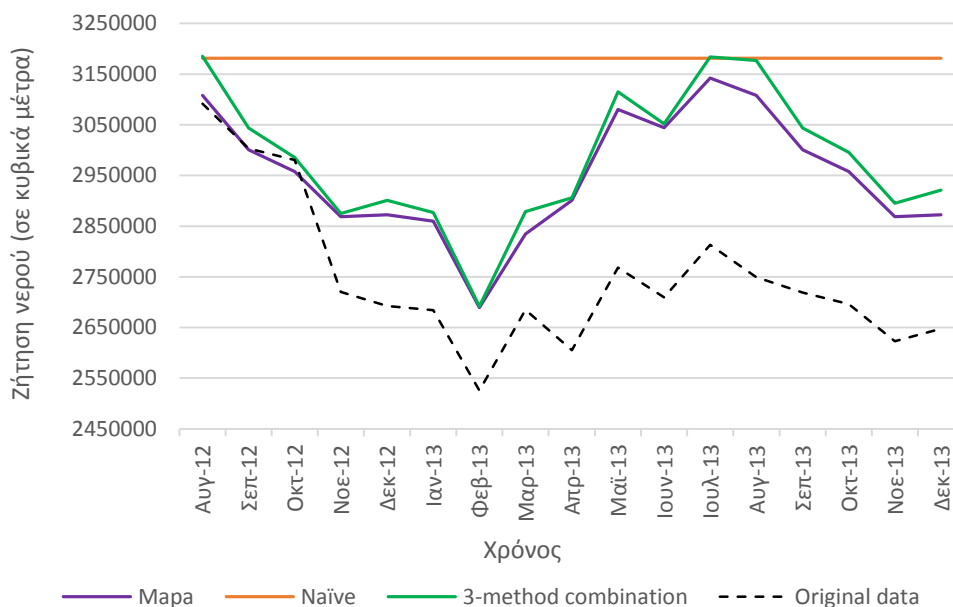
Εικόνα 5.8-Προσαρμογή των μοντέλων πρόβλεψης στο διάστημα A για το cluster 5



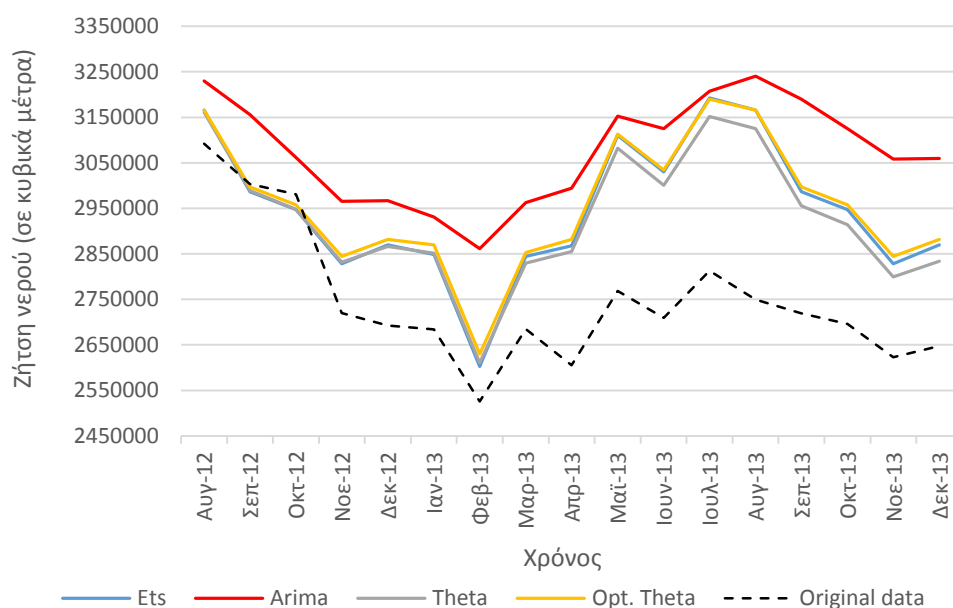
Εικόνα 5.9-Προσαρμογή των υπόλοιπων μοντέλων πρόβλεψης στο διάστημα A για το cluster 5

Στις εικόνες 5.10 και 5.11 έχουμε τις γραφικές παραστάσεις της προσαρμογής των μεθόδων πρόβλεψης για το 5^ο cluster στο διάστημα A (in-sample forecasting). Από τα γραφήματα καταλαβαίνουμε ότι τα μοντέλα Arima, Ets, theta, optimized Theta και συνδυασμένες μέθοδοι αποδίδουν πάρα πολύ καλά. Από την άλλη η μέθοδος MAPA φαίνεται να δυσκολεύεται περισσότερο να προσαρμοστεί στα δεδομένα σε σχέση με τις άλλες μεθόδους.

Οι επόμενες γραφικές παραστάσεις που θα παρουσιαστούν αφορούν την εφαρμογή των επτά μεθόδων πρόβλεψης που εξετάζουμε στο διάστημα B (out-of-sample forecasting) για το 5^ο cluster.



Εικόνα 5.10-Παραγωγή προβλέψεων στο διάστημα B για το πέμπτο cluster



Εικόνα 5.11-Παραγωγή των υπόλοιπων προβλέψεων στο διάστημα B για το πέμπτο cluster

Παρατηρούμε ότι καλύτερες προβλέψεις δίνουν οι μέθοδοι Ets, Arima και Theta.

Μέχρις στιγμής έχουμε δείξει τις χρονοσειρές ζήτησης νερού για το πρώτο και πέμπτο cluster(το 30% από τις 274 χρονοσειρές ζήτησης ανήκουν σε αυτά τα δύο cluster) της μεθοδολογίας No1 και πως σε αυτές αποτυπώνονται οι μέθοδοι πρόβλεψης στο διάστημα A και B.

Παρακάτω θα παρουσιάσουμε τον πίνακα σφαλμάτων για κάθε cluster και για όλες τις μεθόδους πρόβλεψης και αφορούν το διάστημα A. Θυμίζουμε ότι για κάθε cluster, όποια μέθοδος δίνει μικρότερο σφάλμα για το κριτήριο MAPE στο διάστημα A, είναι η βέλτιστη μέθοδος η οποία τελικά είναι αυτή που θα δώσει την πρόβλεψη στο διάστημα B.

	Ets MAPE	Arima MAPE	Theta MAPE	Opt.Theta MAPE	MAPA MAPE	Naive MAPE	Comb.Method MAPE
cluster 1	1.101098093	1.364149661	1.101620259	1.108093331	2.423761534	4.743916415	1.09948821
cluster 2	1.324262849	3.167414557	1.29444236	1.293313207	2.520219376	4.843011294	1.297876886
cluster 3	0.752452572	1.94001006	0.72709939	0.733983798	1.220619954	3.862968319	0.731606929
cluster 4	0.903472326	3.520518648	0.869634946	0.871821511	1.59167985	3.868653794	0.881012825
cluster 5	1.933262197	2.907127569	1.933750333	1.95205738	2.590581178	4.335907324	1.938036304
cluster 6	1.431586311	4.119229479	1.433156003	1.431200801	3.518971017	6.098044795	1.425608148
cluster 7	18.0495794	9.269619538	216.6362962	17.62216919	151.8786926	17.88670518	8.906945748
cluster 8	14.4324575	12.57972944	16.93333448	16.1463706	20.34086404	16.4546447	13.31417003
cluster 9	0.760963186	2.876375727	2.548727695	2.628205713	1.102952177	3.63881424	1.30183955
cluster 10	10.09368933	7.547723465	9.933006151	9.530599005	12.69094437	18.40624997	8.78298364

Πίνακας 5.2-Πίνακας σφαλμάτων MAPE(%) για κάθε cluster και για όλες τις εξεταζόμενες μεθόδους στο διάστημα A

Από τον παραπάνω πίνακα(βλ. Π-5 του παραρτήματος για πληρότητα) παρατηρούμε ότι η μέθοδος με τις καλύτερες προσαρμογές στο διάστημα A είναι η συνδυαστική μέθοδος των τριών καλύτερων μεθόδων πρόβλεψης που παρουσιάσαμε στο κεφάλαιο 3, η οποία στη περίπτωση του cluster 7 βελτιώνει την επίδοση της προσαρμογής της καλύτερης μεμονωμένης μεθόδου(Arima) έως περίπου 4%. Ακόμα παρατηρούμε ότι η κλασική μέθοδος Θ αποδίδει πολύ καλά σε εποχιακά δεδομένα όπως και η οικογένεια των μεθόδων εκθετικής εξομάλυνσης. Επιπλέον βλέπουμε ότι για τις δέκα αυτές χρονοσειρές, η βελτιστοποιημένη μέθοδος Θ (ενότητα 3.5.6) στις μισές περιπτώσεις βελτιώνει την απόδοση της κλασικής μεθόδου Theta αλλά σε μία μόνο περίπτωση έχει την καλύτερη επίδοση όλων (cluster 2) σε αντίθεση με την κλασική μέθοδο Theta που διακρίνεται σε δύο περιπτώσεις (cluster 3 και 4). Από την άλλη η μέθοδος Arima φαίνεται να έχει καλύτερες επιδόσεις σε χρονοσειρές με εμφανή στασιμότητα(cluster 8) αντί για χρονοσειρές με εμφανή εποχικότητα (π.χ. cluster 1).

Συναθροίζοντας τις βέλτιστες προβλέψεις για κάθε cluster με την βοήθεια του πακέτου hts για την γλώσσα R και συγκρίνοντας την συναθροισμένη χρονοσειρά με τα

πραγματικά διαθέσιμα δεδομένα στο διάστημα B καταλήγουμε στο τελικό σφάλμα της μεθόδου 1:

Σφάλμα μεθοδολογίας No1 (cleaned data)

ME	RMSE	MAE	MPE	MAPE
433441.7	534578.5	495661.4	2.577791	2.791232

Στον παραπάνω πίνακα παρουσιάζονται και τα σφάλμα ME(Μέσο σφάλμα), RMSE(Ρίζα μέσου τετραγωνικού σφάλματος), MAE(Μέσο απόλυτο σφάλμα), MPE(Μέσο ποσοστιαίο σφάλμα) για πληρότητα αλλά σαν κριτήριο καλύτερης επίδοσης μεταξύ των τριών μεθοδολογιών που εξετάζει η διπλωματική είναι το σφάλμα MAPE.

Για την περίπτωση που τα cluster μας απεικονίζονταν από την εικόνα 5.1.2 δηλαδή προέρχονταν από τα αρχικά πρωτόλεια και αφιλτράριστα δεδομένα, τότε ακολουθώντας τα ίδια βήματα με πριν, το τελικό σφάλμα της μεθοδολογίας 1 είναι:

Σφάλμα μεθοδολογίας No1 (raw data)

ME	RMSE	MAE	MPE	MAPE
493971.8	572798.9	524439.5	2.939331	3.111149

Παρατηρούμε ότι στην πρώτη περίπτωση όπου έχουμε χρησιμοποιήσει τα δεδομένα $\frac{\text{ζήτησης}}{\text{μετρητή}}$ για τον καθορισμό των cluster, αφού πρώτα έχουμε απομακρύνει ακραίες και κενές τιμές, το σφάλμα MAPE είναι *βελτιωμένο* κατά περίπου 11,5% σε σχέση με την χρησιμοποίηση των πρωτόλειων δεδομένων $\frac{\text{ζήτησης}}{\text{μετρητή}}$.

5.2.2 Αποτελέσματα Μεθοδολογίας No2

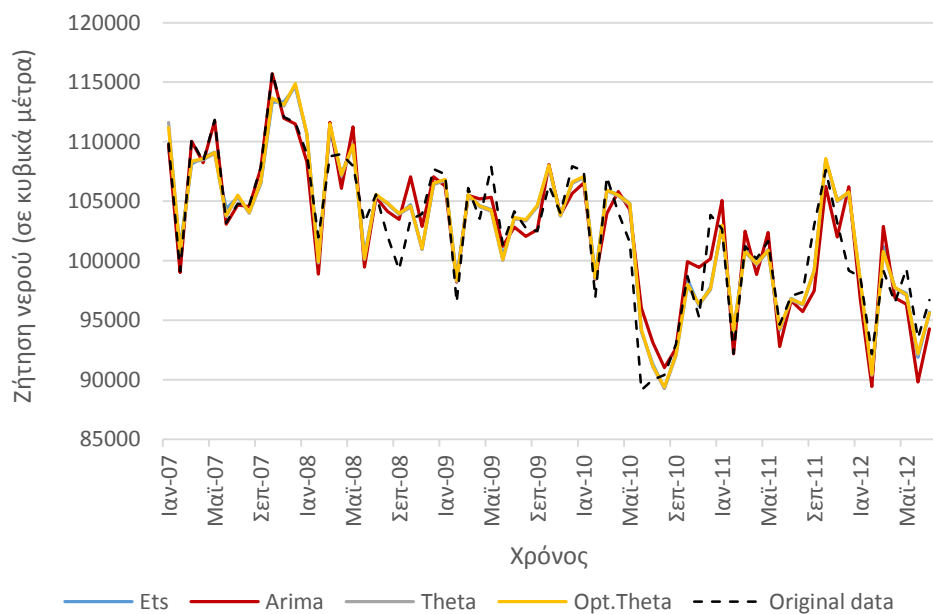
Ξεκινώντας από τα δεδομένα ζήτησης για κάθε ένα διαθέσιμο ταχυδρομικό κώδικα παράγουμε προβλέψεις για αυτούς τους ταχυδρομικούς κώδικες και αφού κρατήσουμε τις βέλτιστες μεθόδους για κάθε μία περίπτωση, τις συναθροίζουμε ώστε να βρούμε την τελική πρόβλεψη της μεθοδολογίας 2. Στις παρακάτω εικόνες θα δείξουμε ενδεικτικά κάποιες προσαρμογές των μοντέλων σε καταναλώσεις διαφόρων ταχυδρομικών κωδικών των δεδομένων μας.

Έπειτα, παρουσιάζουμε τις προσαρμογές των μεθόδων προβλέψεων στον ταχυδρομικό κώδικα 15771(Ζωγράφου) για τα διαστήματα A και B. Παρατηρούμε ότι στην εικόνα 5.12 τα μοντέλα MAPA και συνδυασμένες μέθοδοι έχουν πολύ καλή προσαρμογή στο διάστημα A όπως και τα μοντέλα Theta και Arima. Να σημειώσουμε ότι η

βελτιστοποιημένη μέθοδος Θ βελτιώνει την απόδοση της κλασικής μεθόδους για αυτήν την χρονοσειρά κατά περίπου 0,6%.

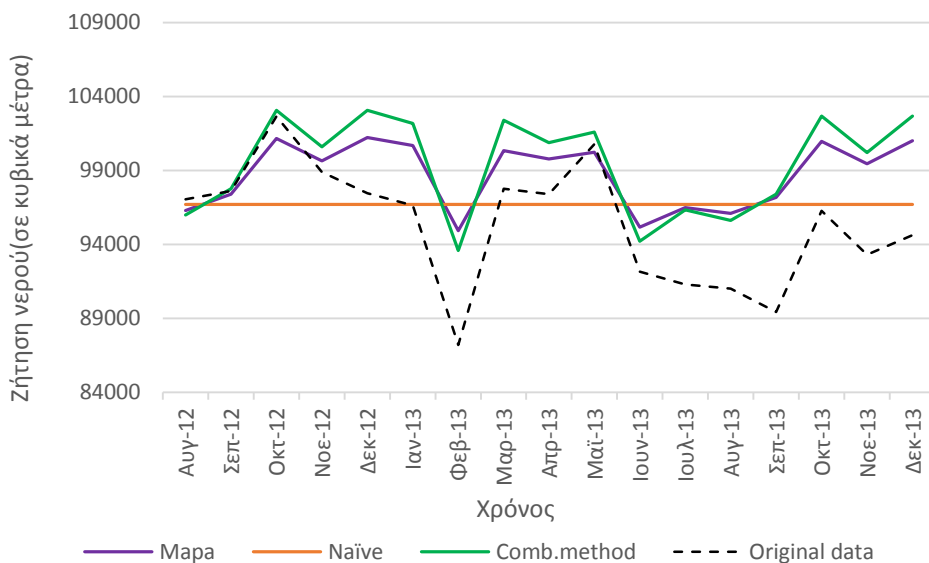


Εικόνα 5.12-Προσαρμογή των μεθόδων πρόβλεψης για τον ταχ.κώδικα 15771,Ζωγράφου

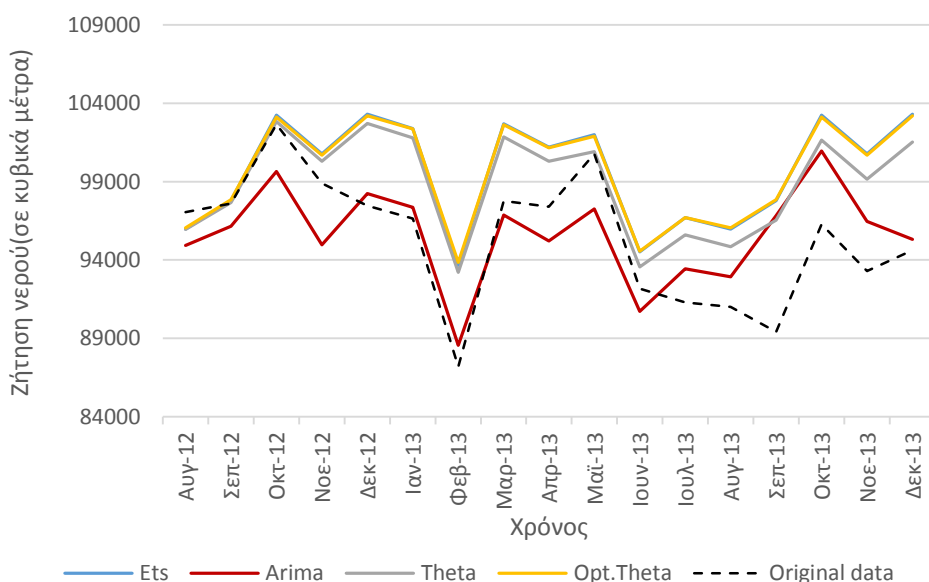


Εικόνα 5.13-Προσαρμογή των υπόλοιπων μεθόδων πρόβλεψης για τον ταχ.κώδικα 15771,Ζωγράφου

Επομένως τα μοντέλα πρόβλεψης αποδεικνύεται ότι μπορούν να προσαρμοσθούν σε χρονοσειρές όπως η παραπάνω που έχει εμφανής διακυμάνσεις και τυχαιότητα.



Εικόνα 5.14-Προσαρμογή των μεθόδων για το διάστημα B

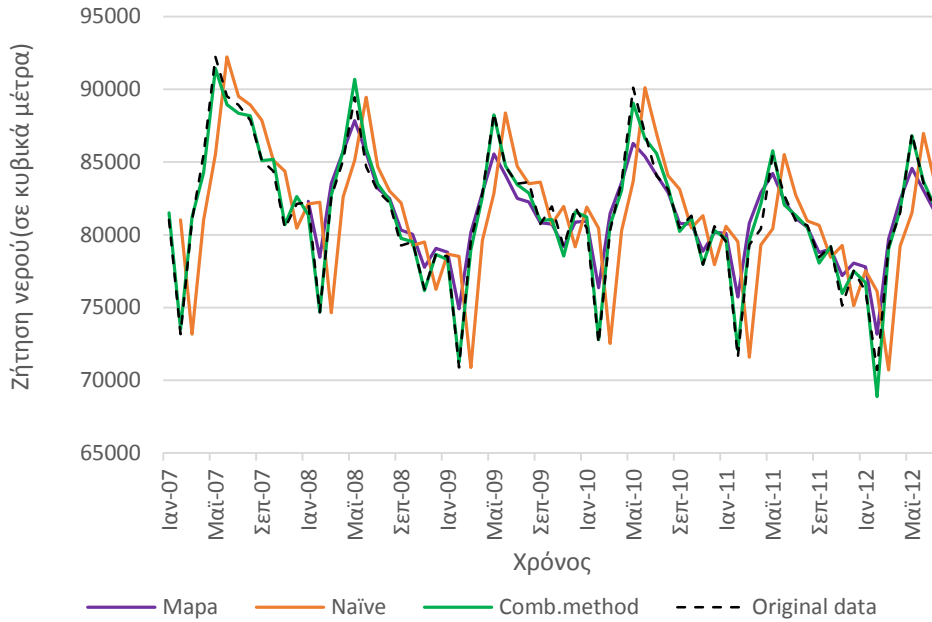


Εικόνα 5.15-Προσαρμογή των μεθόδων για το διάστημα B

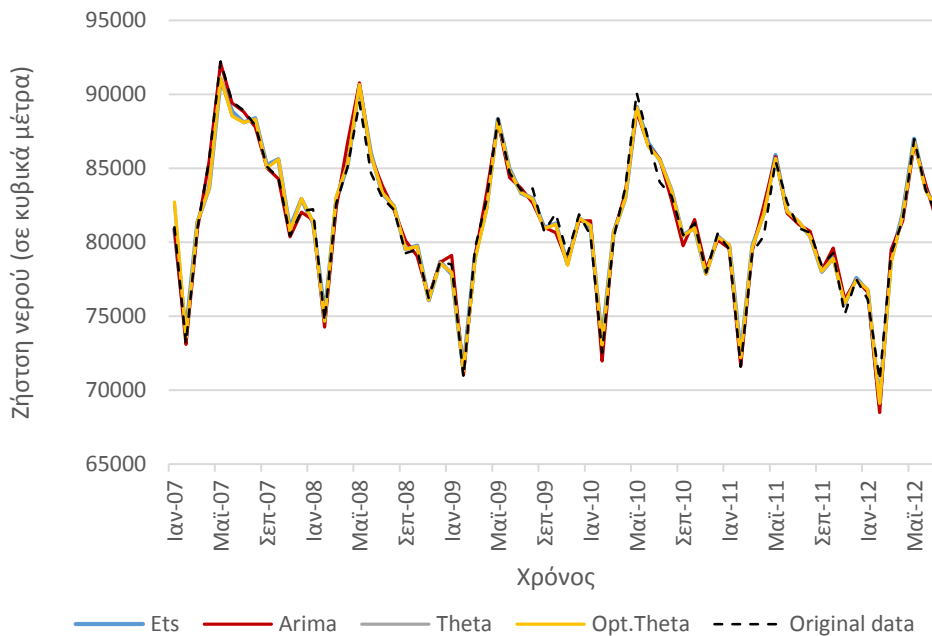
Από τις παραπάνω γραφικές παραστάσεις παρατηρούμε ότι τα μοντέλα Ets(A,A,N), Theta, Mapa και συνδυασμένες μέθοδοι αποδίδουν με παρόμοιο τρόπο στο διάστημα B. Το μοντέλο Arima όμως προσαρμόζεται καλύτερα στις διακυμάνσεις των δεδομένων και διαχειρίζεται πιο εύκολα τις τοπικές καμπυλότητες αυτών. Να σημειώσουμε ότι η κλασική μέθοδος Θ δίνει ακριβέστερες προβλέψεις από την βελτιστοποιημένη μέθοδο της στο αναφερόμενο διάστημα B, παρατηρούμε δηλαδή ότι παρόλο που η βελτιστοποιημένη μέθοδος Θ δίνει καλύτερη προσαρμογή στο διάστημα A, εντούτοις στο διάστημα B η κλασική Theta δίνει καλύτερα αποτελέσματα. Αυτό επιβεβαιώνει την αντίληψη ότι αν ένα μοντέλο προσαρμόζεται καλύτερα στα ιστορικά δεδομένα δεν

σημαίνει απαραίτητα ότι θα δίνει τις καλύτερες προβλέψεις για αυτά σε σύγκριση με οποιαδήποτε άλλη υλοποίηση.

Στην επόμενη σελίδα δίνονται οι γραφικές παραστάσεις ενός ακόμα ταχυδρομικού κώδικα της Αττικής(13652) που εντοπίζεται στην περιοχή των Αγίων Αναργύρων.



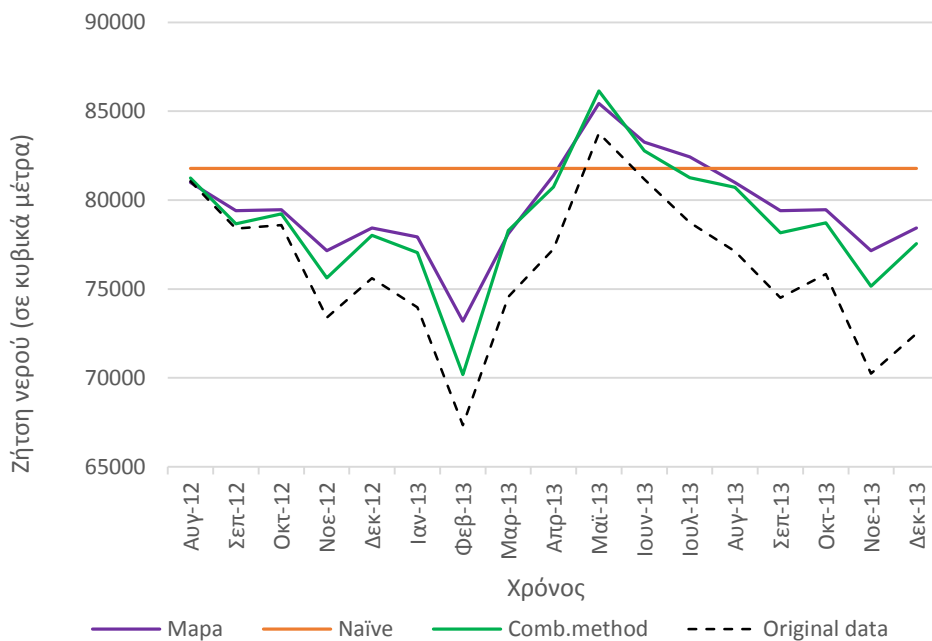
Εικόνα 5.16-Προσαρμογή των μεθόδων στο διάστημα A



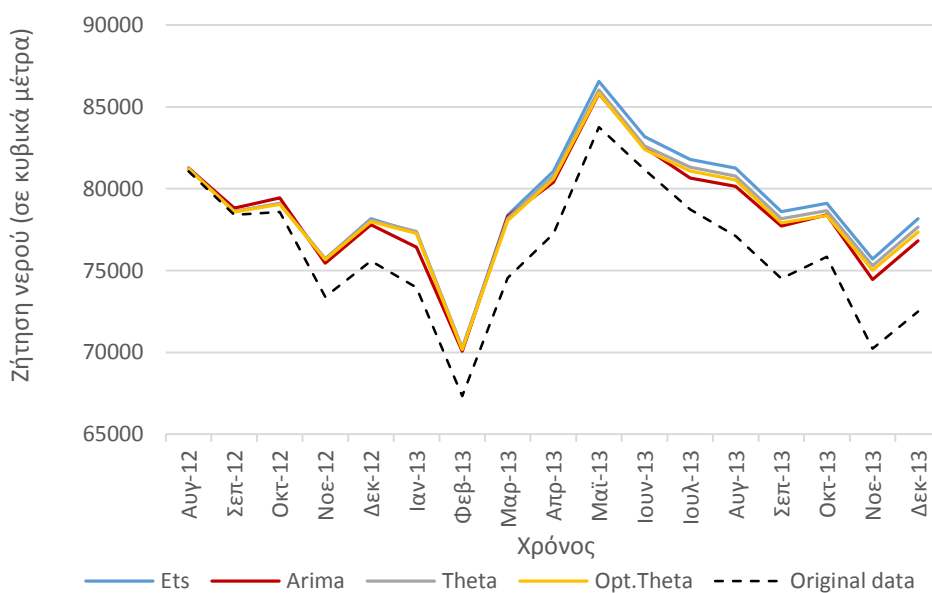
Εικόνα 5.17-Προσαρμογή των υπόλοιπων μεθόδων στο διάστημα A

Στο παραπάνω πλήρως εποχικό μοτίβο βλέπουμε ότι όλες οι μέθοδοι αποδίδουν πολύ καλά αποτελέσματα με σφάλματα της τάξεως <1% με εξαίρεση την μέθοδο Mara που δείχνει να δυσκολεύεται λίγο περισσότερο σε σχέση με τις άλλες μεθόδους. Σε αυτήν την χρονοσειρά όμως ξεχωρίζει η προσαρμογή της μεθόδου συνδυασμού των τριών βέλτιστων μεθόδων(combined methods) η οποία δίνει σχεδόν τέλεια προσαρμογή στα δεδομένα του διαστήματος A με σφάλμα Mare ίσο με 0,023%.

Ύστερα παρουσιάζονται οι γραφικές παραστάσεις για τις προβλέψεις των μεθόδων στο διάστημα B οι οποίες είναι ικανοποιητικές από άποψη ακρίβειας.



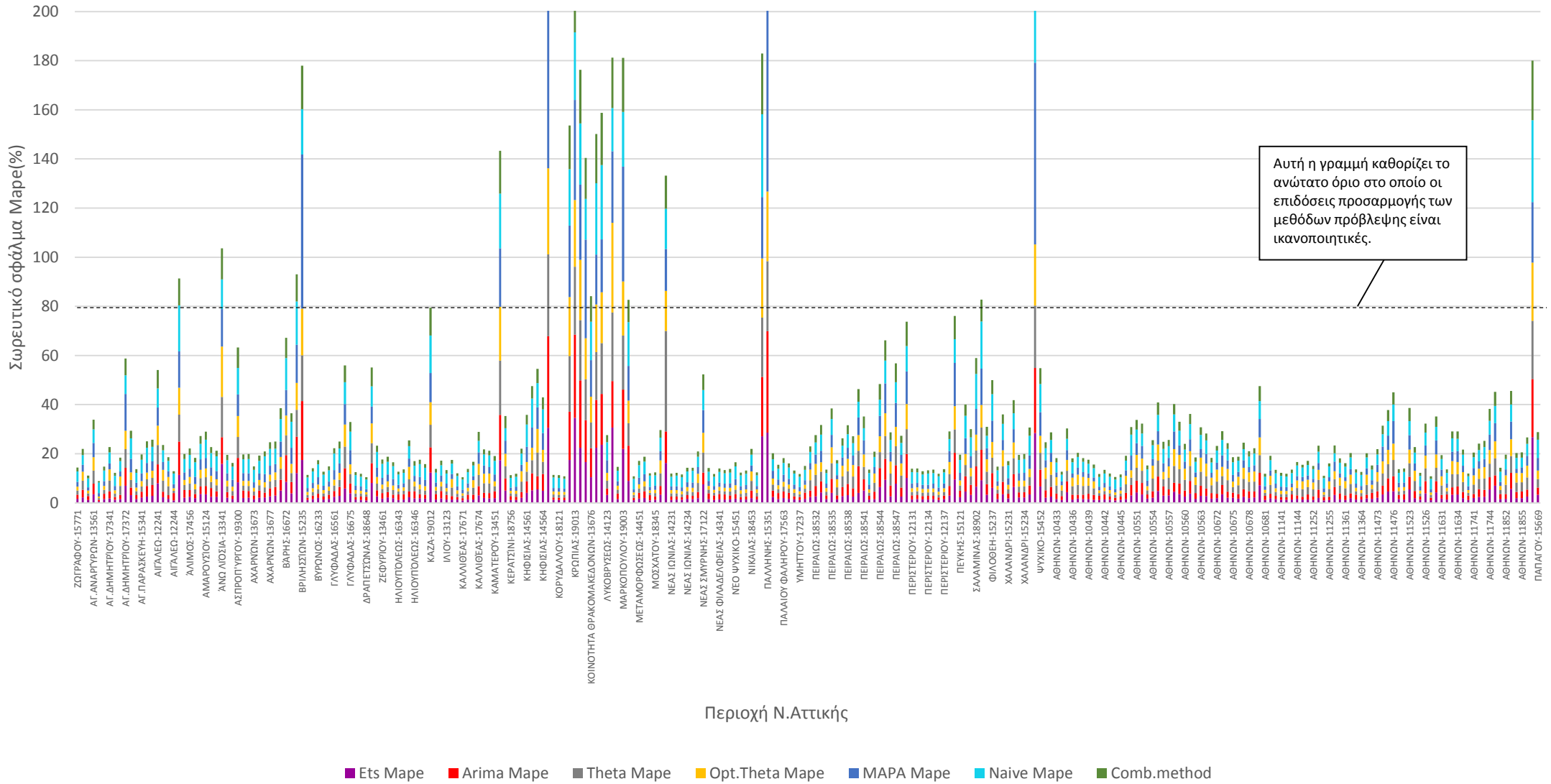
Εικόνα 5.18-Προβλέψεις των μεθόδων στο διάστημα B



Εικόνα 5.19-Προβλέψεις των υπόλοιπων μεθόδων στο διάστημα B

Περιοχή	Ets MAPE	Arima MAPE	Theta MAPE	Opt.Theta MAPE	MAPA MAPE	Naive MAPE	Comb. Method MAPE
ΚΑΛΛΙΘΕΑΣ-17672	1.402719597	1.976076913	1.378729929	1.377962142	1.766115144	4.568370533	1.37731560
ΚΑΛΛΙΘΕΑΣ-17673	1.40020235	1.912698054	2.975583315	2.934199421	1.704164906	4.277753448	1.535445663
ΚΑΛΛΙΘΕΑΣ-17674	3.03338440	3.790334878	4.524392666	4.527258582	4.72967739	4.685648942	3.574477539
ΚΑΛΛΙΘΕΑΣ-17675	2.05281056	2.067575893	3.863934957	3.874740967	3.347083236	4.536632662	2.085273757
ΚΑΛΛΙΘΕΑΣ-17676	1.872153365	2.104859444	3.985387219	3.982172343	2.67334722	4.719261373	1.904990104
ΚΑΜΑΤΕΡΟΥ-13451	1.92748670	2.853156251	1.992724881	1.987991804	2.825790236	5.544139142	1.954009136
ΚΕΡΑΤΕΑΣ-19001	17.3483382	18.50696048	22.07806868	21.94772801	23.7737019	22.27872555	17.38230292
ΚΕΡΑΤΣΙΝΙ-18755	5.020770647	4.991216165	4.983932597	5.002383239	5.304064458	5.063366499	4.99027010
ΚΕΡΑΤΣΙΝΙ-18756	1.138926671	1.628712691	1.09472457	1.100699821	1.296020748	3.950018574	1.109050513
ΚΕΡΑΤΣΙΝΙ-18757	1.214427792	1.245922244	1.215511549	1.21285148	1.68246984	4.063910282	1.214015476
ΚΕΡΑΤΣΙΝΙ-18758	2.107748669	2.300355459	4.250699532	4.250733859	2.57136028	4.582984574	2.01354835
ΚΗΦΙΣΙΑΣ-14561	4.063656463	4.699364587	3.926615993	3.926615993	6.228765468	9.060473828	3.91250116
ΚΗΦΙΣΙΑΣ-14562	5.079956561	7.000648455	5.263288609	5.098993026	8.232339853	11.82945092	5.05579553
ΚΗΦΙΣΙΑΣ-14563	5.52026541	5.33190206	9.629016315	9.632097639	8.828224621	9.782552467	5.801701678
ΚΗΦΙΣΙΑΣ-14564	4.885525015	6.875022871	4.923499053	4.906048564	6.868276734	9.654494329	4.86167454
ΚΗΦΙΣΙΑΣ-14671	30.54591024	37.31342241	33.33585606	35.00786157	154.233016	31.46210408	30.0646586
ΚΟΥΡΥΔΑΛΛΟΥ-18120	1.257201662	0.95542141	1.208833065	1.205688598	1.466052993	4.258548595	1.059376769
ΚΟΥΡΥΔΑΛΛΟΥ-18121	1.201074556	0.95458452	1.198881903	1.198597479	1.517911189	4.148088253	1.083094135
ΚΟΥΡΥΔΑΛΛΟΥ-18122	1.121190798	1.080885151	1.108748172	1.107511825	1.40276823	3.952166261	1.05355329
ΚΡΩΠΙΑΣ-19010	17.4544487	19.69676167	22.61494528	23.97649938	29.10266884	22.97413642	17.81840367
ΚΡΩΠΙΑΣ-19013	34.59159053	33.92504425	27.6158142	27.15098641	40.69675715	27.56188314	26.8429613
ΚΡΩΠΙΑΣ-19400	22.4768788	27.39730006	24.51884527	24.5362389	30.62144019	24.90585177	21.7986426
ΚΟΙΝΟΤΗΤΑ ΓΛΥΚΩΝ ΝΕΡΩΝ-15354	16.66043726	17.02426568	16.68170667	16.66043711	40.1767434	16.5071954	16.66477297
ΚΟΙΝΟΤΗΤΑ ΘΡΑΚΟΜΑΚΕΔΟΝΩΝ-13676	10.49452357	11.71507053	10.4864449	10.49635652	14.94272315	15.63225119	10.3746785

Πίνακας 5.3-Παρουσίαση σφαλμάτων Μape στο διάστημα Α τυχαία επιλεγμένων περιοχών της Αττικής για όλες τις μεθόδους



Εικόνα 5.20-Σωρευτική απεικόνιση σφαλμάτων MAPE όλων των μεθόδων για όλες τις περιοχές της Αττικής

Ο πίνακας 5.3 απεικονίζει ένα μικρό μέρος των σφαλμάτων του διαστήματος A για τις διάφορες περιοχές της Αττικής που είχαμε διαθέσιμες στη διπλωματική μας. Ο πλήρης πίνακας των δεδομένων μπορεί να βρεθεί στο τέλος της διπλωματικής στα παραρτήματα (βλ.πίνακα Π-6). Με έντονο χρώμα σημειώνονται οι μέθοδοι που δίνουν το μικρότερο σφάλμα *Μαρε* ανά περιοχή. Παρατηρούμε ότι τη μερίδα του λέοντος ως προς την ακρίβεια των προσαρμογών έχει το μοντέλο του συνδυασμού των τριών βέλτιστων μεθόδων (3-method combination) με καλύτερη επίδοση στο 46% των περιπτώσεων του πίνακα 5.3. Αμέσως μετά ακολουθεί η οικογένεια των συναρτήσεων εκθετικής εξομάλυνσης (Ets) και οι αυτοπαλινδρομικές συναρτήσεις κινητού μέσου όρου(Arima) με καλύτερη επίδοση στο 25% και 12,5% των περιπτώσεων του πίνακα 5.3 αντίστοιχα. Αξίζει να σημειωθεί ότι η μέθοδος Θ αποδίδει επίσης πολύ καλά παρόλο που διακρίνεται σε μία μόνο περίπτωση.

Στην εικόνα 5.20 παρουσιάζεται η σωρευτική απεικόνιση των σφαλμάτων MAPE για όλες τις διαθέσιμες περιοχές του νομού Αττικής. Μπορούμε να παρατηρήσουμε ότι στη συντριπτική πλειοψηφία των περιπτώσεων, τα μοντέλα πρόβλεψης προσαρμόζονται πολύ καλά στα αρχικά δεδομένα του διαστήματος A καθώς το σφάλμα MAPE είναι μικρότερο του 10% ,τιμή η οποία θεωρείται ικανοποιητική σε γενικές γραμμές. Σε διάφορες μεμονωμένες περιπτώσεις όπου η βέλτιστη επίδοση είναι αρκετά μεγαλύτερη του 10% (π.χ. Δήμος Λυκόβρυσης, Δήμος Μαρκόπουλου κ.α.) οφείλεται στη φύση των δεδομένων που χαρακτηρίζονται από μεγάλη διακύμανση και τυχαιότητα κάτι που δυσκολεύει τις μεθόδους χρονοσειρών να προσαρμοστούν επαρκώς στα δεδομένα.

	Ets MAPE	Arima MAPE	Theta MAPE	Opt.Theta MAPE	MAPA MAPE	Naive MAPE	Comb.Method MAPE
Max	34.59159053	41.3070542	40.94485676	36.5641179	154.233016	33.690339	30.06465867
Average	4.064812722	4.448923654	4.754825427	4.721308024	6.289925841	6.756338679	3.867013512
Median	2.053028304	2.470113795	3.282805406	3.270283513	2.837217328	4.784050377	2.082550655
Min	0.656210395	0.588208487	0.672653122	0.674114783	1.248771725	2.653482741	0.562196734

Πίνακας 5.4-Υπολογισμός μέσου, μέγιστου, διάμεσου και ελάχιστου όρου των συνολικών σφαλμάτων MAPE για το διάστημα A

Από τον πίνακα 5.4 βλέπουμε ότι η μέθοδος συνδυασμού τριών βέλτιστων μεθόδων (3-method combination) είναι αυτή που δίνει το μικρότερο μέσο σφάλμα MAPE βελτιωμένο κατά 4,9% από την αμέσως λιγότερο ακριβή μέθοδο(Ets) και κατά 42,7% από την μέθοδο Naive. Εντύπωση προκαλεί το γεγονός ότι η μέθοδος με το μεγαλύτερο σφάλμα είναι η MAPA. Αυτό σημαίνει ότι υπάρχει μία τουλάχιστον χρονοσειρά στην οποία η μέθοδος Naive αποδίδει καλύτερα από την μέθοδο MAPA (βλ.πίνακα 5.3, κοινότητα Γλυκών νερών). Επίσης παρατηρούμε ότι η κλασική μέθοδος Θ σε σχέση με την βελτιστοποιημένη μέθοδο της, έχει μικρότερη ελάχιστη τιμή αλλά και μεγαλύτερη μέγιστη και μέση τιμή. Όλα αυτά τα συμπεράσματα μας προϊδεάζουν για το γεγονός ότι η συνδυασμένη μέθοδος θα είναι η βέλτιστη για τις πιο πολλές χρονοσειρές στο

διάστημα A και ότι η κλασική μέθοδος Θ με τη βελτιστοποιημένη μέθοδο της αποδίδουν παρόμοια.

Με τη βοήθεια του πακέτου hts για τη γλώσσα προγραμματισμού R, συναθροίζοντας τις βέλτιστες προβλέψεις για κάθε ταχ.κώδικα διαθέσιμο πρώτα στα δέκα cluster της μεθοδολογίας No1 και ύστερα στην τελική χρονοσειρά ζήτησης νερού, καταλήγουμε στο σφάλμα της μεθοδολογίας No2 για το διάστημα B:

Σφάλμα μεθοδολογίας No2 (cleaned data)

ME	RMSE	MAE	MPE	MAPE
541346.9	622051.3	562824.7	3.207426	3.329027

Στον παραπάνω πίνακα παρουσιάζονται και τα σφάλμα ME(Μέσο σφάλμα), RMSE(Ρίζα μέσου τετραγωνικού σφάλματος), MAE(Μέσο απόλυτο σφάλμα), MPE(Μέσο ποσοστιαίο σφάλμα) για πληρότητα αλλά σαν κριτήριο καλύτερης επίδοσης μεταξύ των τριών μεθοδολογιών που εξετάζει η διπλωματική είναι το σφάλμα MAPE.

Για την περίπτωση που χρησιμοποιούσαμε τα αρχικά αφιλτράριστα δεδομένα τότε ο πίνακας σφαλμάτων θα διαμορφώνονταν ως εξής:

Σφάλμα μεθοδολογίας No2 (raw data)

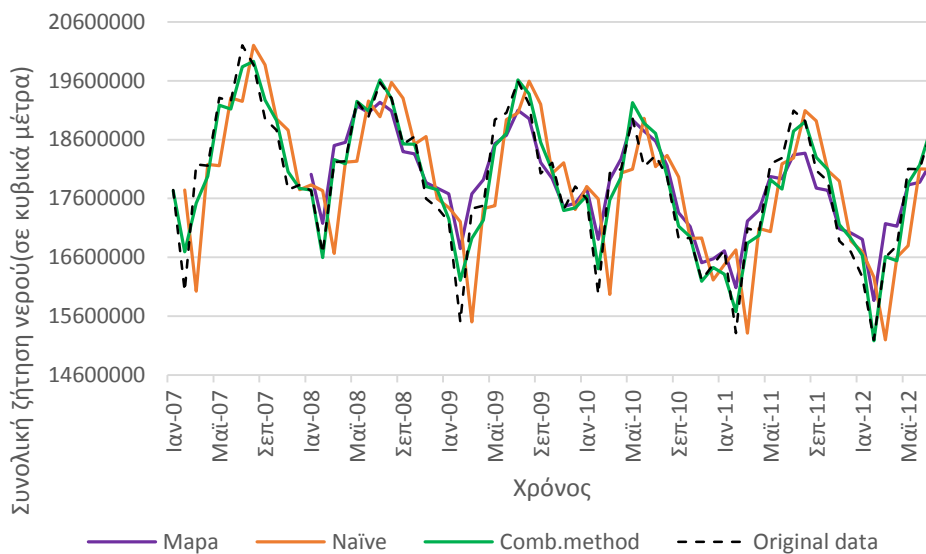
ME	RMSE	MAE	MPE	MAPE
631194.2	717927.4	651448.5	3.655534	3.730011

Παρατηρούμε δηλαδή για ακόμα μία φορά ότι τα φιλτραρισμένα δεδομένα δίνουν κατά 12,1% μικρότερο σφάλμα MAPE σε σχέση με τη χρησιμοποίηση των αρχικών πρωτόλειων δεδομένων. Αποδεικνύεται και εδώ, όπως και στη προηγούμενη μεθοδολογία, η ανάγκη «καθαρισμού» των δεδομένων από κενές και ακραίες τιμές καθώς μία βελτίωση των προβλέψεων της τάξης του 10% δεν είναι αμελητέα, ιδιαίτερα σε επιχειρησιακό επίπεδο.

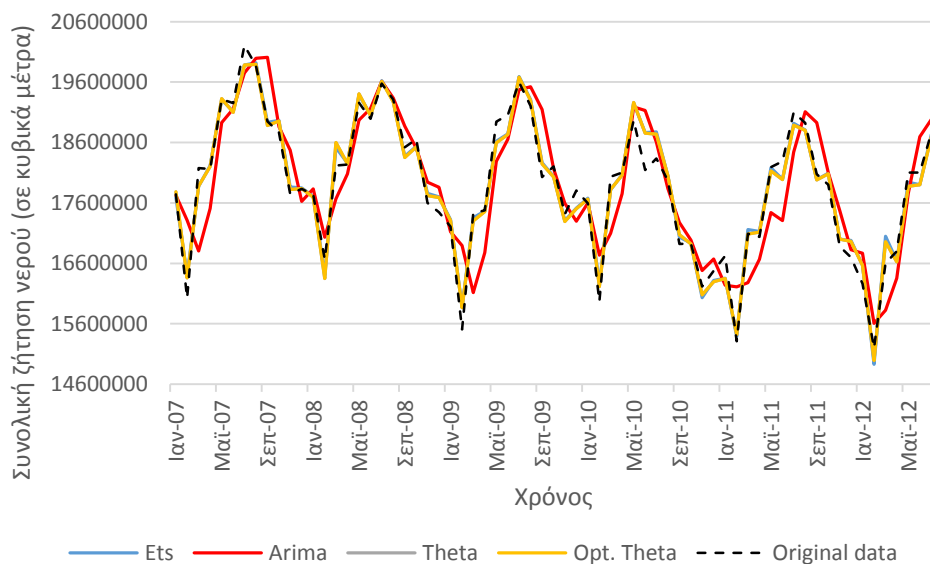
5.2.3 Αποτελέσματα Μεθοδολογίας No3

Στη μέθοδο αυτή οι προβλέψεις βγαίνουν με απλούστερο τρόπο. Έχοντας υπολογίσει τις χρονοσειρές μηνιαίας ζήτησης για κάθε διαθέσιμο ταχυδρομικό κώδικα, τις αθροίζουμε συνολικά ανά μήνα και καταλήγουμε στη συνολική μηνιαία χρονοσειρά ζήτησης για όλους τους ταχυδρομικούς κώδικες της Αττικής.

Παρακάτω δείχνουμε τις χρονοσειρές προσαρμογής των διαφόρων μεθόδων στο διάστημα A.



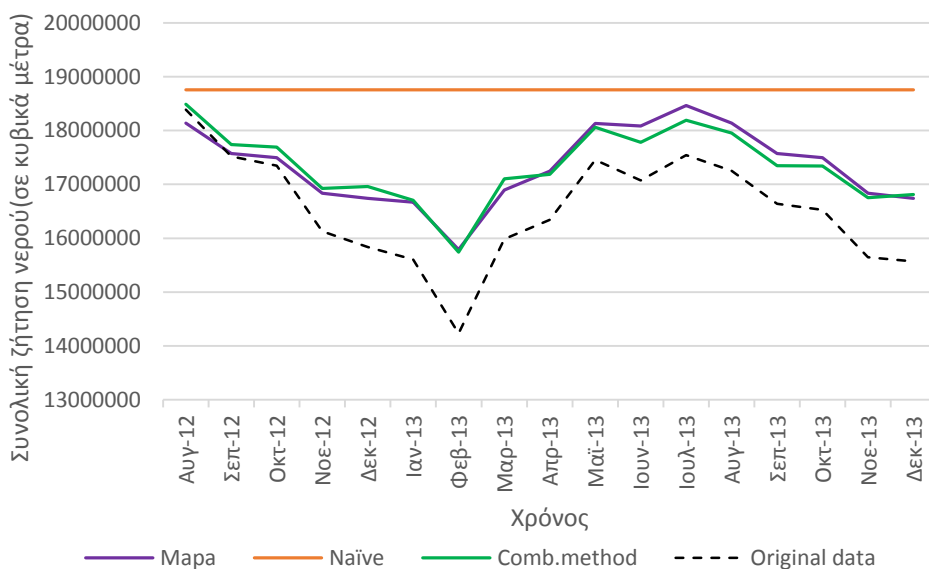
Εικόνα 5.21-Προσαρμογή μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα A



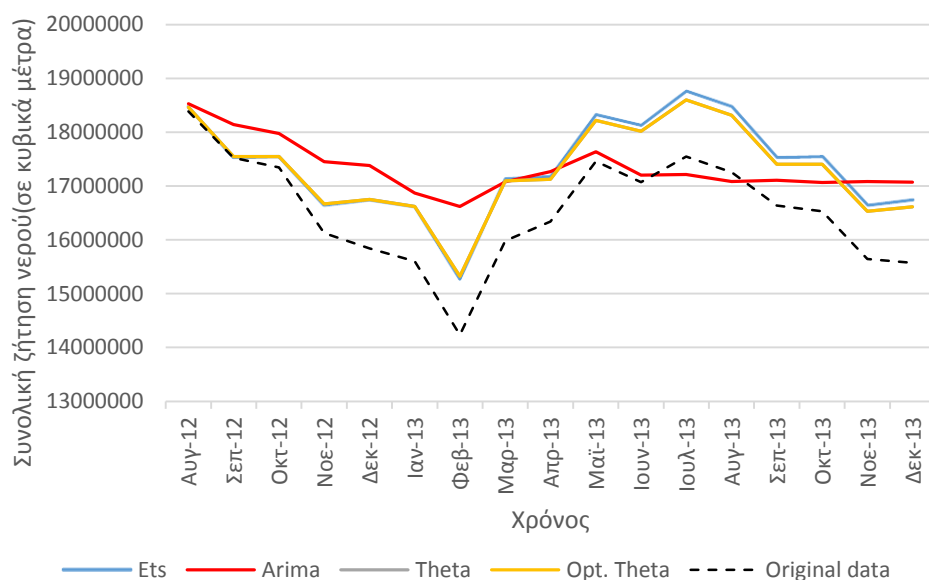
Εικόνα 5.22-Προσαρμογή υπόλοιπων μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα A

Παρατηρούμε ότι στο πλήρως εποχικό μοτίβο της συνολικής μηνιαίας χρονοσειράς ζήτησης νερού, όλες οι μέθοδοι αποδίδουν πολύ καλά και καταφέρνουν να προσαρμοστούν σωστά στα δεδομένα. Ξεχωρίζουν όμως σε ακρίβεια οι μέθοδοι Theta και Ets(A,N,A)¹ με την τελευταία να είναι κατά 0,2% ακριβέστερη. Επομένως, αυτή η μέθοδος επειδή ακριβώς δίνει το μικρότερο σφάλμα MAPE στο διάστημα A, θα επιλεγεί να προεκταθεί και να δώσει τις προβλέψεις για τη συνολική ζήτηση στο διάστημα B (εικόνα 5.24).

¹ Σύμφωνα με τους συμβολισμούς των διαφόρων μοντέλων που δίνονται στο πακέτο forecast της γλώσσας R, το μοντέλο Ets(A,N,A) είναι μοντέλο εκθετικής εξομάλυνσης με αθροιστικό σφάλμα και εποχικότητα και καμία τάση.



Εικόνα 5.23-Προβλέψεις μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα



Εικόνα 5.24-Προβλέψεις υπόλοιπων μεθόδων στα δεδομένα συνολικής ζήτησης στο διάστημα B

Παρατηρούμε από τα παραπάνω γραφήματα ότι οι μέθοδοι Ets και Theta δίνουν τις καλύτερες προβλέψεις, δηλαδή πιο κοντά στις πραγματικές τιμές που σημειώνονται με μαύρη διακεκομμένη γραμμή.

Στη συνέχεια δίνεται ο πίνακας σφαλμάτων των μεθόδων για τη τρίτη μεθοδολογία και αφορούν το διάστημα A. Τονίζεται ότι, σύμφωνα με την μέθοδο που χρησιμοποιήσαμε στη παρούσα διπλωματική και περιγράφεται εκτενώς στο κεφάλαιο 4, η μέθοδος η οποία έχει την καλύτερη προσαρμογή (σύμφωνα με το κριτήριο MAPE) στα δεδομένα παρατήρησης για την συνολική ζήτηση νερού για το διάστημα A, θα επιλεγεί από τον αλγόριθμο που αναπτύξαμε ώστε να προεκταθεί για τις επόμενες 17 παρατηρήσεις. Οι

τιμές αυτές θα καθορίσουν και την ακρίβεια της μεθοδολογίας No3 όπως θα δειχθεί σε επόμενο πίνακα.

	Ets MAPE	Arima MAPE	Theta MAPE	Opt.Theta MAPE	MAPA MAPE	Naive MAPE	Comb.method MAPE
Χρονοσειρά συνολικής ζήτησης	1.558066	2.346119	1.557879	1.557981	2.283824	3.921564	1.711381

Πίνακας 5.5-Πίνακας σφαλμάτων MAPE(%) για τη χρονοσειρά συνολικής ζήτησης και για όλες τις εξεταζόμενες μεθόδους στο διάστημα A

Όπως έχουμε σημειωμένα, η κλασική μέθοδος Θ προσαρμόζεται καλύτερα στα δεδομένα του διαστήματος A έχοντας κατά 0.006% καλύτερη επίδοση από τη βελτιστοποιημένη μέθοδο Θ και κατά 0,012% από τη μέθοδο Ets(A,N,A) , για αυτό και επιλέγεται να επεκταθεί στο διάστημα B. Άρα, το συνολικό σφάλμα της μεθοδολογίας No3 είναι:

Σφάλμα μεθοδολογίας No3 (cleaned data)

ME	RMSE	MAE	MPE	MAPE
807147.6	914487.7	828617.3	4.716521	4.789004

Στον παραπάνω πίνακα παρουσιάζονται και τα σφάλμα ME(Μέσο σφάλμα), RMSE(Ρίζα μέσου τετραγωνικού σφάλματος), MAE(Μέσο απόλυτο σφάλμα), MPE(Μέσο ποσοστιαίο σφάλμα) για πληρότητα αλλά σαν κριτήριο καλύτερης επίδοσης μεταξύ των τριών μεθοδολογιών που εξετάζει η διπλωματική είναι το σφάλμα MAPE όπως έχουμε τονίσει πολλές φορές στα προηγούμενα.

Σφάλμα μεθοδολογίας No3 (raw data)

ME	RMSE	MAE	MPE	MAPE
923445.2	1034533	951232.7	5.310269	5.461488

Αν χρησιμοποιήσουμε τα αρχικά δεδομένα χωρίς να τα καθαρίσουμε από κενές και ακραίες τιμές, το σφάλμα MAPE της μεθοδολογίας θα ήταν 5,461488 δηλαδή κατά 14% μεγαλύτερο από το σφάλμα που θα είχαμε αν χρησιμοποιούσαμε τα «καθαρισμένα» δεδομένα σύμφωνα με τα βήματα 1 και 2 της ενότητας 4.3. Αποδεικνύεται για ακόμη μία φορά η αξία του φιλτραρίσματος των δεδομένων και η αντικατάσταση ακραίων και κενών τιμών με τις ανάλογες τιμές που υποδεικνύονται από τις διάφορες μεθοδολογίες.

6 Συμπεράσματα και Προεκτάσεις

6.1 Συμπεράσματα

Με βάση την ανάλυση των αποτελεσμάτων που προηγήθηκε στο κεφάλαιο 5, μπορούμε να εξάγουμε κάποια χρήσιμα συμπεράσματα για την μεθόδους πρόβλεψης που χρησιμοποιήσαμε αλλά και για τη φύση των δεδομένων και το κοινωνικό αντίκτυπο που αποκαλύπτουν.

6.1.1 Επιλογή Βέλτιστης Μεθοδολογίας

Συγκεντρώνοντας τα σφάλματα MAPE και των τριών μεθοδολογιών του διαστήματος B στο πίνακα 6.1, συμπεραίνουμε ότι η μεθοδολογία Νο1 είναι αυτή που δίνει αποτελέσματα με μικρότερο σφάλμα, δηλαδή πιο κοντά στα ιστορικά δεδομένα.

	Μεθοδολογία No1	Μεθοδολογία No2	Μεθοδολογία No3
Σφάλμα MAPE	2.791232	3.329027	4.789004

Πίνακας 6.1-Συγκεντρωτικός πίνακας σφαλμάτων μεθοδολογιών

Επομένως, με βάση τα δεδομένα που έχουμε, θα ήταν προτιμότερο να ακολουθήσουμε τα εξής βήματα για την παραγωγή των προβλέψεων μας:

1. Έχοντας τα αρχικά δεδομένα μηνιαίας ζήτησης, διαιρούμε με τον αριθμό των μετρητών για κάθε μήνα και για κάθε μετρητή και παίρνουμε τις μηνιαίες παρατηρήσεις $\frac{\text{ζήτηση νερού}}{\text{αριθμό μετρητών}}$
2. Έχοντας τα δεδομένα του πρώτου βήματος, τα καθαρίζουμε από ακραίες και/ή κενές τιμές
3. Εύρεση των cluster
4. Έχοντας τα «καθαρισμένα» δεδομένα $\frac{\text{ζήτηση νερού}}{\text{αριθμό μετρητών}}$ τα πολλαπλασιάζουμε με τον μέσο όρο των μετρητών ανά περιοχή

5. Έχοντας τα δεδομένα από το προηγούμενο βήμα σε μονάδες m^3 κάνουμε προβλέψεις για το διάστημα B
6. Έχοντας τις τελικές βέλτιστες προβλέψεις για κάθε cluster, συναθροίζουμε τα δεδομένα bottom-up (10 cluster \rightarrow 1 συνολική χρονοσειρά ζήτησης) και καταλήγουμε στην τελική πρόβλεψη της μεθόδου No1

Με βάση το συγκεντρωτικό πίνακα, η μεθοδολογία No1 έχει κατά 16,1% καλύτερη επίδοση από τη μεθοδολογία No2 και κατά 41,7% από τη μεθοδολογία No3. Οπότε, για την παραγωγή προβλέψεων σε μελλοντικούς χρόνους για τους οποίους δεν έχουμε ιστορικά δεδομένα, είναι προτιμότερο να χρησιμοποιήσουμε τη μεθοδολογία No1 όπως έχει περιγραφεί σε αυτή την ενότητα αλλά και σε προηγούμενες.

6.1.2 Επιπλέον Συμπεράσματα

Συνεχίζοντας τον σχολιασμό των αποτελεσμάτων, έχει τονιστεί ιδιαίτερα η σημασία του «φιλτραρίσματος» των δεδομένων πριν την παραγωγή προβλέψεων. Συνήθως στα αρχικά δεδομένα πάντα θα υπάρχουν τιμές-παρατηρήσεις οι οποίες δεν συμβαδίζουν με τη μεγέθη των γειτονικών τιμών. Για παράδειγμα μία απότομη αύξηση ή μείωση στις τιμές της εξεταζόμενης μεταβλητής είναι κάτι σύνηθες για αρχικά δεδομένα, αλλά όχι κάτι αποδεκτό στην επιστήμη των προβλέψεων. Και αυτό γιατί η συσσώρευση τέτοιων τιμών μπορεί να αποπροσανατολίσει την προσαρμογή των μοντέλων χρονοσειρών και να δώσει αναξιόπιστες προβλέψεις. Το ίδιο ισχύει και σε περίπτωση που υπάρχουν κενές τιμές. Σε περίπτωση που οι κενές τιμές καλύπτουν μεγάλο μέγεθος των παρατηρήσεων, χρησιμοποιούνται εξειδικευμένοι μέθοδοι διακοπτόμενης ζήτησης (μέθοδοι Croston, ADIDA κ.α.) αλλά στη παρούσα διπλωματική αυτό δεν χρειάστηκε.

Όπως αποδείχτηκε στο προηγούμενο κεφάλαιο, και οι τρεις μεθοδολογίες δίνουν καλύτερα αποτελέσματα σε σύγκριση με τα αποτελέσματα που θα έδιναν αν είχαν χρησιμοποιηθεί τα αρχικά δεδομένα χωρίς προεπεξεργασία.

	Μεθοδολογία No1	Μεθοδολογία No2	Μεθοδολογία No3
Cleaned data	2.791232	3.329027	4.789004
Raw data	3.111149	3.730011	5.461488
Διαφορά(%)	11.46%	12.05%	14.04%

Πίνακας 6.2-Σφάλμα MAPE για τις τελικές χρονολογίες αρχικών και καθαρισμένων χρονοσειρών, και η ποσοστιαία διαφορά τους

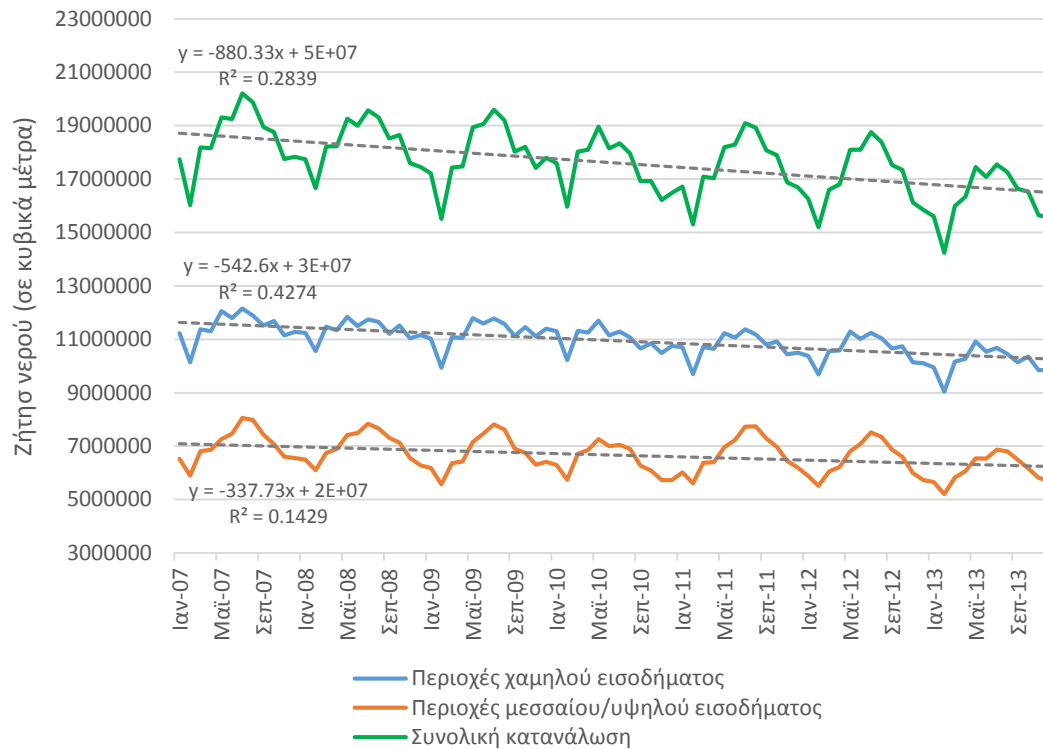
Ο παραπάνω πίνακας αφορά το διάστημα Β στο οποίο γίνεται η τελική κρίση για το ποιά μεθοδολογία είναι η προτιμότερη (δηλαδή ποια δίνει μικρότερο σφάλμα MAPE). Είναι εμφανές ότι στη περίπτωση των επεξεργασμένων δεδομένων η βελτίωση των αποτελεσμάτων είναι της τάξης του 10%-15%. Για να το περιγράψουμε περισσότερο κατανοώντας τα μεγέθη, ένα σφάλμα της τάξης του 15% στις τιμές πρόβλεψης των μοντέλων για το διάστημα Β (17 μήνες συνολικά), σημαίνει ότι υπάρχει απόκλιση ζήτησης 43,5 hm³ νερού που αντιστοιχεί περίπου στην ωφέλιμη χωρητικότητα του φράγματος του Μαραθώνα για ένα έτος. Τα μεγέθη αυτά είναι πολύ μεγάλα και οι οικονομικές επιπτώσεις που θα είχε να αντιμετωπίσει ο διαχειριστής των υδάτινων πόρων της περιοχής (εδώ ΕΥΔΑΠ) θα ήταν δύσκολα αντιμετωπίσιμες. Σε περίπτωση που η εταιρία διαχείρισης δεν μπορούσε να απορροφήσει τις ζημιές του λάθος υπολογισμού, τότε προκειμένου να διασφαλίσει τη μακροπρόθεσμη βιωσιμότητα της, θα μετακυλούσε τα κόστη στη πλάτη των καταναλωτών με αλυσιδωτές αντιδράσεις σε όλη την τοπική και εγχώρια οικονομία.

Ένα άλλο συμπέρασμα που θα μπορούσαμε να εξάγουμε κοιτώντας τις γραφικές παραστάσεις της ζήτησης νερού σε διάφορες περιοχές της Αττικής αλλά και συνολικά, είναι ότι παρουσιάζουν εμφανή εποχικότητα στην πλειοψηφία τους. Φυσικά υπάρχουν και χρονοσειρές που δεν έχουν εποχικότητα αλλά χαρακτηρίζονται από τυχαιότητα, αριθμητικά όμως μειοψηφούν. Επιπλέον, αξίζει να σημειώσουμε ότι η συνολική ζήτηση νερού που προκύπτει από τα δεδομένα είναι τελείως εποχική και έχει πτωτική τάση. Το «μοτίβο της πυραμίδας» επαναλαμβάνεται κάθε έτος με σταθερή συχνότητα αλλά με μειούμενη μέση τιμή. Η ζήτηση νερού αφού κάνει μία μικρή βουτιά τον Φεβρουάριο (πιθανότατα επειδή ο Φεβρουάριος υπολείπεται ημερών από τους άλλους μήνες) αρχίζει και αυξάνεται προοδευτικά μέχρι τους καλοκαιρινούς μήνες όπου ύστερα η ζήτηση εξομαλύνεται σε όλο και χαμηλότερα επίπεδα μέχρι το τέλος του έτους.

Στην εικόνα 6.1 απεικονίζονται εκτός από την συνολική ζήτηση, και οι καταναλώσεις με ομαδοποίηση ανάλογα με το μέσο εισόδημα ανά περιοχή. Με αυτό τον τρόπο θέλαμε να καταδείξουμε το αν και πόσο μπορούν να επηρεαστούν οι καταναλωτικές συνήθειες ανάλογα με το μέσο κατά κεφαλήν εισόδημα ανά περιοχή², με βάση τις υπάρχουσες κοινωνικό-οικονομικές συνθήκες. Είναι φανερό ότι τόσο η συνολική κατανάλωση αλλά και οι επιμέρους καταναλώσεις των περιοχών με χαμηλό και μεσαίο/υψηλό εισόδημα έχουν επηρεαστεί από τις υπάρχουσες καταστάσεις καθώς παρατηρείται σταδιακή μείωση της ζήτησης νερού καθ'όλο το χρονικό εύρος των δεδομένων μας. Επιπλέον, όπως σημειώνεται και στο γράφημα, η μείωση στη ζήτηση νερού είναι ηπιότερη στις περιοχές με μεσαίο/υψηλό κατά κεφαλήν μέσο εισόδημα από ότι στις περιοχές με υψηλό μέσο κατά κεφαλήν εισόδημα. Έχει παρουσιαστεί πολλάκις στην επιστημονική βιβλιογραφία [Billing, Jones Forecasting Urban Water Demand 2nd Ed.] και επαληθεύεται και από την παρούσα διπλωματική ότι η κατανάλωση νερού επηρεάζεται από τις

² Πηγή: ΕΛ.ΣΤΑΤ.

διάφορες κοινωνικό-οικονομικές συγκυρίες, καθώς σε καταστάσεις οικονομικής δυσπραγίας και κοινωνικών αναταραχών οι άνθρωποι τείνουν να προσαρμόζουν ανάλογα τις καθημερινές τους συνήθειες και να γίνονται πιο προσεκτικοί σε αχρείαστες σπατάλες.



Εικόνα 6.1-Γραφήματα Συνολικής ζήτησης και ζήτησης ανά περιοχή εισοδήματος

Χρήσιμα συμπεράσματα μπορούμε να βγάλουμε και για τις επιδόσεις των μεθόδων στις διάφορες μεθοδολογίες που χρησιμοποιήσαμε. Για το σκοπό αυτό, φτιάχνουμε τους παρακάτω πίνακες συχνότητας θέσης κατάταξης, και υπολογίζουμε τη μέση επίδοση των μεθόδων πρόβλεψης ανά μεθοδολογία.

Θέση	Ets	Arima	Theta	Opt. Theta	MAPA	Naive	Comb. method
1 ^η	2	2	2	1	0	0	3
2 ^η	1	1	2	2	1	0	3
3 ^η	2	0	1	3	0	0	4
4 ^η	3	0	3	3	0	1	0
5 ^η	2	1	0	1	5	1	0
6 ^η	0	6	1	0	3	0	0
7 ^η	0	0	1	0	1	8	0
Μέση επίδοση	3.2	4.5	3.4	3.1	5.2	6.5	2.1

Πίνακας 6.3-Επίδοση μεθόδων για το διάστημα A και τη μεθοδολογία Νο1

Από τον παραπάνω πίνακα είναι ξεκάθαρο ότι καλύτερη επίδοση συνολικά για την μεθοδολογία No1 έχει η μέθοδος του συνδυασμού των τριών βέλτιστων μεθόδων (Combined method ή 3-method combination). Είναι άξιο παρατήρησης ότι στις χρονοσειρές των cluster που εφαρμόσθηκε, η χειρότερη επίδοση που είχε ήταν η τρίτη θέση(4 φορές) ενώ είχε και τις πιο πολλές βέλτιστες προσαρμογές για το διάστημα A (3 φορές). Αμέσως μετά ακολουθεί η βελτιστοποιημένη μέθοδος Θ, έχοντας επίδοση οριακά καλύτερη από αυτή της οικογένειας συναρτήσεων εκθετικής εξομάλυνσης ενώ κατά πόδας ακολουθεί και η κλασική μέθοδος Θ με καλύτερη επίδοση σε δύο cluster αλλά με λιγότερη καλή παρουσία σε άλλα δύο cluster όπου καταλαμβάνει την τελευταία και προτελευταία θέση. Έπειτα ακολουθεί η μέθοδος Arima η οποία τείνει να προσαρμόζεται καλύτερα από άλλες μεθόδους σε χρονοσειρές με απότομες διακυμάνσεις. Στην έκτη θέση επίδοσης έρχεται η μέθοδος MAPA η οποία δεν καταφέρνει να διακριθεί σε καμία χρονοσειρά, ακολουθούμενη από την μέθοδο Naive η οποία, όπως αναμένονταν, είναι η λιγότερο αποδοτική στο 80% των περιπτώσεων για την συγκεκριμένη μεθοδολογία.

Θέση	Ets	Arima	Theta	Opt. Theta	MAPA	Naive	Comb. method
1 ^η	91	47	23	21	4	2	86
2 ^η	29	37	26	35	0	1	147
3 ^η	47	50	47	56	35	4	35
4 ^η	57	38	76	49	47	1	6
5 ^η	37	50	63	64	49	11	0
6 ^η	12	51	36	45	107	23	0
7 ^η	1	1	3	5	32	232	0
Μέση επίδοση	2.854	3.598	3.909	3.927	5.138	6.704	1.857

Πίνακας 6.4-Επίδοση μεθόδων για το διάστημα A και τη μεθοδολογία No2

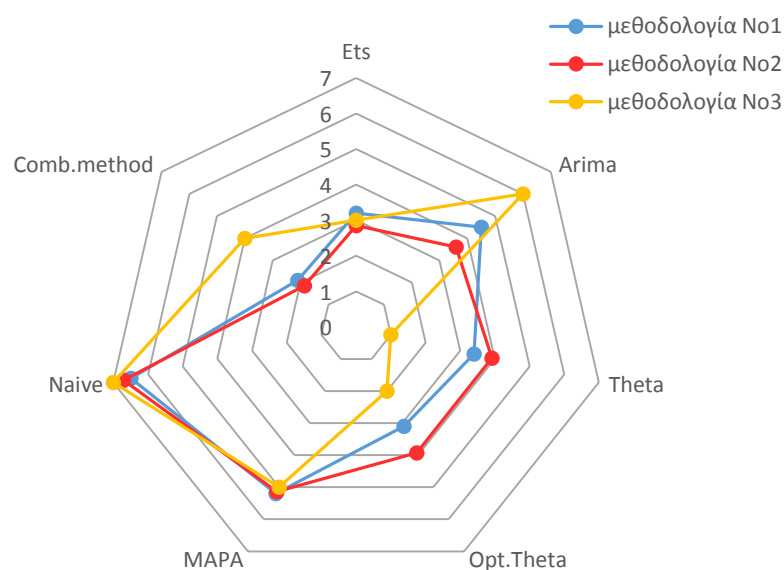
Περνώντας στη μεθοδολογία No2, ο πίνακας 6.4 δείχνει τις επιδόσεις των επιμέρους μεθόδων στις προσαρμογές των μοντέλων στο διάστημα A. Η εικόνα όσον αφορά την καλύτερη και την δεύτερη καλύτερη επίδοση δεν αλλάζει καθώς η μέθοδος των συνδυασμένων μεθόδων δίνει καλύτερα αποτελέσματα ακολουθούμενη από την οικογένεια συναρτήσεων εκθετικής εξομάλυνσης, όπως ακριβώς και στον πίνακα 6.3. Η ανατροπή έρχεται στην τρίτη θέση την οποία καταλαμβάνει η μέθοδος Arima. Παρατηρούμε ότι αυτή η μέθοδος έχει μία ομοιόμορφη επίδοση καθώς καταλαμβάνει όλες τις θέσεις από την πρώτη μέχρι την έκτη με περίπου ίσες συχνότητες. Έπειτα, στην τέταρτη θέση έρχεται η κλασική μέθοδος Θ σε πολύ μικρή απόσταση από τη βελτιστοποιημένη μέθοδο της. Παρατηρούμε ότι σε γενικές γραμμές αυτές οι δύο μέθοδοι αποδίδουν με παρόμοιο ικανοποιητικό τρόπο σε πλήθος χρονοσειρών και σε

αυτήν την περίπτωση η κλασική μέθοδος Θ έχει ένα μικρό προβάδισμα καθώς έχει ελαφρώς μεγαλύτερη συχνότητα επίδοσης στην πρώτη και δεύτερη θέση. Σε αρκετή απόσταση από την πέμπτη θέση έρχεται η μέθοδος MAPA η οποία παρ'όλα αυτά καταφέρνει να διακριθεί στο 1,5% των χρονοσειρών(4 χρονοσειρές). Τέλος, αξίζει να σημειώσουμε ότι με βάση τα δεδομένα του πίνακα 6.4, η μέθοδος Naive είναι η μέθοδος με την καλύτερη επίδοση σε 2 χρονοσειρές ενώ συνήθως καταλαμβάνει την έκτη και έβδομη θέση επίδοσης στο 93% των περιπτώσεων, κάτι που ήταν αναμενόμενο.

Θέση	Ets	Arima	Theta	Opt. Theta	MAPA	Naive	Comb. method
1 ^η	0	0	1	0	0	0	0
2 ^η	0	0	0	1	0	0	0
3 ^η	1	0	0	0	0	0	0
4 ^η	0	0	0	0	0	0	1
5 ^η	0	0	0	0	1	0	0
6 ^η	0	1	0	0	0	0	0
7 ^η	0	0	0	0	0	1	0
Μέση επίδοση	3	6	1	2	5	7	4

Πίνακας 6.5-Επίδοση μεθόδων για το διάστημα A και τη μεθοδολογία Νο3

Στον πίνακα 6.5 έχουμε τις επιδόσεις των μεθόδων για την μεθοδολογία Νο3 στο διάστημα A. Να θυμίσουμε ότι στη μέθοδο αυτή τα μοντέλα έπρεπε να προσαρμοστούν σε μία χρονοσειρά συνολικής ζήτησης νερού και όχι σε πολλές επιμέρους χρονοσειρές όπως έγινε στις προηγούμενες μεθοδολογίες.



Εικόνα 6.2-Συγκεντρωτική παρουσίαση επίδοσης μεθόδων για τις τρεις μεθοδολογίες

Εδώ την καλύτερη προσαρμογή την κάνει το μοντέλο της κλασικής μεθόδου Θ (με καλύτερη επίδοση ακόμα και από το μοντέλο συνδυασμένων μεθόδων) ακολουθούμενο από την βελτιστοποιημένη μορφή της. Εντύπωση προκαλεί η επίδοση της μεθόδου Arima η οποία αν και είναι πολύ καλή (μόλις 2,35% σφάλμα Mape) είναι η έκτη καλύτερη πίσω και από την μέθοδο MAPA.

Για την συνολική επίδοση των μεθόδων και στις τρεις μεθοδολογίες σε σύνολο 285 χρονοσειρών (10 χρονοσειρές από 1^η μεθοδολογία + 274 χρονοσειρές από 2^η μεθοδολογία + 1 χρονοσειρά από 3^η μεθοδολογία) έχουμε τον παρακάτω πίνακα:

	Ets	Arima	Theta	Opt. Theta	MAPA	Comb. method	Naive
Επίδοση μεθ. Νο1	2.85	3.6	3.92	3.93	5.14	1.86	6.71
Επίδοση μεθ. Νο2	3.2	4.5	3.4	3.1	5.2	2.1	6.5
Επίδοση μεθ. Νο3	3	6	1	2	5	4	7
Συνολική επίδοση	2.87	3.64	3.89	3.892	5.14	1.87	6.69
Διαφορά(%)	57.2%	45.7%	41.9%	41.9%	23.3%	72.03%	-

Πίνακας 6.6-Συνολική επίδοση μεθόδων στο διάστημα A

Βλέπουμε επομένως ότι η μέθοδος συνδυασμού βέλτιστων μεθόδων αποτελεί και τη μέθοδο με την καλύτερη συνολική επίδοση στις 285 συνολικά χρονοσειρές της διπλωματικής αυτής. Στη δεύτερη θέση βρίσκεται η οικογένεια των μεθόδων εκθετικής εξομάλυνσης(με κατά 35% περίπου χειρότερη επίδοση σε σχέση με την πρώτη) με μία αρκετά ισορροπημένη παρουσία στο σύνολο των μεθοδολογιών όπως μπορεί να φανεί και από την εικόνα 6.2. Μετά ακολουθεί η μέθοδος Arima και σε πολύ μικρή απόσταση η κλασική μέθοδος Θ η οποία δίνει ελαφρώς καλύτερα αποτελέσματα από την βελτιστοποιημένη μέθοδο Θ . Στη πέμπτη θέση βρίσκεται σταθερά η μέθοδος πολλαπλής συνάθροισης ενώ στην τελευταία θέση, με διαφορά από την προηγούμενη, η μέθοδος Naive. Στην τελευταία σειρά του πίνακα 6.6 αναγράφεται και η βελτίωση των μεθόδων που παρέχουν στη προσαρμογή των χρονοσειρών σε σύγκριση με την απλοϊκή μέθοδο Naive.

6.2 Προεκτάσεις

Η βραχυπρόθεσμη πρόβλεψη ζήτησης νερού αποτελεί σημαντικό επιστημονικό αντικείμενο για την επιστήμη των προβλέψεων, καθώς είναι ιδιαίτερα χρήσιμη για τις

εταιρίες διαχείρισης υδάτινων πόρων. Έτσι, η παραγωγή αξιόπιστων προβλέψεων είναι το ζητούμενο ώστε η εταιρία διαχείρισης να μπορεί να κάνει τις σωστές επενδύσεις για επέκταση και συντήρηση του υπάρχοντος συστήματος διανομής υδάτων χωρίς εκτροχιασμό του κόστους ανάπτυξης και υλοποίησης του προτζεκτ.

Πρώτα απ' όλα, επιτακτική κρίνεται η ανάγκη εξέτασης πιο εξεζητημένων μεθόδων πρόβλεψης από αυτές που χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη. Οι μέθοδοι αυτές μπορεί να περιλαμβάνουν μοντέλα γραμμικής και μη γραμμικής παλινδρόμησης, τεχνητά νευρωνικά δίκτυα, συνδυασμοί μεθόδων και άλλα εξειδικευμένα μοντέλα. Αξίζει, ακόμα, να εξεταστούν συνδυασμοί μεθόδων επιλεγμένων σύμφωνα με τα χαρακτηριστικά της εκάστοτε χρονοσειράς, καθώς σε πολλές εφαρμογές έχει αποδειχθεί ότι λειτουργούν.

Όπως αναφέρθηκε στα πρώτα κεφάλαια της παρούσας διπλωματικής, υπάρχουν περισσότεροι του ενός παράγοντες που επηρεάζουν την ζήτηση νερού και η ενσωμάτωση τους στα μοντέλα πρόβλεψης είναι επιθυμητή προκειμένου να εκμεταλλευτούμε όσο το δυνατόν περισσότερα δεδομένα. Στην παρούσα διπλωματική εξετάστηκαν κατά κύριο λόγο μοντέλα χρονοσειρών (time series models) και οι προβλέψεις των μοντέλων αυτών πραγματοποιούνται αφού πρώτα αναγνωριστεί ένα συγκεκριμένο πρότυπο στα δεδομένα το οποίο επαναλαμβάνεται με το χρόνο και παραμένει σταθερό. Έπειτα της αναγνώρισης του προτύπου γίνεται η προέκταση του σε τόσες περιόδους όσοι και το επιθυμητό χρονικό εύρος πρόβλεψης. Γίνεται επομένως αντιληπτό ότι μία από τις αδυναμίες του συγκεκριμένου μοντέλου είναι ότι δεν μπορεί να παράξει αξιόπιστες προβλέψεις σε περιπτώσεις που υπάρχουν παράγοντες εξωγενείς που επηρεάζουν το υπό πρόβλεψη μέγεθος και δεν μπορούν να προσδιοριστούν με βάση την φύση των ιστορικών στοιχείων. Κατ'επέκταση, μία άλλη πτυχή μπορεί να δειχθεί χρησιμοποιώντας μοντέλα που λαμβάνουν υπόψιν παράγοντες που επηρεάζουν την ζήτηση νερού, τα λεγόμενα αιτιοκρατικά μοντέλα.

Οι κυριότεροι παράγοντες που επηρεάζουν τη ζήτηση νερού είναι οι καιρικές συνθήκες, η τιμή πώλησης της μονάδας νερού, και η κατάσταση της οικονομίας. Είναι γεγονός ότι η ενσωμάτωση των καιρικών παραγόντων είναι πολύ ενδιαφέρουσα στην πρόβλεψη ζήτησης νερού, έχει όμως ορισμένα έντονα μειονεκτήματα η μελέτη της. Οι καιρικές συνθήκες μιας περιοχής δεν είναι διαθέσιμες τη στιγμή που παράγονται οι προβλέψεις ζήτησης του νερού και πρέπει να περιοριστεί κανείς στη χρήση προβλέψεων των καιρικών συνθηκών για το ίδιο χρονικό διάστημα που πρόκειται να γίνει η πρόβλεψη της ζήτησης νερού. Ακόμη, η πρόβλεψη των καιρικών συνθηκών δε γίνεται καθολικά για μια ευρεία περιοχή, αλλά για τα επιμέρους τμήματά της. Χρειάζεται λοιπόν, να μελετηθεί προσεκτικά ο τρόπος συμπερίληψης των καιρικών συνθηκών στην παραγωγή προβλέψεων. Σημειώνουμε τέλος ότι βασικός και με διαφορά καιρικός παράγοντας όσον αφορά στην επιρροή της ζήτησης νερού είναι η θερμοκρασία. Επομένως, στόχος είναι

αρχικά η ενσωμάτωση αυτής στις προβλέψεις και μελλοντικά και των λοιπών καιρικών παραγόντων (υγρασία, άνεμοι, ηλιοφάνεια, κλπ).

Η ενσωμάτωση ορισμένων από τους οικονομικούς παράγοντες αποτελεί ενδιαφέρον για προέκταση, καθώς το εξωτερικό περιβάλλον επηρεάζει σημαντικά τη συμπεριφορά του καταναλωτικού κοινού. Ένα πολύ καλό οικονομικό επίπεδο και περίοδοι εφορίας μιας περιοχής οδηγούν σε αυξημένη ζήτηση νερού και το αντίστροφο. Έχει παρατηρηθεί ότι σε οικογένειες με άνεργα μέλη(και άρα μικρότερο εισόδημα), αναπροσαρμόζεται η συμπεριφορά τους ως προς την κατανάλωση νερού σε σύγκριση με προηγούμενες καταστάσεις όπου είχαν κάποια απασχόληση. Ακόμη, οι οικονομικοί παράγοντες προκαλούν και δημογραφικές μεταβολές. Για παράδειγμα, σε περιόδους ανάπτυξης παρατηρείται αύξηση του αστικού πληθυσμού και αντίστοιχη μείωση του αγροτικού, γεγονός που επηρεάζει τη ζήτηση νερού κατά τόπους (και κατ' επέκταση στο σύνολο της περιοχής). Στην παρούσα εργασία δε μελετήθηκαν οι οικονομικοί παράγοντες, καθώς είναι ιδιαίτερα πολύπλοκη υπόθεση η ενσωμάτωσή τους στα μοντέλα πρόβλεψης.

Τελειώνοντας, τα συμπεράσματα που βγήκαν από τα την παρούσα διπλωματική θα μπορούσαν να χρησιμοποιηθούν σαν βάση για έναν ευρύτερο σχεδιασμό διαχείρισης υδάτων. Παρουσιάσαμε στο κεφάλαιο 2 κάποιες Ευρωπαϊκές μελέτες και οδηγίες οι οποίες μεταξύ άλλων αναφέρονται στην ανάγκη αποτελεσματικής επίβλεψης και διαχείρισης των υδάτινων αποθεμάτων, ειδικά στις χώρες της νότιας Ευρώπης (μεταξύ αυτών και η Ελλάδα), οι οποίες αναμένεται να αντιμετωπίσουν και το μεγαλύτερο πρόβλημα αποθεμάτων ως το 2030. Θα μπορούσε λοιπόν αυτή η διπλωματική να αποτελέσει ένα μικρό κομμάτι ενός ευρύτερου πρότζεκτ για την Αττική στο οποίο θα γίνει προσπάθεια να θεμελιωθεί και να υποστηριχθεί σε βάθος χρόνου ένα ολοκληρωμένο σύστημα εποπτείας και διαχείρισης των αττικών υδάτινων πόρων στο οποίο θα έχει ξεχωριστή θέση μία αξιόπιστη μεθοδολογία πρόβλεψης της ζήτησης νερού, σαν αυτή που αναπτύξαμε στη παρούσα διπλωματική. Η τεχνογνωσία που θα αποκτηθεί από αυτή τη μελέτη, θα μπορούσε να χρησιμοποιηθεί για την καθορισμό ενός οδικού χάρτη για τη δημιουργία ενός ενοποιημένου συστήματος διαχείρισης υδάτων σε πανελλαδικό επίπεδο.

Παράρτημα

A) Εδώ θα παρουσιαστούν αναλυτικοί πίνακες αποτελεσμάτων για την πληρότητα της παρουσίασης της διπλωματικής

Method	Data	Time Scale	Model Scale	Sources	Characteristics	Limitations
Multiple regression	time series data	daily or monthly	municipal	Maidment et al. [1985], Maidment and Miao [1986], Gato et al. [2007], Adamowski [2008], Ghiassi et al. [2008], Cutado [2010], Agthe and Billings [1980], Al-Qunaybi et al. [1985], Maidment et al. [1985], Miao [1990]	(1) short-term forecasting, (2) identify peak and off-peak daily demand to ensure necessary treatment and distribution capacity to meet demand, (3) estimate price and income elasticities, (4) assess the differential effect of demand determinants during peak versus off-peak periods	(1) lacking or highly aggregated spatial data, (2) aggregated data used instead of microlevel data, (3) OLS assumption of independence in the error term is easily violated, (4) endogeneity bias due to price based on consumed quantity under multipart tariff pricing, (5) serial correlation bias
Multiple regression	time series data	seasonal/bimonthly	municipal	Maidment et al. [1985], Maidment and Miao [1986], Miao [1990], Zhou et al. [2000], Syme et al. [2004], Gutzler and Nims [2005], Gato et al. [2007], Praskiewicz and Chang [2009], Wong et al. [2010]	(1) short-term forecasting, (2) separate water use into two components: weather-insensitive base use (winter or indoor) and weather-sensitive seasonal use (summer or outdoor), (3) examine climatic effects on demand	(1) lacking or highly aggregated spatial data, (2) difficult to determine the correct functional form of the model, (3) conventional models may underestimate water use response to climate variables because of the influence of stochastic events on seasonal use
Multiple regression	dynamic panel data	monthly/bimonthly	household level or census tract scale	Nauger and Thomas [2003], Arfaoui et al. [2004], Arfaoui et al. [2010], Folebiński and Palmer [2010]	(1) incorporates both temporal and subject-based variability into coefficient estimates, (2) more efficient and consistent parameter estimates than OLS, (3) fine spatial scale data, (4) integrate lagged independent variables (e.g., price)	(1) analysis of both disaggregated temporal and spatial data is more time and data intensive than traditional methods, (2) determining the regression method (e.g., fixed effects versus random effects) to account for spatial variability remains uncertain
Piecewise linear regression	time series data; cross-sectional data	daily or monthly	municipal or census block group scale	Maidment et al. [1985], Maidment and Miao [1986], Chang et al. [2010]	(1) determine structural or temporal regime shift in a regression model, (2) create discrete linear segments connected at a point of change, (3) capture nonlinear responses in slope when thresholds are passed, (4) simple to interpret	(1) difficult to determine a priori the knot (point of change), (2) statistical testing required to ensure that the slopes are statistically significantly different before and after the knot
Spatially explicit ordinary least squares (OLS) regression	cross-sectional data; geotagged data	monthly/bimonthly	census block group scale or census tract scale	Chang et al. [2010], House-Peters et al. [2010], Shandas and Parasoulas [2010], Gañanika and Gober [2007], Wentz and Gober [2007], Balling et al. [2008], Lee and Wentz [2008], Lee et al. [2010]	(1) visualize and quantify water use patterns at fine spatial scales, (2) elucidate spatial patterns of clustering and dispersion of high and low water users, (3) model individual household level consumption data, (4) correct for heterogeneity due to spatial autocorrelation, which otherwise causes biased parameter estimation	(1) water provider service areas do not match administrative boundaries (e.g., census block), (2) data usually must be aggregated to protect customer privacy, (3) no consistency between water providers regarding collection of water use data
Geographically weighted regression (GWR)	cross-sectional data; geotagged data	monthly	individual water consumption data aggregated to census tract scale	Wentz and Gober [2007]	(1) forecast small-area water consumption, (2) accounts for spatial autocorrelation, (3) calculates a set a unique regression equation for each observation defined by geographic coordinates, (4) improvement over OLS ($R^2 = 0.64$) versus GWR (mean $R^2 = 0.85$)	(1) computationally and data intensive, (2) each sample has its own unique regression, difficult to interpret results for a large sample, (3) availability of geotagged data continues to be lacking, (4) model demonstrates spatial variation but does not have explanatory power
Simultaneous equation demand model	panel data	monthly	municipal	Agthe et al. [1986], Espey et al. [1997], Torregrona et al. [2010]	(1) simultaneously and jointly determines the endogenous dependent variables on the basis of exogenous variables, (2) corrects for multicollinearity and serial autocorrelation	(1) data aggregation at large spatial scales, (2) violation of the economic assumption that households are perfectly informed of water price
Autoregressive integrated moving average (ARIMA)	time series data	daily	municipal	Bougadis et al. [2005], Adamowski [2008], Praskiewicz and Chang [2009]	(1) accounts for the autocorrelation in the water demand time series, (2) uses the previous day's water use as an independent variable	(1) data aggregation at large spatial scales, (2) difficult to determine a priori correct model form and parameters
State-space forecasting model	time series data	monthly	municipal	Billings and Agthe [1998]	(1) Computes forecasts based on the dependence of a variable upon its	(1) The values of the independent variables must be forecasted in order to compute

Εικόνα Π.1: Παρουσίαση και σχολιασμός των σημαντικότερων ερευνητικών μεθόδων στη πρόβλεψη ζήτησης νερού

Method	Data	Time Scale	Model Scale	Sources	Characteristics	Limitations
Bayesian maximum entropy (BME)	cross-sectional data; soft data	monthly	census tract scale	Lee and Wentz [2008], Lee et al. [2010]	own lags and the cross lags of the independent variable, (2) simpler than ARIMA	a forecast of the dependent variable
Artificial neural networks (ANNs)	time series data	daily or hourly	municipal	Adya and Collopy [1998], Bougadis et al. [2005], Adamowski [2008], Ghiassi et al. [2008], Bárdossy et al. [2009], Firat et al. [2009], Adamowski and Karapatakis [2010], Herrera et al. [2010]	(1) highly effective for forecasting short-term demand (99% accuracy), (2) alternative to traditional linear modeling approach, (3) explicitly analyze nonlinear time series events, (4) minimize relative error, (5) maximize robustness	(1) complex, data and computationally intensive training and testing requirements, (2) loss of parsimony, (3) lack of explanatory power of the results, (4) sensitive to misspecification error
System dynamics models (SDMs)	time series data	monthly	regional	Rosenberg et al. [2007], Winz et al. [2009], Ahmad and Prasha [2010]	(1) incorporate diverse variables and submodels, (2) visualize the effects of intervention strategies, (3) continuously test assumptions and system sensitivity under scenarios	(1) unlike ABMs, the behavior of neighbors and the influence of this behavior cannot be simulated, (2) data, software, and computationally intensive

Εικόνα Π.2: Συνέχεια από Π.1

Παρακάτω παρουσιάζεται ο ολοκληρωμένος πίνακας των 274 χρονοσειρών που αποτελεί ουσιαστικά την βάση των δεδομένων μας.

Περιοχή	min	Q1(25%)	Median (50%)	Q2(75%)	max	mean	std
ΖΩΓΡΑΦΟΥ-15771	87188.95	96841.6775	101581.2364	105706.0875	115816.06	101024.8703	6237.885807
ΖΩΓΡΑΦΟΥ-15772	112800.78	128829.3875	134551.175	142009.8	152623.14	134583.2774	8905.12832
ΖΩΓΡΑΦΟΥ-15773	81697.2	92027.3225	97079.51	102231.6725	126390.94	97430.90536	7164.675909
ΑΓ.ΑΝΑΓΥΡΩΝ-13561	43476.18	66319.53	71766.65	74878.2475	83389.93	69735.71879	8761.051424
ΑΓ.ΑΝΑΓΥΡΩΝ-13562	67333.85	77836.4875	80616.09181	83513.695	92232.09	80371.41447	4934.50871
ΑΓ.ΒΑΡΒΑΡΑΣ-12351	94082.98	111191.8725	115256.73	118122.215	130484.24	114573.4707	6555.859199
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17341	86211.64	102336.285	106069	109379.7025	122207.72	105841.971	6435.050639
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17342	108184.97	124383.145	128121.18	131530.36	139807.24	127352.2349	6353.782678
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17343	75958.93	90281.02	93909.755	96555.4475	111763.78	93491.20291	5789.317808
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17372	15.47	32.74193571	48.14	60.8875	110.34	50.6967258	25.93128876
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18232	88.91	134.745	151.115	169.6169851	217.34	153.113461	29.61146892
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18233	66321.75	78680.345	84120.36	86379.6775	97254.07	82832.82727	6322.379878
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15341	91884.6	113139.695	120313.595	127537.5525	141923.89	119834.2439	10557.38726
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15342	65206.43	82053.8375	87763.645	95421.555	114381.34	88325.22304	10405.82543
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15343	95753.09	112971.4775	121161.02	128384.3675	144342.01	120545.2653	11583.63047
ΑΙΓΑΛΕΩ-12241	58870.77	70063.6525	74163.565	89641.6375	114492.1	79531.82661	13074.09574
ΑΙΓΑΛΕΩ-12242	56441.24	68260.5025	74845.905	81701.885	90554.5	74930.36526	7992.709479
ΑΙΓΑΛΕΩ-12243	87150.12	102690.3475	106723.85	109866.2025	120540.21	106336.5361	5934.636755
ΑΙΓΑΛΕΩ-12244	68461.25	82396.51	88564.145	91970.19	100554.88	87420.46588	6364.231333
ΆΛΙΜΟΣ-17436	41.03	64.73	84.995	111.37	209.52	94.79788978	40.91246875
ΆΛΙΜΟΣ-17455	114826.24	138995.4425	145477.93	154214.0175	168306.62	145489.4892	12178.28314
ΆΛΙΜΟΣ-17456	80099.41	93264.2375	101426.475	107182.955	120355.5	100448.8715	9564.109582
ΑΜΑΡΟΥΣΙΟΥ-15122	69102.98	80420.2775	86234.305	90497.4675	100713.53	85319.1694	7219.296699
ΑΜΑΡΟΥΣΙΟΥ-15123	26175.54	29626.825	33141.285	35505.67	39240.22	32758.99188	3554.220028
ΑΜΑΡΟΥΣΙΟΥ-15124	82814.56	97514.705	106370.465	112606.9325	134511.14	105714.7085	11167.96163
ΑΜΑΡΟΥΣΙΟΥ-15125	86172.96	100741.985	108408.38	117583.0725	125187.6	108301.0232	10222.27735
ΑΜΑΡΟΥΣΙΟΥ-15126	79484.93	91204.835	99020.46	105233.6025	117852.43	98189.27679	9193.106918
ΆΝΩ ΛΙΨΟΣΙΑ-13341	90.12	112.2975	611.745	1583.75	2258.93	826.2387862	734.0704242
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16451	70307.33	79357.0475	84230.43	89251.64	100377.13	84205.42798	6549.89995
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16452	73661.64	82349.5475	88132.325	93682.66166	103443.66	87857.61625	7146.952252
ΑΣΠΡΟΠ'ΥΡΓΟΥ-19300	169704.76	337816.0275	370664.14	426717.6	609332.2	381067.7032	79385.33661
ΑΧΑΡΝΩΝ-13671	52637.04	63044.4075	67477.595	71710.07	78773.07	67319.9751	6278.360981
ΑΧΑΡΝΩΝ-13672	48871.3	59250.12	63108.165	67927.4175	74269.12	63370.73499	6249.923774

ΑΧΑΡΝΩΝ-13673	35418.58	42480.9725	45195.07	47282.0825	51301.79	44646.15634	3586.93741
ΑΧΑΡΝΩΝ-13674	73119.56	85750.8	89933.56	94976.2	111900.26	90122.23256	7732.317263
ΑΧΑΡΝΩΝ-13675	46663.29	55070.98	61489.49	67901.98	75688.65	61509.98612	7665.030536
ΑΧΑΡΝΩΝ-13677	24413.33	35200.5525	39281.88	43524.7675	50693.79	39404.95821	5773.908441
ΑΧΑΡΝΩΝ-13678	22338.97	27159.2825	29244.915	32341.6525	35788.08	29493.18684	3208.50636
ΑΧΑΡΝΩΝ-13679	13978.42	18750.8175	21885.97	25751.8175	31110.81	22355.51738	4287.361628
ΒΑΡΗΣ-16672	66754.29	91903.745	112440.59	136640.4075	173992.42	116715.6764	29662.28887
ΒΟΥΛΑΣ-16673	151007.51	197760.165	230155.625	272374.645	328331.91	235189.4949	46974.05627
ΒΟΥΛΙΑΓΜΕΝΗΣ-16671	15771.66	47621.905	60924.49	86570.77	119024.78	65393.89938	24927.26995
ΒΡΙΑΗΣΣΙΩΝ-15235	2987.32	19029.2443	195125.17	201456.46	211680.07	142836.7948	86572.05436
ΒΥΡΩΝΟΣ-16231	99698.35	108463.1325	112741.635	115202.5525	126535.91	111996.7273	5363.393168
ΒΥΡΩΝΟΣ-16232	94029.32	105837.22	110999.935	114336.0075	121515.24	110032.0781	5844.822537
ΒΥΡΩΝΟΣ-16233	52740.58	60441.7	63316.68	66999.03	73523.57	63583.14583	4604.929204
ΓΑΛΑΤΣΙΟΥ-11146	142258.8	158604.3775	167809.695	175976.545	188609.83	167683.2536	10607.10923
ΓΑΛΑΤΣΙΟΥ-11147	85537.58	95842.61	100042.145	104549.715	111724.92	99876.362	5610.447131
ΓΛΥΦΑΔΑΣ-16561	108163.9	127448.2025	142690.5	154950.725	170162.11	141733.756	16034.2357
ΓΛΥΦΑΔΑΣ-16562	80930.21	97222.31	105134.43	110929.9625	125094.9	104820.8158	9946.783499
ΓΛΥΦΑΔΑΣ-16674	110628.13	150564.3025	174118.615	194326.3175	235452.88	174215.048	30192.20289
ΓΛΥΦΑΔΑΣ-16675	92734.71	109320.2625	121644.47	138038.6775	162201.32	123757.8302	17574.25808
ΔΑΦΝΗΣ-17234	30744.85	37091.535	38295.42	39522.5325	42214.46	38093.31796	2119.485082
ΔΑΦΝΗΣ-17235	37525.07	41050.3375	42249.69	43506.175	45952.12	42196.19281	1870.006483
ΔΡΑΠΕΤΣΩΝΑΣ-18648	50750.77	59621.91	61227.415	62637.7575	65322.43	60819.12638	2725.3783
ΕΛΕΥΣΙΝΟΣ-19200	95729.43	125201.335	140694.635	156218.1525	190556.34	141355.6013	20634.41451
ΕΛΛΗΝΙΚΟΥ-16777	73787.31	83910.96	93554.415	100601.8375	117731.8	93064.42643	9558.695422
ΖΕΦΥΡΙΟΥ-13461	30101.21	34849.46	36814.66	39161.75	43485.8	36896.52114	3207.640347
ΗΛΙΟΥΠΟΛΕΩΣ-16341	56858.02	63477.6575	68189.155	71735.02	79478.91	67661.04202	5493.756183
ΗΛΙΟΥΠΟΛΕΩΣ-16342	49072.44	55784.4125	59885.3	63696.7475	69225.03	59500.42169	4955.906484
ΗΛΙΟΥΠΟΛΕΩΣ-16343	44707.45	51384.755	52501.895	54161.315	59395.73	52495.907	2872.07976
ΗΛΙΟΥΠΟΛΕΩΣ-16344	41516.75	46586.1325	48535.29	49489.91	54824.74	48220.61689	2407.001462
ΗΛΙΟΥΠΟΛΕΩΣ-16345	60382.02	68910.015	72008.32	75654.1725	98433.38	73616.14716	8002.996719
ΗΛΙΟΥΠΟΛΕΩΣ-16346	62818.5	70313.0025	74661.905	77549.7175	86620.23	74146.75615	4947.280346
ΗΡΑΚΛΕΙΟΥ-14121	102413.117	115298.005	122157.595	128758.1331	141276.27	122297.9692	9504.925041
ΗΡΑΚΛΕΙΟΥ-14122	105164.18	117880.735	124078.855	130504.67	143656.61	123842.2409	9106.192818
ΚΑΤΖΑ-19012	4.67	6.6075	9.81	16.2	69	16.39175102	16.07054157
ΙΛΙΟΥ-13121	98918.69	116094.2425	120022.205	123958.045	132313.14	119774.949	6175.168544
ΙΛΙΟΥ-13122	119199.91	141900.5025	147008.7	152984.5325	163159.17	146094.0955	9384.884091
ΙΛΙΟΥ-13123	93555.06	108482.74	111191.69	114035.1625	123960.76	110714.2833	5418.154813
ΚΑΙΣΑΡΙΑΝΗΣ-16121	84703.38	95723.175	98278.195	100693.255	125314.05	99066.40512	7034.16509
ΚΑΙΣΑΡΙΑΝΗΣ-16122	31283.09	35002.8175	36140.37	36965.7625	38703.69	35924.4349	1422.353692
ΚΑΛΛΙΘΕΑΣ-17671	66074.52	71789.96	76026.535	78383.675	82578.37	75034.91524	4140.082839
ΚΑΛΛΙΘΕΑΣ-17672	68244.38	78028.4025	83615.31	87154.5325	92635.78	82558.81393	6011.479244
ΚΑΛΛΙΘΕΑΣ-17673	92707.7	107567.245	112199.595	117632.29	127441.25	112499.8029	7401.490656
ΚΑΛΛΙΘΕΑΣ-17674	38684.55	56113.7325	59236.09	62165.525	65857.24	57591.04661	6544.908993
ΚΑΛΛΙΘΕΑΣ-17675	68938.82	86130.0525	91188.94	95480.9975	100233.58	89673.8503	7536.819138

ΚΑΛΛΙΘΕΑΣ-17676	66464.32	74211.01	80228.79	83245.58	91440.68	79098.86647	6073.242677
ΚΑΜΑΤΕΡΟΥ-13451	107515.03	125261.0175	133833.1625	142407.235	155784.94	133565.459	11951.17947
ΚΕΡΑΤΕΑΣ-19001	150.12	300.9725	526.745	804.945	1564.63	590.7564286	369.4189503
ΚΕΡΑΤΣΙΝΙ-18755	68491.1	82021.18	85466.56	91978.6525	114098.85	88337.3525	9970.717914
ΚΕΡΑΤΣΙΝΙ-18756	57549	67174.1825	69061.415	72617.71	77648.38	69701.09199	3918.402105
ΚΕΡΑΤΣΙΝΙ-18757	93714.1	107971.3025	110510.445	113151.4875	120555.47	110055.2706	4658.492522
ΚΕΡΑΤΣΙΝΙ-18758	62414.55	69573.885	71405.055	74210.2525	90441.25	73212.99212	6517.836421
ΚΗΦΙΣΙΑΣ-14561	74930.49	93767.23	104716.315	121021.43	147591.12	108157.0652	18380.23116
ΚΗΦΙΣΙΑΣ-14562	62645.38	86078.97	98633.565	122904.1275	159750.65	104769.4913	24736.05686
ΚΗΦΙΣΙΑΣ-14563	19543.69	49811.5725	55077.4	62153.895	77260.54	54672.50388	12461.26217
ΚΗΦΙΣΙΑΣ-14564	74872.36	93438.0275	108711.29	126347.8958	152891.6	111090.3479	22453.79435
ΚΗΦΙΣΙΑΣ-14671	1842.83	30941.91	34233.31	145598.715	154707.5	73259.03278	59167.91291
ΚΟΥΡΥΔΑΛΛΟΥ-18120	106397.84	119960.2375	127087.61	131921.2225	142782.8	126511.9919	7457.462437
ΚΟΥΡΥΔΑΛΛΟΥ-18121	58340.16	68106.4725	70970.38	73508.61	77224.97	70503.35071	3756.892563
ΚΟΥΡΥΔΑΛΛΟΥ-18122	59751.55	65876.4525	67620.735	69471.8975	72390.11	67285.61624	2881.822575
ΚΡΩΠΙΑΣ-19010	1318.63	3798.425	5455.175	7667.6575	12788.64	6052.46793	2980.244069
ΚΡΩΠΙΑΣ-19013	134.39	811.9075	1262.375	1832.4525	2660.15	1300.814453	671.1674255
ΚΡΩΠΙΑΣ-19400	1810.07	4384.7525	6863.32	10143.325	14568.57	7537.864167	3598.448656
ΚΟΙΝΟΤΗΤΑ ΓΛΥΚΩΝ ΝΕΡΩΝ-15354	3651.21	24567.7825	27954.58	32488.595	38505.92	26993.78911	8631.608942
ΚΟΙΝΟΤΗΤΑ ΘΡΑΚΟΜΑΚΕΔΟΝΩΝ-13676	27982.42	36314.4975	48847.365	65050.2275	100744.78	52196.19745	17867.12833
ΚΟΙΝΟΤΗΤΑ ΠΑΛΑΙΑΣ ΠΕΝΤΕΛΗΣ-19009	101	101	196.595	278.525	435.07	163.7299311	138.5321563
ΛΑΥΡΕΩΤΙΚΗΣ-19500	110	203	464.8208791	656.5	1059.42	403.6961844	336.3609611
ΛΥΚΟΒΡΥΣΕΩΣ-14123	37919.37	48525.5125	53544.19	61159.68	67742.73	53998.10728	7435.945955
ΜΑΝΔΡΑΣ-19600	5.3	11.2175	17.7667956	318.35	513.38	148.572156	165.0501511
ΜΑΡΑΘΩΝΑ-19007	15	25.4925	84.9713003	202.795	934.9	185.3808251	238.4462541
ΜΑΡΚΟΠΟΥΛΟΥ-19003	6.49	42.4425	53.115	70.5925	108.81	54.36196989	22.74892739
ΜΕΓΑΡΕΩΝ-19100	2935.83	3874.945	5217.19	7702.97	9417.09	5752.99073	2069.34617
ΜΕΛΙΣΣΙΩΝ-11527	54499.57	61954.5625	65454.76	68655.1625	75073.7	65224.34524	4041.946223
ΜΕΤΑΜΟΡΦΩΣΕΩΣ-14451	55089.15	65268.46	68674.27	72592.8825	79814.81	68774.31054	5122.811867
ΜΕΤΑΜΟΡΦΩΣΕΩΣ-14452	68652.6	83363.18	87657.875	92900.4925	101162.74	87448.90357	6970.1676
ΜΟΣΧΑΤΟΥ-18344	43182.09	49099.5075	50667.835	52207.9025	57036.38	50449.77634	2891.224557
ΜΟΣΧΑΤΟΥ-18345	60539.45	67522.91	70364.89	72487.305	77118.28	69916.9175	3626.901708
ΜΟΣΧΑΤΟΥ-18346	6919.53	8674.57	9471.605	10425.3525	13522.29	9689.375476	1545.362914
ΝΕΑΣ ΠΕΡΑΜΟΥ-19006	1992.32	37897.8875	50864.47	60543.3	113194.95	49945.81259	18821.31728
ΝΕΑΣ ΙΩΝΙΑΣ-14231	73945.28	84774.9	87983.95	90482.2175	99393.95	87641.51812	4728.642208
ΝΕΑΣ ΙΩΝΙΑΣ-14232	50841.38	56678.5725	58718.15	60616.39	68430.99	58751.22709	3172.650666
ΝΕΑΣ ΙΩΝΙΑΣ-14233	61128.97	69062.7275	71512.515	72942.665	81029.04	71059.15952	3708.20081
ΝΕΑΣ ΙΩΝΙΑΣ-14234	41039.87	47227.92	48955.66	50837.025	55547.45	48815.17869	3126.194934
ΝΕΑΣ ΙΩΝΙΑΣ-14235	45071.43	51615.47464	53420.625	56633.79	62538.4	53653.34059	3632.777747
ΝΕΑΣ ΣΜΥΡΝΗΣ-17121	70785.23	86438.27925	91255.885	95051.2925	102564.33	90215.42854	7078.451091

ΝΕΑΣ ΣΜΥΡΝΗΣ-17122	19290.28	81069.3375	84202.885	88837.71	95471.41	81424.52457	14384.32788
ΝΕΑΣ ΣΜΥΡΝΗΣ-17123	70726.2	80431.4825	82210.21	85044.7375	89314.99	82103.83286	3959.04903
ΝΕΑΣ ΣΜΥΡΝΗΣ-17124	94733.19	107725.8925	110484.175	113978.37	119646.25	110591.8432	4749.013386
ΝΕΑΣ ΦΙΛΑΔΕΛΦΕΙΑΣ-14341	37781.79	44398.6875	45963.835	47849.56	51032.88	45873.61476	2519.964526
ΝΕΑΣ ΦΙΛΑΔΕΛΦΕΙΑΣ-14342	62397.77	74725.1175	77213.995	80546.8825	86322.31	77359.8044	4517.553975
ΝΕΑΣ ΧΑΛΚΗΔΟΝΑΣ-14343	39655.97	46608.2475	48573.16	50944.7725	56328.92	48863.84726	3218.054159
ΝΕΟ ΨΥΧΙΚΟ-15451	53608.57	62119.12	65352.93	69920.34	77370.48	65708.36536	5403.405408
ΝΙΚΑΙΑΣ-18451	68664.57	77248.195	81180.02	83161.665	89251.33	80496.38662	4250.407069
ΝΙΚΑΙΑΣ-18452	75688.61	84476.885	87174.105	90259.09	96344.01	87050.0275	4528.543961
ΝΙΚΑΙΑΣ-18453	48089.36	55646.785	58111.135	60720.33512	72660.53	58064.15185	4026.528277
ΝΙΚΑΙΑΣ-18454	75006.01	83530.42535	86121.14	88127.625	96586.93	85735.80191	4283.478176
ΠΑΙΑΝΙΑΣ-19002	234	327.945	764.28	1654.9375	3410.094413	1065.154629	941.8078214
ΠΑΛΛΗΝΗΣ-15351	2208.92	67354.7725	75250.815	83894.44	93381.19	70407.93667	23063.40357
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17561	69791.26	85217.6125	89556.19	95108.8775	105554.98	89619.87768	7509.408309
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17562	63491.22	75956.305	79265.25	82970.4125	91245.72	79082.8156	5627.041115
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17563	91483	107983.9025	113792.33	118877.95	146465.98	113989.2442	9717.231787
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17564	64385.05	75690.25	79022.96	81772.1225	90421.24	78519.93864	5309.670301
ΥΜΗΤΤΟΥ-17236	31092.54	37161.3975	38600.48	39953.385	43086.64	38467.79869	2270.565077
ΥΜΗΤΤΟΥ-17237	44598.52	51961.03	54027.33719	55647.3475	60381.6	53674.32374	2966.363053
ΠΕΙΡΑΙΩΣ-18450	70568.79	83098.8275	86812.06	89734.1425	100047.46	86217.1297	5523.447818
ΠΕΙΡΑΙΩΣ-18531	11146.18	12423.1175	13821.24	15122.135	18089.02	13930.76048	1704.807762
ΠΕΙΡΑΙΩΣ-18532	34980.43	38626.0575	41743.305	44659.405	59314.4	42962.37265	5822.636101
ΠΕΙΡΑΙΩΣ-18533	43595.5	49727.335	53513.145	55280.93	82961.46	53958.795	6967.433185
ΠΕΙΡΑΙΩΣ-18534	44561.64	50280.16	53742.105	56292.7975	60959.27	53449.84232	3748.268688
ΠΕΙΡΑΙΩΣ-18535	21938.55	25388.2025	28282.13	30922.255	46730.3	28895.89534	4361.230237
ΠΕΙΡΑΙΩΣ-18536	39456.61	44260.7775	47628.185	50066.4	55714.7	47237.56711	3666.694863
ΠΕΙΡΑΙΩΣ-18537	47304.2	52829.485	56463.11	59786.08	73931.3	56957.06664	5348.690563
ΠΕΙΡΑΙΩΣ-18538	39545.85	47006.53	49568.69	52197.6875	67443.36	50107.70531	5435.728423
ΠΕΙΡΑΙΩΣ-18539	83240.7	92475.295	96749.94	102555.695	123160.29	98907.53909	8898.441132
ΠΕΙΡΑΙΩΣ-18540	12762.1	15304.55177	17154.41	19719.52171	34256.53	19352.35797	6070.057077
ΠΕΙΡΑΙΩΣ-18541	60269.01	70744.925	82762.9	88344.3675	106247.89	80678.35226	11117.50369
ΠΕΙΡΑΙΩΣ-18542	30161.21	34081.9825	35333.775	36633.46	41108.99	35359.15994	2013.414214
ΠΕΙΡΑΙΩΣ-18543	46118.16	53275.0625	55593.67	57647.51	67813.96	55741.36209	4189.264567
ΠΕΙΡΑΙΩΣ-18544	35713.65	42473.195	45080.3	47275.87	123802.43	50604.68833	18922.84146
ΠΕΙΡΑΙΩΣ-18545	30336.37	36835.42	41103.105	58114.7925	73331.17	46410.7619	11672.56603
ΠΕΙΡΑΙΩΣ-18546	63362.63	72611.1625	75512.82	78363.64	115169.46	77564.70231	9376.75366
ΠΕΙΡΑΙΩΣ-18547	41795.01	58558.8975	66212.61491	83518.95	98130.29	69724.32345	14493.89229
ΠΕΡΑΜΑΤΟΣ-18863	91298.62	108121.41	113787.095	118992.5625	140801.76	113998.6425	9886.076482
ΠΕΡΑΜΑΤΟΣ-18883	9.77	19.08	23.915	31.3225	45.62	25.74455349	9.172876507

ΠΕΡΙΣΤΕΡΙΟΥ-12131	66857.74	76430.5825	80464.56	82869.015	92324.62	79773.665	5202.843399
ΠΕΡΙΣΤΕΡΙΟΥ-12132	52447.67	60047.58	63290.575	65986.065	75439.95	63228.88359	4853.595669
ΠΕΡΙΣΤΕΡΙΟΥ-12133	58718.48	66343.91	69782.025	72658.845	81683.58	69551.43988	4599.753201
ΠΕΡΙΣΤΕΡΙΟΥ-12134	52119.03	58232.3375	60320.435	62602.845	69457.4	60285.56417	3319.408604
ΠΕΡΙΣΤΕΡΙΟΥ-12135	95314.6	110236.41	114959.27	118949.205	126840.95	114220.4526	6173.293237
ΠΕΡΙΣΤΕΡΙΟΥ-12136	102691.52	116307.6275	120769.055	125550.2925	134356.63	120482.3721	6761.567668
ΠΕΡΙΣΤΕΡΙΟΥ-12137	111038.73	120443.4225	125033.905	129298.44	139717.66	124851.4196	6685.732618
ΠΕΤΡΟΥΠΟΛΕΩΣ-13231	116417.97	161973.7407	169459.46	174992.215	186127.04	165280.5445	15969.72743
ΠΕΤΡΟΥΠΟΛΕΩΣ-13232	25290.17	87337.875	91859.345	95655.1025	131699.2	86866.055	20359.61345
ΠΕΥΚΗΣ-15121	86356.53	105283.465	109170.23	116425.9	128809.76	109776.1116	8648.412779
ΣΑΛΑΜΙΝΑΣ-18900	47372.12	59755.165	69130.705	78471.9325	92677.19	69575.95064	11951.43302
ΣΑΛΑΜΙΝΑΣ-18901	11966.69	14093.06	15795.52	17303.4875	21043.53	15810.01571	2174.01173
ΣΑΛΑΜΙΝΑΣ-18902	20469.1	28246.64	34787.425	43324.7775	56838.81	36526.93343	10147.55896
ΣΑΛΑΜΙΝΑΣ-18903	15678.18	23540.9875	33341.645	43759.02	63412.76	35247.66349	14036.06101
ΤΑΥΡΟΥ-17778	65553.15	79203.0275	84130.29	117754.3375	142622.03	96656.86634	22025.33896
ΦΙΛΟΘΕΗ-15237	51312.06	58050.2425	71214.495	86026.9	104942.52	73043.40738	15725.64258
ΧΑΙΔΑΡΙ-12461	103416.88	122894.01	128398.12	132532.805	142872.77	127403.7486	8311.74437
ΧΑΙΔΑΡΙ-12462	77927.18	91825.92	99917.775	108438.1763	141358.34	100362.61	12197.06634
ΧΑΛΑΝΔΡΙ-15231	69290.71	79529.8375	85447.06	91513.8475	100485	85558.11501	7629.441097
ΧΑΛΑΝΔΡΙ-15232	74447.28	84467.365	90532.76	96974.63	133610.95	91754.4481	10975.97064
ΧΑΛΑΝΔΡΙ-15233	72549.21	81327.52	86287.11	92987.935	101983.74	86721.84057	7561.200662
ΧΑΛΑΝΔΡΙ-15234	101071.86	112001.265	120112.265	129679.075	144519.92	120844.299	11082.72536
ΧΑΛΑΝΔΡΙ-15238	27974.15	38955.01662	44218.2157	51230.0125	65515.82	44892.73415	8646.786033
ΧΑΛΑΝΔΡΙ-15344	429.55	14156.1125	15408.35	16985.9925	37769.43	17024.83556	7873.417293
ΨΥΧΙΚΟ-15452	60103.65	79671.0225	89594.02	121577.2975	148674.35	100612.9382	24536.09668
ΑΘΗΝΩΝ-10431	5041.97	6017.275516	7116.1	8268.0975	9471.63	7151.496453	1191.705888
ΑΘΗΝΩΝ-10432	9255.28	12052.4525	14311.185	15230.5025	17534.35	13782.50071	2119.46671
ΑΘΗΝΩΝ-10433	15790.77	20214.9575	22425.33	24001.7475	25572.97	21812.91152	2601.613571
ΑΘΗΝΩΝ-10434	29556.68	35999.805	39739.31	42143.6625	44299.2	38653.03305	4330.797143
ΑΘΗΝΩΝ-10435	19615.94	23981.8675	25533.545	27065.8875	35101.81	25908.54564	3483.957271
ΑΘΗΝΩΝ-10436	13019.84	14862.865	16534.22	17582.6525	19679.19	16379.03196	1695.154222
ΑΘΗΝΩΝ-10437	15900.1	19851.39097	24354.235	25768.1975	28452.6	23141.56391	3463.884079
ΑΘΗΝΩΝ-10438	19619	24253.225	26924.355	28888.1675	30903.22	26401.84667	2916.888678
ΑΘΗΝΩΝ-10439	21449.12	27045.3525	31036.79	32745.58	35547.26	29842.90484	3664.16355
ΑΘΗΝΩΝ-10440	36170.03	46615.3175	51547.77	53456.685	56923.74	49642.25317	5158.893055
ΑΘΗΝΩΝ-10441	44464.46	52083.2	54857.74	56593.365	59926.77	54235.08712	3417.835294
ΑΘΗΝΩΝ-10442	34997.86	41646.94912	43399.625	45906.8175	49888.99	43701.68676	3199.575532
ΑΘΗΝΩΝ-10443	103909.08	122264.3275	129086.03	134199.485	144243.16	127977.4426	7979.928113
ΑΘΗΝΩΝ-10444	67764.64	79745.6025	85078.38	88128.9025	94332.61	84107.01996	5311.182848
ΑΘΗΝΩΝ-10445	90605.38	109283.085	115796.18	124345.0975	131445.58	115893.1254	9305.240225
ΑΘΗΝΩΝ-10446	73373.78	91125.09	100240.0109	105554.5425	117306.91	97508.60672	10416.06747
ΑΘΗΝΩΝ-10447	7924.02	10473.3375	13031.895	15380.28748	21550.7	13252.23488	3173.037755
ΑΘΗΝΩΝ-10551	5542.52	6575.848287	7697.775	8309.655	11031.64	7552.579437	1161.739767
ΑΘΗΝΩΝ-10552	4638.61	7243.38	8283.875	9569.82	14284.51	8321.730714	2028.274891

ΑΘΗΝΩΝ-10553	5574.49	6947.845	7368.96	7733.915	8398.91	7272.381786	596.7068368
ΑΘΗΝΩΝ-10554	6381.07	7344.8425	7845.095	9190.5175	11365.85	8307.334415	1300.580971
ΑΘΗΝΩΝ-10555	2819.11	3589.69	3839.57	4108.185	5296.35	3871.54619	472.0213681
ΑΘΗΝΩΝ-10556	3670.65	4405.0725	4791.69	5115.9375	6134.63	4789.648929	578.6843083
ΑΘΗΝΩΝ-10557	8472.75	10087.975	10870.445	11741.9225	13463.25	10923.89048	1150.048687
ΑΘΗΝΩΝ-10558	8492.29	10992.3575	11779.87	13655.25	20254.9	12649.29244	2769.303602
ΑΘΗΝΩΝ-10559	4174.17	5345.1975	5919.09	6509.0325	7852.1	5901.773628	774.1765924
ΑΘΗΝΩΝ-10560	2697.16	3074.21	3234.74	3466.51	4129.34	3288.546606	297.1939073
ΑΘΗΝΩΝ-10561	1873.36	2398.51	2789.905	3157.18	3601.37	2771.770357	438.6421653
ΑΘΗΝΩΝ-10562	4541.3	4905.525	5389.915	5899.8725	6641.41	5440.402619	593.8958485
ΑΘΗΝΩΝ-10563	5708.8	6906.49	7480.425	8394.0225	9525.58	7660.58948	919.3131634
ΑΘΗΝΩΝ-10564	8758.41	12703.9925	13999.04	16302.3125	25911.16	14647.37261	3208.735838
ΑΘΗΝΩΝ-10671	8509.55	9432.8875	10147.475	11103.7225	12307.28	10293.96369	1027.577693
ΑΘΗΝΩΝ-10672	10481.6	11795.5725	12255.835	13447.92	14669.97	12500.17405	1151.583061
ΑΘΗΝΩΝ-10673	12048.41	14085.52	16134.07661	17850.985	19935.03	16006.81763	2066.165589
ΑΘΗΝΩΝ-10674	15038.82	16894.54	18039.66	19079.9125	31616.47	19052.85778	3845.412808
ΑΘΗΝΩΝ-10675	15337.53	17748.2925	18990.885	19757.34	21402.78	18821.80896	1393.005698
ΑΘΗΝΩΝ-10676	15099.14	16954.7175	17911.595	19293.3675	20682.52	18063.67868	1402.920318
ΑΘΗΝΩΝ-10677	3915.46	4811	5111.81	5783.135	6829.11	5290.52668	700.7539273
ΑΘΗΝΩΝ-10678	6043.06	6770.987399	7388.335739	8213.13	9525.34	7537.58691	879.1129842
ΑΘΗΝΩΝ-10679	3548.21	4364.11	4660.105	5107.1725	6169.39	4741.484233	612.2988458
ΑΘΗΝΩΝ-10680	13970.21886	16777.04616	18626.835	19738.88	32902.02	18906.58303	3660.091813
ΑΘΗΝΩΝ-10681	14882.04	17650.36	18944.75	19634.73038	20728.41	18594.69833	1365.316541
ΑΘΗΝΩΝ-10682	10364.14	12391.4575	13762.21	14264.595	15695.14	13401.74753	1289.554452
ΑΘΗΝΩΝ-10683	12199.88	13842.57058	14674.515	15506.175	16399.48	14624.73635	1039.027348
ΑΘΗΝΩΝ-11141	79491.31	90686.7025	96172.225	100888.085	108327.72	95766.74542	6443.661849
ΑΘΗΝΩΝ-11142	71667.68	80922.395	84650.37	88307.9325	95440.99	84614.05048	4697.602835
ΑΘΗΝΩΝ-11143	59152.79	68009.5975	70963.375	75125.7825	81488.62	71375.50524	4637.783327
ΑΘΗΝΩΝ-11144	65173.15	76579.7975	82289.755	86951.6675	92185.73	81524.78282	6314.58049
ΑΘΗΝΩΝ-11145	54398.09	62758.6175	65411.535	69821.0725	73359.25	65886.00008	4349.070874
ΑΘΗΝΩΝ-11251	39190.35	50300.925	56033.15	58257.0125	63269.53	53881.09852	6066.703426
ΑΘΗΝΩΝ-11252	37072.37	46002.42	49531.525	51777.6175	55927.05	48405.23159	4628.919246
ΑΘΗΝΩΝ-11253	40458.98	50419.52	54887.25	57697.685	72387.24	53550.56606	5918.462678
ΑΘΗΝΩΝ-11254	36526.06	43677.1225	46840.84	49374.265	51695.6	46075.07488	4038.765778
ΑΘΗΝΩΝ-11255	42234.69	49766.02	53665.715	55862.7575	60588.67	52710.67972	4359.922817
ΑΘΗΝΩΝ-11256	21141.45	25550.0425	28477.875	30031.23	38013.09	28594.00893	4182.153058
ΑΘΗΝΩΝ-11257	31924.45	37959.935	41867.795	44011.145	47150.95	40812.8194	3979.880559
ΑΘΗΝΩΝ-11361	34142.67	39892.915	43394.225	45464.86	49016.52	42556.83274	3717.989708
ΑΘΗΝΩΝ-11362	46621.6	54563	60417.2	63202.385	70264.46	58931.92578	5761.027627
ΑΘΗΝΩΝ-11363	81321.09	91335.97	99304.855	104034.6775	114250.66	98042.78762	7589.290203
ΑΘΗΝΩΝ-11364	75531.2	86055.415	93449.35	97682.4475	107263.21	92221.94277	7127.722728
ΑΘΗΝΩΝ-11471	35783.99	40212.9975	42616.82	45139.545	50211.08	42841.88792	3477.073506
ΑΘΗΝΩΝ-11472	29949.11	32534.3075	34598.775	36518.035	43186.34	34793.39738	2848.657933
ΑΘΗΝΩΝ-11473	36059.66	41798.78	45718.515	48188.76	59957.67701	45600.19008	4984.796282
ΑΘΗΝΩΝ-11474	32699.21	57625.445	61463.83	65119.115	70926.68	59625.56628	8901.233164

ΑΘΗΝΩΝ-11475	22786.91	37090.475	39480.135	41817.46	54664.07364	39189.19716	5270.239068
ΑΘΗΝΩΝ-11476	15834.61	40131.77	42654.335	44932.2875	62414.20071	41665.91232	7824.64502
ΑΘΗΝΩΝ-11521	29521.22	34069.9825	36279.26	38344.0175	42585.59	36341.00681	2725.358807
ΑΘΗΝΩΝ-11522	44549.65	49334.2525	51302.685	54140.5775	59183.22	51521.09357	3055.896338
ΑΘΗΝΩΝ-11523	53293.34	59662.1375	62544.585	65953.145	92843.38	63854.18832	6804.029389
ΑΘΗΝΩΝ-11524	57113.28	70726.535	73392.58	77662.055	85886.41	73465.8388	5830.805037
ΑΘΗΝΩΝ-11525	49008.48	53919.03	56043.675	58108.8525	63590.27	55983.8206	2999.547751
ΑΘΗΝΩΝ-11526	65346.9	72456.475	75684.06	80276.28	118022.7882	78127.48857	9962.574307
ΑΘΗΝΩΝ-11527	54499.57	61954.5625	65454.76	68655.1625	75073.7	65224.34524	4041.946223
ΑΘΗΝΩΝ-11528	33931.29	41790.7	46279.83	49756.7825	63681.12	46760.55988	7346.154546
ΑΘΗΝΩΝ-11631	39932.4	43655.155	45823.655	47112.7125	61639.84	46436.68946	4710.616042
ΑΘΗΝΩΝ-11632	42849.72	47490.3625	49583.39	51151.06	54525.9	49214.77393	2426.709354
ΑΘΗΝΩΝ-11633	53186.6	61673.2275	65699	68356.075	101472.57	66209.29401	7553.033487
ΑΘΗΝΩΝ-11634	45043.68	53473.36	56682.97	63164.5675	90270.19	60546.83718	11663.87103
ΑΘΗΝΩΝ-11635	45941.8	51468.72493	56406.80085	58504.455	67145.79	55755.95445	4559.974401
ΑΘΗΝΩΝ-11636	32759.03	37309.84	38807.17	40570.735	43284.03	38914.01595	2217.970093
ΑΘΗΝΩΝ-11741	44846.41	52915.205	55563.945	58070.71	82735.84	56675.13673	6987.438296
ΑΘΗΝΩΝ-11742	29436.38	34703.9225	36084.055	37468.98	47730.96	36745.25474	3524.332165
ΑΘΗΝΩΝ-11743	70092.45	78583.9075	81069.9	84958.185	107880.64	83956.32417	9122.773822
ΑΘΗΝΩΝ-11744	68527.76	75948.82	79987.225	103303.9425	135292.29	90295.20946	17204.83082
ΑΘΗΝΩΝ-11745	54092.2	66636.1125	71115.43	76161.51	114176.09	74015.54753	12244.94241
ΑΘΗΝΩΝ-11851	38394.36	43301.11	46408.835	48197.52	53654.7	45841.8168	3085.780529
ΑΘΗΝΩΝ-11852	32645.62	37461.3525	39381.98	41822.325	47872.79	39675.6104	3510.500889
ΑΘΗΝΩΝ-11853	51656.54	58256.4775	61630.37	65074.24	102247	65392.80855	12085.76155
ΑΘΗΝΩΝ-11854	26723.41	29054.31	30910.345	33291.9325	38770.6	31385.54637	2761.838173
ΑΘΗΝΩΝ-11855	20082.14	23555.66	25250.11	27061.895	30855.41	25260.2571	2288.223505
ΧΟΛΑΡΓΟΥ-15562	68038.47	84249.88	89150.765	95497.6525	104352.73	88681.10489	8494.250801
ΣΠΑΤΩΝ-19019	808.90	17118.75	23752.11143	33170.11	41760	25165.11702	10259.12669
ΠΑΠΑΓΟΥ-15669	62860.09	78581.17	87867.875	95270.6125	111071.45	87455.81376	11636.56653

Πίνακας Π-3: Τα δεδομένα της διπλωματικής για την ζήτηση νερού συγκεντρωτικά

Στην επόμενη σελίδα παρουσιάζεται ο μέσος όρος του αριθμού των μετρητών ανά περιοχή. Προτιμήσαμε να παρουσιάσουμε τον μέσο όρο των τιμών και όχι την εξέλιξη τους σε βάθος χρόνου κατά την διάρκεια των 84 μηνών των διαθέσιμων μετρήσεων που είχαμε στα δεδομένα μας, γιατί εκτός των πλεονεκτημάτων από άποψη χώρου(θα χρειαζόμασταν ένα πίνακα 274×84), πολύ απλά οι διακυμάνσεις γύρω από την μέση τιμή είναι πολύ μικρές και πρακτικά μπορούν να αντιπροσωπευθούν από το μέσο όρο των τιμών τους για την περίοδο 2007-2013. Επομένως, κάποιος ο οποίος θέλει να καταλάβει τα μεγέθη των μετρητών ανά περιοχή, θα μπορέσει μέσω του επόμενου πίνακα να έχει μία σαφή εικόνα για αυτά.

Περιοχή	Μ.Ο μετρητών	Περιοχή	Μ.Ο μετρητών
ΖΩΓΡΑΦΟΥ-15771	15778.45238	ΒΥΡΩΝΟΣ-16233	6469.452381
ΖΩΓΡΑΦΟΥ-15772	19580.29762	ΓΑΛΑΤΣΙΟΥ-11146	21508.46429
ΖΩΓΡΑΦΟΥ-15773	14579.72619	ΓΑΛΑΤΣΙΟΥ-11147	12321.96429
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13561	7522.988095	ΓΛΥΦΑΔΑΣ-16561	13471.03571
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13562	8353.52381	ΓΛΥΦΑΔΑΣ-16562	9984.928571
ΑΓ.ΒΑΡΒΑΡΑΣ-12351	11327.67857	ΓΛΥΦΑΔΑΣ-16674	15577.63095
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17341	10459.70238	ΓΛΥΦΑΔΑΣ-16675	11668.70238
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17342	12865.11905	ΔΑΦΝΗΣ-17234	4548.297619
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17343	9636.666667	ΔΑΦΝΗΣ-17235	5056.071429
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17372	3.404761905	ΔΡΑΠΕΤΣΩΝΑΣ-18648	7441.011905
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18232	10	ΕΛΕΥΣΙΝΟΣ-19200	11498.29762
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18233	8570.142857	ΕΛΛΗΝΙΚΟΥ-16777	8482.690476
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15341	11961.17857	ΖΕΦΥΡΙΟΥ-13461	3098.27381
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15342	8426.404762	ΗΛΙΟΥΠΟΛΕΩΣ-16341	7028.190476
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15343	12178.03571	ΗΛΙΟΥΠΟΛΕΩΣ-16342	6236.964286
ΑΙΓΆΛΕΩ-12241	7052.595238	ΗΛΙΟΥΠΟΛΕΩΣ-16343	5885.630952
ΑΙΓΆΛΕΩ-12242	7667.404762	ΗΛΙΟΥΠΟΛΕΩΣ-16344	5154.154762
ΑΙΓΆΛΕΩ-12243	11217	ΗΛΙΟΥΠΟΛΕΩΣ-16345	7757.619048
ΑΙΓΆΛΕΩ-12244	9075.928571	ΗΛΙΟΥΠΟΛΕΩΣ-16346	8029.166667
ΆΛΙΜΟΣ-17436	4	ΗΡΑΚΛΕΙΟΥ-14121	12444.19048
ΆΛΙΜΟΣ-17455	14788.88095	ΗΡΑΚΛΕΙΟΥ-14122	12962.47619
ΆΛΙΜΟΣ-17456	8687.285714	ΚΑΤΖΑ-19012	0.988095238
ΑΜΑΡΟΥΣΙΟΥ-15122	8788.511905	ΙΛΙΟΥ-13121	12388.53571
ΑΜΑΡΟΥΣΙΟΥ-15123	2953.535714	ΙΛΙΟΥ-13122	15250.57143
ΑΜΑΡΟΥΣΙΟΥ-15124	9845.5	ΙΛΙΟΥ-13123	11160.55952
ΑΜΑΡΟΥΣΙΟΥ-15125	9825.261905	ΚΑΙΣΑΡΙΑΝΗΣ-16121	14410.02381
ΑΜΑΡΟΥΣΙΟΥ-15126	9239.416667	ΚΑΙΣΑΡΙΑΝΗΣ-16122	4376.940476
ΆΝΩ ΛΙΟΣΙΑ-13341	12.78571429	ΚΑΛΛΙΘΕΑΣ-17671	10776.58333
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16451	8957.77381	ΚΑΛΛΙΘΕΑΣ-17672	12661.91667
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16452	8408.714286	ΚΑΛΛΙΘΕΑΣ-17673	15956.36905
ΑΣΠΡΟΠ'ΥΡΓΟΥ-19300	12007.28571	ΚΑΛΛΙΘΕΑΣ-17674	7187.25
ΑΧΑΡΝΩΝ-13671	5979.22619	ΚΑΛΛΙΘΕΑΣ-17675	11656.5119
ΑΧΑΡΝΩΝ-13672	5327.369048	ΚΑΛΛΙΘΕΑΣ-17676	11700.66667
ΑΧΑΡΝΩΝ-13673	4409.797619	ΚΑΜΑΤΕΡΟΥ-13451	12012.02381
ΑΧΑΡΝΩΝ-13674	9287.238095	ΚΕΡΑΤΕΑΣ-19001	307.0357143
ΑΧΑΡΝΩΝ-13675	4720.452381	ΚΕΡΑΤΣΙΝΙ-18755	9298.285714
ΑΧΑΡΝΩΝ-13677	3227.738095	ΚΕΡΑΤΣΙΝΙ-18756	8069.666667
ΑΧΑΡΝΩΝ-13678	2458.642857	ΚΕΡΑΤΣΙΝΙ-18757	12440.03571
ΑΧΑΡΝΩΝ-13679	1795.833333	ΚΕΡΑΤΣΙΝΙ-18758	7525.52381
ΒΆΡΗΣ-16672	10883.71429	ΚΗΦΙΣΙΑΣ-14561	8798.452381
ΒΟΥΛΑΣ-16673	17732.86905	ΚΗΦΙΣΙΑΣ-14562	6920.035714
ΒΟΥΛΙΑΓΜΕΝΗΣ-16671	4756.178571	ΚΗΦΙΣΙΑΣ-14563	3129.904762

ΒΡΙΑΗΣΣΙΩΝ-15235	9630.059524	ΚΗΦΙΣΙΑΣ-14564	7843.011905
ΒΥΡΩΝΟΣ-16231	14798.04762	ΚΗΦΙΣΙΑΣ-14671	4499.964286
ΒΥΡΩΝΟΣ-16232	14087.79762	ΚΟΥΡΥΔΑΛΛΟΥ-18120	16114.90476
ΚΟΥΡΥΔΑΛΛΟΥ-18121	7349.77381	ΥΜΗΤΤΟΥ-17236	4030.72619
ΚΟΥΡΥΔΑΛΛΟΥ-18122	6699.595238	ΥΜΗΤΤΟΥ-17237	6339.02381
ΚΡΩΠΙΑΣ-19010	428.8452381	ΠΕΙΡΑΙΩΣ-18450	10429.21429
ΚΡΩΠΙΑΣ-19013	96.63095238	ΠΕΙΡΑΙΩΣ-18531	1865.035714
ΚΡΩΠΙΑΣ-19400	474.9761905	ΠΕΙΡΑΙΩΣ-18532	6709.714286
ΚΟΙΝΟΤΗΤΑ ΓΛΥΚΩΝ ΝΕΡΩΝ-15354	2488.988095	ΠΕΙΡΑΙΩΣ-18533	6300.297619
ΚΟΙΝΟΤΗΤΑ ΘΡΑΚΟΜΑΚΕΔΟΝΩΝ-13676	2924.571429	ΠΕΙΡΑΙΩΣ-18534	7550.535714
ΚΟΙΝΟΤΗΤΑ ΠΑΛΑΙΑΣ ΠΕΝΤΕΛΗΣ-19009	12.5952381	ΠΕΙΡΑΙΩΣ-18535	4534.047619
ΛΑΥΡΕΩΤΙΚΗΣ-19500	1.095238095	ΠΕΙΡΑΙΩΣ-18536	7649.630952
ΛΥΚΟΒΡΥΣΕΩΣ-14123	5113.559524	ΠΕΙΡΑΙΩΣ-18537	8430.392857
ΜΑΝΔΡΑΣ-19600	2.416666667	ΠΕΙΡΑΙΩΣ-18538	6399.928571
ΜΑΡΑΘΩΝΑ-19007	9.119047619	ΠΕΙΡΑΙΩΣ-18539	12665.21429
ΜΑΡΚΟΠΟΥΛΟΥ-19003	9.869047619	ΠΕΙΡΑΙΩΣ-18540	2118.666667
ΜΕΓΑΡΕΩΝ-19100	493.6428571	ΠΕΙΡΑΙΩΣ-18541	8063.845238
ΜΕΛΙΣΣΙΩΝ-11527	10991.71429	ΠΕΙΡΑΙΩΣ-18542	4314.238095
ΜΕΤΑΜΟΡΦΩΣΕΩΣ-14451	6378.547619	ΠΕΙΡΑΙΩΣ-18543	6462.595238
ΜΕΤΑΜΟΡΦΩΣΕΩΣ-14452	8245.345238	ΠΕΙΡΑΙΩΣ-18544	5243.5
ΜΟΣΧΑΤΟΥ-18344	5666.821429	ΠΕΙΡΑΙΩΣ-18545	4498.630952
ΜΟΣΧΑΤΟΥ-18345	7761.988095	ΠΕΙΡΑΙΩΣ-18546	9408.333333
ΜΟΣΧΑΤΟΥ-18346	737.5	ΠΕΙΡΑΙΩΣ-18547	7020.761905
ΝΕΑΣ ΠΕΡΑΜΟΥ-19006	4680.619048	ΠΕΡΑΜΑΤΟΣ-18863	11854.66667
ΝΕΑΣ ΙΩΝΙΑΣ-14231	10920.28571	ΠΕΡΑΜΑΤΟΣ-18883	6.285714286
ΝΕΑΣ ΙΩΝΙΑΣ-14232	6814.892857	ΠΕΡΙΣΤΕΡΙΟΥ-12131	8438.02381
ΝΕΑΣ ΙΩΝΙΑΣ-14233	8119.75	ΠΕΡΙΣΤΕΡΙΟΥ-12132	6399.452381
ΝΕΑΣ ΙΩΝΙΑΣ-14234	5238.083333	ΠΕΡΙΣΤΕΡΙΟΥ-12133	7151.404762
ΝΕΑΣ ΙΩΝΙΑΣ-14235	5802.166667	ΠΕΡΙΣΤΕΡΙΟΥ-12134	7037.845238
ΝΕΑΣ ΣΜΥΡΝΗΣ-17121	12673.42857	ΠΕΡΙΣΤΕΡΙΟΥ-12135	12097.16667
ΝΕΑΣ ΣΜΥΡΝΗΣ-17122	10568.14286	ΠΕΡΙΣΤΕΡΙΟΥ-12136	12226.32143
ΝΕΑΣ ΣΜΥΡΝΗΣ-17123	10288.84524	ΠΕΡΙΣΤΕΡΙΟΥ-12137	12361.90476
ΝΕΑΣ ΣΜΥΡΝΗΣ-17124	13178.41667	ΠΕΤΡΟΥΠΟΛΕΩΣ-13231	17984.58333
ΝΕΑΣ ΦΙΛΑΔΕΛΦΕΙΑΣ-14341	5679.130952	ΠΕΤΡΟΥΠΟΛΕΩΣ-13232	9067.964286
ΝΕΑΣ ΦΙΛΑΔΕΛΦΕΙΑΣ-14342	8773.714286	ΠΕΥΚΗΣ-15121	11240.85714
ΝΕΑΣ ΧΑΛΚΗΔΟΝΑΣ-14343	5658.619048	ΣΑΛΑΜΙΝΑΣ-18900	6738.452381
ΝΕΟ ΨΥΧΙΚΟ-15451	6861.119048	ΣΑΛΑΜΙΝΑΣ-18901	1550.583333
ΝΙΚΑΙΑΣ-18451	9311.130952	ΣΑΛΑΜΙΝΑΣ-18902	3434.666667
ΝΙΚΑΙΑΣ-18452	9355.642857	ΣΑΛΑΜΙΝΑΣ-18903	3789.02381
ΝΙΚΑΙΑΣ-18453	7136.97619	ΤΑΥΡΟΥ-17778	8294.154762
ΝΙΚΑΙΑΣ-18454	10317.32143	ΦΙΛΟΘΕΗ-15237	4187.404762
ΠΑΙΑΝΙΑΣ-19002	0.928571429	ΧΑΙΔΑΡΙ-12461	12944.27381
ΠΑΛΛΗΝΗΣ-15351	7509.809524	ΧΑΙΔΑΡΙ-12462	8411.02381
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17561	10607.46429	ΧΑΛΑΝΔΡΙ-15231	8225.547619

ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17562	8998.178571	ΧΑΛΑΝΔΡΙ-15232	9271.047619
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17563	11853.19048	ΧΑΛΑΝΔΡΙ-15233	8932.75
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17564	8038.190476	ΧΑΛΑΝΔΡΙ-15234	12570.89286
ΧΑΛΑΝΔΡΙ-15344	1849.714286	ΧΑΛΑΝΔΡΙ-15238	3544.833333
ΨΥΧΙΚΟ-15452	6037.5	ΑΘΗΝΩΝ-10683	2893.238095
ΑΘΗΝΩΝ-10431	655.9642857	ΑΘΗΝΩΝ-11141	13164.54762
ΑΘΗΝΩΝ-10432	1633.845238	ΑΘΗΝΩΝ-11142	10394.92857
ΑΘΗΝΩΝ-10433	3747.797619	ΑΘΗΝΩΝ-11143	9336.72619
ΑΘΗΝΩΝ-10434	6061.690476	ΑΘΗΝΩΝ-11144	12563.16667
ΑΘΗΝΩΝ-10435	3271.309524	ΑΘΗΝΩΝ-11145	8898.845238
ΑΘΗΝΩΝ-10436	2289.821429	ΑΘΗΝΩΝ-11251	8570.535714
ΑΘΗΝΩΝ-10437	2704.928571	ΑΘΗΝΩΝ-11252	7509.47619
ΑΘΗΝΩΝ-10438	3892.857143	ΑΘΗΝΩΝ-11253	8275.559524
ΑΘΗΝΩΝ-10439	4376.559524	ΑΘΗΝΩΝ-11254	7102.238095
ΑΘΗΝΩΝ-10440	7200.988095	ΑΘΗΝΩΝ-11255	7775.654762
ΑΘΗΝΩΝ-10441	7577.571429	ΑΘΗΝΩΝ-11256	4468.654762
ΑΘΗΝΩΝ-10442	5042.5	ΑΘΗΝΩΝ-11257	6797.702381
ΑΘΗΝΩΝ-10443	16417.35714	ΑΘΗΝΩΝ-11361	6886.869048
ΑΘΗΝΩΝ-10444	11858.75	ΑΘΗΝΩΝ-11362	9694.22619
ΑΘΗΝΩΝ-10445	16918.59524	ΑΘΗΝΩΝ-11363	14475.94048
ΑΘΗΝΩΝ-10446	14665.36905	ΑΘΗΝΩΝ-11364	14025.96429
ΑΘΗΝΩΝ-10447	1059.964286	ΑΘΗΝΩΝ-11471	7006.738095
ΑΘΗΝΩΝ-10551	847.5952381	ΑΘΗΝΩΝ-11472	5912.02381
ΑΘΗΝΩΝ-10552	839.3333333	ΑΘΗΝΩΝ-11473	7631.547619
ΑΘΗΝΩΝ-10553	947.6428571	ΑΘΗΝΩΝ-11474	10373.40476
ΑΘΗΝΩΝ-10554	1124.035714	ΑΘΗΝΩΝ-11475	6325.595238
ΑΘΗΝΩΝ-10555	326.8095238	ΑΘΗΝΩΝ-11476	5321.130952
ΑΘΗΝΩΝ-10556	580.2857143	ΑΘΗΝΩΝ-11521	6090.880952
ΑΘΗΝΩΝ-10557	1301.642857	ΑΘΗΝΩΝ-11522	7727.404762
ΑΘΗΝΩΝ-10558	1565.392857	ΑΘΗΝΩΝ-11523	9888.297619
ΑΘΗΝΩΝ-10559	350.3452381	ΑΘΗΝΩΝ-11524	9975.309524
ΑΘΗΝΩΝ-10560	516.4880952	ΑΘΗΝΩΝ-11525	7956.321429
ΑΘΗΝΩΝ-10561	252.25	ΑΘΗΝΩΝ-11526	12415.36905
ΑΘΗΝΩΝ-10562	524.7380952	ΑΘΗΝΩΝ-11527	10991.71429
ΑΘΗΝΩΝ-10563	913.7619048	ΑΘΗΝΩΝ-11528	6757.559524
ΑΘΗΝΩΝ-10564	477.8333333	ΑΘΗΝΩΝ-11631	6297.559524
ΑΘΗΝΩΝ-10671	1781.297619	ΑΘΗΝΩΝ-11632	6843.083333
ΑΘΗΝΩΝ-10672	2287.869048	ΑΘΗΝΩΝ-11633	10953.92857
ΑΘΗΝΩΝ-10673	2852.416667	ΑΘΗΝΩΝ-11634	9203.595238
ΑΘΗΝΩΝ-10674	2075.535714	ΑΘΗΝΩΝ-11635	9636.404762
ΑΘΗΝΩΝ-10675	2905.5	ΑΘΗΝΩΝ-11636	5881.202381
ΑΘΗΝΩΝ-10676	3137.345238	ΑΘΗΝΩΝ-11741	8545.845238
ΑΘΗΝΩΝ-10677	745.2738095	ΑΘΗΝΩΝ-11742	5135.380952
ΑΘΗΝΩΝ-10678	1129.607143	ΑΘΗΝΩΝ-11743	11728.35714
ΑΘΗΝΩΝ-10679	675.75	ΑΘΗΝΩΝ-11744	10280.52381

ΑΘΗΝΩΝ-10680	3567.404762	ΑΘΗΝΩΝ-11745	8927.416667
ΑΘΗΝΩΝ-10681	3480.988095	ΑΘΗΝΩΝ-11851	6624.488095
ΑΘΗΝΩΝ-10682	2811.77381	ΑΘΗΝΩΝ-11852	4869.428571
ΑΘΗΝΩΝ-11854	3530.178571	ΑΘΗΝΩΝ-11853	8068.642857
ΑΘΗΝΩΝ-11855	2934.857143	ΣΠΑΤΩΝ-19019	1.011904762
ΧΟΛΑΡΓΟΥ-15562	9630.630952	ΠΑΠΑΓΟΥ-15669	7124.619048

Πίνακας Π-4: Μέσος όρος μετρητών ανά ταχυδρομικό κώδικα

Στον παρακάτω πίνακα παρουσιάζεται η αντίστοιχη μορφή του πίνακα 5.2.1 αλλά για τα αρχικά-αφιλτράριστα δεδομένα. Παρόλο που τα κάθε cluster σε αυτούς τους δύο πίνακες δεν είναι συγκρίσιμα γιατί εμπεριέχουν διαφορετικά cluster (βλ εικόνα 5.1.1 και 5.1.2) εντούτοις θεωρήσαμε σημαντικό να παρουσιάσαμε τον παρακάτω πίνακα για λόγους πληρότητας.

	Ets MAPE	Arima MAPE	Theta MAPE	Opt.Theta MAPE	MAPA MAPE	Naive MAPE	Comb.Method MAPE
cluster 1	0.624453117	2.640342701	0.579604149	0.577970679	1.014137389	3.811737796	0.58907418
cluster 2	0.855324567	3.670255166	0.837746066	0.841217613	1.839581683	4.011224896	0.839204291
cluster 3	3.805418874	4.162233582	3.733611505	3.74211341	5.980353433	8.522541355	3.751369351
cluster 4	14.4324575	12.72801294	16.93333448	16.1463706	20.34086404	16.4546447	13.29015773
cluster 5	68.46949842	81.47052213	62.6900126	57.23646329	73.23693705	58.0985821	58.76056047
cluster 6	18.0495794	8.604254982	216.6362962	17.62216919	151.8786926	17.88670518	7.713904398
cluster 7	Inf	Inf	Inf	Inf	Inf	Inf	Inf
cluster 8	10.85757908	10.36357169	10.37800371	10.35003177	15.02187534	14.10593663	9.592703847
cluster 9	8.972126162	6.984475596	8.839756099	10.08901195	8.238779885	8.334758638	6.500948481
cluster 10	10.09323941	7.667823465	9.984406182	10.67702063	12.81009091	18.43531298	8.78298364

Πίνακας Π-5: Πίνακας σφαλμάτων mape(%) για κάθε cluster και για όλες τις εξεταζόμενες μεθόδους στο διάστημα A για τα αφιλτράριστα δεδομένα

Παρακάτω παρουσιάζεται ο πλήρης πίνακας σφαλμάτων MAPE για κάθε μέθοδο και διαθέσιμο ταχυδρομικό κώδικα, σαν επέκταση του πίνακα 5.2.3 της διπλωματικής.

ΖΩΓΡΆΦΟΥ-15771	1.61848339	1.734387776	1.634034742	1.620040023	1.97138983	4.221709104	1.621505158
ΖΩΓΡΆΦΟΥ-15772	2.221561589	3.093429185	3.697291304	3.696082324	2.836125468	3.904713494	2.559958432
ΖΩΓΡΆΦΟΥ-15773	1.163453246	1.631852579	1.117747092	1.117747092	1.290331149	3.641627412	1.128242067
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13561	3.516242937	4.287736613	5.385222797	5.403853657	5.8086943	5.485514013	4.012003564
ΑΓ.ΑΝΑΡΓΥΡΩΝ-13562	0.656210395	0.588208487	0.672653122	0.674114783	1.400824732	4.298569683	0.562196734
ΑΓ.ΒΑΡΒΆΡΑΣ-12351	1.589227743	2.085489189	1.606662382	1.625369556	1.97145447	4.322195098	1.600038781

ΑΓ.ΔΗΜΗΤΡΙΟΥ-17341	1.846905937	2.911335601	4.388388173	4.4608721	2.436018618	4.502997328	2.17483684
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17342	1.288380446	1.156758563	1.318150995	1.311825428	1.848195164	4.0638337	1.202368309
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17343	1.557599268	1.642909301	3.74068139	3.772272291	2.11057641	4.143462889	1.455696879
ΑΓ.ΔΗΜΗΤΡΙΟΥ-17372	7.002566375	7.319136218	7.597034969	7.477701898	14.91710498	7.590647673	6.865678048
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18232	3.108055243	3.147895021	3.066409019	3.066315703	5.602786832	8.279740157	3.06378171
ΑΓ.ΙΩΑΝΝΗΣ ΡΕΝΤΗ-18233	1.597491916	1.529724593	1.573870502	1.573113403	1.892995151	4.074382274	1.489731471
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15341	2.203248249	2.33033364	2.149119519	2.147687866	3.239522235	5.478771953	2.164999311
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15342	2.631999009	4.295970545	2.662286883	2.70015146	3.944594359	6.208785267	2.631271555
ΑΓ.ΠΑΡΑΣΚΕΥΗ-15343	2.905648713	4.354130785	2.921322154	2.922076217	3.855306803	5.829445865	2.913873311
ΑΙΓΑΛΕΩ-12241	8.011146975	7.648522824	7.889654456	7.922416754	7.345232953	7.76815222	7.541085807
ΑΙΓΑΛΕΩ-12242	2.302889778	2.28845201	4.431207695	4.43213583	3.189760165	4.755687636	2.139286025
ΑΙΓΑΛΕΩ-12243	1.639063733	1.487438861	3.722908877	3.722908104	2.308954893	4.20247033	1.49160477
ΑΙΓΑΛΕΩ-12244	1.027035491	2.929373037	1.012017557	1.010181445	1.677115743	4.263157732	1.012038592
ΆΛΙΜΟΣ-17436	11.38719105	13.55973937	10.97087771	10.99103287	14.77007943	18.76732452	10.94041723
ΆΛΙΜΟΣ-17455	2.122275468	3.04653182	2.140661549	2.067507744	3.031069184	5.407521251	2.106162831
ΆΛΙΜΟΣ-17456	2.536309972	2.687525459	2.658730344	2.659255073	3.379511253	5.602624247	2.610972179
ΑΜΑΡΟΥΣΙΟΥ-15122	1.978027523	1.647126586	1.996876103	1.998373412	3.059248999	5.839743584	1.776850377
ΑΜΑΡΟΥΣΙΟΥ-15123	3.592692117	3.14733379	3.370506145	3.371337234	4.262063863	6.249263601	3.204345956
ΑΜΑΡΟΥΣΙΟΥ-15124	3.526901343	3.409230308	3.54043817	3.555289594	4.310300454	7.339032526	3.311294247
ΑΜΑΡΟΥΣΙΟΥ-15125	2.542085373	3.474305866	2.486031874	2.49620389	3.48379529	5.791442146	2.452838839
ΑΜΑΡΟΥΣΙΟΥ-15126	2.465158087	2.238033002	2.473996937	2.473694955	3.336916019	5.947971403	2.24325321
ΆΝΘ ΛΙΘΟΣΙΑ-13341	15.90597117	10.83196	16.3655859	20.49404987	15.31309851	12.02588652	12.65860346
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16451	2.159940653	2.229270509	2.328440506	2.329537873	3.003643993	5.371084834	2.079827552
ΑΡΓΥΡΟΥΠΟΛΕΩΣ-16452	1.633425066	1.368010482	1.794888075	1.824414017	2.768805598	5.373689001	1.483962719
ΑΣΠΡΟΠ'ΥΡΓΟΥ-19300	8.417316456	9.870044254	8.544696184	8.516677091	8.814347975	10.77777013	8.364986628
ΑΧΑΡΝΩΝ-13671	2.031881264	3.010655092	1.987443185	1.956366185	3.13646404	5.655851647	1.980749687
ΑΧΑΡΝΩΝ-13672	2.09176774	2.703322434	2.064945189	2.109879115	3.060234005	5.782547645	2.087450905
ΑΧΑΡΝΩΝ-13673	1.447935787	1.629571774	1.526998913	1.5447222	2.274396929	5.018130966	1.426301422
ΑΧΑΡΝΩΝ-13674	2.123722627	2.923242566	2.141471393	2.140710296	2.617449942	5.190392617	2.134999954
ΑΧΑΡΝΩΝ-13675	2.176037282	2.039253384	2.093528674	2.094479741	3.366327265	7.192975264	1.983706212
ΑΧΑΡΝΩΝ-13677	2.681525971	3.32359101	2.803358184	2.721610038	3.578776422	6.909739608	2.712378958
ΑΧΑΡΝΩΝ-13678	2.921288067	3.612749689	2.919925833	2.921606194	3.405397554	6.317212998	2.915102832
ΑΧΑΡΝΩΝ-13679	5.023830292	4.370734776	4.541176142	4.720600485	6.287386025	9.244436262	4.347082547
Β'ΑΡΗΣ-16672	8.580732078	10.85231918	8.071338789	8.075878808	10.36478136	13.05893818	8.168024317
ΒΟΥΛΑΣ-16673	3.883884578	4.575134911	3.696335278	3.697560153	6.623319854	10.59195455	3.387986417
ΒΟΥΛΙΑΓΜΕΝΗΣ-16671	12.1430064	14.83336559	10.93765493	10.95187968	15.54678546	17.54603943	11.06585566
ΒΡΙΛΗΣΣΙΩΝ-15235	17.45474681	23.98159614	18.66851139	18.97450483	62.7373235	18.45676235	17.69207391
ΒΥΡΩΝΟΣ-16231	1.145543544	0.980657746	1.143352479	1.151190138	1.63411137	4.320893779	1.012941368
ΒΥΡΩΝΟΣ-16232	1.537223442	1.409380661	1.51208491	1.522470785	2.09236145	4.699530685	1.349528497

ΒΥΡΩΝΟΣ-16233	2.067094163	1.646171834	1.859168057	1.848015373	3.041814611	5.227514589	1.658778556
ΓΑΛΑΤΣΙΟΥ-11146	1.357961643	1.191109676	1.338546797	1.338546797	1.768306385	4.477362805	1.211188931
ΓΑΛΑΤΣΙΟΥ-11147	1.675603344	1.626373955	1.656688849	1.650227243	2.095121909	4.569432844	1.553822211
ΓΛΥΦΑΔΑΣ-16561	2.518461339	2.303678792	2.306008674	2.255599933	3.9599937	6.737127379	2.202073578
ΓΛΥΦΑΔΑΣ-16562	2.885120727	3.65038396	2.844601704	2.852503277	3.844431198	6.077476722	2.837271123
ΓΛΥΦΑΔΑΣ-16674	5.84391089	8.286143565	8.752061866	9.059738294	8.249720398	8.90530387	6.823875068
ΓΛΥΦΑΔΑΣ-16675	3.912792595	4.010727776	3.885146355	3.91095176	5.76259248	7.550617735	3.888652974
ΔΑΦΝΗΣ-17234	1.329846638	1.116835203	1.369970551	1.384580361	1.814814319	4.311179638	1.198331894
ΔΑΦΝΗΣ-17235	1.224708808	1.351660744	1.196216522	1.205125316	1.703851109	3.978314251	1.207049058
ΔΡΑΠΕΤΣΩΝΑΣ-18648	1.000647388	1.493395109	0.930372373	0.929753969	1.264779157	3.894799253	0.946135396
ΕΛΕΥΣΙΝΟΣ-19200	8.103938154	8.10257033	8.096867268	8.091215229	6.80109248	8.239345935	7.639378875
ΕΛΛΗΝΙΚΟΥ-16777	2.692426605	2.401519573	2.742136997	2.752698012	3.710082475	6.417254431	2.55833786
ΖΕΦΥΡΙΟΥ-13461	1.758268071	2.212817843	1.742047083	1.789112776	2.67330185	5.715925294	1.749389429
ΗΛΙΟΥΠΟΛΕΩΣ-16341	2.247269641	2.100049655	2.216973537	2.186922793	2.817029985	5.176213182	2.002937046
ΗΛΙΟΥΠΟΛΕΩΣ-16342	1.922736574	1.415311233	1.914653521	1.816484828	2.633767727	5.215743622	1.607698725
ΗΛΙΟΥΠΟΛΕΩΣ-16343	1.195088328	2.275480517	1.178925057	1.177065348	1.504021922	4.210360231	1.180652074
ΗΛΙΟΥΠΟΛΕΩΣ-16344	1.366639163	2.190751446	1.390221131	1.37260829	1.687816544	4.237151105	1.372478001
ΗΛΙΟΥΠΟΛΕΩΣ-16345	2.328154441	2.311031232	5.005428181	5.007058207	3.416820444	5.025306675	2.304587611
ΗΛΙΟΥΠΟΛΕΩΣ-16346	1.881869333	2.88464729	1.944975858	1.950425361	2.340402724	4.14263452	1.909412332
ΗΡΑΚΛΕΙΟΥ-14121	1.959101128	1.598034592	1.987277779	1.958584874	3.021241147	5.222906477	1.779222214
ΗΡΑΚΛΕΙΟΥ-14122	1.75509379	1.589607578	1.715526562	1.720874314	2.565633685	4.961029147	1.479473388
ΚΑΖΑ-19012	12.3343456	10.3098901	9.1359761	9.198105	11.908004	15.288948	11.288948
ΙΛΙΟΥ-13121	1.409146342	2.189007379	1.445897957	1.432087391	1.804128175	4.087777895	1.42673833
ΙΛΙΟΥ-13122	1.819970343	3.051886462	1.834180958	1.831989049	2.565114339	4.238677346	1.82160037
ΙΛΙΟΥ-13123	1.396097523	1.85821454	1.380544075	1.38207413	1.84423127	4.123785681	1.383500931
ΚΑΙΣΑΡΙΑΝΗΣ-16121	1.389673212	1.380744763	3.470820676	3.471950725	2.040847099	4.263003849	1.426518724
ΚΑΙΣΑΡΙΑΝΗΣ-16122	1.349601457	1.066133346	1.334421774	1.334279247	1.599423135	4.232517546	1.138400358
ΚΑΛΛΙΘΕΑΣ-17671	1.023597403	0.912236502	0.991448097	0.981279868	1.517484641	4.058084756	0.94468475
ΚΑΛΛΙΘΕΑΣ-17672	1.402719597	1.976076913	1.378729929	1.377962142	1.766115144	4.568370533	1.377315602
ΚΑΛΛΙΘΕΑΣ-17673	1.400202358	1.912698054	2.975583315	2.934199421	1.704164906	4.277753448	1.535445663
ΚΑΛΛΙΘΕΑΣ-17674	3.033384408	3.790334878	4.524392666	4.527258582	4.72967739	4.685648942	3.574477539
ΚΑΛΛΙΘΕΑΣ-17675	2.052810564	2.067575893	3.863934957	3.874740967	3.347083236	4.536632662	2.085273757
ΚΑΛΛΙΘΕΑΣ-17676	1.872153365	2.104859444	3.985387219	3.982172343	2.67334722	4.719261373	1.904990104
ΚΑΜΑΤΕΡΟΥ-13451	1.927486701	2.853156251	1.992724881	1.987991804	2.825790236	5.544139142	1.954009136
ΚΕΡΑΤΕΑΣ-19001	17.34833823	18.50696048	22.07806868	21.94772801	23.7737019	22.27872555	17.38230292
ΚΕΡΑΤΣΙΝΙ-18755	5.020770647	4.991216165	4.983932597	5.002383239	5.304064458	5.063366499	4.990270103
ΚΕΡΑΤΣΙΝΙ-18756	1.138926671	1.628712691	1.094724572	1.100699821	1.296020748	3.950018574	1.109050513
ΚΕΡΑΤΣΙΝΙ-18757	1.214427792	1.245922244	1.215511549	1.212851486	1.68246984	4.063910282	1.214015476
ΚΕΡΑΤΣΙΝΙ-18758	2.107748669	2.300355459	4.250699532	4.250733859	2.57136028	4.582984574	2.013548355
ΚΗΦΙΣΙΑΣ-14561	4.063656463	4.699364587	3.926615993	3.926615993	6.228765468	9.060473828	3.912501168
ΚΗΦΙΣΙΑΣ-14562	5.079956561	7.000648455	5.263288609	5.098993026	8.232339853	11.82945092	5.055795531
ΚΗΦΙΣΙΑΣ-14563	5.52026541	5.331902066	9.629016315	9.632097639	8.828224621	9.782552467	5.801701678
ΚΗΦΙΣΙΑΣ-14564	4.885525015	6.875022871	4.923499053	4.906048564	6.868276734	9.654494329	4.861674541
ΚΗΦΙΣΙΑΣ-14671	30.54591024	37.31342241	33.33585606	35.00786157	154.233016	31.46210408	30.06465867
ΚΟΥΡΥΔΑΛΛΟΥ-18120	1.257201662	0.955421414	1.208833065	1.205688598	1.466052993	4.258548595	1.059376769

ΚΟΡΥΔΑΛΛΟΥ-18121	1.201074556	0.954584529	1.198881903	1.198597479	1.517911189	4.148088253	1.083094135
ΚΟΡΥΔΑΛΛΟΥ-18122	1.121190798	1.080885151	1.108748172	1.107511825	1.40276823	3.952166261	1.053553292
ΚΡΩΠΙΑΣ-19010	17.45444876	19.69676167	22.61494528	23.97649938	29.10266884	22.97413642	17.81840367
ΚΡΩΠΙΑΣ-19013	34.59159053	33.92504425	27.6158142	27.15098641	40.69675715	27.56188314	26.84296132
ΚΡΩΠΙΑΣ-19400	22.4768788	27.39730006	24.51884527	24.5362389	30.62144019	24.90585177	21.79864266
ΚΟΙΝΟΤΗΤΑ ΓΛΥΚΩΝ ΝΕΡΩΝ-15354	16.66043726	17.02426568	16.68170667	16.66043711	40.1767434	16.50719545	16.66477297
ΚΟΙΝΟΤΗΤΑ ΘΡΑΚΟΜΑΚΕΔΟΝΩΝ-13676	10.49452357	11.71507053	10.4864449	10.49635652	14.94272315	15.63225119	10.37467856
ΚΟΙΝΟΤΗΤΑ ΠΑΛΛΙΑΣ ΠΕΝΤΕΛΗΣ-19009	22.456011	19.456011	19.436011	19.436011	20.345121	29.009281	20.009281
ΛΑΥΡΕΩΤΙΚΗΣ-19500	23.686011	20.686011	20.666011	20.666011	21.575121	30.239281	21.239281
ΛΥΚΟΒΡΥΣΕΩΣ-14123	3.203976272	2.681221024	3.172468952	3.203571299	5.157924096	7.299688421	2.873635412
ΜΑΝΔΡΑΣ-19600	30.6966391	18.89828094	27.86060267	36.5641179	29.09960358	17.58305089	20.51144749
ΜΑΡΑΘΩΝΑ-19007	1.533056342	2.312917379	1.569807957	1.555997391	1.928038175	4.211687895	1.55064833
ΜΑΡΚΟΠΟΥΛΟΥ-19003	21.99542342	24.13810074	22.02119811	21.99364233	46.67541136	22.32542067	21.9867136
ΜΕΓΑΡΕΩΝ-19100	8.949011597	14.32572286	9.172922898	9.146308364	14.2127602	17.78891386	9.017908262
ΜΕΛΙΣΣΙΩΝ-11527	1.072866356	1.06169552	1.085080124	1.046547701	1.248771725	4.299599293	1.020934454
ΜΕΤΑΜΟΡΦΩΣΕΩΣ-14451	1.722223834	2.388972789	1.724030277	1.732134119	2.691053256	5.103506995	1.722827589
ΜΕΤΑΜΟΡΦΩΣΕΩΣ-14452	2.005141028	2.745260377	2.001141857	2.004348584	2.899559075	5.141704091	1.999528336
ΜΟΣΧΑΤΟΥ-18344	0.973213395	2.1805507	0.963389903	0.964387996	1.45089411	4.548896814	0.965876529
ΜΟΣΧΑΤΟΥ-18345	1.322679233	1.317292362	1.320106417	1.317854925	1.714883948	4.199584798	1.260398533
ΜΟΣΧΑΤΟΥ-18346	3.332756006	3.525572355	5.191189563	5.252753121	3.692801446	5.542099394	3.104237161
ΝΕΑΣ ΠΕΡΑΜΟΥ-19006	16.24108706	12.85835849	40.94485676	16.23185016	17.04665049	16.44654204	13.44456511
ΝΕΑΣ ΙΩΝΙΑΣ-14231	1.231277028	1.326172056	1.303596896	1.314568803	1.506956129	3.951102234	1.27808757
ΝΕΑΣ ΙΩΝΙΑΣ-14232	1.363114515	1.197950241	1.322863768	1.318553714	1.630626213	4.24958433	1.18315855
ΝΕΑΣ ΙΩΝΙΑΣ-14233	1.243679394	1.066952275	1.232154794	1.2195033	1.577571169	4.154958749	1.05817832
ΝΕΑΣ ΙΩΝΙΑΣ-14234	1.666228511	1.463807558	1.65319909	1.656106024	1.97968609	4.384661937	1.496763987
ΝΕΑΣ ΙΩΝΙΑΣ-14235	1.521512254	1.348677905	1.636080818	1.635828755	2.167440047	4.755132148	1.263399083
ΝΕΑΣ ΣΜΥΡΝΗΣ-17121	1.963737584	2.495019889	3.687755619	3.697491434	2.774553297	4.1728153	2.112775047
ΝΕΑΣ ΣΜΥΡΝΗΣ-17122	7.577229536	4.760807999	8.098142407	8.111828563	9.236335282	8.205250827	6.279082864
ΝΕΑΣ ΣΜΥΡΝΗΣ-17123	1.564004582	2.299185323	1.521787966	1.528096863	2.014124313	3.739769813	1.537523654
ΝΕΑΣ ΣΜΥΡΝΗΣ-17124	1.157534403	1.883865012	1.153276962	1.148862784	1.575898425	3.718032096	1.153204414
ΝΕΑΣ ΦΙΛΑΔΕΛΦΕΙΑΣ-14341	1.464563533	2.156664662	1.429230893	1.438999646	1.851932085	4.2287515	1.443342293
ΝΕΑΣ ΦΙΛΑΔΕΛΦΕΙΑΣ-14342	1.340061575	2.096788262	1.349167688	1.347933974	1.738244981	4.139377217	1.34429187
ΝΕΑΣ ΧΑΛΚΗΔΟΝΑΣ-14343	1.509928127	1.593850044	1.532250159	1.533674026	1.846331097	4.323491331	1.521240534
ΝΕΟ ΨΥΧΙΚΟ-15451	1.768555205	1.822874575	1.776510005	1.76921104	2.559319538	5.15038632	1.717818332
ΝΙΚΑΙΑΣ-18451	1.37918862	1.010640132	1.365811866	1.368930657	1.646545524	4.395077973	1.198710004
ΝΙΚΑΙΑΣ-18452	1.394569212	1.256125752	1.410287444	1.411819366	1.7941472	4.619682837	1.259269015

ΝΙΚΑΙΑΣ-18453	1.939209017	3.041154352	3.930400205	3.907591897	2.439123515	4.416031855	2.316134636
ΝΙΚΑΙΑΣ-18454	1.397929066	1.233275689	1.413888505	1.408232672	1.43950202	4.280463816	1.231865493
ΠΑΙΑΝΙΑΣ-19002	27.137069	24.137069	24.117069	24.117069	25.026179	33.690339	24.690339
ΠΑΛΛΗΝΗΣ-15351	28.62503083	41.3070542	28.40542152	28.38230897	115.3724611	28.8119762	28.47084474
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17561	2.446697185	2.598967145	2.408877915	2.41303492	3.1897329	4.72174302	2.417314784
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17562	1.579028297	2.441628233	1.573501659	1.566722052	2.331418662	4.449941713	1.559917263
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17563	2.053246045	2.635889262	1.949823474	1.970835783	2.842358165	4.783359893	1.987529312
ΠΑΛΑΙΟΥ ΦΑΛΗΡΟΥ-17564	1.646395235	2.451367726	1.688094478	1.688513415	2.317183803	4.608852949	1.673073847
ΥΜΗΤΤΟΥ-17236	1.433195406	1.220851601	1.441801798	1.447135677	1.95260725	4.433929772	1.258588238
ΥΜΗΤΤΟΥ-17237	1.19886151	1.169475389	1.189473282	1.194884046	1.702972735	4.187054235	1.160068613
ΠΕΙΡΑΙΩΣ-18450	1.642809541	2.541881681	1.589872658	1.58770769	1.883385137	4.194278503	1.595358505
ΠΕΙΡΑΙΩΣ-18531	2.448355942	2.531616057	4.270603845	4.264970047	2.733219115	4.508353736	2.272428385
ΠΕΙΡΑΙΩΣ-18532	2.688570657	4.180997424	4.491738264	4.49218379	3.791532846	4.982475657	2.982298182
ΠΕΙΡΑΙΩΣ-18533	4.274311223	4.519988768	3.886257244	3.895216196	4.650768554	6.506071461	3.948947465
ΠΕΙΡΑΙΩΣ-18534	1.566636594	1.485290199	1.532402608	1.514165939	1.75985822	4.278373646	1.450154179
ΠΕΙΡΑΙΩΣ-18535	4.060028119	5.736902248	6.399389098	6.38706839	5.094108816	6.483953179	4.288955572
ΠΕΙΡΑΙΩΣ-18536	1.420327219	1.597892754	3.294485694	3.294485694	1.671180215	4.784740862	1.344498701
ΠΕΙΡΑΙΩΣ-18537	2.744036467	3.71956648	4.363555946	4.350633639	3.310573036	4.792447846	2.989351211
ΠΕΙΡΑΙΩΣ-18538	3.165984441	4.611457061	5.082346356	5.096425785	4.475933028	5.520453961	3.6730518
ΠΕΙΡΑΙΩΣ-18539	2.553026061	3.436911799	4.702771325	4.703242604	3.821889619	5.130442809	2.77322851
ΠΕΙΡΑΙΩΣ-18540	4.081726099	10.8332486	6.345495777	6.305346007	7.232541491	6.399906121	5.156330976
ΠΕΙΡΑΙΩΣ-18541	5.043098935	4.601935211	5.042025647	5.041571975	5.67324844	5.044037883	4.766274282
ΠΕΙΡΑΙΩΣ-18542	1.357353118	1.510419303	1.326894227	1.325641541	1.983146886	4.672123341	1.336477071
ΠΕΙΡΑΙΩΣ-18543	1.56614676	2.951064191	3.727725689	3.726853497	2.570057306	4.191443035	2.161913036
ΠΕΙΡΑΙΩΣ-18544	6.440019661	7.759898984	6.44721119	6.440019756	8.378578276	6.507968485	6.44162371
ΠΕΙΡΑΙΩΣ-18545	9.467760539	8.500026661	9.23630851	9.269648313	12.20233117	9.289250718	8.230016176
ΠΕΙΡΑΙΩΣ-18546	2.700279948	4.412106588	4.835848043	4.835754322	3.94035845	4.928427669	3.087296431
ΠΕΙΡΑΙΩΣ-18547	7.875663568	7.443772008	7.783763104	7.87564785	9.723604625	8.388765647	7.678809887
ΠΕΡΑΜΑΤΟΣ-18863	2.566171115	3.667609525	4.881560534	4.944274894	3.435620564	4.894782166	2.894591815
ΠΕΡΑΜΑΤΟΣ-18883	10.16078013	9.733808563	10.17041296	10.16077938	13.32961718	10.2955672	9.90388021
ΠΕΡΙΣΤΕΡΙΟΥ-12131	1.472322302	1.605939388	1.418670068	1.418670068	1.994028674	4.586339827	1.378128869
ΠΕΡΙΣΤΕΡΙΟΥ-12132	1.531896987	1.658639921	1.57519741	1.57233373	1.820022147	4.266734791	1.553049179
ΠΕΡΙΣΤΕΡΙΟΥ-12133	1.458715796	1.307401668	1.450175496	1.442379102	1.721613806	4.201434619	1.313189408
ΠΕΡΙΣΤΕΡΙΟΥ-12134	1.398830284	1.482432045	1.37620734	1.386757785	1.74071427	4.530074283	1.33396039
ΠΕΡΙΣΤΕΡΙΟΥ-12135	1.514256144	1.403077985	1.483290611	1.470780788	1.970241376	4.341482987	1.377853381
ΠΕΡΙΣΤΕΡΙΟΥ-12136	1.172670443	1.047103632	1.155503287	1.152483123	1.716577646	4.590777156	1.036223974
ΠΕΡΙΣΤΕΡΙΟΥ-12137	1.514621933	1.479378615	1.461493798	1.463379443	2.040618359	4.456700971	1.326099631
ΠΕΤΡΟΥΠΟΛΕΩΣ-13231	2.958895833	3.956866755	4.871521098	4.866011658	4.438173542	4.871861619	3.167281582
ΠΕΤΡΟΥΠΟΛΕΩΣ-13232	9.537429955	10.83057407	9.536561962	9.453817331	17.63753836	9.596481843	9.509269749
ΠΕΥΚΗΣ-15121	2.266350925	2.694903593	2.15857824	2.070383522	3.044488683	5.255920446	2.136936633
ΣΑΛΑΜΙΝΑΣ-18900	4.630161545	6.172115149	4.576946798	4.562548885	6.414932774	9.064844124	4.586192359
ΣΑΛΑΜΙΝΑΣ-18901	3.282217673	4.086002265	3.427992069	3.4166567	4.724740063	7.71770997	3.366741861

ΣΑΛΑΜΙΝΑΣ-18902	6.635922905	8.234428961	6.478387801	6.39907177	10.54120821	14.20665015	6.467710056
ΣΑΛΑΜΙΝΑΣ-18903	9.195013362	12.4020471	9.096800185	9.129679841	14.82259777	19.16278427	8.961858979
ΤΑΥΡΟΥ-17778	3.365748175	4.294897443	5.072413477	5.034625787	4.451427845	5.07687087	3.683441702
ΦΙΛΟΘΕΗ-15237	6.132228898	6.049811002	6.269674998	6.347541199	8.682612148	10.73754766	5.703352713
ΧΑΙΔΑΡΙ-12461	1.531338953	2.315108215	1.513175761	1.509903328	1.995802404	4.356530111	1.517956931
ΧΑΙΔΑΡΙ-12462	3.520002006	5.054889031	6.415665231	6.423311216	4.243074716	6.520422479	3.849843216
ΧΑΛΑΝΔΡΙ-15231	1.861986477	1.820716017	1.822190414	1.864748741	2.68604207	5.244108106	1.726095516
ΧΑΛΑΝΔΡΙ-15232	6.520199505	5.14330509	6.520982171	6.387710561	5.470120429	6.47622759	5.269966967
ΧΑΛΑΝΔΡΙ-15233	2.229105173	2.000975846	2.420245918	2.424491774	3.058609925	5.320702913	2.078733815
ΧΑΛΑΝΔΡΙ-15234	2.240837234	2.170505394	2.332021258	2.326214537	3.200587885	5.605097022	2.025351348
ΧΑΛΑΝΔΡΙ-15238	3.444798466	3.600966171	3.379004608	3.384163373	5.522218371	8.00441796	3.360611368
ΧΑΛΑΝΔΡΙ-15344	28.40948495	26.54882375	25.34896913	24.87468631	73.97362591	25.24789257	24.22388975
ΨΥΧΙΚΟ-15452	6.686091615	6.837287772	6.886755823	6.940708507	9.607763318	11.38853338	6.453387127
ΑΘΗΝΩΝ-10431	2.14876434	2.966652081	4.63853849	4.629365971	3.030175547	4.798519629	2.459499729
ΑΘΗΝΩΝ-10432	2.620831138	3.880183357	5.028750906	5.073675585	3.752293715	5.254934124	3.074897818
ΑΘΗΝΩΝ-10433	1.56437158	2.243268466	3.314196554	3.327054625	1.960439892	4.062296556	1.794262568
ΑΘΗΝΩΝ-10434	1.202688813	1.554700892	2.231682756	2.194424162	1.607533279	2.653482741	1.302959935
ΑΘΗΝΩΝ-10435	4.374952669	4.374560187	4.36276721	4.362735477	4.011847453	4.598782151	4.24289363
ΑΘΗΝΩΝ-10436	1.339010117	2.00366011	3.471172582	3.460922696	1.70301151	4.598892668	1.559729945
ΑΘΗΝΩΝ-10437	1.725462294	1.571040728	4.083612247	4.092278508	2.838309188	4.371084548	1.700728719
ΑΘΗΝΩΝ-10438	1.530184188	1.764488511	3.615504681	3.601744121	1.97078811	4.254767484	1.511286505
ΑΘΗΝΩΝ-10439	1.543513708	2.163705347	3.244073898	3.239465557	1.830086832	3.891493818	1.695992179
ΑΘΗΝΩΝ-10440	1.268140448	1.958159649	2.766288941	2.804156889	1.535095394	3.759934625	1.487371438
ΑΘΗΝΩΝ-10441	1.143265029	1.681523982	1.151158304	1.163478005	1.489884467	3.953193038	1.150223544
ΑΘΗΝΩΝ-10442	1.503076454	1.901015609	1.435277947	1.433121077	1.821613829	3.854104324	1.457155861
ΑΘΗΝΩΝ-10443	1.284577199	1.318895621	1.291824441	1.307854938	1.575009022	3.893105247	1.290531989
ΑΘΗΝΩΝ-10444	1.037685153	1.091647813	1.071913473	1.073639103	1.440852877	3.825442102	1.054426547
ΑΘΗΝΩΝ-10445	1.270157143	1.285013982	1.226830852	1.225911498	1.505407701	3.84180558	1.231179217
ΑΘΗΝΩΝ-10446	1.595696415	2.447788793	3.491478332	3.51062425	2.190400287	3.977992392	1.926264031
ΑΘΗΝΩΝ-10447	3.014898568	4.229557642	3.021611455	2.971350612	6.055268542	8.530153685	2.998799006
ΑΘΗΝΩΝ-10551	4.25164593	4.837341927	5.196043125	5.227397959	4.75555717	5.555570859	3.981392373
ΑΘΗΝΩΝ-10552	3.449351633	4.657008973	4.944942754	4.956044795	5.176927848	5.03108348	4.057881596
ΑΘΗΝΩΝ-10553	1.742849577	1.684890678	1.69952684	1.702683765	2.192304472	4.659099518	1.569059026
ΑΘΗΝΩΝ-10554	2.105295197	2.648531985	5.179583001	5.191837496	2.865494135	5.488209305	2.091878744
ΑΘΗΝΩΝ-10555	4.841031394	5.886320074	6.476534823	6.592088398	5.406342971	6.586142297	5.053489413
ΑΘΗΝΩΝ-10556	2.954832556	3.751860287	2.950340055	2.945917275	3.564131111	5.788747229	2.939110923
ΑΘΗΝΩΝ-10557	3.026797292	2.453663864	4.699391998	4.716154753	3.453485175	4.77538598	2.469673547
ΑΘΗΝΩΝ-10558	4.570075432	3.628186717	6.85954939	6.831347619	7.154482194	6.934822874	4.277270323
ΑΘΗΝΩΝ-10559	3.63238479	4.337600634	5.777448025	5.772761103	3.923992084	5.905183805	3.692832472
ΑΘΗΝΩΝ-10560	2.579479951	2.409202936	4.482721241	4.483043108	2.830100722	4.932772315	2.297603727
ΑΘΗΝΩΝ-10561	3.803249292	5.175986856	6.038337623	6.038260535	4.876370605	6.138314147	4.165776615
ΑΘΗΝΩΝ-10562	1.628765369	2.347077189	3.271125119	3.246081332	1.949684786	4.25584126	1.824078483
ΑΘΗΝΩΝ-10563	2.903554967	4.103375592	5.695366937	5.703458207	3.4808455	5.836522278	3.090715285
ΑΘΗΝΩΝ-10564	2.801865443	3.026038873	5.025199421	5.102741152	3.983289442	5.487038263	2.778035058
ΑΘΗΝΩΝ-10671	1.729907477	2.227649246	3.20949021	3.093263641	1.994360335	4.213535611	1.839833362

ΑΘΗΝΩΝ-10672	1.968781741	1.778155807	4.943402307	4.891455503	2.817334093	4.950444392	1.877527365
ΑΘΗΝΩΝ-10673	3.536002151	3.724367337	3.672795176	3.694203236	4.742384511	6.225587627	3.631494163
ΑΘΗΝΩΝ-10674	2.113415689	2.486563725	4.601617001	4.601837475	3.503075567	4.671706095	2.310458187
ΑΘΗΝΩΝ-10675	2.143100597	2.017850868	2.236526666	2.316034301	2.640549241	5.197436595	2.003691565
ΑΘΗΝΩΝ-10676	1.733763444	1.863888932	3.55639383	3.504483766	1.971622991	4.564500435	1.55033874
ΑΘΗΝΩΝ-10677	2.496813402	3.04635687	4.295444151	4.452228336	3.065818777	4.494115996	2.605918726
ΑΘΗΝΩΝ-10678	1.955954767	2.780785549	3.825353006	3.833450642	2.14983519	4.373496427	2.103947829
ΑΘΗΝΩΝ-10679	2.242645621	2.385255271	4.075142812	4.088529516	2.512562306	4.791696601	2.158007927
ΑΘΗΝΩΝ-10680	7.187213217	5.062105837	7.209324524	7.168193967	7.502035482	7.276588484	6.144523172
ΑΘΗΝΩΝ-10681	1.195668288	1.701048282	1.222222507	1.222756211	1.456878278	3.882329643	1.207018719
ΑΘΗΝΩΝ-10682	1.949762684	1.784896517	3.62155424	3.641127022	2.220738003	4.153389463	1.753903938
ΑΘΗΝΩΝ-10683	1.545234632	1.362002731	1.536568534	1.539077596	1.858948809	4.230646885	1.36165464
ΑΘΗΝΩΝ-11141	1.287070622	1.02832451	1.313908745	1.321618511	1.688728812	4.413852795	1.148238138
ΑΘΗΝΩΝ-11142	1.299726088	1.05377436	1.240944237	1.239805621	1.618355465	4.326952341	1.064380508
ΑΘΗΝΩΝ-11143	1.360556659	2.091518011	1.45569449	1.451377318	1.574426348	4.180266766	1.417339191
ΑΘΗΝΩΝ-11144	1.610067557	1.404217165	3.051432974	3.086968399	1.76352371	4.311338371	1.403486993
ΑΘΗΝΩΝ-11145	1.353661715	1.146742502	3.120275673	3.143831864	1.819537062	3.778754344	1.179133183
ΑΘΗΝΩΝ-11251	1.357950271	2.118767591	2.907618607	3.00217014	1.772208531	4.300379873	1.61854776
ΑΘΗΝΩΝ-11252	1.222641332	1.459036033	2.784683728	2.803388768	1.551730658	3.945911604	1.242802489
ΑΘΗΝΩΝ-11253	2.051920864	3.333925451	4.03284475	4.035546246	2.760457265	4.614689178	2.462985856
ΑΘΗΝΩΝ-11254	0.997601766	1.415737551	0.964492598	0.962879666	1.586819306	4.162625179	0.966160646
ΑΘΗΝΩΝ-11255	1.21123318	1.275844356	3.155716701	3.144982574	1.679680174	4.344964995	1.219479951
ΑΘΗΝΩΝ-11256	3.209891003	3.237798906	3.242842545	3.220627125	3.923646453	3.332210844	3.220966053
ΑΘΗΝΩΝ-11257	1.511174741	2.344846899	3.212549105	3.212549105	1.781848801	4.287345598	1.764093797
ΑΘΗΝΩΝ-11361	1.214166411	1.811350188	2.918737672	2.940362494	1.528556963	4.324544719	1.412144203
ΑΘΗΝΩΝ-11362	1.979767209	1.915946368	3.820410459	3.804157557	2.239098737	4.757192402	1.797714886
ΑΘΗΝΩΝ-11363	1.415749309	1.945713993	1.379130858	1.376582601	1.745249304	4.445864085	1.375924531
ΑΘΗΝΩΝ-11364	1.346834124	1.320921159	1.288023264	1.300574857	1.635311785	4.53126814	1.205450335
ΑΘΗΝΩΝ-11471	1.795460005	1.512998674	4.070485533	4.069193464	2.66768668	4.427432816	1.623503076
ΑΘΗΝΩΝ-11472	1.624654008	2.436838501	1.638367244	1.643361917	1.889580596	4.343680478	1.626130835
ΑΘΗΝΩΝ-11473	2.084658354	2.660412231	3.888073441	3.905060574	2.994910093	4.231189466	2.096396262
ΑΘΗΝΩΝ-11474	3.459010891	3.491302166	4.962703054	4.990149982	5.963994104	5.065697265	3.47768194
ΑΘΗΝΩΝ-11475	4.607209059	5.401075713	5.768207351	5.810573664	5.836739393	5.859176546	4.510955812
ΑΘΗΝΩΝ-11476	4.775310797	6.023591207	6.663394126	6.683435014	9.111135728	6.782095216	4.958722224
ΑΘΗΝΩΝ-11521	1.528777669	1.708898775	1.534267035	1.534868893	1.933382144	3.96079914	1.531500651
ΑΘΗΝΩΝ-11522	1.410905337	2.16210639	1.429613302	1.429634888	1.859083844	4.324160852	1.420486662
ΑΘΗΝΩΝ-11523	5.44938777	5.229765612	5.514135923	5.511628737	5.903225147	5.685676319	5.325118635
ΑΘΗΝΩΝ-11524	1.979971166	2.871244056	4.095625022	4.094821502	3.007844811	4.398201848	2.283446504
ΑΘΗΝΩΝ-11525	1.235097065	1.334432939	1.234828829	1.229162884	1.642974319	4.312321845	1.230515734
ΑΘΗΝΩΝ-11526	3.931283567	4.956773616	5.178245575	5.170233801	4.114787009	5.257393391	3.65083274
ΑΘΗΝΩΝ-11527	1.072866356	1.06169552	1.085080124	1.046547701	1.248771725	4.299599293	1.020934454
ΑΘΗΝΩΝ-11528	4.274533747	4.580563107	5.428810501	5.429766006	5.747974045	5.493361371	4.195595339
ΑΘΗΝΩΝ-11631	1.733401111	1.556914853	3.788395402	3.7881764	2.674080716	4.276596674	1.599282169
ΑΘΗΝΩΝ-11632	1.388485366	2.047330054	1.350433571	1.350086395	1.75551946	4.358106888	1.362990146
ΑΘΗΝΩΝ-11633	3.312902038	2.997579916	5.37075915	5.36837435	3.799136054	5.425740783	2.822660405

ΑΘΗΝΩΝ-11634	3.418590684	3.30340058	4.99305552	4.956684806	4.386665105	5.047811627	2.96409035
ΑΘΗΝΩΝ-11635	1.890329988	2.555083071	4.024839011	4.017347719	2.668909756	4.451153382	2.043002051
ΑΘΗΝΩΝ-11636	1.138929531	1.859417694	1.120369621	1.118623682	1.548015968	4.030217571	1.124897445
ΑΘΗΝΩΝ-11741	1.991982327	1.613805338	3.818213558	3.818369357	3.00242284	4.366285312	1.900441445
ΑΘΗΝΩΝ-11742	2.396135543	3.170824285	4.193691055	4.189096647	2.962429933	4.638004997	2.531747697
ΑΘΗΝΩΝ-11743	2.498687087	2.820670931	4.628426894	4.629663024	3.377489761	4.776063457	2.546363523
ΑΘΗΝΩΝ-11744	5.499368475	4.984221676	5.523845175	5.499368653	6.304549892	5.564849008	4.858937489
ΑΘΗΝΩΝ-11745	6.945912444	4.256253081	6.944443963	6.94682859	7.379940337	7.036276556	5.670033422
ΑΘΗΝΩΝ-11851	1.576378017	2.010622227	1.638724184	1.635046982	1.759868339	4.082532018	1.614987932
ΑΘΗΝΩΝ-11852	1.66911229	1.918549653	3.806107168	3.820193523	2.428190042	4.249430917	1.699622968
ΑΘΗΝΩΝ-11853	5.692017959	6.545617941	6.816354549	6.823388071	7.255406938	6.903132355	5.53545122
ΑΘΗΝΩΝ-11854	1.94938109	2.551993219	3.807565897	3.788870687	1.853242666	4.400456625	1.980084484
ΑΘΗΝΩΝ-11855	1.962294362	2.732039002	3.73256461	3.738352741	2.271003625	3.951740182	2.215046511
ΧΟΛΑΡΓΟΥ-15562	2.558018571	2.884226135	4.90725322	4.808540942	4.008512979	4.885434743	2.623574272
ΣΠΑΤΩΝ-19019	26.7242382	23.7242382	23.7042382	23.7042382	24.6133482	33.2775082	24.2775082
ΠΑΠΑΓΟΥ-15669	3.352085351	2.87383239	3.454927	3.454927	5.147399524	7.462854468	3.061028624

Πίνακας Π-6: Πλήρης πίνακας σφαλμάτων MAPE

περιοχή	Αυγ-12	Σεπ-12	Οκτ-12	Νοε-12	Δεκ-12	Ιαν-13	Φεβ-13	Μαρ-13	Απρ-13	Μαϊ-13	Ιουν-13	Ιουλ-13	Αυγ-13	Σεπ-13	Οκτ-13	Νοε-13	Δεκ-13
ΖΩΓΡΑΦ ΟΥ-15771	95975.52 065	97793.31 569	103250.1 383	100765.3 112	103298.7 559	102390.2 242	93667.31 693	102690.4 945	101208.3 195	101989.5 457	94516.04 099	96702.94 932	95975.52 065	97793.31 569	103250.1 383	100765.3 112	103298.7 559
ΖΩΓΡΑΦ ΟΥ-15772	128180.9 244	130812.7 394	136469.3 654	133018.2 463	134773.8 86	133739.0 381	122596.5 964	128698.0 885	124989.5 828	128315.9 43	127345.2 035	128530.9 8	128180.9 714	130812.7 873	136469.4 153	133018.2 949	134773.9 353
ΖΩΓΡΑΦ ΟΥ-15773	91070.59 307	91218.29 334	95341.17 013	92938.50 265	94311.20 785	94011.30 028	86004.06 909	93406.91 809	90851.76 133	93398.08 147	89411.89 991	90605.51 675	89814.87 724	89959.09 411	94023.54 198	91652.59 875	93004.80 481
ΑΓ.ΑΝΑΡ ΓΥΡΩΝ-13561	71539.81 008	69145.67 595	70264.84 967	67713.23 117	68694.69 325	69917.56 447	67789.04 151	74739.91 934	78318.63 852	80664.03 533	75914.85 877	71768.44 977	71539.81 008	69145.67 595	70264.84 967	67713.23 117	68694.69 325
ΑΓ.ΑΝΑΡ ΓΥΡΩΝ-13562	81245.49 071	78673.76 732	79215.90 516	75624.59 498	78015.99 933	77041.28 72	70187.27 533	78293.92 333	80743.96 152	86140.07 226	82764.69 173	81254.96 948	80724.42 653	78160.31 413	78720.20 323	75149.80 274	77548.11 99
ΑΓ.ΒΑΡΒ ΑΡΑΣ-12351	114930.0 311	110370.9 869	112814.3 358	107357.5 047	110567.5 059	109295.0 905	98662.94 929	110257.7 038	112495.6 017	118408.5 752	114094.2 237	115575.1 146	114930.0 311	110370.9 869	112814.3 358	107357.5 047	110567.5 059
ΑΓ.ΔΗΜ ΗΤΡΙΟΥ-17341	108897.5 531	104159.7 363	107361.0 697	101947.4 782	103635.8 895	102684.1 86	94757.91 22	108828.7 846	107564.5 906	111280.0 403	106432.7 129	109452.7 249	108897.5 531	104159.7 363	107361.0 697	101947.4 782	103635.8 895
ΑΓ.ΔΗΜ ΗΤΡΙΟΥ-17342	130295.3 671	124800.8 931	128018.1 064	121736.4 04	120601.8 43	119826.5 167	109029.8 384	124599.5 416	124198.5 258	128919.8 442	125343.3 608	129018.5 953	128096.2 032	123348.1 892	126903.5 958	120952.6 92	119181.9 986
ΑΓ.ΔΗΜ ΗΤΡΙΟΥ-17343	96266.29 011	94237.86 146	97737.27 896	93705.72 489	91861.70 526	91374.37 56	85180.21 276	94189.68 978	91518.77 582	94588.61 171	93632.32 911	95888.57 125	95806.09 315	93998.40 342	97674.15 894	93636.81 373	91467.32 845
ΑΓ.ΔΗΜ ΗΤΡΙΟΥ-17372	60.72424 644	61.46917 192	62.74807 309	63.54204 586	64.58627 711	65.38334 476	66.29561 987	67.08062 614	67.91725 763	68.68755 086	69.48000 381	70.23738 133	71.00358 019	71.75093 049	72.50128 122	73.24132 121	73.98197 374
ΑΓ.ΙΩΑΝ ΝΗΣ ΡΕΝΤΗ-18232	150.5100 829	145.2915 151	131.0366 543	115.2050 416	117.3417 346	111.0237 703	99.00467 925	108.5797 694	117.3980 272	135.2671 305	134.8808 707	150.6880 038	150.4773 318	145.2598 951	131.0081 748	115.1799 901	117.3161 444
ΑΓ.ΙΩΑΝ ΝΗΣ ΡΕΝΤΗ-18233	79522.90 053	77341.81 649	78854.23 148	75353.88 288	77243.07 273	75961.51 268	69376.66 011	76359.23 159	76376.29 072	81149.88 55	78691.47 113	79738.51 15	78258.21 611	76001.45 922	77430.08 625	73924.04 081	75764.77 257
ΑΓ.ΠΑΡΑ ΣΚΕΥΗ-15341	135043.4 105	123671.7 987	121871.3 887	116134.3 246	115140.0 967	113102.9 264	104129.5 12	123511.5 21	123018.1 632	127940.0 97	127621.3 926	134227.5 446	134395.0 305	123077.7 792	121285.7 825	115576.0 621	114586.3 916
ΑΓ.ΠΑΡΑ ΣΚΕΥΗ-15342	95467.08 764	87960.06 685	85446.47 122	80255.77 872	77684.76 783	74940.48 51	67659.42 012	79247.96 234	81380.87 208	85438.00 762	89015.42 144	95694.66 069	94892.55 598	87430.28 578	84931.78 918	79771.90 727	77216.27 768

ΑΓ.ΠΑΡΑ ΣΚΕΥΗ- 15343	131796.5 57	123150.5 815	119152.5 678	113417.4 776	112313.8 855	112399.7 825	103799.7 896	119281.9 734	124735.8 508	131454.2 852	130113.7 666	136655.9 883	131796.5 57	123150.5 815	119152.5 678	113417.4 776	112313.8 855
ΑΙΓΑΛΕΩ- 12241	70129.97 806	70129.97 806	70129.97 806	70129.97 806	70042.82 712	69933.59 51	69933.59 51	69933.59 51	69846.44 416	69846.44 416	69737.21 214	69737.21 214	69650.06 12	69650.06 12	69650.06 12	69540.82 917	69453.67 824
ΑΙΓΑΛΕΩ- 12242	67780.15 387	70950.32 949	73289.39 655	70753.65 519	68643.41 254	68167.06 483	63246.20 768	70844.37 644	69368.60 763	71721.82 985	67999.45 536	68017.12 748	66258.81 471	68531.80 572	70832.67 544	67957.91 518	65929.77 819
ΑΙΓΑΛΕΩ- 12243	104631.7 568	104808.2 868	108738.5 319	105107.4 535	102680.0 171	102093.0 796	92462.36 197	105800.5 827	104643.8 214	109037.0 257	104853.5 006	106797.1 424	106358.0 982	104606.6 681	108566.9 813	105093.1 638	108263.3 566
ΑΙΓΑΛΕΩ- 12244	83437.31 287	80602.75 292	82911.93 114	79971.21 243	79922.20 367	78322.96 314	70196.66 846	80623.11 018	80353.30 671	84079.98 891	81182.36 552	82610.12 72	81753.96 842	78973.85 703	81233.54 268	78349.61 752	78298.85 944
ΆΛΙΜΟΣ- 17436	110.1655 852	109.2038 032	88.27296 72	49.98001 616	47.93805 624	47.58536 95	48.19952 209	56.31191 755	64.99729 341	100.6209 618	101.9432 731	105.0137 075	107.7103 002	106.7927 311	86.38559 273	48.91201 331	46.90157 093
ΆΛΙΜΟΣ- 17455	159423.6 703	150438.1 883	144565.3 939	138661.1 661	135671.2 519	131014.3 138	120301.7 971	136863.7 532	142040.1 16	149861.3 626	150836.1 255	159718.6 903	158649.1 316	149707.0 083	143862.4 73	137986.6 8	135011.0 42
ΆΛΙΜΟΣ- 17456	112970.9 238	106182.2 65	103971.0 734	93590.33 624	94871.14 773	93745.27 719	86066.05 776	96235.80 633	98796.36 503	109652.7 099	109851.9 649	113191.5 936	112970.9 299	106182.2 708	103971.0 79	93590.34 131	94871.15 287
ΑΜΑΡΟΥ ΣΙΟΥ- 15122	90891.81 216	83656.56 059	85992.23 923	80346.70 012	77763.87 514	78758.19 628	72294.70 388	84388.92 992	82570.36 215	88720.11 59	92547.07 54	95779.05 381	94686.51 483	88880.03 089	90756.38 905	83092.93 686	78765.40 173
ΑΜΑΡΟΥ ΣΙΟΥ- 15123	36719.35 415	34788.44 767	34948.63 473	29874.64 276	29793.87 32	29612.37 532	28674.78 669	30024.82 786	29748.91 395	35549.83 736	36584.74 73	37972.79 937	36426.65 996	34754.80 695	34914.99 401	29841.00 204	29760.23 249
ΑΜΑΡΟΥ ΣΙΟΥ- 15124	119544.5 448	110879.8 227	110235.9 148	99486.69 673	97504.33 7	105486.0 259	88774.99 191	100055.4 752	99106.89 419	109835.7 865	113805.7 98	123333.7 637	119267.8 723	110929.5 225	110251.3 744	99399.37 047	97366.62 29
ΑΜΑΡΟΥ ΣΙΟΥ- 15125	101278.2 85	100869.8 705	99777.42 744	91306.99 358	103841.9 403	100784.7 931	115302.2 639	117591.7 021	122352.6 744	118778.7 789	113431.7 37	113020.9 368	101138.2 029	100730.3 076	101278.2 85	100869.8 705	99777.42 744
ΑΜΑΡΟΥ ΣΙΟΥ- 15126	93080.15 421	89790.56 882	88694.87 671	82562.23 934	91662.91 672	88919.71 025	102023.2 927	107588.0 231	111266.2 777	108044.5 467	102066.5 982	105548.9 197	93880.23 84	90313.40 501	93080.15 421	89790.56 882	88694.87 671
ΆΝΩ ΛΙΟΣΙΑ- 13341	1587.119 677	1869.967 461	1962.260 958	2032.057 59	2202.562 365	2239.589 295	2205.193 623	2511.349 345	2541.185 251	2862.916 443	2231.441 017	2196.885 446	2103.854 281	2359.741 842	1587.119 677	1869.967 461	1962.260 958
ΑΡΓΥΡΟΥ ΠΟΛΕΩΣ- 16451	76926.55 055	79873.96 116	79502.83 577	72844.29 041	80365.29 704	78950.37 25	90752.88 516	89513.67 449	91790.05 703	87081.45 641	83720.34 108	84440.86 221	76575.83 547	79292.68 219	76926.55 055	79873.96 116	79502.83 577
ΑΡΓΥΡΟΥ ΠΟΛΕΩΣ- 16452	79605.47 832	79212.57 561	79030.53 077	73588.99 541	81863.97 385	80634.87 412	90966.72 876	91134.49 226	94290.99 536	89059.35 489	85037.78 409	87940.49 555	77409.09 686	78393.76 002	79605.47 832	79212.57 561	79030.53 077
ΑΣΠΡΟΠ' ΥΡΓΟΥ- 19300	372168.7 553	386553.2 736	404171.0 037	380124.2 899	429525.7 186	419501.1 874	526414.8 244	524502.1 235	513971.4 806	520944.6 127	452715.5 677	396786.0 902	376180.9 166	390649.3 512	372168.7 553	386553.2 736	404171.0 037

ΑΧΑΡΝΩ N-13671	64435.23 692	62433.82 801	60853.23 15	55973.43 509	64827.22 336	67051.34 31	70775.15 61	72137.10 885	75883.66 23	74880.87 778	69472.13 548	66602.04 757	64204.22 364	62209.92 329	64435.23 692	62433.82 801	60853.23 15
ΑΧΑΡΝΩ N-13672	59309.94 066	58302.66 62	57607.99 179	52724.86 215	60784.96 861	62532.88 873	66938.27 263	67515.80 852	71772.25 46	71093.83 508	65655.66 477	61864.58 421	59154.91 263	58150.23 784	59309.94 066	58302.66 62	57607.99 179
ΑΧΑΡΝΩ N-13673	44110.92 11	43526.32 289	42294.38 072	38674.91 208	44420.78 062	45898.55 754	47361.17 207	47716.71 103	50351.15 903	50025.51 408	46960.30 174	46026.12 383	44178.86 056	43593.00 516	44110.92 11	43526.32 289	42294.38 072
ΑΧΑΡΝΩ N-13674	86010.49 929	84529.01 384	80955.98 759	73712.71 104	83654.23 074	86420.21 947	92692.17 715	91195.54 076	96708.79 1	96971.46 271	92103.28 535	89811.20 252	86010.49 929	84529.01 384	86010.49 929	84529.01 384	80955.98 759
ΑΧΑΡΝΩ N-13675	57224.47 391	57581.73 189	55700.11 979	51248.38 773	57791.57 759	65317.80 553	67736.58 415	67255.86 407	75172.20 292	75550.50 867	71963.53 478	60920.17 783	57645.07 999	57931.06 11	57224.47 391	57581.73 189	55700.11 979
ΑΧΑΡΝΩ N-13677	37791.22 582	36984.28 283	36814.13 903	34746.48 709	39727.24 231	41566.18 661	45174.23 887	48012.41 055	51740.38 062	50542.32 973	45358.40 944	41829.23 318	38476.89 257	37633.93 362	37791.22 582	36984.28 283	36814.13 903
ΑΧΑΡΝΩ N-13678	28676.47 112	28142.99 13	27286.56 431	25296.47 512	28929.73 848	30970.88 911	32076.88 446	33031.03 816	35151.94 149	35066.20 555	32939.70 692	30165.84 372	28786.46 403	28250.83 738	28676.47 112	28142.99 13	27286.56 431
ΑΧΑΡΝΩ N-13679	20575.79 763	20486.92 887	19604.29 403	18026.18 467	20405.61 908	23520.55 245	25940.14 42	25936.21 228	29176.05 686	29355.73 094	27358.89 941	22561.35 075	21023.11 85	20931.92 099	20575.79 763	20486.92 887	19604.29 403
ΒΑΡΗΣ- 16672	105626.1 072	97457.84 321	96023.50 237	89255.05 85	110590.1 474	125160.4 647	140467.0 175	142387.0 637	167727.1 749	164798.1 125	145441.4 509	114794.8 755	105253.2 581	97113.72 6	105626.1 072	97457.84 321	96023.50 237
ΒΟΥΛΑΣ- 16673	216533.3 989	194966.9 335	183217.9 553	169425.4 128	222491.4 664	230272.7 027	243034.5 153	287723.0 341	313452.7 697	310032.1 034	264035.6 011	229434.5 951	215939.8 492	194432.2 609	216533.3 989	194966.9 335	183217.9 553
ΒΟΥΛΙΑΓ ΜΕΝΗΣ- 16671	53433.93 483	47844.70 363	52076.05 133	49644.30 617	65228.36 244	76853.95 716	83963.47 343	93572.87 858	108625.7 308	106377.7 673	82873.07 893	61511.64 042	52046.30 427	46599.52 562	53433.93 483	47844.70 363	52076.05 133
ΒΡΙΛΗΣΣΙ ΩΝ- 15235	48833.62 345	57894.58 784	77329.70 354	103791.2 345	97511.87 965	115433.0 859	119553.0 81	136279.1 346	139330.4 488	149671.2 511	165392.9 891	172178.5 459	192127.8 519	201188.8 163	48833.62 345	57894.58 784	77329.70 354
ΒΥΡΩΝΟ Σ-16231	103910.6 97	107646.5 196	107357.6 2	97213.77 501	106359.6 996	103267.4 964	111260.3 608	107693.1 743	109758.4 994	103857.5 063	99686.51 205	103621.3 433	102125.6 829	105861.5 055	103910.6 97	107646.5 196	107357.6 2
ΒΥΡΩΝΟ Σ-16232	106555.7 059	110230.7 27	107222.7 68	97600.51 742	106935.6 154	104922.2 592	113574.4 227	110111.0 351	111236.7 558	105452.1 982	101762.0 016	105727.3 401	105757.0 775	109422.9 703	106555.7 059	110230.7 27	107222.7 68
ΒΥΡΩΝΟ Σ-16233	58237.12 751	59687.72 991	59589.18 07	53904.62 42	59730.97 424	58019.59 606	66787.39 73	66250.95 703	67990.49 036	64249.20 518	61429.15 678	62518.61 232	57448.24 381	58898.84 621	58237.12 751	59687.72 991	59589.18 07
ΓΑΛΑΤΣΙ ΟΥ- 11146	154061.7 432	159601.3 566	159516.9 582	145337.2 165	159396.1 661	154370.9 604	166070.3 537	161489.4 016	165312.4 932	153946.0 961	147501.3 036	153307.5 976	150075.5 429	155615.1 564	154061.7 432	159601.3 566	159516.9 582
ΓΑΛΑΤΣΙ ΟΥ- 11147	96572.62 055	99212.87 384	98781.83 89	90393.54 607	98821.36 654	96120.58 74	102886.8 408	101397.0 878	103424.4 983	96304.15 474	93668.31 009	98022.18 674	95475.98 733	98072.68 792	96572.62 055	99212.87 384	98781.83 89
ΓΛΥΦΑΔ ΑΣ-16561	132707.2 66	124275.9 587	124029.4 986	115532.6 623	132685.1 339	130405.1 971	145879.7 177	158711.5 411	163486.5 298	158082.7 833	146834.7 575	148911.6 952	132090.5 093	123438.0 752	132707.2 66	124275.9 587	124029.4 986
ΓΛΥΦΑΔ ΑΣ-16562	99450.56 915	95491.62 748	95170.71 655	88335.00 849	101590.3 883	100794.1 095	111888.7 519	113888.6 433	118442.7 144	114666.3 795	106483.0 86	108394.6 402	99298.39 678	95345.70 347	99450.56 915	95491.62 748	95170.71 655
ΓΛΥΦΑΔ ΑΣ-16674	165012.6 441	160499.9 307	159250.6 75	146387.4 405	174854.3 547	179400.3 636	200494.8 448	218374.7 03	223893.7 117	213861.6 552	190675.2 06	186812.1 571	165012.7 482	160500.0 32	165012.6 441	160499.9 307	159250.6 75

ΓΛΥΦΑΔ ΑΞ-16675	110763.1 056	106886.9 441	106894.3 369	99515.62 719	117606.2 67	117055.8 538	126992.6 928	135851.3 843	143476.1 508	137991.4 711	128865.2 065	126494.3 179	109056.6 218	105238.0 619	110763.1 056	106886.9 441	106894.3 369
ΔΑΦΝΗΣ -17234	36343.63 85	36897.65 607	37009.34 888	33322.55 893	37228.18 197	36152.53 408	37462.71 004	36348.73 017	37392.43 357	37351.57 982	36533.77 56	37803.25 24	36030.66 763	36004.52 266	36343.63 85	36897.65 607	37009.34 888
ΔΑΦΝΗΣ -17235	41295.04 129	41203.48 207	41067.51 98	37748.63 255	42107.78 566	41242.65 313	42580.94 195	41138.46 752	41933.71 473	41895.30 869	41447.87 114	42848.60 43	41018.43 092	40927.33 086	41295.04 129	41203.48 207	41067.51 98
ΔΡΑΠΕΤΣ ΩΝΑΣ- 18648	58661.55 323	58954.84 805	58969.61 267	53900.43 665	59674.80 92	57908.16 521	60601.86 818	59182.10 806	61218.48 837	61158.96 937	59557.21 435	61295.67 495	58324.49 96	58615.94 696	58661.55 323	58954.84 805	58969.61 267
ΕΛΕΥΣΙΝ ΟΣ- 19200	138213.5 607	137567.7 108	137567.7 108	136391.4 881	136391.4 881	138482.3 476	140496.9 944	141354.4 288	141354.4 288	141861.5 036	141861.5 036	139964.5 802	138213.5 607	137567.7 108	138213.5 607	137567.7 108	137567.7 108
ΕΛΛΗΝΙΚ ΟΥ- 16777	83883.50 721	83901.15 136	84191.14 578	76559.82 989	85526.56 885	84425.25 705	98393.31 926	100314.0 362	104600.0 818	99547.95 764	95462.88 041	96845.65 153	82907.70 774	83009.03 801	83883.50 721	83901.15 136	84191.14 578
ΖΕΦΥΡΙΟ Υ-13461	34451.62 886	34734.98 23	33866.91 563	30421.38 883	35231.39 572	37043.27 549	38715.25 946	39192.24 065	41626.95 325	40797.86 453	37399.91 872	36227.79 605	34362.04 175	34644.63 878	34451.62 886	34734.98 23	33866.91 563
ΗΛΙΟΥΠ ΌΛΕΩΣ- 16341	60973.09 851	62821.96 828	61868.49 25	56648.45 635	62311.63 332	64353.16 369	70539.86 539	69289.69 942	71292.74 502	69963.07 611	66990.93 72	66345.55 778	60200.79 972	62034.90 174	60973.09 851	62821.96 828	61868.49 25
ΗΛΙΟΥΠ ΌΛΕΩΣ- 16342	54053.50 803	55407.62 736	54405.40 547	49606.78 566	55484.68 231	56026.80 158	62489.87 479	61352.61 523	63242.04 523	62688.00 931	59919.37 367	58781.56 626	53288.52 929	54806.58 295	54053.50 803	55407.62 736	54405.40 547
ΗΛΙΟΥΠ ΌΛΕΩΣ- 16343	49811.27 609	50889.78 176	50339.17 95	45629.97 111	51225.63 519	51373.73 61	54579.88 67	53262.16 27	54051.21 405	53337.34 733	51382.47 243	52137.90 255	49811.27 609	50889.78 176	49811.27 609	50889.78 176	50339.17 95
ΗΛΙΟΥΠ ΌΛΕΩΣ- 16344	46125.05 693	47584.42 05	46903.71 993	42818.36 75	47532.86 224	48183.45 871	50687.28 862	49003.74 677	48681.96 51	48189.08 331	46573.05 379	48195.91 541	46125.05 693	47584.42 05	46125.05 693	47584.42 05	46903.71 993
ΚΑΖΑ- 19012	71594.87 189	72353.96 193	71138.99 661	65714.04 975	72490.74 778	73683.81 509	78419.14 322	76431.04 495	76722.17 675	75636.25 552	72730.65 626	73071.24 193	71594.87 65	72353.96 659	71594.87 189	72353.96 193	71138.99 661
ΙΛΙΟΥ- 13121	110464.6 903	113121.6 774	112777.0 066	103770.0 268	114730.8 509	112775.2 181	130908.9 157	129044.1 525	131528.9 177	125384.5 364	120250.3 469	121527.9 599	109158.8 558	111815.8 429	110464.6 903	113121.6 774	112777.0 066
ΙΛΙΟΥ- 13122	115369.7 863	118887.0 947	118638.0 392	108948.4 851	120790.8 226	120468.0 129	131271.0 18	130401.9 364	134132.5 898	130033.8 899	125402.9 052	127436.2 499	114794.0 443	118227.7 785	115369.7 863	118887.0 947	118638.0 392
ΙΛΙΟΥ- 13123	29.74279 961	21.75160 145	19.18329 798	18.87445 509	20.11918 689	23.25655 508	22.23826 8	22.84524 342	35.00823 473	55.75953 499	53.79808 239	42.50552 181	29.74279 961	21.75160 145	19.18329 798	18.87445 509	20.11918 689
ΚΑΙΣΑΡΙΑ ΝΗΣ- 16121	114936.7 173	115686.5 902	115355.7 712	105865.4 389	118994.9 934	117381.0 006	122707.1 7	119692.2 946	122439.5 505	120719.4 37	116850.1 441	120913.0 39	114936.7 173	115686.5 902	114936.7 173	115686.5 902	115355.7 712

Πίνακας Π-7: Πίνακας βέλτιστων προβλέψεων καταναλώσεων με βάση τη μεθοδολογία Νο2 για ορισμένες περιοχές των διαθέσιμων Τ.Κ.

	Αυγ-12	Σεπ-12	Οκτ-12	Νοε-12	Δεκ-12	Ιαν-13	Φεβ-13	Μαρ-13	Απρ-13	Μαϊ-13	Ιουν-13	Ιουλ-13	Αυγ-13	Σεπ-13	Οκτ-13	Νοε-13	Δεκ-13
cluster 1	2441323.528	2310911.228	2277819.066	2092374.779	2068721.611	2053102.457	1887727.8	2131955.266	2152930.643	2325047.207	2330343.136	2439580.247	2408422.379	2279732.6	2247052.323	2064081.004	2040716.124
cluster 2	2982618.049	2787843.079	2727363.122	2572253.245	2538400.058	2486931.808	2283958.941	2584635.724	2649083.114	2824322.326	2837611.701	2957105.796	2915151.347	2724663.086	2665436.815	2513738.076	2480545.326
cluster 3	4196119.991	4079980.434	4216877.761	4092926.622	4176798.079	4148501.105	3783748.011	4156482.93	4102999.332	4342887.611	4198291.755	4253269.818	4145339.535	4030555.627	4165742.963	4043244.678	4126046.725
cluster 4	2071149.457	1998277.116	2036570.62	1947425.437	1941424.467	1920316.356	1760792.605	1971527.396	1957388.175	2055379.783	2004512.373	2061891.652	2036727.944	1965020.648	2002629.777	1914925.123	1908979.18
cluster 5	3165785.913	2986632.304	2946982.541	2828073.786	2870023.709	2848461.051	2602613.414	2844674.44	2867840.832	3110528.673	3030122.431	3192681.154	3165785.913	2986632.304	2946982.541	2828073.786	2870023.709
cluster 6	2248841.223	2046380.118	2003571.65	1845534.181	1788065.259	1750614.45	1600354.867	1861247.734	1891362.214	2108560.126	2194253.027	2317950.347	2248841.223	2046380.118	2003571.65	1845534.181	1788065.259
cluster 7	379.0148973	338.0293251	309.4092855	235.6558454	190.6201602	179.4728239	186.0241899	236.5657562	250.720474	332.1939929	375.9647992	387.3036309	382.7938235	353.1393128	338.8441011	277.0206207	244.0535775
cluster 8	302.2533306	288.2494285	278.9674511	272.3179053	266.4052232	259.9131357	252.2259826	243.332713	233.6028634	223.5400683	213.5929693	204.0561217	195.0526379	186.5696334	178.5152175	170.7734023	163.2436507
cluster 9	1038566.311	1014208.985	1053631.659	1014743.37	1026332.889	1010865.718	918508.1403	1023415.273	1011526.552	1058285.931	1022267.955	1049118.437	1038566.311	1014208.985	1053631.659	1014743.37	1026332.889
cluster 10	31708.37071	28382.5383	21862.08995	15651.69159	12205.36286	10019.0854	9460.159522	11684.93613	12936.82648	19597.56262	24354.17873	31509.58213	32257.98524	29079.60202	19878.90779	14419.56559	11676.4462

Πίνακας Π-8: Πίνακας βέλτιστων προβλέψεων καταναλώσεων με βάση τη μεθοδολογία Νο1

B) Σε αυτό το μέρος θα παρουσιαστεί ένα τμήμα των αλγορίθμων που αναπτύχθηκαν για τους σκοπούς της διπλωματικής αυτής.

```
#ConsInClust:amount of consumption in every cluster
#ErrInClust:MAE errors produced by fitted values for every forecasting technique
#TkInClust:postal codes in every cluster
#ForecastInClust:out of sample optimal forecasts for every cluster

library(stringr)
library(forecTheta)
library(MAPA)
library(hts)
library(Metrics)

file<-"D:/Users/Akis/(...) .csv"
data<-read.csv(file,header=T)
file2<-"D:/Users/Akis/(...) .csv"
data2<-read.csv2(file2,header=T)
file3<-"D:/Users/Akis/(...) .csv"
data3<-read.csv(file3,header=T)
n_tk<-as.numeric(str_extract(data2$Area_Tk,"[0-9]+")) #removing letters from the
Area_tk frame
data$tk<-n_tk #creating a column with the tks'
nClust=10 #number of clusters

TkInClust<-data.frame(one,two,three,four,five,six,seven,eight,nine,ten)
ConsInClust=as.data.frame(matrix(0,ncol=ncol(data3),nrow=nClust)) #initializing
consumption in every clust.
d<-numeric(ncol(data3))
meanCnt<-apply(data3, 1, function(x) mean(x[x>0]))

for (j in 1:nClust)
  for (i in 1:length(data$tk))
    if (is.element(data$tk[i],TkInClust[[j]]))
    {
      d<-data[i,1:84]*round(meanCnt[i])
      ConsInClust[j,]<-ConsInClust[j,]+d
    }

ForecastInClust=as.data.frame(matrix(0,ncol=17,nrow=10))
ErrInClust=as.data.frame(matrix(0,ncol=7,nrow=10))
colnames(ErrInClust)<-c("Ets Mae","Arima Mae","Theta Mae","Opt.Theta Mae","MAPA
Mae","Naive Mae","Comb.Method")
combo=as.data.frame(matrix(0,ncol=84,nrow=5))

for (i in 1:nClust)
{
  #take each row of ConsInClust and make predictions using ets/auto.arima function
  #put the mape errors in the ErrInClust data frame for every method
  cons<-as.numeric(ConsInClust[i,])
  insample=head(cons,0.8*length(cons))
  outsample=tail(cons,0.2*length(cons))

  fit.ets<-ets(ts(insample,freq=12,start=c(2007,1)),model="ZZZ")
  forecast.ets=forecast(fit.ets,h=17)
  ErrInClust[i,1]=accuracy(fit.ets[["fitted"]],insample)[5]
  min=ErrInClust[i,1]
}
```

```

if ((i==7)|| (i==8))
  fit.arima<-auto.arima(ts(insample,freq=12,start=c(2007,1)),lambda=0.45)
else
  fit.arima<-auto.arima(ts(insample,freq=12,start=c(2007,1)))
forecast.arima=forecast(fit.arima,h=17)
ErrInClust[i,2]=accuracy(forecast.arima[["fitted"]],insample)[5]
  if (ErrInClust[i,2]<min)
    min=ErrInClust[i,2]

fit.theta<-thetaM(ts(insample,freq=12,start=c(2007,1)),h=17)
ErrInClust[i,3]=accuracy(fit.theta[["fitted"]],insample)[5]
  if (ErrInClust[i,3]<min)
    min=ErrInClust[i,3]

fit.Otheta<-otm(ts(insample,freq=12,start=c(2007,1)),h=17)
ErrInClust[i,4]=accuracy(fit.Otheta[["fitted"]],insample)[5]
  if (ErrInClust[i,4]<min)
    min=ErrInClust[i,4]

fit.mapa<-mapa(ts(insample,freq=12,start=c(2007,1)),fh=17,outplot=0,model="ZZZ")
ErrInClust[i,5]=accuracy(as.numeric(fit.mapa[["infor"]]),insample)[5]
  if (ErrInClust[i,5]<min)
    min=ErrInClust[i,5]

fit.naive<-naive(ts(insample,freq=12,start=c(2007,1)),h=17)
ErrInClust[i,6]=accuracy(fit.naive[["fitted"]],insample)[5]
  if (ErrInClust[i,6]<min)
    min=ErrInClust[i,6]

vec=as.numeric(ErrInClust[i,1:5])
col=colnames(ErrInClust)[1:5]
no_passes = 0
while(1) {
no_swaps = 0
for (j in 1 : (length(vec) - 1 - no_passes)) {
  if (vec[j] > vec[j + 1]) {
    tmp1 = vec[j]
    tmp2=col[j]
    vec[j] = vec[j+1]
    col[j]=col[j+1]
    vec[j+1] = tmp1
    col[j+1]=tmp2
    no_swaps = no_swaps + 1
  }
}
no_passes = no_passes + 1
if(no_swaps == 0) break
}

q=1
while (q<=4){
  if (col[q]=="Ets Mae"){
    combo[q,1:67]<-fit.ets[["fitted"]]
    combo[q,68:84]<-forecast.ets[["mean"]]
  }
  else if (col[q]== "Arima Mae"){
    combo[q,1:67]<-forecast.arima[["fitted"]]
    combo[q,68:84]<-forecast.arima[["mean"]]
  }
}

```



```

else if (col[q]=="Theta Mae"){
  combo[q,1:67]<-fit.theta[["fitted"]]
  combo[q,68:84]<-fit.theta[["mean"]]
}
else if (col[q]=="Opt.Theta Mae"){
  combo[q,1:67]<-fit.Otheta[["fitted"]]
  combo[q,68:84]<-fit.Otheta[["mean"]]
}
else if (col[q]=="MAPA Mae"){
  combo[q,1:67]<-fit.mapa[["infor"]]
  combo[q,68:84]<-fit.mapa[["outfor"]]
}
q=q+1
}
comb.forec<-apply(combo[1:3,], 2, function(x) mean(x,na.rm=T))
ErrInClust[i,7]=accuracy(comb.forec[1:67],insample)[5]
if (ErrInClust[i,7]<min)
  min=ErrInClust[i,7]

if (min==ErrInClust[i,1])
  ForecastInClust[i,]<-forecast.ets[["mean"]]
else if (min==ErrInClust[i,2])
  ForecastInClust[i,]<-forecast.arima[["mean"]]
else if (min==ErrInClust[i,3])
  ForecastInClust[i,]<-fit.theta[["mean"]]
else if (min==ErrInClust[i,4])
  ForecastInClust[i,]<-fit.Otheta[["mean"]]
else if (min==ErrInClust[i,6])
  ForecastInClust[i,]<-fit.naive[["mean"]]
else if (min==ErrInClust[i,7])
  ForecastInClust[i,]<-comb.forec[68:84]
else
  ForecastInClust[i,]<-fit.mapa[["outfor"]]
}

df<-t(ForecastInClust)
row.names(df)<-NULL
bts<-ts(df,freq=12,start=c(2012,8))
colnames(bts)<-
c("cl.1","cl.2","cl.3","cl.4","cl.5","cl.6","cl.7","cl.8","cl.9","cl.10")
y2<-hts(bts,nodes=list(ncol(bts)))
total2<-aggts(y2,levels=c(0))

```

Κώδικας 1-Εύρεση πρόβλεψη κατανάλωσης σε κάθε cluster(ForecastInClust), και τελική πρόβλεψη κατανάλωσης (total2) μεθοδολογίας Νο1

```

library(stringr)
library(forecTheta)
library(MAPA)

file<-"D:/Users/Akis/(...) .csv"
data<-read.csv(file,header=T)
ConsInTk<-data

ForecastInTk=as.data.frame(matrix(0,ncol=17,nrow=nrow(ConsInTk)))
ErrInTk=as.data.frame(matrix(0,ncol=7,nrow=nrow(ConsInTk)))
colnames(ErrInTk)<-c("Ets Mape","Arima Mape","Theta Mape","Opt.Theta Mape","MAPA
Mape","Naive Mape","Comb.method")
combo=as.data.frame(matrix(0,ncol=84,nrow=5))

for (i in 1:nrow(ConsInTk)){
  cons<-as.numeric(ConsInTk[i,])
  insample=head(cons,0.8*length(cons))
  outsample=tail(cons,0.2*length(cons))

  fit.ets<-ets(ts(insample,freq=12,start=c(2007,1)),model="ZZZ")
  forecast.ets=forecast(fit.ets,h=17)
  ErrInTk[i,1]=accuracy(fit.ets[["fitted"]],insample)[5]
  min=ErrInTk[i,1]

  fit.arima<-auto.arima(ts(insample,freq=12,start=c(2007,1)))
  forecast.arima=forecast(fit.arima,h=17)
  ErrInTk[i,2]=accuracy(forecast.arima[["fitted"]],insample)[5]
  if (ErrInTk[i,2]<min)
    min=ErrInTk[i,2]

  fit.theta<-thetaM(ts(insample,freq=12,start=c(2007,1)),h=17)
  ErrInTk[i,3]=accuracy(fit.theta[["fitted"]],insample)[5]
  if (ErrInTk[i,3]<min)
    min=ErrInTk[i,3]

  fit.Otheta<-otm(ts(insample,freq=12,start=c(2007,1)),h=17)
  ErrInTk[i,4]=accuracy(fit.Otheta[["fitted"]],insample)[5]
  if (ErrInTk[i,4]<min)
    min=ErrInTk[i,4]

  fit.mapa<-
  mapa(ts(insample,freq=12,start=c(2007,1)),fh=17,outplot=0,model="ZZZ")
  ErrInTk[i,5]=accuracy(as.numeric(fit.mapa[["infor"]]),insample)[5]
  if (ErrInTk[i,5]<min)
    min=ErrInTk[i,5]

  fit.naive<-naive(ts(insample,freq=12,start=c(2007,1)),h=17)
  ErrInTk[i,6]=accuracy(fit.naive[["fitted"]],insample)[5]
  if (ErrInTk[i,6]<min)
    min=ErrInTk[i,6]

  vec=as.numeric(ErrInTk[i,1:5])
  col=colnames(ErrInTk)[1:5]
  no_passes = 0
  while(1) {
    no_swaps = 0
    for (j in 1 : (length(vec) - 1 - no_passes)) {
      try({if (vec[j] > vec[j + 1]) {
        tmp1 = vec[j]
        tmp2=col[j]
        vec[j] = vec[j+1]

```

```

        vec[j+1] = tmp1
        col[j+1]=tmp2
        no_swaps = no_swaps + 1
    })
}
no_passes = no_passes + 1
if(no_swaps == 0) break
}
q=1
while (q<=4){
  if (col[q]=="Ets Mape"){
    combo[q,1:67]<-fit.ets[["fitted"]]
    combo[q,68:84]<-forecast.ets[["mean"]]
  }
  else if (col[q]== "Arima Mape"){
    combo[q,1:67]<-forecast.arima[["fitted"]]
    combo[q,68:84]<-forecast.arima[["mean"]]
  }
  else if (col[q]=="Theta Mape"){
    combo[q,1:67]<-fit.theta[["fitted"]]
    combo[q,68:84]<-fit.theta[["mean"]]
  }
  else if (col[q]=="Opt.Theta Mape"){
    combo[q,1:67]<-fit.Otheta[["fitted"]]
    combo[q,68:84]<-fit.Otheta[["mean"]]
  }
  else if (col[q]=="MAPA Mape"){
    combo[q,1:67]<-fit.mapa[["infor"]]
    combo[q,68:84]<-fit.mapa[["outfor"]]
  }
  q=q+1
}
comb.forec<-apply(combo[1:3,], 2, function(x) mean(x,na.rm=T))
ErrInTk[i,7]=accuracy(comb.forec[1:67], insample) [5]
if (ErrInTk[i,7]<min)
  min=ErrInTk[i,7]

if (min==ErrInTk[i,1])
  ForecastInTk[i,]<-forecast.ets[["mean"]]
else if (min==ErrInTk[i,2])
  ForecastInTk[i,]<-forecast.arima[["mean"]]
else if (min==ErrInTk[i,3])
  ForecastInTk[i,]<-fit.theta[["mean"]]
else if (min==ErrInTk[i,4])
  ForecastInTk[i,]<-fit.Otheta[["mean"]]
else if (min==ErrInTk[i,6])
  ForecastInTk[i,]<-fit.naive[["mean"]]
else if (min==ErrInTk[i,5])
  ForecastInTk[i,]<-fit.mapa[["outfor"]]
else
  ForecastInTk[i,]<-comb.forec[68:84]
}

```

Κώδικας 2-Εύρεση πρόβλεψης κατανάλωσης για όλους τους διαθέσιμους ταχυδρομικούς κώδικες της μεθοδολογίας Νο2

```

library(stringr)
library(forecTheta)
library(MAPA)

file<-"D:/Users/(...) .csv"
data<-read.csv(file,header=T)

TotalCons<-as.data.frame(matrix(0,ncol=ncol(data),nrow=1))
TotalForec<-as.data.frame(matrix(0,ncol=17,nrow=1))
combo=as.data.frame(matrix(0,ncol=84,nrow=5))

for (i in 1:nrow(data))
  TotalCons<-TotalCons+data[i,]

TotalErr<-as.data.frame(matrix(0,ncol=7,nrow=1))
colnames(TotalErr)<-c("Ets Mape", "Arima Mape", "Theta Mape", "Opt.Theta
Mape", "MAPA Mape", "Naive Mape", "Combined Mape")

cons<-as.numeric(TotalCons[1,])
insample=head(cons,0.8*length(cons))
outsample=tail(cons,0.2*length(cons))

fit.ets<-ets(ts(insample,freq=12,start=c(2007,1)),model="ZZZ")
forecast.ets=forecast(fit.ets,h=17)
TotalErr[1,1]=accuracy(fit.ets[["fitted"]],insample)[5]
min=TotalErr[1,1]

fit.arima<-auto.arima(ts(insample,freq=12,start=c(2007,1)))
forecast.arima=forecast(fit.arima,h=17)
TotalErr[1,2]=accuracy(forecast.arima[["fitted"]],insample)[5]
if (TotalErr[1,2]<min)
  min=TotalErr[1,2]

fit.theta<-thetaM(ts(insample,freq=12,start=c(2007,1)),h=17)
TotalErr[1,3]=accuracy(fit.theta[["fitted"]],insample)[5]
if (TotalErr[1,3]<min)
  min=TotalErr[1,3]

fit.Otheta<-otm(ts(insample,freq=12,start=c(2007,1)),h=17)
TotalErr[1,4]=accuracy(fit.Otheta[["fitted"]],insample)[5]
if (TotalErr[1,4]<min)
  min=TotalErr[1,4]

fit.mapa<-
mapa(ts(insample,freq=12,start=c(2007,1)),fh=17,outplot=0,model="ZZZ")
TotalErr[1,5]=accuracy(as.numeric(fit.mapa[["infor"]]),insample)[5]
if (TotalErr[1,5]<min)
  min=TotalErr[1,5]

fit.naive<-naive(ts(insample,freq=12,start=c(2007,1)),h=17)
TotalErr[1,6]=accuracy(fit.naive[["fitted"]],insample)[5]
if (TotalErr[1,6]<min)
  min=TotalErr[1,6]

vec=as.numeric(TotalErr[i,1:5])
col=colnames(TotalErr)[1:5]
no_passes = 0

```

```

while(1) {
  no_swaps = 0
  for (j in 1 : (length(vec) - 1 - no_passes)) {
    if (vec[j] > vec[j + 1]) {
      tmp1 = vec[j]
      tmp2=col[j]
      vec[j] = vec[j+1]
      col[j]=col[j+1]
      vec[j+1] = tmp1
      col[j+1]=tmp2
      no_swaps = no_swaps + 1
    }
  }
  no_passes = no_passes + 1
  if(no_swaps == 0) break
}
q=1
while (q<=4){
  if (col[q]=="Ets Mape"){
    combo[q,1:67]<-fit.ets[["fitted"]]
    combo[q,68:84]<-forecast.ets[["mean"]]
  }
  else if (col[q]== "Arima Mape"){
    combo[q,1:67]<-forecast.arima[["fitted"]]
    combo[q,68:84]<-forecast.arima[["mean"]]
  }
  else if (col[q]=="Theta Mape"){
    combo[q,1:67]<-fit.theta[["fitted"]]
    combo[q,68:84]<-fit.theta[["mean"]]
  }
  else if (col[q]=="Opt.Theta Mape"){
    combo[q,1:67]<-fit.Otheta[["fitted"]]
    combo[q,68:84]<-fit.Otheta[["mean"]]
  }
  else if (col[q]=="MAPA Mape"){
    combo[q,1:67]<-fit.mapa[["infor"]]
    combo[q,68:84]<-fit.mapa[["outfor"]]
  }
  q=q+1
}
comb.forec<-apply(combo[1:3,], 2, function(x) mean(x,na.rm=T))
TotalErr[1,7]=accuracy(comb.forec[1:67],insample)[5]
if (TotalErr[1,7]<min)
  min=TotalErr[1,7]
if (min==TotalErr[1,1])
  TotalForec[1,]<-forecast.ets[["mean"]]
else if (min==TotalErr[1,2])
  TotalForec[1,]<-forecast.arima[["mean"]]
else if (min==TotalErr[1,3])
  TotalForec[1,]<-fit.theta[["mean"]]
else if (min==TotalErr[1,4])
  TotalForec[1,]<-fit.Otheta[["mean"]]
else if (min==TotalErr[1,6])
  TotalForec[1,]<-fit.naive[["mean"]]
else if (min==TotalErr[i,7])
  TotalForec[1,]<-comb.forec[68:84]
else
  TotalForec[1,]<-fit.mapa[["outfor"]]

```

Κώδικας 3- Εύρεση πρόβλεψης συνολικής κατανάλωσης (TotalForec) για τη μεθοδολογία Νο3

```

library(stringr)
File<- "D:/Users/Akis/Documents/R/R_diploma/diploma/stat_zipc.txt"
data<-read.table(File, sep=";", header=TRUE)
file2<-"D:/Users/Akis/Desktop/Most recently used/Area_TK2.csv"
data2<-read.csv2(file2, header=T)
n_tk<-as.numeric(str_extract(data2$Area_Tk, "[0-9]+"))

#initializing data frame
df<-as.data.frame(matrix(NA, ncol=84, nrow=274))

selectedYear <- c(2007:2013) #Select Year
d=0
for (i in n_tk)
{
  selectedTK<-i
  d=d+1

  #Initialization
  data$Year <- as.numeric(substring(data$GMD_MONTH_CODE, 1, 4))
  data$Month <- as.numeric(substring(data$GMD_MONTH_CODE, 5, 6))
  cons <- numeric(12*length(selectedYear))
  cnt <- numeric(12*length(selectedYear))

  #Subsetting
  selectedDataYear <- NA
  selectedData <- NA
  for (Etos in 1:length(selectedYear))
    selectedDataYear <-
rbind(selectedDataYear, subset(data, data$Year==selectedYear[Etos]))
  for (TK in 1:length(selectedTK))
    selectedData <-
rbind(selectedData, subset(selectedDataYear, selectedDataYear$PAR_TK==selectedTK[TK]))
  selectedData <- na.omit(selectedData)

  #CNT find
  for (i in 1:nrow(selectedData))
    for (k in 0:(length(selectedYear)-1))
      for (j in 1:12)
        if (selectedData$Month[i]==j && selectedData$Year[i]==selectedYear[k+1])
          cnt[j+(k*12)] <- cnt[j+(k*12)]+selectedData$CNT[i]

  df[d,]<-cnt
}

```

Κώδικας 4- Εύρεση των μετρητών ανά ταχυδρομικό κώδικα

```

library(stringr)
#Select Number of Clusters
n_clusters = 10
#Data
data <- read.csv(file="D:/Users/Akis/Desktop/Most recently
used/ConsPerCnt(clean).csv")
file2<-"D:/Users/Akis/Desktop/Most recently used/Area_TK2.csv"
data2<-read.csv2(file2,header=T)
n_tk<-as.numeric(str_extract(data2$Area_Tk,"[0-9]+")) #removing letters from the
Area_tk frame
data$TK<-n_tk #creating a column with the tk's
data_clust <- as.matrix(data[1:85])

#data$TK <- sub(".*-", "", data$Area_Tk)
#data_clust <- as.matrix(data[1:85])
row.names(data_clust) <- data$TK
#Clustering
d<-dist(na.omit(data_clust), method = "euclidean") #Cluster after NA removal
fit<-hclust(d, method="ward")
groups <- cutree(fit, k=n_clusters) # cut tree into clusters
#Plot
plot(fit)
rect.hclust(fit, k=n_clusters, border="red") # draw dendogram with red borders
around the clusters

```

Κώδικας 5 -Εύρεση των cluster για την μεθοδολογία Νο1

Βιβλιογραφία

Papers

- Armstrong, J. S. (2001a). "Evaluating forecasting methods", In: J. S. Armstrong (Ed.), Principles of forecasting: A handbook for researchers and practitioners, Boston, MA: Kluwer Academic Publishing, pp. 443 – 472.
- Zhou,S.,McMahon,T.,Walton,A., and Lewis, J. (2000). " Forecasting daily urban water demand: A case study of Melbourne." J. Hydrol., 236(3-4), 153-164.
- Alhumoud J.(2008). "Freshwater consumption in Kuwait: Analysis and forecasting" J.water supply:Res.Technol.-AQUA,57(4),279-288.
- Arbues, F., Garcia-Valinas, M., A., and Martinez-Espineira, R.(2003). Estimation of residential water demand: A state of the art review." J.Soc.Econ.,32(1),81-102.
- Bougadis, J., Adamowski, K. , and Diduch, R. (2005). "Short term municipal water demand forecasting." Hydrolog.Process.,19(1),137-148.
- Firat, M., Yurdusev, M., and Turan, M.(2009). "Evaluation of artificial neural network techniques for municipal water consumption modeling." Water Resour.Manage., 23(4), 617-632.
- Jain,A.,Varshney, A., and Joshi, U.(2001). "Short term water demand forecast modelling techniques versus AI". J.Am. Water Works Assoc.,94(7), 64-72.
- Armstrong, J. S. (2001b). "Combining forecasts", In: J. S. Armstrong (Ed.), Principles of forecasting: A handbook for researchers and practitioners, Boston, MA: Kluwer Academic Publishing, pp. 417 – 439.
- Assimakopoulos, V. and Nikolopoulos, N. (2000) "The theta model: a decomposition approach to forecasting', International Journal of Forecasting, Vol. 16, No. 4, pp. 521-530.
- Chatfield, C. (1988) "Apples, Oranges and Mean Square Error", International Journal of Forecasting, Vol. 4, pp. 515-518.
- Clemen, R.T. (1989) "Combining forecasts: A review and annotated biography (with discussion)" International Journal of Forecasting, Vol. 5, pp. 559-583.
- De Gooijer, J. G. and Hyndman, R. J. (2006) "25 years of time series forecasting", International Journal of Forecasting, Vol. 22, pp. 443-473.
- Fildes, R., Goowin, P., Lawrence, M. and Nikolopoulos, K. (2009) "Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning", International Journal of Forecasting, Vol. 25, pp. 3-23.
- Gardner, E.S. (1985) "Exponential Smoothing: The State of the Art" Journal of Forecasting, Vol. 4, pp. 1-28.

- Gardner, E.S. (2006) "Exponential Smoothing: The State of the Art – Part II", *International Journal of Forecasting*, Vol. 22, pp. 637-666.
- Goodwin, P. and Lawton, R. (1999) "On the asymmetry of the symmetric MAPE", *International Journal of Forecasting*, Vol. 15, pp. 405-408.
- Hyndman, R.J. and Koehler, A.B. (2006) "Another look at measures of forecast accuracy", *International Journal of Forecasting*, Vol. 22, pp. 679-688.
- Makridakis, S. (1993) "Accuracy measures - theoretical and practical concerns", *International Journal of Forecasting*, Vol. 9, pp. 527-529.
- Makridakis, S. and Hibon, M. (2000) "The M3-Competition: Results, conclusions and implications", *International Journal of Forecasting*, Vol. 16, No. 4, pp. 451-476.
- Nikolopoulos, K., Assimakopoulos, V., Bougioukos, N. and Petropoulos F. (2008) "Advances in Theta model", Working Paper No. 0023, University of Peloponnese, Department of Economics.
- Yokum, J. T. and Armstrong, J. S. (1995) "Beyond accuracy: Comparison of criteria used to select forecasting methods", *International Journal of Forecasting*, Vol. 11, pp. 591-597.
- Cort J. Willmott, Kenji Matsuura "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance"
- Rob J Hyndman, Roman A Ahmed, George Athanasopoulos, Han Lin Shang (2010) "Optimal combination forecasts for hierarchical time series"
- Lily A. House-Peters and Heejun Chang (2011) "Urban water demand modeling: Review of concepts, methods, and organizing principles"
- Nikolaos Kourentzes, Fotios Petropoulos, Juan R. Trapero (2013) "Improving forecasting by estimating time series structural components across multiple frequencies"
- Jorge Caiado (2009) "Performance of combined double seasonal univariate time series models for forecasting water consumption"
- José Augusto Fioruci, Tiago Ribeiro Pellegrini, Francisco Louzada and Fotios Petropoulos (2013) "The Optimised Theta Method"
- Rob J Hyndman, George Athanasopoulos, Han Lin Shang (2014) "hts: An R Package for Forecasting Hierarchical or Grouped Time Series"
- Javier López-de-Lacalle (2014) "tsoutliers R Package for Detection of Outliers in Time Series"
- D.Kofinas,N.Mellios,E.Papageorgiou,C.Laspidou (2014) "Urban Water Demand Forecasting for the Island of Skiathos", *Procedia Engineering* 89 (2014) 1023-1030
- Sadeh Behboudian, Massoud Tabesh, Maliheh Falahnezhad and Farrokh Alavian Ghavanini "A long-term prediction of domestic water demand using preprocessing in artificial neural network", *Journal of Water Supply: Research and Technology-AQUA* (2014)

- Rob J. Hyndman and Yeasmin Khandakar. Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, 27(3):1–22, 2008.
- D. S. G. Pollock. *A Handbook of Time-Series Analysis Signal Processing and Dynamics*. Academic Press, 1999.
- Chung Chen and Lon-Mu Liu. Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association*, 88(421):284–297, 1993.
- Athanasopoulos G, Ahmed RA, Hyndman RJ (2009). “Hierarchical forecasts for Australian domestic tourism.” *International Journal of Forecasting*, 25(1), 146–166.
- Wright, D.J., Capon, G., Page, R., Quiroga, J., Taseen, A.A. and Tomasini, F. (1986) “Evaluation of forecasting methods for decision support”, *International Journal of Forecasting*, Vol. 2, No. 2, pp. 139–153.
- D Plous, S. (1995), “A comparison of strategies for reducing interval overconfidence in group judgments,” *Journal of Applied Psychology*, 80, 443-454.
- D Meade, N. & T. Islam (1998), “Technological forecasting - model selection, model stability, and combining models,” *Management Science*, 44, 1115-1130.
- D Lobo, G. J. & R. D. Nair (1990), “Combining judgmental and statistical forecasts: An application of earnings forecasts,” *Decision Sciences*, 21, 446-460.
- D Winkler, R. L. & S. Makridakis (1983), “The combination of forecasts,” *Journal of the Royal Statistical Society, Series A*, 146, Part 2, 150-157.
- Tian, Y. & Wiens, D. P. (2006), “On equality and proportionality of ordinary least squares, weighted least squares and best linear unbiased estimators in the general linear model”, *Statistics and Probability Letters* 76(12), 1265–1272.
- Shing, N. K. (1993), “A Study of Bottom-up and Top-down Forecasting Methods”, M.Sc. thesis, Royal Melbourne Institute of Technology.
- Schwarzkopf, A. B., Tersine, R. J. & Morris, J. S. (1988), ‘Top-down versus bottom-up forecasting strategies’, *International Journal of Production Research* 26(11), 1833–1843.

Βιβλία

- Πετρόπουλος Φ, Ασημακόπουλος Β, (2011). “Επιχειρησιακές Προβλέψεις”, εκδόσεις συμμετρία, Αθήνα.
- R. Bruce Billigs, Clive V. Jones (2008) .“Forecasting Urban Water Demand” SE , American Water Works Association.
- Makridakis, S., Hogarth R. and Gaba A. (2010). *Dance with Chance: Making Luck Work for You*, Oneworld Publications.