

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
National Technical University of Athens



ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧ. & ΜΗΧ. ΥΠΟΛΟΓΙΣΤΩΝ
School of Electrical and Computer Engineering

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜ/ΚΩΝ ΔΙΑΤΑΞΕΩΝ & ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ
Division of Industrial Electric Devices and Decision Systems

Εκπόνηση Διπλωματικής Εργασίας
στη
ΜΟΝΑΔΑ ΠΡΟΒΛΕΨΕΩΝ & ΣΤΡΑΤΗΓΙΚΗΣ

ΑΞΙΟΛΟΓΗΣΗ ΕΠΙΔΡΑΣΗΣ ΤΗΣ ΙΕΡΑΡΧΗΣΗΣ ΔΕΔΟΜΕΝΩΝ
ΣΤΗΝ ΑΠΟΔΟΣΗ ΠΡΟΒΛΕΨΗΣ ΠΩΛΗΣΕΩΝ

συγγραφή:
ΙΩΑΝΝΗΣ ΤΖΑΝΑΚΑΚΗΣ
υπ. διπλωματούχος Ε.Μ.Π.

επιβλέπων:
ΒΑΣΙΛΕΙΟΣ ΑΣΗΜΑΚΟΠΟΥΛΟΣ
καθηγητής Ε.Μ.Π.

υπεύθυνος:
ΕΥΑΓΓΕΛΟΣ ΣΠΗΛΙΩΤΗΣ
διδάκτωρ Ε.Μ.Π.

ΑΘΗΝΑ, ΝΟΕΜΒΡΙΟΣ 2021

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
National Technical University of Athens

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧ. & ΜΗΧ. ΥΠΟΛΟΓΙΣΤΩΝ
School of Electrical and Computer Engineering



ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜ/ΚΩΝ ΔΙΑΤΑΞΕΩΝ & ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ
Division of Industrial Electric Devices and Decision Systems

Εκπόνηση Διπλωματικής Εργασίας
στη
ΜΟΝΑΔΑ ΠΡΟΒΛΕΨΕΩΝ & ΣΤΡΑΤΗΓΙΚΗΣ

ΑΞΙΟΛΟΓΗΣΗ ΕΠΙΔΡΑΣΗΣ ΤΗΣ ΙΕΡΑΡΧΗΣΗΣ ΔΕΔΟΜΕΝΩΝ
ΣΤΗΝ ΑΠΟΔΟΣΗ ΠΡΟΒΛΕΨΗΣ ΠΩΛΗΣΕΩΝ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 3η Νοεμβρίου 2021.

.....
ΒΑΣΙΛΕΙΟΣ ΑΣΗΜΑΚΟΠΟΥΛΟΣ
καθηγητής Ε.Μ.Π.

.....
ΙΩΑΝΝΗΣ ΨΑΡΡΑΣ
καθηγητής Ε.Μ.Π.

.....
ΔΗΜΗΤΡΙΟΣ ΑΣΚΟΤΗΝΗΣ
καθηγητής Ε.Μ.Π.

ΑΘΗΝΑ, ΝΟΕΜΒΡΙΟΣ 2021

Ιωάννης Τζανακάκης

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

Copyright © - All rights reserved | Ιωάννης Τζανακάκης, 2021

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Ατυχείς Προβλέψεις #1

«Δεν υπάρχει λόγος να θέλει κανείς
έναν υπολογιστή στο σπίτι του.»

Ken Olson
πρόεδρος της *DEC Computers*
(*Digital Equipment Corporation*), 1977

Περίληψη

Η ιεραρχική πρόβλεψη έχει αποτελέσει για σχεδόν έναν αιώνα, χρήσιμο εργαλείο για ευθυγραμμισμένη λήψη απόφασης. Εμπειρικά αποτελέσματα της τελευταίας δεκαετίας έχουν δείξει πως μέθοδοι ιεραρχικής πρόβλεψης μπορούν να βελτιώνουν την ακρίβεια πρόβλεψης, καθώς ενσωματώνουν πληροφορία από διαφορετικά επίπεδα της ιεραρχίας, που δεν αποτυπώνεται σε κάθε ξεχωριστή χρονοσειρά. Ωστόσο, όσο οι ιεραρχίες γίνονται συνθετότερες και οι χρονοσειρές στα επίπεδα συνάθροισης πολυπληθέστερες, αυξάνεται και το κόστος πρόβλεψης. Αν, για παράδειγμα, οργανωθούν εξαντλητικά πωλήσεις χιλιάδων μονάδων σε ιεραρχίες για τις κατηγορίες προϊόντων και για τις γεωγραφικές διαμερίσεις των πωλήσεων, προκύπτουν συλλογές χρονοσειρών με υψηλό κόστος μοντελοποίησης και χωρίς εγγύηση για την προβλεπτική αξία κάθε επιπέδου συνάθροισης, όταν ζητούμενο είναι η ενίσχυση της ακρίβειας πρόβλεψης ενός υποσυνόλου των δεδομένων.

Στόχος της παρούσας εργασίας είναι η αξιολόγηση της επίδρασης που έχουν διαφορετικές ιεραρχήσεις των δεδομένων, στην απόδοση των προβλέψεων. Αναπτύσσεται μια μεθοδολογία για τη σύγκριση και αυτόματη διαλογή μιας ή περισσότερων ιεραρχιών, ως κατάλληλες παραμέτρους βελτιστοποίησης τόσο της ακρίβειας όσο και του κόστους των προβλέψεων ενός ή περισσότερων επιπέδων συνάθροισης. Η μεθοδολογία αυτή εφαρμόζεται και αξιολογείται πάνω στα δεδομένα που τέθηκαν υπό πρόβλεψη στο διαγωνισμό *M5* και περιλαμβάνουν χρονοσειρές πωλήσεων με γεωγραφική και ανά προϊόν οργάνωση.

Από τη μελέτη περίπτωσης προκύπτει πως τα επίπεδα συνάθροισης μιας συλλογής χρονοσειρών έχουν μη κοινή προβλεπτική αξία. Η συμπερίληψη επιπέδων συνάθροισης που περιέχουν πολλές χρονοσειρές, δεν οδηγεί απαραίτητα σε ακριβέστερες προβλέψεις, ενώ αυξάνει σημαντικά το κόστος της πρόβλεψης, ιδιαίτερα όσο ψηλότερα τοποθετούνται στην ιεραρχία. Αντίθετα, ιεραρχίες που περιέχουν -αποκλειστικά- επίπεδα υψηλής συνάθροισης, έχουν χαμηλό κόστος πρόβλεψης, αλλά και χαμηλότερη μεσοσταθμική ακρίβεια στις προβλέψεις των άλλων επιπέδων. Επίπεδα συνάθροισης, τα οποία βρίσκονται στο μέσο της ή των ξεχωριστών ιεραρχιών που συνδέουν τα δεδομένα, προκύπτει πως διαθέτουν μεσοσταθμικά, την υψηλότερη προβλεπτική ικανότητα, ενώ το κόστος της μοντελοποίησης παραμένει χαμηλό. Τέλος, οι μέθοδοι που συνδυάζουν προβλέψεις διαφόρων επιπέδων συνάθροισης, παράγουν ακριβέστερες προβλέψεις σε σχέση με τις παραδοσιακές μεθόδους ιεραρχικής πρόβλεψης *bottom up* και *top down*.

Λέξεις Κλειδιά: Χρονοσειρά, Τεχνικές Προβλέψεων, Ιεραρχικές Χρονοσειρές, Ιεραρχική Πρόβλεψη, Εναρμόνιση Προβλέψεων, Ακρίβεια Πρόβλεψης, Κόστος Πρόβλεψης, Σύνολο δεδομένων *M5*, Διαγωνισμοί *M*.

Abstract

Hierarchical forecasting has been for almost a century, a useful tool for aligned decision-making. Empirical results over the past decade have shown that hierarchical forecasting methods can improve forecast accuracy, as they incorporate information from different levels of the hierarchy that is not captured in the data of every single time series. However, as the hierarchies become more complex and the time series in the aggregation levels increase, so does the cost of forecasting such collections. When sales of thousands of units are exhaustively organized by product categories and geographic divisions, they result in collections of millions of time series, with high modelling cost and without guarantees for the contribution of every aggregation level in the forecast accuracy, when the goal is the enhancement of the forecasts for a subset of the collection.

The goal of this study is the evaluation of the impact that different groupings of the data have upon sales forecasts. A methodology is proposed, that evaluates forecasts that have been reconciled with different hierarchies. The goal of the comparison is the selection one or more hierarchies, as optimal parameters in terms of accuracy and cost of forecasts for one or more aggregation levels. The methodology is evaluated in the data provided in the M5 Competition and contain time series of unit sales, organized by product categories and geographic divisions.

From the case study occurs that different aggregation levels in a collection of time series have different prediction value. The inclusion of aggregation levels that contain many time series, does not necessarily result in more accurate forecasts, while it upsurges forecasting cost significantly. On the other hand, hierarchies that include -exclusively- levels of high aggregation, show reduced forecasting cost but result in worse accuracy on average. Levels that are located in the middle of one or more hierarchies that connect the data, occurs that contain the biggest prediction value on average, while the modelling cost remains low. At last, linear combination methods outperform the traditional single-level approaches bottom up and top down.

Key Words: Time Series, Forecasting Techniques, Hierarchical Time Series, Hierarchical Forecasting, Forecast Reconciliation, Forecast Accuracy, Forecast Cost, *M5* Dataset, *M* Competitions.

Ευχαριστίες

Η εκπόνηση του παρόντα τόμου σηματοδοτεί την ολοκλήρωση του κύκλου των προπτυχιακών μου σπουδών και την αφετηρία νέων παραστάσεων, προκλήσεων και εμπειριών. Το τέλος αυτού του ταξιδιού με αφήνει διαφορετικό από τον -μέχρι πρότινος- μαθητή λυκείου, που μπήκε για πρώτη φορά στα αμφιθέατρα του Πολυτεχνείου. Έμαθα να πετυχαίνω εκεί που είχα αποτύχει, να υπερπηδώ ψηλά εμπόδια, να θέτω μακρόπνοους στόχους και να συνεργάζομαι με ταλαντούχους ανθρώπους. Με αποτυχίες και απογοητεύσεις, στιγμές υψηλής πίεσης και άγχους που με έβγαλαν από τη ζώνη άνεσης και με εξέλιξαν, αλλά και επιτυχίες που μου χάρισαν δυνατά συναισθήματα ικανοποίησης και αυτοπεποίθησης.

Συμπαραστάτη σε αυτήν, όπως και όλες τις σημαντικές εμπειρίες της ζωής μου, είχα την οικογένειά μου. Στους γονείς μου Στέφανο και Γιώτα και την αδερφή μου Εύα, οφείλω το μεγαλύτερο ευχαριστώ, για την αδιαπραγμάτευτη υποστήριξη και ανοχή τους. Κυρίως για τα διδάγματα της αγάπης για τη ζωή και των ανοιχτών κεραιών στα ερεθίσματα του κόσμου γύρω μας, που κουβαλώ και θα με συντροφεύουν στην υπόλοιπη ζωή μου.

Ένα μεγάλο ευχαριστώ θέλω να εκφράσω στον επιβλέποντα της παρούσας εργασίας, καθηγητή Βασίλειο Ασημακόπουλο, η πορεία και το έργο του οποίου, έχουν τροφοδοτήσει το ενδιαφέρον μου για τον τομέα της πρόβλεψης και αποτελούν οδηγό για τους νεότερους επιστήμονες. Θερμές ευχαριστίες οφείλω και στον υπεύθυνο της παρούσας εργασίας, Ευάγγελο Σπηλιώτη, ερευνητικό συνεργάτη και συντονιστή της Μονάδας Προβλέψεων και Στρατηγικής. Η καθοδήγησή του για την εκπόνηση της παρούσας εργασίας, καθώς και οι συμβουλές του εντός κι εκτός των επιστημονικών ορίων, υπήρξαν πολύτιμες. Αποτέλεσε πηγή γνώσης και πρότυπο στη διαμόρφωση της ακαδημαϊκής μου προσωπικότητας.

Ευχαριστώ θερμά, τους καθηγητές Ιωάννη Ψαρρά και Δημήτριο Ασκούνη, τόσο για τη συμμετοχή τους στην επιτροπή εξέτασης, όσο κυρίως για τη γνώση και την εμπειρία που έχω αποκτήσει, ακούγοντας τους από έδρας και συμμετέχοντας στα μαθήματά τους. Στον απολογισμό που κάνω -κλείνοντας τον κύκλο των προπτυχιακών σπουδών μου-, όλοι οι προαναφερθέντες ανήκουν στις φυσιογνωμίες που έχουν επιδράσει σημαντικά στον τρόπο σκέψης μου κι έχω εξελιχθεί μέσα από την αλληλεπίδραση μαζί τους.

Τί αξία θα είχε άραγε ένα τέτοιο μοναδικό ταξίδι, αν δεν υπήρχαν συνοδοιπόροι; Οι σχέσεις που ανέπτυξα κατά τη διάρκεια των σπουδών μου στο ΕΜΠ, αποτελούν τη μεγαλύτερη επιβράβευση. Ευδοκία, Στέλιο, Νίκο σας ευχαριστώ για τις όμορφες στιγμές που μοιραστήχαμε, τα καλύτερα έρχονται!

Γιάννης Τζανακάκης
Αθήνα, Οκτώβριος 2021

Πίνακας Περιεχομένων

| | |
|--|-----------|
| Λίστα Σχημάτων | 14 |
| Λίστα Πινάκων | 16 |
| 1 Ευρεία Περίληψη | 19 |
| 1.1 Εισαγωγή | 19 |
| 1.2 Στόχος της Εργασίας | 21 |
| 1.3 Πρόβλεψη | 22 |
| 1.4 Ιεραρχική Πρόβλεψη | 24 |
| 1.5 Μελέτη Περίπτωσης | 25 |
| 1.6 Ευρήματα & Συμπεράσματα | 26 |
| 2 Μοντέλα Πρόβλεψης | 29 |
| 2.1 Απλές Μέθοδοι Πρόβλεψης | 29 |
| 2.1.1 Μέσος Όρος | 29 |
| 2.1.2 Αφελής Μέθοδος (<i>Naive</i>) | 30 |
| 2.1.3 Εποχιακή <i>Naive</i> | 30 |
| 2.1.4 Τάση | 30 |
| 2.2 Αποσύνθεση | 31 |
| 2.2.1 Κινητοί Μέσοι Όροι | 31 |
| 2.2.2 Κλασσική Μέθοδος Αποσύνθεσης | 31 |
| 2.2.3 Άλλες Μέθοδοι Αποσύνθεσης | 33 |
| 2.3 Παλινδρόμηση | 33 |
| 2.3.1 Απλή Γραμμική Παλινδρόμηση | 33 |
| 2.3.2 Πολλαπλή Γραμμική Παλινδρόμηση | 33 |
| 2.3.3 Εκτίμηση μέσω Ελαχίστων Τετραγώνων | 34 |
| 2.4 Εκθετική Εξομάλυνση | 34 |
| 2.4.1 Απλή Εκθετική Εξομάλυνση | 34 |
| 2.4.2 Εκθετική Εξομάλυνση με Τάση | 35 |
| 2.4.3 Εκθετική Εξομάλυνση με Εποχιακότητα | 36 |
| 2.4.4 Ταξινόμηση Μεθόδων Εκθετικής Εξομάλυνσης | 36 |
| 2.4.5 Η μέθοδος ETS | 37 |
| 2.4.6 Επιλογή Μοντέλου | 39 |
| 2.5 Η μέθοδος <i>Theta</i> | 40 |
| 2.5.1 Γραμμές <i>Theta</i> | 40 |
| 2.5.2 Κλασσική Μέθοδος <i>Theta</i> | 41 |
| 2.6 Μοντέλα ARIMA | 42 |
| 2.6.1 Μοντέλα AR | 42 |
| 2.6.2 Μοντέλα MA | 43 |
| 2.6.3 Μοντέλα ARIMA | 43 |
| 2.7 Μέθοδοι Πρόβλεψης Διακοπτόμενης Ζήτησης | 45 |

| | | |
|----------|---|-----------|
| 2.7.1 | Μέθοδος Croston | 45 |
| 2.7.2 | Μέθοδος SBA | 45 |
| 2.7.3 | Μέθοδος ADIDA | 45 |
| 3 | Δείκτες Ακριβείας | 49 |
| 3.1 | Σφάλματα Εξαρτώμενα από την Κλίμακα | 49 |
| 3.1.1 | Μέσο Σφάλμα | 50 |
| 3.1.2 | Μέσο Απόλυτο Σφάλμα | 50 |
| 3.1.3 | Μέσο Τετραγωνικό Σφάλμα | 50 |
| 3.1.4 | Ρίζα Μέσου Τετραγωνικού Σφάλματος | 51 |
| 3.2 | Ποσοστιαία Σφάλματα | 51 |
| 3.2.1 | Μέσο Απόλυτο Ποσοστιαίο Σφάλμα | 51 |
| 3.2.2 | Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα | 51 |
| 3.3 | Σχετικά Σφάλματα | 52 |
| 3.3.1 | Μέσο Απόλυτο Κανονικοποιημένο Σφάλμα | 52 |
| 3.3.2 | Ρίζα Μέσου Τετραγωνικού Κανονικοποιημένου Σφάλματος | 52 |
| 3.3.3 | Theil's U-Statistic | 53 |
| 4 | Ιεραρχική Πρόβλεψη | 55 |
| 4.1 | Ιεραρχικές & Ομαδοποιημένες Χρονοσειρές | 55 |
| 4.2 | Μέθοδοι Ιεραρχικής Πρόβλεψης | 56 |
| 4.2.1 | Bottom-Up Μέθοδος | 56 |
| 4.2.2 | Top-Down Μέθοδοι | 57 |
| 4.2.3 | Μέθοδοι Γραμμικού Συνδυασμού | 58 |
| 5 | Μελέτη Περίπτωσης | 61 |
| 5.1 | Στόχοι | 62 |
| 5.2 | Σύνολο Δεδομένων | 63 |
| 5.3 | Επεξηγηματική Ανάλυση Δεδομένων | 66 |
| 5.3.1 | Διάρκεια Παρατηρήσεων | 66 |
| 5.3.2 | Συνολικές Πωλήσεις | 67 |
| 5.3.3 | Τιμές των Παρατηρήσεων | 68 |
| 5.3.4 | Πωλήσεις ανά Πολιτεία | 69 |
| 5.3.5 | Πωλήσεις ανά Κατάστημα & Κατηγορία Προϊόντος | 70 |
| 5.3.6 | Κατανομές Αφειτηρίας & Λήξης Παρατηρήσεων | 71 |
| 5.4 | Προδιαγραφές & Πειραματική Διάταξη | 72 |
| 5.4.1 | Επιλογή Ιεραρχιών & Εξεταζόμενα Επίπεδα Συνάθροισης | 73 |
| 5.4.2 | Προδιαγραφές Προβλέψεων | 74 |
| 5.4.3 | Προδιαγραφές Σύγκρισης Απόδοσης | 76 |
| 5.5 | Πειραματική Διαδικασία | 78 |
| 5.5.1 | Προεπεξεργασία Δεδομένων | 78 |
| 5.5.2 | Δειγματοληψία Χρονοσειρών | 80 |
| 5.5.3 | Κατασκευή Ιεραρχιών & Μέθοδοι Ιεραρχικής Πρόβλεψης | 81 |

| | | |
|----------|---|------------|
| 5.5.4 | Εξαγωγή Μετρήσεων Απόδοσης | 83 |
| 5.6 | Ανάπτυξη σε Κώδικα | 83 |
| 6 | Αποτελέσματα | 87 |
| 6.1 | Μέση Απόδοση σε όλα τα Επίπεδα Συνάθροισης | 88 |
| 6.1.1 | Κατάταξη ως προς την Ακρίβεια Πρόβλεψης | 89 |
| 6.1.2 | Κατάταξη ως προς το Κόστος Πρόβλεψης | 92 |
| 6.2 | Μέση Απόδοση σε Ομάδες Επιπέδων Συνάθροισης | 96 |
| 6.2.1 | Επίπεδα Υψηλής Συνάθροισης | 96 |
| 6.2.2 | Επίπεδα Μέσης Συνάθροισης | 98 |
| 6.2.3 | Επίπεδα Χαμηλής Συνάθροισης | 100 |
| 6.3 | Αλληλοσυσχέτιση Μέτρων Κόστους Πρόβλεψης | 101 |
| 6.4 | Συστάδες Κόστους Πρόβλεψης | 102 |
| 6.5 | Κατατάξεις Ακρίβειας Επιπέδων Συνάθροισης | 104 |
| 7 | Συμπεράσματα | 109 |
| 7.1 | Ευρήματα | 109 |
| 7.2 | Μελλοντικές Προεκτάσεις | 110 |
| | Βιβλιογραφία | 113 |
| A | Διαγωνισμοί M | 123 |
| A.1 | Διαγωνισμός M1 | 123 |
| A.2 | Διαγωνισμός M2 | 124 |
| A.3 | Διαγωνισμός M3 | 124 |
| A.4 | Διαγωνισμός M4 | 125 |
| A.5 | Διαγωνισμός M5 | 126 |
| B | Λοιπά Διαγράμματα Μελέτης Περίπτωσης | 129 |

Λίστα Σχημάτων

| | | |
|----|---|----|
| 1 | Ιεραρχική Χρονοσειρά 2 Επιπέδων | 55 |
| 2 | Ομαδοποιημένη Χρονοσειρά Συνόλου Δεδομένων <i>M5</i> | 65 |
| 3 | Διάρκεια & Περίοδοι <i>M5 Dataset</i> | 66 |
| 4 | Συνολικές Πωλήσεις (πορτοκαλί) & Smoothing Fit (μπλε) <i>M5 Dataset</i> | 67 |
| 5 | Heatmap Εβδομαδιαίας & Μηνιαίας Εποχιακότητας Συνολικών Πωλήσεων (Επίπεδο 1) <i>M5 Dataset</i> | 67 |
| 6 | Πυκνότητες Πιθανοτήτων Τιμών Παρατηρήσεων - 42840 σειρές <i>M5 Dataset</i> | 68 |
| 7 | Μηνιαίες Πωλήσεις Μονάδων ανά Πολιτεία & Smoothing Fit <i>M5 Dataset</i> | 69 |
| 8 | Εβδομαδιαία & Μηνιαία Εποχιακότητα Πωλήσεων ανά Πολιτεία <i>M5 Dataset</i> | 69 |
| 9 | Πωλήσεις ανά Κατάστημα & Κατηγορία <i>M5 Dataset</i> | 70 |
| 10 | Ποσοστά Νέων Καταχωρήσεων ανά Μήνα & Έτος <i>M5 Dataset</i> | 71 |
| 11 | Heatmaps Νέων Καταχωρήσεων ανά Μήνα & Έτος <i>M5 Dataset</i> | 72 |
| 12 | Μέση Απόδοση Benchmarks <i>M5 Dataset</i> | 75 |
| 13 | Παράδειγμα Διακοπτόμενης Ζήτησης <i>id μονάδας: Hobbies/2/010/Wi/1</i> | 79 |
| 14 | Πρώτες 10 Γραμμές & Στήλες <i>sales_train_validation.csv</i> | 82 |
| 15 | 5 Πρώτες Σειρές του Εξαγόμενου Αρχείου Μετρήσεων | 83 |
| 16 | Κατάταξη των 50 Επιλογών Πρόβλεψης ως προς τη Μέση Ακρίβεια Δείγμα <i>1531 Μη Συναθροισμένων Χρονοσειρών - 25 Προϊόντα Ανά Υποκατηγορία</i> | 89 |
| 17 | Ανάλυση της Κατάταξης του Σχ. 16 στις Ακρίβειες Πρόβλεψης των Ιεραρχιών πάνω στα 12 Επίπεδα Συνάθροισης Δείγμα <i>1531 Μη Συναθροισμένων Χρονοσειρών - Boxplots χωρίς Outliers</i> | 90 |
| 18 | Κατάταξη των 48 Ιεραρχιών ως προς το Πλήθος των Περιεχόμενων Χρονοσειρών Δείγμα <i>1531 Μη Συναθροισμένων Χρονοσειρών Χρωματική Κλίμακα για τη Μέση Ακρίβεια</i> | 92 |
| 19 | Συσχέτιση Μέσης Ακρίβειας Πρόβλεψης με Πλήθος Περιεχόμενων Χρονοσειρών των 48 Ιεραρχιών και των 2 Benchmarks + Ευθεία Παλινδρόμησης Δείγμα <i>1531 Μη Συναθροισμένων Χρονοσειρών - Χρώμα για Χρόνο Μοντελοποίησης</i> | 93 |
| 20 | Συσχέτιση Μέσης Ακρίβειας & Κόστους σε Μνήμη - Ανάλυση στα 12 Επίπεδα Συνάθροισης Δείγμα <i>1531 Μη Συναθροισμένων Χρονοσειρών - Χρωματική Κλίμακα για το Πλήθος των Περιεχόμενων Χρονοσειρών κάθε Ιεραρχίας</i> | 94 |
| 21 | Κατάταξη των 48 Ιεραρχιών και των Benchmarks ως προς τη Μέση Ακρίβεια των Προβλέψεων στα 5 Υψηλότερα Επίπεδα Συνάθροισης Δείγμα <i>1531 Μη Συναθροισμένων Χρονοσειρών - Χρωματική Κλίμακα για το Πλήθος των Περιεχόμενων Χρονοσειρών σε κάθε Ιεραρχία</i> | 95 |
| 22 | Κατάταξη των 48 Ιεραρχιών και των Benchmarks ως προς τη Μέση Ακρίβεια των Προβλέψεων στα 4 Μεσαία Επίπεδα Συνάθροισης | 97 |

| | | |
|----|---|-----|
| 23 | Κατάταξη των 48 Ιεραρχιών και των Benchmarks ως προς τη Μέση Ακρίβεια των Προβλέψεων στα 3 Χαμηλότερα Επίπεδα Συνάθροισης <i>Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών - Χρωματική Κλίμακα για το Πλήθος των Περιεχόμενων Χρονοσειρών σε κάθε Ιεραρχία</i> | 99 |
| 24 | Αλληλοσυσχέτιση των Μέτρων Κόστους Πρόβλεψης στη Μελέτη Περίπτωσης | 101 |
| 25 | Συσταδοποίηση Ιεραρχιών ανά Κόστος Πρόβλεψης κ-Κοντινότεροι Γείτονες Χρωματική Κλίμακα ανά Συστάδα και Αποτύπωση Κέντρων Κόστους <i>Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών</i> | 102 |
| 26 | Μέση Απόδοση ανά Cluster Κόστους Πρόβλεψης των Ιεραρχιών | 103 |
| 27 | Μέση Κατάταξη Εμφάνισης Επιπέδων στις 50 Εξεταζόμενες Μεθόδους | 104 |
| 28 | Μέση Κατάταξη Επιπέδων στις 10 Ακριβέστερες Εμφανίσεις | 105 |
| 29 | Κατάταξη ως προς τη Μέση Ακρίβεια των 50 Επιλογών Πρόβλεψης <i>Δείγμα 2672 Μη Συναθροισμένων Χρονοσειρών - 25 προϊόντα ανά υποκατηγορία</i> | 129 |
| 30 | Κατάταξη και Boxplots Δείγματος στο οποίο τα Benchmarks προέβλεψαν με Υψηλότερη Μεσοσταθμική Ακρίβεια από τις Εξεταζόμενες Ιεραρχίες <i>Δείγμα 2976 Μη Συναθροισμένων Χρονοσειρών - 50 προϊόντα ανά υποκατηγορία</i> | 130 |

Λίστα Πινάκων

| | | |
|----|---|-----|
| 1 | Μοντέλα Εκθετικής Εξομάλυνσης | 37 |
| 2 | Μοντέλα με Αθροιστικό Σφάλμα ε_t | 38 |
| 3 | Μοντέλα με Πολλαπλασιαστικό Σφάλμα ε_t | 39 |
| 4 | Ειδικές Περιπτώσεις Μοντέλων <i>ARIMA</i> | 43 |
| 5 | Δεδομένα & Προβλέψεις Χρονοσειράς | 49 |
| 6 | Ταυτότητα Συνόλου Δεδομένων Διαγωνισμού <i>M5</i> | 64 |
| 7 | Περιγραφή Επιπέδων <i>M5</i> Dataset | 64 |
| 8 | Λίστα Ιεραρχιών Συνόλου Δεδομένων <i>M5</i> | 73 |
| 9 | Εξεταζόμενα Επίπεδα | 74 |
| 10 | Απόδοση Επιλεγμένων Μεθόδων σε όλα τα Επίπεδα Συνάθροισης | 91 |
| 11 | Απόδοση Επιλεγμένων Μεθόδων στα Επίπεδα Συνάθροισης 1-5 | 96 |
| 12 | Απόδοση Επιλεγμένων Μεθόδων στα Μεσαία Επίπεδα 6-9 | 98 |
| 13 | Απόδοση Επιλεγμένων Μεθόδων στα Χαμηλά Συνάθροισης 10-12 | 100 |
| 14 | Μέση Κατάταξη Ακριβείας & Συχνότητες Εμφάνισης Επιπέδων Συνάθροισης στις Εξεταζόμενες Ιεραρχίες | 105 |

Λίστα Αλγορίθμων

| | | |
|---|--|----|
| 1 | Σχεδιάγραμμα Αλγορίθμου Επιλογής Βέλτιστης Ιεραρχίας | 84 |
|---|--|----|

Ατυχείς Προβλέψεις #2

«Στο μέλλον θα είναι απλό να μετατρέψει κανείς,
ένα φορτηγό σιδερένιες ράβδους σε καθαρό χρυσό.»

Thomas Edison
Αμερικανός πρωτοπόρος εφευρέτης
και επιχειρηματίας, 1911

1 Ευρεία Περίληψη

1.1 Εισαγωγή

Η πρόβλεψη του μέλλοντος αποτελεί κρίσιμη είσοδο στη λήψη απόφασης και τη χάραξη πολιτικών, ιδιαίτερα σε περιβάλλοντα με αβεβαιότητα. Κατασκευάζοντας μοντέλα από παρατηρήσεις φυσικών, οικονομικών, επιχειρησιακών και άλλων μεγεθών, μπορούμε να προβλέψουμε τη μελλοντική τους εξέλιξη και να υποστηρίξουμε την απόφαση σε διάφορες πτυχές της ανθρώπινης δραστηριότητας. Σε πολλές εφαρμογές, οι παρατηρήσεις οργανώνονται σε επίπεδα συνάθροισης, τα οποία ιεραρχούνται πλήρως ή μερικώς. Οι πωλήσεις μιας επιχείρησης, για παράδειγμα, μπορούν να αποσυναθροιστούν σε χρονοσειρές ανά χώρα, περιφέρεια, πόλη και κατάσταση. Η συλλογή αυτών των χρονοσειρών υπακούει σε μια δομή συνάθροισης που ορίζει μια φυσική ιεραρχία, για αυτό και τέτοιες συλλογές καλούνται *ιεραρχικές χρονοσειρές*. Η επιχείρηση μπορεί να εμπορευείται μεγάλο αριθμό προϊόντων που οργανώνονται σε κατηγορίες και υποκατηγορίες, ορίζοντας ακόμη μια φυσική ιεραρχία κατηγοριών. Πλέον οι παρατηρήσεις οργανώνονται σε περισσότερες ιεραρχίες. Τέτοιες συλλογές καλούνται *ομαδοποιημένες χρονοσειρές*.

Ο αποφασίζων κατά κανόνα επιθυμεί προβλέψεις, εναρμονισμένες με την ιεραρχική δομή των παρατηρήσεων. Λόγω της μη γραμμικής φύσης των μοντέλων, η πρόβλεψη μιας συναθροισμένης χρονοσειράς δεν είναι ίση με το άθροισμα των προβλέψεων των μη συναθροισμένων χρονοσειρών που την διαμερίζουν [2011]. Συνεπώς, η ανεξάρτητη πρόβλεψη των επιπέδων συνάθροισης δε θα υπακούει -κατά κανόνα- στους περιορισμούς που επιβάλλει η ιεραρχική δομή. Η ανάγκη για λήψη απόφασης που είναι συνεπής στα διάφορα επίπεδα της ιεραρχίας, οδήγησε στην ανάπτυξη της *ιεραρχικής πρόβλεψης*, μιας συλλογής τεχνικών με τις οποίες παράγονται προβλέψεις που υπακούν στην έμφυτη ιεραρχία των παρατηρήσεων. Η ιεραρχική πρόβλεψη αποτέλεσε για δεκαετίες χρήσιμο εργαλείο για πρόβλεψη ιεραρχικών ή ομαδοποιημένων χρονοσειρών, αλλά ιδιαίτερα την τελευταία δεκαετία, παρατηρείται αυξανόμενη έρευνα για το πεδίο, σε ένα ευρύ φάσμα εφαρμογών [2009, 2017, 2018, 2018, 2019, 2019, 2019, 2019, 2020, 2021, 2021].

Κλασσικές προσεγγίσεις ιεραρχικής πρόβλεψης περιλαμβάνουν bottom-up, top-down ή middle-out πρόβλεψη. Η bottom-up μέθοδος, προβλέπει όλες τις μη συναθροισμένες χρονοσειρές -το χαμηλότερο επίπεδο της ιεραρχίας- και στη συνέχεια τις προσθέτει με κατάλληλο τρόπο, ώστε να παράξει προβλέψεις για τα υψηλότερα στην ιεραρχία επίπεδα. Η top-down μέθοδος προβλέπει την πιο συναθροισμένη χρονοσειρά, δηλαδή το υψηλότερο επίπεδο της ιεραρχίας και στη συνέχεια διαμερίζει τις προβλέψεις βάσει ποσοστών που προκύπτουν με διάφορες παραλλαγές [2009]. Μια τρίτη προσέγγιση, γνωστή και ως middle-out μέθοδος, μοντελοποιεί ένα μεσαίο επίπεδο συνάθροισης και στη συνέχεια παράγει τις ευθυγραμμισμένες προβλέψεις αθροίζοντας και διαμερίζοντας για τις προβλέψεις των υψηλότερων και χαμηλότερων επιπέδων στην ιεραρχία, αντίστοιχα.

Τέτοιες προσεγγίσεις εστιάζουν στην πληροφορία ενός επιπέδου συνάθροισης για να προβλέψουν όλη την ιεραρχία. Όσο οι ιεραρχίες γίνονται πιο σύνθετες, τόσο περισσότερη είναι η πληροφορία που αγνοούν οι διάφορες προσεγγίσεις που μοντελοποιούν ένα μοναδικό επίπεδο συνάθροισης, για αυτό τα τελευταία χρόνια έχουν προταθεί μέθοδοι

που συνδυάζουν γραμμικά, μοντέλα από περισσότερα επίπεδα συνάθροισης, για να παράξουν ευθυγραμμισμένες προβλέψεις, τόσο μέσω στατιστικών μεθόδων όσο και μεθόδων μηχανικής μάθησης [2011, 2019, 2021, 2021, 2017, 2019, 2021, 2021]. Ο αλγόριθμος που προτάθηκε από τους Hyndman et al. [2011] ως *Optimal Combination*, γενικεύτηκε και βελτιώθηκε υπολογιστικά από τους Hyndman et al. [2014] και επαναδιατυπώθηκε από τους Wickramasuriya et al. [2019] ως *Minimum Trace* εναρμόνιση, υποθέτει μια συλλογή αμερόληπτων προβλέψεων για τις χρονοσειρές σε κάθε επίπεδο της ιεραρχίας και χρησιμοποιεί ένα μοντέλο παλινδρόμησης για να εναρμονίσει τις προβλέψεις με την ιεραρχική δομή. Οι μέθοδοι που συνδυάζουν γραμμικά τις προβλέψεις όλων των επιπέδων συνάθροισης, άλλαξαν τη μέχρι πρότινος προσέγγιση της ιεραρχικής πρόβλεψης. Αγνοώντας αρχικά τη φυσική ιεραρχία των δεδομένων, μοντελοποιούνται ξεχωριστά όλες οι χρονοσειρές και στη συνέχεια μειώνεται από κοινού στα μοντέλα, το σφάλμα που οφείλεται στη μη εναρμόνιση με τη φυσική ιεραρχία των παρατηρήσεων, έτσι αξιοποιείται πληροφορία που αγνοούσαν μέθοδοι όπως η bottom up, η top down ή η middle out.

Οι μέθοδοι ιεραρχικής πρόβλεψης γεννήθηκαν και εξελίχθηκαν πάνω στην ανάγκη για λήψη απόφασης, που δεν παρουσιάζει ασυνέπειες μεταξύ των επιπέδων συνάθροισης στην ιεραρχία· ασυνέπειες που οδηγούν σε αποκλίνουσες αποφάσεις και μεταφράζονται σε κόστος. Πέρα όμως από το βασικό στόχο που είναι η εναρμόνιση των προβλέψεων με την εκάστοτε ιεραρχία, σε πολλές εφαρμογές ιεραρχικές μέθοδοι φαίνεται να βελτιώνουν τις προβλέψεις των επιπέδων συνάθροισης. Παραδείγματα τέτοιων εφαρμογών περιλαμβάνουν ατυχήματα και επείγουσες εισαγωγές Athanasopoulos et al. [2017], δείκτες θνητότητας Shang and Hyndman [2017], πληθυσμό φυλακισμένων George Athanasopoulos and Weatherburn [2019], προβλέψεις λιανικής Villegas García and Pedregal [2018], ηλιακή ενέργεια Yang et al. [2018] και Yagli et al. [2019], τουριστική ζήτηση Athanasopoulos et al. [2009], Hyndman et al. [2011], Wickramasuriya et al. [2019] και παραγωγή αιολικής ενέργειας Pan et al. [2019]. Η εναρμόνιση των προβλέψεων με την εκάστοτε ιεραρχική δομή που τις συνδέει, μπορεί υπό συνθήκες να μεταφέρει πληροφορία μεταξύ των επιπέδων συνάθροισης, που δεν αποτυπώνεται στο κάθε ανεξάρτητο μοντέλο και συνεπώς να επιδρά θετικά στις προβλέψεις τόσο ολόκληρης της ιεραρχίας, όσο και υποσυνόλων της, όπως ένα top-level ΑΕΠ [2020] ή πωλήσεις σε μια γεωγραφική περιφέρεια [2019].

Όταν ζητούμενο είναι η πρόβλεψη μιας ή περισσότερων χρονοσειρών που ανήκουν σε μια συλλογή ομαδοποιημένων χρονοσειρών και επιλέγουμε να συνδυάσουμε γραμμικά κι άλλα επίπεδα συνάθροισης της συλλογής, για να ενισχύσουμε την πρόβλεψή μας, δεν είναι ξεκάθαρο ποια επίπεδα θα συνεισφέρουν θετικά στις προβλέψεις, ενώ είναι σίγουρο το κόστος για τη μοντελοποίηση των χρονοσειρών που θα εισάγουμε στην ιεραρχία. Αν μας ζητηθεί να προβλέψουμε, για παράδειγμα, τις πωλήσεις μιας συγκεκριμένης κατηγορίας προϊόντων στα καταστήματα μιας πόλης, για μια συλλογή με ιεραρχίες προϊόντων και γεωγραφικών τμημάτων, η αποσυνάθροιση σε ποια επίπεδα και με ποιά σειρά των δεδομένων που έχουμε στη διάθεσή μας θα εξυπηρετούσε καλύτερα τη ζητούμενη πρόβλεψη;

Το πλαίσιο στο οποίο αναπτύχθηκε η μελέτη των ιεραρχικών χρονοσειρών, αφορούσε -κατά κανόνα- δεδομένα που είχαν οργανωθεί *ad hoc* από τον αποφασίζοντα (επιχείρηση, οργανισμό, φορέα κλπ) σε επίπεδα συνάθροισης, σύμφωνα με κριτήρια που είχαν εκ των προτέρων τεθεί, σε σχέση με το ποια επίπεδα συνάθροισης θα περιλαμβάνονται στα δε-

δομένα και ποια όχι. Οι προβλέψεις επομένως, εναρμονίζονται με τα διαθέσιμα επίπεδα συνάθροισης, χωρίς να έχει προηγηθεί ανάλυση πάνω στην προβλεπτική τους ικανότητα και ταυτόχρονα το κόστος που υπαγορεύει η συμπερίληψη και θέση τους στην ιεραρχία.

Για μεθόδους όπως οι bottom up, top down και middle out, αυτό έχει ενδεχομένως μικρή σημασία, αφού οι χρονοσειρές που μοντελοποιούνται, βρίσκονται αποκλειστικά σε ένα επίπεδο συνάθροισης. Οι μεταβολές στην οργάνωση των δεδομένων σε επίπεδα συνάθροισης αφήνουν πρακτικά ανεπηρέαστες τόσο τη διαμέριση των προβλέψεων στα χαμηλότερα επίπεδα της top down πρόβλεψης, όσο και τη συνάθρυσή τους στα υψηλότερα επίπεδα κατά την bottom up. Ωστόσο, για τις μεθόδους γραμμικού συνδυασμού των μοντέλων ολόκληρης της ιεραρχίας, τα επίπεδα συνάθροισης που θα συμπεριληφθούν στη μοντελοποίηση, καθώς και ο τρόπος με τον οποίο θα διαταχθούν, διαδραματίζουν ενεργό ρόλο τόσο στην ακρίβεια όσο και στο κόστος των εξαγόμενων προβλέψεων.

1.2 Στόχος της Εργασίας

Γίνεται επομένως κατανοητό, πως η εναρμόνιση των προβλέψεων με κατάλληλες ιεραρχίες -μέσω μεθόδων γραμμικού συνδυασμού όλων των επιπέδων τους- μπορεί να επιδρά θετικά στην ακρίβεια της πρόβλεψης και κατά συνέπεια να αποτελεί η ίδια παράμετρο απόδοσης των προβλέψεων. Ωστόσο, η προβλεπτική ικανότητα που μπορεί να εμφανίζει μια «βέλτιστη» οργάνωση σε επίπεδα συνάθροισης, περιορίζεται από τη θεώρηση της ιεραρχίας ως μιας εξωγενούς ή αμετάβλητης οντότητας που προϋπάρχει της ανάλυσης και ικανοποιεί κριτήρια που δεν αφορούν πάντα τη βελτιστοποίηση των προβλέψεων. Εντοπίζεται επομένως ένα κενό στην έρευνα της ιεραρχικής πρόβλεψης, που αφορά στην μελέτη της επίδρασης που έχουν διαφορετικές οργανώσεις των δεδομένων στην ακρίβεια και το κόστος των προβλέψεων κάθε επιπέδου συνάθροισης, σε ένα πλαίσιο ελευθερίας του αναλυτή να επιλέξει ανάμεσα σε διαφορετικές οργανώσεις των επιπέδων συνάθροισης.

Στόχο της παρούσας εργασίας θα αποτελέσει η εύρεση της «βέλτιστης» οργάνωσης των δεδομένων για ενίσχυση της πρόβλεψης διαφορετικών επιπέδων συνάθροισης μιας συλλογής ομαδοποιημένων χρονοσειρών. Με τον όρο «βέλτιστη» θα εννοείται ένας συνδυασμός κριτηρίων τόσο ακρίβειας πρόβλεψης για ένα ή περισσότερα επίπεδα συνάθροισης, όσο και κόστους για τη μοντελοποίηση των περιεχόμενων χρονοσειρών σε κάθε εξεταζόμενη ιεραρχία. Πιο συγκεκριμένα, θα αναπτυχθεί μια μεθοδολογία, μέσω της οποίας θα συγκρίνονται με αυτόματο τρόπο ως προς την ακρίβεια και το κόστος πρόβλεψης, προβλέψεις που έχουν εναρμονιστεί με διαφορετικές ιεραρχικές δομές πάνω σε δείγματα του συνόλου δεδομένων, με σκοπό την εύρεση βέλτιστης σε απόδοση ιεραρχίας προς εναρμόνιση των προβλέψεων μιας ή περισσότερων χρονοσειρών. Η παραπάνω μεθοδολογία θα εφαρμοστεί και αξιολογηθεί σε μια μελέτη περίπτωσης, η οποία θα αφορά το σύνολο δεδομένων που τέθηκε προς πρόβλεψη στο διαγωνισμό M5 και περιλαμβάνει 42840 χρονοσειρές πωλήσεων σε 12 επίπεδα γεωγραφικής και ανά προϊόν συνάθροισης.

Η παρούσα εργασία, έχει ως ευρύτερο σκοπό την ανάδειξη της ιεράρχησης δεδομένων σε επίπεδα συνάθροισης, ως ενεργητικό στοιχείο της ανάλυσης και πρόβλεψής τους. Μετατοπίζεται δηλαδή το κέντρο ενδιαφέροντος, από την αναζήτηση βέλτιστων προβλέψεων με περιορισμό την εναρμόνιση με όλα τα *ad hoc* παρεχόμενα επίπεδα συνάθροισης,

στην αναζήτηση της βέλτιστης οργάνωσης σε επίπεδα, με στόχο την πρόβλεψη με την καλύτερη απόδοση (ακρίβεια και κόστος πρόβλεψης) - τόσο συνολικά όσο και σε επιμέρους επίπεδα συνάνθρωισης. Η ενίσχυση της πρόβλεψης μέσα από εναρμόνιση με μια κατάλληλη ιεραρχία μπορεί να αποτελέσει χρήσιμο εργαλείο για τον αναλυτή, σε βαθμό που κανείς θα αναζητά τα κατάλληλα δεδομένα για να εναρμονίσει τις προβλέψεις, ξεπερνώντας τη μέχρι τώρα πρακτική να παρέχονται από τον αποφασίζοντα, στο προαναφερθέν πλαίσιο. Η οργάνωση των δεδομένων που θα οδηγεί σε βελτιστοποιημένες προβλέψεις μιας χρονοσειράς θα μπορούσε να αποτελεί ένα ακόμη υπολογίσιμο χαρακτηριστικό της εκάστοτε χρονοσειράς, όπως η τάση ή η εποχιακότητα στις παρατηρήσεις της.

1.3 Πρόβλεψη

Ανάγκη για πρόβλεψη προκύπτει σε πολλές καταστάσεις. Το αν και που θα εγκατασταθούν μονάδες παραγωγής ενέργειας τα επόμενα 10 χρόνια, προαπαιτεί πρόβλεψη της μελλοντικής ζήτησης ενέργειας. Η ομαλή ροή στην εφοδιαστική αλυσίδα μιας επιχείρησης, απαιτεί προβλέψεις των βραχυπρόθεσμων και μακροπρόθεσμων απαιτήσεων που πρέπει να ικανοποιηθούν. Ο προγραμματισμός των βαρδιών της επόμενης εβδομάδας σε μια μεταφορική εταιρεία, απαιτεί προβλέψεις των δρομολογίων που θα χρειαστεί να γίνουν σε αυτό το διάστημα. Πρόβλεψη του μέλλοντος μπορεί να χαρακτηριστεί ένα ευρύ σύνολο ενεργειών που σχετίζονται με τις ανθρώπινες δραστηριότητες - από την πρόβλεψη του παγκόσμιου πληθυσμού μέχρι την εκτίμηση του πότε θα αδειάσει το ψυγείο μας.

Φυσικά, οι τρόποι με τους οποίους εξάγουμε προβλέψεις ποικίλλουν ανάλογα με το περιεχόμενο και τη σημασία της εκάστοτε εφαρμογής. Όταν διαθέτουμε μια χρονοσειρά¹ για την ενδιαφερόμενη ποσότητα και μπορούμε να συμπεράνουμε πως η εξάρτηση των παρατηρήσεων είναι -και θα συνεχίσει να είναι- σε κάποιο βαθμό αιτιοκρατική, μπορούμε να εφαρμόσουμε ποσοτικές μεθόδους πρόβλεψης. Οι μέθοδοι αυτές, αποτελούν μαθηματικές εξισώσεις που τροφοδοτούνται με τα δεδομένα και παράγουν μοντέλα τους. Ένα κατάλληλα επιλεγμένο μοντέλο -που έχει προκύψει από μια συστηματική διαδικασία διαλογής- είναι ικανό να αποτυπώσει σημαντικό τμήμα της αιτιοκρατίας που συνδέει τις παρατηρήσεις. Το εκάστοτε μοντέλο, προεκτεινόμενο, μπορεί να δώσει προβλέψεις για το μελλοντικό διάστημα που επιθυμούμε (ορίζοντας πρόβλεψης). Υπάρχει ένα ευρύ φάσμα ποσοτικών μεθόδων πρόβλεψης, οι οποίες ανταποκρίνονται σε δεδομένα με διαφορετικά χαρακτηριστικά ή εφαρμογές διαφορετικού ενδιαφέροντος. Περισσότερα για τα μοντέλα πρόβλεψης, τις ιδιότητες και τα πεδία εφαρμογής τους παρουσιάζονται στο κεφάλαιο 2.

Όπως είδαμε, οι μέθοδοι ιεραρχικής πρόβλεψης που συνδυάζουν γραμμικά τα μοντέλα ολόκληρης της συλλογής ομαδοποιημένων χρονοσειρών, όπως η *Minimum Trace*, προαπαιτούν τη μοντελοποίηση όλων των περιεχόμενων χρονοσειρών σε μια ιεραρχία. Η *Minimum Trace* παρέχει ελευθερία επιλογής μεθόδου ή μεθόδων για τη μοντελοποίηση και την πρόβλεψη βάσης ή ανεξάρτητη πρόβλεψη² των χρονοσειρών που περιέχονται στην

¹Χρονοσειρά καλείται ένα σύνολο χρονικά διαδοχικών και ισαπεχουσών τιμών ενός μεγέθους. Σε κάποιες εφαρμογές, τα δεδομένα χρειάζεται να μετασχηματιστούν για να καταλήξουν σε μορφή χρονοσειράς, που αποτελεί και την αναγκαία μορφή εισόδου για τις περισσότερες μεθόδους πρόβλεψης.

²Προβλέψεις βάσης ή ανεξάρτητες προβλέψεις (*base/independent forecasts*) μιας χρονοσειράς καλούν-

εξεταζόμενη ιεραρχία. Η μεθοδολογία που θα αναπτυχθεί στο πλαίσιο της παρούσας εργασίας, έχει ως στόχο τη σύγκριση προβλέψεων που έχουν εναρμονιστεί με διαφορετικές ιεραρχικές δομές. Την επίτευξη του παραπάνω στόχου κρίθηκε πως θα ενισχύσει η θεώρηση των ανεξάρτητων προβλέψεων κάθε χρονοσειράς ως σταθερά της πειραματικής διάταξης. Επομένως, θα χρειαστεί να γίνει η επιλογή μιας κατάλληλης μεθόδου για τη μοντελοποίηση και παραγωγή των ανεξάρτητων προβλέψεων των χρονοσειρών ολόκληρης της εκάστοτε εξεταζόμενης συλλογής. Οι χρονοσειρές πωλήσεων που θα εξετάσουμε, παρουσιάζουν ορισμένα κοινά χαρακτηριστικά: ισχυρή εποχιακότητα πολλαπλών περιόδων (εβδομαδιαία, μηνιαία, ετήσια), ύπαρξη τάσης στην πλειοψηφία των χρονοσειρών, καθώς και υψηλή διακοπτότητα (*intermittency*) στα επίπεδα χαμηλής συνάθροισης, δηλαδή διαστήματα μηδενικών πωλήσεων ανάμεσα στις μη μηδενικές. Για τα παραπάνω χαρακτηριστικά μια παραδοσιακά καλή επιλογή μεθόδου για την παραγωγή ανεξάρτητων προβλέψεων, είναι η εκθετική εξομάλυνση (*exponential smoothing*).

Η εκθετική εξομάλυνση αποτελεί μια οικογένεια μεθόδων που αναπτύχθηκε τη δεκαετία του '50 και έκτοτε αποτελεί μια αρκετά δημοφιλή -ίσως τη δημοφιλέστερη- μέθοδο στον τομέα των επιχειρησιακών προβλέψεων και ιδιαίτερα στις προβλέψεις πωλήσεων, λόγω του συνδυασμού αξιόπιστων και γρήγορων προβλέψεων. Βασική ιδέα στις μεθόδους εκθετικής εξομάλυνσης είναι πως το μέλλον εξαρτάται από το παρελθόν με μια εκθετική πτώση όσο πιο πίσω πηγαίνουμε στο χρόνο. Η πρόβλεψη με εκθετική εξομάλυνση, προκύπτει μέσω σταθμισμένων μέσων όρων από παρελθοντικές παρατηρήσεις, με βάρη που φθίνουν εκθετικά όσο κινούμαστε προς το παρελθόν. Η οικογένεια των μεθόδων εκθετικής εξομάλυνσης εμπλουτίστηκε σταδιακά με παραλλαγές που μπορούν να προβλέπουν αποτελεσματικά χρονοσειρές με διαφορετικά ποιοτικά χαρακτηριστικά (γραμμική ή μη γραμμική τάση και εποχιακότητα). Επιπλέον, η επιλογή της κατάλληλης εκ των παραπάνω παραλλαγών γίνεται με αυτόματο τρόπο, που αποτελεί σημαντικό πλεονέκτημα για τη μοντελοποίηση συλλογών χρονοσειρών με διαφορετικά χαρακτηριστικά. Περισσότερα για τις μεθόδους εκθετικής εξομάλυνσης, την ταξινόμησή τους και τη μέθοδο ETS που θα αξιοποιηθεί στη μελέτη περίπτωσης αναφέρονται στην παράγραφο 2.4.

Η επιλογή της εκθετικής εξομάλυνσης δεν έγινε μόνο γιατί αποτελεί μια παραδοσιακά αξιόπιστη επιλογή στις προβλέψεις πωλήσεων. Οι Makridakis, Spiliotis, and Assimakopoulos [2021], κατά την παρουσίαση των ευρημάτων του διαγωνισμού *M5*, έδειξαν ότι η bottom up συνάθροιση προβλέψεων που έχουν εξαχθεί μέσω εκθετικής εξομάλυνσης, δηλαδή μια φαινομενικά απλή και οικονομική επιλογή, προέβλεψε με μεγαλύτερη ακρίβεια από όλα τα benchmarks με τα οποία συγκρίθηκε. Από αυτή την ανάλυση, προέκυψε ότι η εκθετική εξομάλυνση με bottom up, παράγαγε ακριβέστερες προβλέψεις από τις τρεις μεθόδους διακοπτόμενης ζήτησης (*Croston*, *SBA*, *ADIDA*), γεγονός που ενίσχυσε την υπεροχή της, ανάμεσα σε μεθόδους που θα μπορούσαν να αποτελέσουν πιθανές επιλογές για τη μοντελοποίηση των χρονοσειρών στη μελέτη περίπτωσης της παρούσας εργασίας, λόγω της υψηλής διακοπτότητας που παρουσιάζουν.

ται οι προβλέψεις που προκύπτουν από κάποια μέθοδο που τροφοδοτείται αποκλειστικά από τις παρατηρήσεις της εκάστοτε χρονοσειράς. Ο όρος θα χρησιμοποιηθεί σε αντιπαράβολή με τις εναρμονισμένες ή συμφιλιωμένες προβλέψεις (*reconciled/coherent forecasts*) που θα εξάγονται μέσα από κάποια μέθοδο ιεραρχικής πρόβλεψης.

Πέρα από την κορυφαία απόδοση της εκθετικής εξομάλυνσης με bottom up συνάνθροιση, σε σχέση με τις υπόλοιπες μεθόδους που αποτέλεσαν τα benchmarks του διαγωνισμού, η μέθοδος πήγε πολύ καλά και συγκρινόμενη με τις υποβολές των διαγωνιζόμενων ομάδων. Πιο συγκεκριμένα, μόνο το 7,5% των υποβληθέντων προβλέψεων ή αλλιώς οι προβλέψεις 415 από τις 5507 ομάδες υπεραπέδωσαν σε ακρίβεια της ES_{bu} 2021. Στην παραπάνω κατάταξη πρέπει να συνυπολογιστεί ότι στις πρώτες 500 ομάδες, η πλειοψηφία χρησιμοποίησε αρκετά πιο σύνθετες μεθόδους, εκπαίδευσης νευρωνικών δικτύων και μηχανικής μάθησης. Αναλυτικότερα η παρουσία της εκθετικής εξομάλυνσης στις προβλέψεις του συνόλου δεδομένων που θα αποτελέσει τη μελέτη περίπτωσης της παρούσας εργασίας, παρουσιάζονται στην παράγραφο 5.4.2.

1.4 Ιεραρχική Πρόβλεψη

Η μελέτη συλλογών δεδομένων που οργανώνονται σε επίπεδα συνάνθροισης, απασχολεί την έρευνα και την πρακτική σε διάφορους τομείς της ανθρώπινης δραστηριότητας εδώ και δεκαετίες. Πρόβλεψη ΑΕΠ, πληθωρισμού, βιομηχανικής παραγωγής, πωλήσεων λιανικής, μονάδων εφοδιαστικής αλυσίδας (SKU), δημογραφικών μεγεθών, δεικτών θνητότητας, κατανάλωσης ενέργειας, φυσικών φαινομένων, τουριστικής ζήτησης είναι μερικές από τις εφαρμογές που εμφανίζεται η ανάγκη για προβλέψεις συλλογών χρονοσειρών που θα ικανοποιούν παράλληλα τους περιορισμούς που θέτει η εκάστοτε ιεραρχική δομή.

Η πληροφορία που βρίσκεται σε μια συλλογή ομαδοποιημένων χρονοσειρών, περιλαμβάνει τα μοντέλα των χρονοσειρών από όλα τα επίπεδα συνάνθροισης, καθώς και τους περιορισμούς συνάνθροισης (γραμμικοί περιορισμοί) που συνδέουν τα επίπεδα μεταξύ τους. Από αυτή την πληροφορία, οι μέθοδοι ιεραρχικής πρόβλεψης bottom up, top down και middle out αξιολογούν τα μοντέλα από ένα μοναδικό επίπεδο συνάνθροισης, και με χρήση των περιορισμών συνάνθροισης, διαμερίζουν ή συναθροίζουν τις προβλέψεις στα υπόλοιπα επίπεδα. Η *Optimal Combination* που προτάθηκε από τους Hyndman et al. [2011] και αργότερα επαναδιατυπώθηκε ως *Minimum Trace* από τους Wickramasuriya et al. [2019], όπως και άλλες μέθοδοι γραμμικού συνδυασμού προβλέψεων, αποτέλεσαν μια καινοτόμο προσέγγιση στην ιεραρχική πρόβλεψη. Οι μέθοδοι αυτές, υποθέτοντας ένα σύνολο ανεξάρτητων προβλέψεων για όλα τα επίπεδα συνάνθροισης, παράγουν βέλτιστες αμερόληπτες προβλέψεις, εναρμονισμένες με τους περιορισμούς συνάνθροισης. Περισσότερα για τις διαφορετικές μεθόδους ιεραρχικής πρόβλεψης θα παρουσιαστούν στο κεφάλαιο 4.

Στις περισσότερες των περιπτώσεων, ο αποφασίζων επιθυμεί οι προβλέψεις να είναι εναρμονισμένες με όλα τα επίπεδα της ιεραρχίας, δηλαδή οι προβλέψεις να υπακούν στους περιορισμούς συνάνθροισης των δεδομένων που παρέχονται ad hoc σε κάθε εφαρμογή. Για δεκαετίες, αυτό αποτέλεσε και το πλαίσιο ανάπτυξης τόσο των single-level μεθόδων ιεραρχικής πρόβλεψης, όσο και των μεθόδων γραμμικού συνδυασμού των μοντέλων ολόκληρης της ιεραρχίας. Από τη σκοπιά του αναλυτή, η παραπάνω θεώρηση καθιστά την ιεράρχηση των δεδομένων σε επίπεδα συνάνθροισης, παθητικό στοιχείο της ανάλυσης. Με άλλα λόγια, όποια επιλογή κι αν γίνει, οι τελικές προβλέψεις θα πρέπει να ευθυγραμμίζονται με όλα τα παρεχόμενα επίπεδα συνάνθροισης της συλλογής ομαδοποιημένων χρονοσειρών.

Υπό το πρίσμα των εμπειρικών αποτελεσμάτων που έδειξαν ότι σε πολλές εφαρμογές,

οι μέθοδοι ιεραρχικής πρόβλεψης μπορούν να προβλέπουν τα επίπεδα συνάθροισης τους με μεγαλύτερη ακρίβεια από ό,τι τα ανεξάρτητα μοντέλα των χρονοσειρών, ή με άλλα λόγια πως η εναρμόνιση με την ιεραρχία μπορεί να μεταφέρει εντός του μοντέλου κάθε χρονοσειράς, εξωγενή πληροφορία που περιέχεται σε μοντέλα άλλων επιπέδων συνάθροισης, γεννιέται η απορία αν οι ίδιες οι ιεραρχίες με τις οποίες οι προβλέψεις εναρμονίζονται είναι σωστά επιλεγμένες, τόσο ως προς την ενίσχυση της ακρίβειας των προβλέψεων, όσο και την αποφυγή μοντελοποίησης επιπέδων που δεν συνεισφέρουν στην ακρίβεια.

Σκοπός της ιεραρχικής πρόβλεψης στην εργασία, είναι η εναρμόνιση των προβλέψεων βάσης μιας συλλογής χρονοσειρών, με διαφορετικές ιεραρχίες και η μελέτη της επίδρασης στην ακρίβεια και το κόστος των παραγόμενων προβλέψεων. Η μέθοδος ιεραρχικής πρόβλεψης με την οποία θα εναρμονίζονται τα μοντέλα των εξεταζόμενων ιεραρχιών θα είναι η *Minimum Trace*. Συνεπώς, η μέθοδος ιεραρχικής πρόβλεψης της κάθε ιεραρχίας θα αποτελέσει σταθερά του πειράματος και παράμετροι θα είναι οι ιεραρχίες με τις οποίες θα τροφοδοτείται η μέθοδος. Όπως ελέγχθη, η *Minimum Trace* εναρμόνιση, σταθμίζει όλα τα επίπεδα της ομαδοποιημένης χρονοσειράς, με τα μοντέλα της οποίας τροφοδοτείται. Η μεθοδολογία που θα αναπτυχθεί θα αξιοποιεί το μοντέλο παλινδρόμησης της *Minimum Trace*, αλλά για την εναρμόνιση δειγμάτων των δεδομένων με διαφορετικές ιεραρχίες, με σκοπό την εύρεση της κατάλληλης ιεραρχίας για εναρμόνιση. Περισσότερα για την ιεραρχική πρόβλεψη στην εργασία παρουσιάζονται στην παράγραφο 5.4 της μελέτης περίπτωσης.

1.5 Μελέτη Περίπτωσης

Η αξιολόγηση της επίδρασης της ιεράρχησης δεδομένων θα προκύψει από μία μελέτη περίπτωσης, πάνω στα δεδομένα που τέθηκαν υπό πρόβλεψη στο διαγωνισμό *M5* και περιλαμβάνουν χρονοσειρές πωλήσεων με γεωγραφική και ανά προϊόν οργάνωση. Το σύνολο δεδομένων του διαγωνισμού *M5* αποτελεί μια συλλογή ομαδοποιημένων χρονοσειρών που περιέχει 34090 μη συναθροισμένες χρονοσειρές και συνολικά 42840 χρονοσειρές, που οργάνωνονται σε 12 επίπεδα συνάθροισης. Αφορούν πωλήσεις της Walmart, επιχείρηση που δραστηριοποιείται στις λιανικές πωλήσεις στις ΗΠΑ και διεθνώς και αποτελεί την επιχείρηση με τα υψηλότερα έσοδα παγκοσμίως. Περισσότερες πληροφορίες για το σύνολο δεδομένων παρουσιάζονται στις παραγράφους 5.2 και 5.3.

Η διάταξη του πειράματος θα θέτει ως σταθερές τόσο τη μέθοδο με την οποία θα μοντελοποιούνται οι χρονοσειρές της συλλογής ξεχωριστά, όσο και τη μέθοδο ιεραρχικής πρόβλεψης με την οποία θα εναρμονίζονται οι προβλέψεις βάσεις. Οι επιλογές θα είναι εκθετική εξομάλυνση και *Minimum Trace* συνάθροιση αντίστοιχα. Για την εναρμόνιση των προβλέψεων με *Minimum Trace* συνάθροιση θα χρησιμοποιηθούν 48 ιεραρχίες δεδομένων που μπορούν να προκύψουν από τα επίπεδα της συλλογής ομαδοποιημένων χρονοσειρών της μελέτης περίπτωσης. Συνεπώς, με την παραπάνω διάταξη, οι 48 ιεραρχίες μαζί με τις μεθόδους σημεία-αναφοράς *bottom up* και *top down* θα αποτελέσουν τις μοναδικές παραμέτρους ακρίβειας και κόστους των προβλέψεων.

Οι προβλέψεις που έχουν εναρμονιστεί με καθεμιά εκ των 48 ιεραρχιών ή των μεθόδων σημείων-αναφοράς, θα συγκρίνονται ως προς την ακρίβεια με το μέτρο *RMSSE (Root Mean Squared Scaled Error)* λόγω ορισμένων χρήσιμων ιδιοτήτων που επεξηγούνται

αναλυτικότερα στο κεφάλαιο 3 που περιγράφονται βασικοί στατιστικοί δείκτες ακρίβειας και στην παράγραφο 5.4.3 που αναλύεται η μέτρηση ακρίβειας στην παρούσα εργασία. Το κόστος πρόβλεψης θα μετρηθεί σε πλήθος χρονοσειρών, απαιτούμενο χρόνο και μνήμη (5.4.3). Οι προδιαγραφές και η πειραματική διάταξη για την κατάταξη των προβλέψεων και την εύρεση της ή των καλύτερων ιεραρχιών για πρόβλεψη του συνόλου δεδομένων του διαγωνισμού *M5* παρουσιάζονται αναλυτικότερα στην παράγραφο 5.4.

Η πειραματική διαδικασία θα περιλαμβάνει την προεπεξεργασία, δειγματοληψία, μοντελοποίηση, εναρμόνιση με τις εξεταζόμενες ιεραρχίες, πρόβλεψη, κατάταξη και τέλος επιλογή μιας ή περισσοτέρων ιεραρχιών ως βέλτιστες παραμέτρους απόδοσης προβλέψεων. Όλες οι παραπάνω λειτουργίες θα αυτοματοποιηθούν μέσω της ανάπτυξης ενός προγράμματος σε γλώσσα R. Περισσότερες πληροφορίες για τα βήματα της μεθοδολογίας εύρεσης βέλτιστης ιεραρχίας που αναπτύχθηκε στο πλαίσιο της εργασίας, αλλά και της υλοποίησης του προγράμματος, δίνονται στην παράγραφο 5.6.

1.6 Ευρήματα & Συμπεράσματα

Από την ανάλυση των αποτελεσμάτων της μελέτης περίπτωσης που παρουσιάζονται στο κεφάλαιο 6, βασικό συμπέρασμα που προέκυψε είναι ότι η προβλεπτική αξία των επιπέδων σε μια συλλογή ιεραρχικών ή ομαδοποιημένων χρονοσειρών, δεν είναι ισοδύναμη για κάθε εξεταζόμενο επίπεδο ή ομάδα επιπέδων συνάθροισης. Πιο συγκεκριμένα, η εναρμόνιση με επίπεδα που περιέχουν περισσότερες χρονοσειρές δεν οδηγεί απαραίτητα σε καλύτερες προβλέψεις ενώ αυξάνει σημαντικά το κόστος πρόβλεψης. Αντίθετα, ιεραρχίες που περιέχουν -αποκλειστικά- επίπεδα υψηλής συνάθροισης, έχουν χαμηλό κόστος πρόβλεψης, αλλά και χαμηλότερη μεσοσταθμική ακρίβεια στις προβλέψεις των άλλων επιπέδων. Τα επίπεδα που βρίσκονται στο μέσο της ή των ξεχωριστών ιεραρχιών μιας συλλογής ομαδοποιημένων χρονοσειρών, παρέχουν υψηλή προβλεπτική αξία τόσο μεσοσταθμικά όσο και ανά επίπεδο συνάθροιση, ενώ η μοντελοποίηση επιπέδων μεσαίας συνάθροισης, διατηρεί το κόστος πρόβλεψης σε χαμηλά επίπεδα.

Οι προβλέψεις που συνδυάζουν γραμμικά, μοντέλα από περισσότερα επίπεδα, προέβλεψαν με μεγαλύτερη ακρίβεια σε σχέση με τις μεθόδους *bottom up* και *top down* στη συντριπτική πλειοψηφία των δειγμάτων του συνόλου δεδομένων. Ωστόσο, ο κανόνας αυτός δεν είναι καθολικός, αφού μπορεί να προκύψουν και δείγματα στα οποία τα *benchmarks* υπεραποδίδουν. Κατά την κατάταξη των ιεραρχιών ως προς την ακρίβεια αλλά και το κόστος πρόβλεψης, προκύπτουν *clusters* ιεραρχιών που εμφανίζουν παρόμοια ακρίβεια και χρόνους, μεσοσταθμικά αλλά και ανά επίπεδο. Συνεπώς η βέλτιστη οργάνωση επιπέδων για την ενίσχυση των προβλέψεων, μπορεί να περιλαμβάνει ένα σύνολο ιεραρχιών με κοινά χαρακτηριστικά, και όχι απαραίτητα μια μοναδική ιεραρχία.

Κύριος παράγοντας εξάρτησης του μεγέθους κάθε ιεραρχίας, είναι το πλήθος των χρονοσειρών στα επίπεδα που περιλαμβάνει. Στη μελέτη περίπτωσης, το επίπεδο συνάθροισης που χαρακτηρίζει σε μεγάλο βαθμό το μέγεθος της συλλογής -και κατ'επέκταση το κόστος της μοντελοποίησης- είναι αυτό των πωλήσεων ανά προϊόν. Τέτοια επίπεδα, μπορούν να ελέγχονται ως προς την προβλεπτική τους ικανότητα και αν δεν προσφέρουν μετρήσιμη πληροφορία μπορούν να μη συμπεριλαμβάνονται στην οργάνωση

που θα επιλεγεί για εναρμόνιση. Ακόμη ένας παράγοντας εξάρτησης του μεγέθους της ιεραρχίας είναι η θέση στην οποία θα καταλαμβάνουν τα πολυπληθή επίπεδα στην ιεραρχία. Όσο ψηλότερα τοποθετείται ένα επίπεδο με μεγάλο πλήθος χρονοσειρών, τόσο υψηλότερος γίνεται και ο παράγοντας διακλάδωσης (*branching factor*) της εκάστοτε ιεραρχίας και άρα το συνολικό πλήθος των περιεχόμενων χρονοσειρών. Παράλληλα η τοποθέτηση ενός πολυπληθούς επιπέδου ψηλότερα στην ιεραρχία δεν έδειξε να βελτιώνει τις προβλέψεις, τόσο μεσοσταθμικά όσο και ανά επίπεδο συνάνθροισης. Ευρήματα και συμπεράσματα που προκύπτουν από τα αποτελέσματα της παρούσας εργασίας, καθώς και μελλοντικές προεκτάσεις, παρουσιάζονται στο κεφάλαιο 7.

Ατυχείς Προβλέψεις #3

«Νομίζω ότι υπάρχει μια παγκόσμια αγορά
για περίπου πέντε υπολογιστές.»

πρόεδρος της IBM, 1943

2 Μοντέλα Πρόβλεψης

Η πρόβλεψη αποτελεί προϋπόθεση για τη λήψη απόφασης. Χάραξη πολιτικών πάνω σε εσφαλμένες προβλέψεις είναι συχνά επικίνδυνη. Διάσημες προβλέψεις όπως «Δεν υπάρχει λόγος να θέλει κανείς έναν υπολογιστή στο σπίτι του» που έγινε από τον πρόεδρο της *Digital Equipment Corporation* το 1977, καταδεικνύουν πως οι παράγοντες που καθορίζουν το μέλλον συχνά μας διαφεύγουν. Τρία χρόνια μετά, η *IBM* ανέπτυξε τον πρώτο οικιακό υπολογιστή και λιγότερο από δύο δεκαετίες μετά η *DEC* αποτελούσε παρελθόν.

Ο τρόπος με τον οποίο θα προβλέψουμε εξαρτάται σε μεγάλο βαθμό από τα διαθέσιμα δεδομένα. Αν δεν έχουμε δεδομένα ή τα δεδομένα δεν είναι σχετικά με την πρόβλεψη, τότε πρέπει να επιλέξουμε κάποια μέθοδο κριτικής πρόβλεψης. Τέτοιες μέθοδοι αποτελούν δομημένες προσεγγίσεις για την εξαγωγή προβλέψεων όταν δεν υπάρχουν διαθέσιμα ιστορικά δεδομένα. Προσεγγίσεις όπως η μέθοδος των Δελφών, πρόβλεψη μέσω αναλογίας, πρόβλεψη σεναρίων ή προβλέψεις πωλητών, προβλέψεις διοικητών, προθέσεις καταναλωτών επιδιώκουν να εφαρμόσουν διαδικασίες παραγωγής ή προσαρμογής ποσοτικών προβλέψεων μέσω ποιοτικών μεθόδων και κριτηρίων.

Όταν έχουμε ιστορικά δεδομένα για την ενδιαφερόμενη ποσότητα και μπορούμε να συμπεράνουμε πως μέρος των μοτίβων που προκύπτουν σε παρελθοντικές παρατηρήσεις, θα συνεχιστούν και στο μέλλον -δηλαδή ότι η εξάρτηση των παρατηρήσεων είναι σε κάποιο βαθμό αιτιοκρατική- μπορούμε να εφαρμόσουμε ποσοτικές μεθόδους πρόβλεψης.

Υπάρχει ένα ευρύ φάσμα ποσοτικών/στατιστικών μεθόδων που παράγουν μοντέλα χρονοσειρών [2021], κατά κανόνα κατασκευασμένα με συγκεκριμένες αρχές και για συγκεκριμένο σκοπό. Κάθε μέθοδος έχει διαφορετικές ιδιότητες, ακρίβεια και κόστος σε πόρους, όλα εκ των οποίων χρειάζεται να ληφθούν υπ' όψη κατά την επιλογή του κατάλληλου μοντέλου για πρόβλεψη.

2.1 Απλές Μέθοδοι Πρόβλεψης

Πέρα από τη χρήση πιο εξειδικευμένων και ενδεχομένως αποτελεσματικότερων μεθόδων, προβλέψεις μπορούν να εξαχθούν και μέσω αρκετά απλών -διαισθητικά και υπολογιστικά- ποσοτικών μεθόδων. Οι μέθοδοι αυτές χρησιμοποιούνται κυρίως ως benchmarks για την επίδοση άλλων μεθόδων, ωστόσο υπό συνθήκες μπορούν να αποδειχθούν αποτελεσματικές.

2.1.1 Μέσος Όρος

Έχοντας κανείς στη διάθεσή του ιστορικές παρατηρήσεις ενός μεγέθους μπορεί να προβλέψει κάθε μελλοντική περίοδο h , ως τον αριθμητικό μέσο των διαθέσιμων παρατηρήσεων:

$$\hat{\mathbf{y}}_{T+h} = \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \quad (1)$$

όπου \mathbf{y}_t οι ιστορικές παρατηρήσεις, $\hat{\mathbf{y}}_{T+h}$ η πρόβλεψη για τη μελλοντική στιγμή h . Με άλλα λόγια, για κάθε μελλοντική στιγμή προβλέπεται η τιμή της μέσης στάθμης των

παρατηρήσεων στις στιγμές $t = 1, \dots, t$. Αποτελεί μια πολύ απλή μέθοδο πρόβλεψης, ενώ δεν πρέπει να συνδέεται με τους κινητούς μέσους όρους που θα παρουσιαστούν στη συνέχεια.

2.1.2 Αφελής Μέθοδος (*Naive*)

Η «αφελής» μέθοδος, προβλέπει -για κάθε μελλοντική στιγμή- την τιμή της τελευταίας παρατήρησης στη χρονοσειρά:

$$\hat{\mathbf{y}}_{T+h} = \mathbf{y}_T \quad (2)$$

οι προβλέψεις της με naïve μέθοδο, χρησιμοποιούνται συχνά ως benchmarks για την απόδοση πιο σύνθετων μεθόδων ή για κανονικοποιημένα μέτρα ακρίβειας όπως τα MASE, RMSSE. Η naïve μέθοδος τα πηγαίνει πολύ καλά σε αρκετές οικονομικές και χρηματοοικονομικές χρονοσειρές [Hyndman, 2021]. Επειδή οι προβλέψεις με naïve είναι βέλτιστες όταν τα δεδομένα ακολουθούν τυχαίο περίπατο, ονομάζονται και «προβλέψεις τυχαίου περιπάτου». Ένα ακόμη ενδιαφέρον χαρακτηριστικό της naïve είναι πως οι προβλέψεις που παράγει για χρονοσειρές που οργανώνονται σε επίπεδα συνάθροισης, είναι εναρμονισμένες με την εκάστοτε ιεραρχία.

2.1.3 Εποχιακή Naive

Η εποχιακή «αφελής» μέθοδος, υποθέτοντας εποχιακή περίοδο n , θέτει κάθε πρόβλεψη ως την τελευταία παρατηρηθείσα τιμή της αντίστοιχης περιόδου:

$$\hat{\mathbf{y}}_{T+h} = \mathbf{y}_{T+h-n(\lfloor \frac{h-1}{n} \rfloor + 1)} \quad (3)$$

Για παράδειγμα, για μηνιαία δεδομένα οι προβλέψεις κάθε Μάρτη θα είναι όσο η τιμή του πιο πρόσφατου Μάρτη στη χρονοσειρά. Αντίστοιχα προκύπτουν οι τιμές για ημερήσια, εβδομαδιαία, ετήσια κ.α. δεδομένα.

2.1.4 Τάση

Μια παραλλαγή της naïve μεθόδου, είναι να επιτρέπεται στις προβλέψεις να αυξάνονται ή μειώνονται γραμμικά κατά τις περιόδους του ορίζοντα πρόβλεψης h , με κλίση τη μέση ιστορική μεταβολή των παρατηρήσεων:

$$\hat{\mathbf{y}}_{T+h} = \mathbf{y}_T + h \cdot \frac{\mathbf{y}_T - \mathbf{y}_1}{T - 1} \quad (4)$$

Η μέθοδος εποπτικά αποτελεί μια ευθεία γραμμή που περνά από την πρώτη και την τελευταία παρατήρηση της χρονοσειράς. Είναι αποτελεσματική κυρίως για μακροπρόθεσμες προβλέψεις.

2.2 Αποσύνθεση

Οι χρονοσειρές μπορούν να παρουσιάζουν ποικιλία μοτίβων και συχνά είναι χρήσιμος ο διαχωρισμός τους σε συνιστώσες, κάθε μια εκ των οποίων ενσωματώνει μια διαφορετική κατηγορία μοτίβων. Αυτές είναι οι βασικές κατηγορίες μοτίβων που παρουσιάζουν οι χρονοσειρές:

- **Τάση (T)** υπάρχει όταν παρατηρούμε μια μακροπρόθεσμη αύξηση ή μείωση της στάθμης των δεδομένων, όχι απαραίτητα γραμμική.
- **Εποχιακότητα (S)** προκύπτει όταν οι παρατηρήσεις επηρεάζονται από ημερολογιακούς παράγοντες όπως η μέρα της εβδομάδας ή ο μήνας του χρόνου και εμφανίζουν περιοδική συμπεριφορά σταθερής συχνότητας.
- **Κυκλικότητα (C)** εμφανίζεται όταν υπάρχουν άνοδοι και πτώσεις στη στάθμη των δεδομένων, χωρίς σταθερή συχνότητα. Αυτές οι διακυμάνσεις οφείλονται συχνά σε οικονομικούς παράγοντες και συνήθως ξεπερνούν τα δύο χρόνια.
- **Τυχειότητα (R)** είναι ό,τι απομένει μετά την εξαγωγή των παραπάνω συνιστωσών, και -στο βαθμό που έχουμε καταφέρει να απομονώσουμε τα αιτιοκρατικά μοτίβα της χρονοσειράς- αποτελεί τυχαίο θόρυβο.

2.2.1 Κινητοί Μέσοι Όροι

Οι κινητοί μέσοι όροι χρησιμοποιούνται για να εξαχθεί η τάση-κυκλικότητα μιας χρονοσειράς και αποτελούν το πρώτο βήμα στην κλασική μέθοδο αποσύνθεσης. Ένας κινητός μέσος όρος τάξης $n = 2k + 1$ γράφεται ως:

$$\hat{TC}_t = \frac{1}{n} \sum_{i=-k}^k y_{t+i} \quad (5)$$

Η εφαρμογή ΚΜΟ(n) εξομαλύνει τη χρονοσειρά και κατά συνέπεια αφαιρεί τυχειότητα και εποχιακότητα. Υπάρχουν διάφορες παραλλαγές των κινητών μέσων όρων. Ένας ΚΜΟ(n) μπορεί να είναι απλός, διπλός, ν-απλός, σταθμισμένος, κεντρικός και άλλα. Το μοντέλο που προκύπτει από την εφαρμογή ΚΜΟ, λόγω της ν προς 1 σχέσης που έχουν οι παρατηρήσεις με τις αντίστοιχες τιμές του μοντέλου έχει κενές πρώτες ή/και τελευταίες παρατηρήσεις. Οι θέσεις αυτές μπορούν να συμπληρωθούν με back-casting. Οι κινητοί μέσοι όροι, αποτελούν μεν το πρώτο βήμα της αποσύνθεσης, ωστόσο μέθοδοι που εφαρμόζουν συνδυασμούς από ΚΜΟ μπορούν να χρησιμοποιηθούν και ως ανεξάρτητες μέθοδοι πρόβλεψης.

2.2.2 Κλασική Μέθοδος Αποσύνθεσης

Η κλασική μέθοδος αποσύνθεσης αναπτύχθηκε στα 1920 και αποτελεί τη βάση για τις περισσότερες σύγχρονες μεθόδους αποσύνθεσης. Υποθέτει μια συναρτησιακή σχέση

ανάμεσα στη χρονοσειρά και τις συνιστώσες της:

$$Y_t = f(T_t, S_t, C_t, R_t) \quad (6)$$

και είναι συνήθως προσθετική ή πολλαπλασιαστική. Στην κλασσική μέθοδο αποσύνθεσης, υποθέτουμε ότι η εποχιακή συνιστώσα παραμένει σταθερή από χρονιά σε χρονιά. Στο πολλαπλασιαστικό μοντέλο κλασσικής αποσύνθεσης, οι m τιμές που αποτελούν την εποχιακή συνιστώσα, καλούνται δείκτες εποχιακότητας.

Προσθετικό Μοντέλο Κλασσικής Αποσύνθεσης Στο προσθετικό μοντέλο της κλασσικής μεθόδου αποσύνθεσης, υποθέτουμε προσθετική σχέση:

$$Y_t = T_t + S_t + C_t + R_t \quad (7)$$

ανάμεσα στις συνιστώσες που συνθέτουν τη χρονοσειρά Y . Για να απομονώσουμε τις παραπάνω συνιστώσες, ακολουθούμε την εξής διαδικασία:

1. Εξάγουμε την $T_t + C_t$ συνιστώσα, με χρήση κεντρικού ΚΜΟ(m) για περιττή εποχιακή περίοδο m ή $2 \times \text{ΚΚΜΟ}(m)$ αν η περίοδος είναι άρτιος αριθμός. Οι ΚΜΟ, όπως αναφέραμε αφαιρούν διακυμάνσεις που προκύπτουν από την εποχιακότητα και την τυχαιότητα.
2. Υπολογίζουμε την χρονοσειρά χωρίς τάση-κυκλικότητα, αφαιρώντας τη χρονοσειρά των ΚΜΟ από την Y , δηλαδή: $Y_t - T_t - C_t$. Αυτό που απομένει είναι η χρονοσειρά: $S_t + R_t$.
3. Αφού εξάγουμε τάση-κυκλικότητα, στην $Y_t - T_t - C_t$ παίρνουμε το μέσο όρο των παρατηρήσεων από τις ισοδύναμες modulo m χρονικές στιγμές (δηλαδή παρατηρήσεις που αντιστοιχούν στις ίδιες στιγμές της εποχιακής περιόδου m) και κανονικοποιούμε τις τιμές. Εξάγουμε την S συνιστώσα τοποθετώντας επαναλαμβανόμενα τις m τιμές για όσες επαναλήψεις της περιόδου έχουμε στα δεδομένα.
4. Η συνιστώσα της τυχαιότητας R θα είναι η $Y_t - T_t - C_t - S_t$.

Πολλαπλασιαστικό Μοντέλο Κλασσικής Αποσύνθεσης Στο πολλαπλασιαστικό μοντέλο εκθετικής εξομάλυνσης, υποθέτουμε προσθετική σχέση:

$$Y_t = T_t \cdot S_t \cdot C_t \cdot R_t \quad (8)$$

ανάμεσα στις συνιστώσες που συνθέτουν τη χρονοσειρά Y . Ακολουθούμε παρόμοια διαδικασία με του προσθετικού μοντέλου, όμως αντί να αφαιρούμε την κάθε συνιστώσα από την Y , διαιρούμε την Y για να απομονώσουμε τις συνιστώσες:

1. Εξάγουμε την $T_t C_t$ συνιστώσα, με χρήση ΚΜΟ(m) για περιττή εποχιακή περίοδο m ή $2 \times \text{ΚΜΟ}(m)$ αν η περίοδος είναι άρτιος αριθμός.

2. Υπολογίζουμε την χρονοσειρά χωρίς τάση-κυκλικότητα: $Y_t \cdot \frac{1}{T_t C_t} = S_t \hat{R}_t$.
3. Στην $\frac{Y_t}{T_t C_t}$ παίρνουμε το μέσο όρο των παρατηρήσεων από τις ισοδύναμες modulo m χρονικές στιγμές και κανονικοποιούμε τις τιμές. Εξάγουμε την S συνιστώσα τοποθετώντας επαναλαμβανόμενα τους δείκτες εποχιακότητας για όσες επαναλήψεις της περιόδου έχουμε στα δεδομένα.
4. Η συνιστώσα R θα είναι η $\frac{Y_t}{T_t C_t S_t}$.

2.2.3 Άλλες Μέθοδοι Αποσύνθεσης

Η ιδέα της εξαγωγής των συνιστωσών μιας χρονοσειράς που εφαρμόζει η κλασσική αποσύνθεση γνώρισε διάφορες τροποποιήσεις και βελτιώσεις. Η μέθοδος Census II που παρουσιάστηκε το 1955 όρισε περισσότερα βήματα που συμπληρώνουν κενά της κλασσικής μεθόδου και αργότερα βελτιώθηκε με τις μεθόδους X-11 και X-12. Η μέθοδος αποσύνθεσης STL (*Seasonal and Trend decomposition using Loess*) αναπτύχθηκε το 1990 από τους R. B. Cleveland, Cleveland, McRae, & Terpenning και αποτελεί μια ευέλικτη και αποτελεσματική εναλλακτική μέθοδο αποσύνθεσης. Η μέθοδος SEATS είναι ακόμη μια μέθοδος αποσύνθεσης που αναπτύχθηκε από την Τράπεζα της Ισπανίας και χρησιμοποιείται σε διάφορους οργανισμούς.

2.3 Παλινδρόμηση

Βασική ιδέα των μοντέλων παλινδρόμησης είναι ότι προβλέπουμε τη χρονοσειρά ενδιαφέροντος \mathbf{y} υποθέτοντας ότι έχει γραμμική σχέση με κάποια άλλη χρονοσειρά \mathbf{x} . Η \mathbf{y} είναι η εξαρτημένη μεταβλητή και \mathbf{x} η ανεξάρτητη.

2.3.1 Απλή Γραμμική Παλινδρόμηση

Στόχος της απλής γραμμικής παλινδρόμησης είναι η έκφραση της σχέσης ανάμεσα στις \mathbf{y} και \mathbf{x} με την εξίσωση μιας ευθείας γραμμής:

$$\mathbf{y}_t = \beta_0 + \beta_1 \mathbf{x}_t + \varepsilon_t \quad (9)$$

όπου β_0 είναι η προβλεπόμενη τιμή για την \mathbf{y} όταν $\mathbf{x} = 0$ και β_1 η κλίση της ευθείας ή αλλιώς η μέση μεταβολή στην προβλεπόμενη τιμή της \mathbf{y} ανά μονάδα μεταβολής της \mathbf{x} . Ο όρος ε_t είναι το σφάλμα-απόκλιση των παρατηρήσεων από την ευθεία και αποτελεί την ποσότητα που καλούμαστε να ελαχιστοποιήσουμε. Θεωρούμε το ε ως μια τυχαία μεταβλητή που ακολουθεί κανονική κατανομή γύρω από το 0.

2.3.2 Πολλαπλή Γραμμική Παλινδρόμηση

Στην περίπτωση που εισάγουμε περισσότερες ανεξάρτητες μεταβλητές $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, το μοντέλο της γραμμικής παλινδρόμησης μπορεί να γενικευθεί ως:

$$\mathbf{y}_t = \beta_0 + \beta_1 \mathbf{x}_{1,t} + \beta_2 \mathbf{x}_{2,t} + \dots + \beta_n \mathbf{x}_{n,t} + \varepsilon_t \quad (10)$$

2.3.3 Εκτίμηση μέσω Ελαχίστων Τετραγώνων

Τους συντελεστές β_0, \dots, β_n συνήθως εξάγουμε μέσω της μεθόδου ελαχίστων τετραγώνων, δηλαδή ελαχιστοποιώντας την ποσότητα:

$$\sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T (\mathbf{y}_t - \beta_0 - \beta_1 \mathbf{x}_{1,t} - \dots - \beta_n \mathbf{x}_{n,t})^2 \quad (11)$$

Η ελαχιστοποίηση του αθροίσματος των τετραγωνισμένων σφαλμάτων των συντελεστών, ελαχιστοποιεί την ευκλείδεια απόσταση του μοντέλου από το σύνολο των παρατηρήσεων. Άλλες μη γραμμικές συναρτησιακές μορφές μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών, μπορούν να μετασχηματιστούν σε γραμμικές μέσω λογαρίθμησης, ριζικών ή άλλων τεχνικών και να χειριστούν όπως οι παραπάνω.

2.4 Εκθετική Εξομάλυνση

Η εκθετική εξομάλυνση προτάθηκε στα τέλη του 1950 [Brown RG., 1956, Holt, 1957, Brown, 1959, 1960, Charles C. Holt and Simon, 1960, Winters, 1960] και ενέπνευσε μια οικογένεια παραλλαγών που χρησιμοποιούνται επιτυχώς έως και σήμερα. Τα μοντέλα εκθετικής εξομάλυνσης αποτελούν σταθμισμένους κινητούς μέσους όρους, με βάρη που φθίνουν εκθετικά όσο συσχετίζουμε παρελθοντικότερες παρατηρήσεις. Παράγει γρήγορες και αξιόπιστες προβλέψεις για ένα ευρύ φάσμα χρονοσειρών και αποτελεί σημείο αναφοράς σε πολλές εφαρμογές και διαγωνισμούς. Απαριθμούμε κάποιες χαρακτηριστικές μεθόδους εκθετικής εξομάλυνσης και μια ταξινόμησή τους. Στη συνέχεια περιγράφουμε τα State Space Innovations και τη μέθοδο ETS [Hyndman et al., 2002] που θα αξιοποιήσουμε στο πείραμά μας.

2.4.1 Απλή Εκθετική Εξομάλυνση

Η απλή εκθετική εξομάλυνση εφαρμόζεται σε χρονοσειρές χωρίς προφανή τάση ή εποχιακότητα. Η πρόβλεψη για $h = 1$ με AEE είναι:

$$\begin{aligned} \hat{\mathbf{y}}_{T+1} &= \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots \\ &= \alpha y_T - (1 - \alpha)\hat{\mathbf{y}}_T \end{aligned} \quad (12)$$

όπου $\alpha \in [0, 1]$ είναι ο συντελεστής εξομάλυνσης. Θεωρούμε την -ισοδύναμη με την εξίσωση 12- διατύπωση:

$$\begin{aligned} \hat{\mathbf{y}}_{t+h} &= \ell_t \\ \ell_t &= \alpha y_t - (1 - \alpha)\ell_{t-1} \end{aligned} \quad (13)$$

όπου ℓ_t είναι το επίπεδο (εξομαλυσμένη τιμή) της χρονοσειράς τη στιγμή t . Οι παράμετροι που πρέπει να ορίσουμε για το αναδρομικό μοντέλο είναι ο συντελεστής εξομάλυνσης α και το αρχικό επίπεδο ℓ_0 . Μπορούμε να επιλέξουμε και τα 2 με ποιοτικά κριτήρια, ωστόσο ένας πιο αντικειμενικός τρόπος είναι ο υπολογισμός από τις παρατηρήσεις.

Συνήθως, οι παράμετροι ενός μοντέλου εκθετικής εξομάλυνσης υπολογίζονται έτσι ώστε να ελαχιστοποιείται το άθροισμα των τετραγωνισμένων σφαλμάτων (ελάχιστα τετράγωνα) των παρατηρήσεων από το μοντέλο, όπως και στην παλινδρόμηση. Αν $e_t = y_t - \hat{y}_t$ τα υπόλοιπα (residuals) χρονοσειράς-μοντέλου, επιδιώκουμε ελαχιστοποίηση της ποσότητας:

$$\sum_{t=1}^T y_t - \hat{y}_t = \sum_{t=1}^T e_t^2 \quad (14)$$

2.4.2 Εκθετική Εξομάλυνση με Τάση

Όπως φαίνεται από την εξίσωση πρόβλεψης της 13, η απλή εκθετική εξομάλυνση δίνει σταθερή πρόβλεψη ℓ_t για κάθε μελλοντική στιγμή. Η εξίσωση του επιπέδου δεν μπορεί να αποτυπώσει την τάση των παρατηρήσεων. Για το λόγο αυτό αναπτύχθηκαν μέθοδοι εκθετικής εξομάλυνσης που επεκτείνουν την AEE κι επιτρέπουν την πρόβλεψη σε χρονοσειρές με τάση.

Μέθοδος Γραμμικής Τάσης του Holt Ο Holt [1957] επέκτεινε τις εξισώσεις πρόβλεψης και επιπέδου της AEE στην εξίσωση 13, με μια εξίσωση τάσης b_t :

$$\begin{aligned} \hat{y}_{t+h} &= \ell_t + hb_t \\ \ell_t &= \alpha y_t - (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \end{aligned} \quad (15)$$

όπου b^* είναι ο συντελεστής της τάσης.

Οι προβλέψεις που παράγονται από τη μέθοδο Holt και έχουν γραμμική τάση υπήρξαν δημοφιλείς, όμως εμπειρικά δεδομένα έδειξαν πως τείνουν να υπερεκτιμούν στις προβλέψεις τους.

Μέθοδοι μη Γραμμικής Τάσης Πάνω σε αυτή την παρατήρηση αναπτύχθηκαν μοντέλα που προσαρμόζονται σε μη γραμμική τάση. Οι Gardner Jr. [1985], Gardner and McKenzie [1989] πρόσθεσαν μια ακόμη παράμετρο $\phi \in [0, 1]$ που κυρτώνει σταδιακά τη γραμμική τάση του μοντέλου Holt προς μια επίπεδη γραμμή, γνωστό και ως «Damped Trend» μοντέλο:

$$\begin{aligned} \hat{y}_{t+h} &= \ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t \\ \ell_t &= \alpha y_t - (1 - \alpha)(\ell_{t-1} + \phi b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1} \end{aligned} \quad (16)$$

Για $\phi = 1$ η μέθοδος ταυτίζεται με τη μέθοδο του Holt, ενώ για $\phi \in (0, 1)$ εφαρμόζεται η μείωση της τάσης, ώστε η πρόβλεψη να προσεγγίζει σταδιακά μια οριζόντια ευθεία. Συνήθως επιλέγονται τιμές στο διάστημα (0.8, 0.98) για την παράμετρο ϕ .

2.4.3 Εκθετική Εξομάλυνση με Εποχιακότητα

Ο Holt το 1957 και ο Winters το 1960 Charles C. Holt and Simon [1960] επέκτειναν την μέθοδο γραμμικής τάσης Holt, ώστε να αποτυπώνει και την εποχιακότητα μιας χρονοσειράς. Πλέον, το σύστημα των αναδρομικών εξισώσεων περιέχει εξίσωση πρόβλεψης \hat{y}_{t+h} , επιπέδου ℓ_t , τάσης b_t και εποχιακότητας s_t με αντίστοιχους συντελεστές εξομάλυνσης α , β^* και γ . Όταν οι εποχιακές μεταβολές είναι σχεδόν σταθερές μεταξύ των περιόδων, επιλέγεται το προσθετικό μοντέλο Holt-Winters:

$$\begin{aligned}\hat{y}_{t+h} &= \ell_t + hb_t + s_{t+h-n(k+1)} \\ \ell_t &= \alpha(y_t - s_{t-n}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \\ s_t &= \gamma(y_t - \ell_{t-1} - \beta_{t-1}) + (1 - \gamma)s_{t-n}\end{aligned}\quad (17)$$

όπου n η εποχιακή περίοδος και $k = \lfloor \frac{h-1}{n} \rfloor$. Όταν οι εποχιακές μεταβολές μεταβάλλονται αναλογικά με τη στάθμη των χρονοσειρών, εφαρμόζεται το πολλαπλασιαστικό μοντέλο Holt-Winters:

$$\begin{aligned}\hat{y}_{t+h} &= (\ell_t + hb_t)s_{t+h-n(k+1)} \\ \ell_t &= \alpha \frac{y_t}{s_{t-n}} + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \\ s_t &= \gamma \frac{y_t}{\ell_{t-1} + \beta_{t-1}} + (1 - \gamma)s_{t-n}\end{aligned}\quad (18)$$

Μπορούμε να εφαρμόσουμε «damping» - μείωση - της τάσης μέσω της παραμέτρου φ για το προσθετικό και το πολλαπλασιαστικό μοντέλο. Μια δημοφιλής επιλογή για πρόβλεψη εποχιακών χρονοσειρών είναι η Holt-Winters με μη γραμμική τάση και πολλαπλασιαστική εποχιακότητα:

$$\begin{aligned}\hat{y}_{t+h} &= [\ell_t + (\varphi + \varphi^2 + \dots + \varphi^h)b_t]s_{t+h-n(k+1)} \\ \ell_t &= \alpha \frac{y_t}{s_{t-n}} + (1 - \alpha)(\ell_{t-1} + \varphi b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\varphi b_{t-1} \\ s_t &= \gamma \frac{y_t}{\ell_{t-1} + \varphi \beta_{t-1}} + (1 - \gamma)s_{t-n}\end{aligned}\quad (19)$$

2.4.4 Ταξινόμηση Μεθόδων Εκθετικής Εξομάλυνσης

Οι μέθοδοι εκθετικής εξομάλυνσης μπορούν να μοντελοποιήσουν -μέσω των διαφόρων παραλλαγών οι κυριότερες από τις οποίες παρουσιάστηκαν στην παράγραφο 2.4- χρονοσειρές με γραμμική, προσθετική, damped, πολλαπλασιαστική, πολλαπλασιαστική damped ή χωρίς τάση και με προσθετική, πολλαπλασιαστική ή χωρίς εποχιακότητα.

Για κάθε ζευγάρι των διαφορετικών μορφών που έχουν τα ποιοτικά χαρακτηριστικά τάσης και εποχιακότητας, μπορούμε να επιλέξουμε το κατάλληλο μοντέλο εκθετικής εξομάλυνσης. Οι 15 συνδυασμοί που μπορούν να προκύψουν από το εξωτερικό γινόμενο των διαφορετικών χαρακτηριστικών μιας χρονοσειράς (5

| Τάση | εποχιακότητα | | |
|------------------------------|-----------------------------------|---|--|
| | Χωρίς | Προσθετική | Πολλαπλ/ική |
| Χωρίς | Απλή Εκθετική Εξομάλυνση | Όχι T Προσθετική S | Όχι T Πολλαπλ/ική S |
| Προσθετική | Γραμμική Μέθοδος Holt | Προσθετική Holt-Winters | Πολλαπλ/ική Holt-Winters |
| Προσθετική <i>Damped</i> | Προσθετική Damped T Μέθοδος | Damped T Προσθετική S | Holt-Winters Damped |
| Πολλαπλ/ική | Πολλαπλ/ική T Όχι S | Πολλαπλ/ική T Προσθετική S | Πολλαπλ/ική T Πολλαπλ/ική S |
| Πολλαπλ/ική <i>Damped</i> | Πολλαπλ/ική Damped T Όχι S | Πολλαπλ/ική Damped T Προσθετική S | Πολλαπλ/ική Damped T Πολλαπλ/ική S |

Πίνακας 1: Μοντέλα Εκθετικής Εξομάλυνσης

μορφές τάσης και 3 μορφές εποχιακότητας) φαίνονται στον πίνακα 1. Η κατηγοριοποίηση αυτή προτάθηκε από τον Pegels [1969] που συμπεριέλαβε ένα μοντέλο πολλαπλασιαστικής τάσης. Στη συνέχεια, επεκτάθηκε από τους Gardner Jr. [1985] και Taylor [2003] που προσέθεσαν τα μοντέλα της αθροιστικής damped και πολλαπλασιαστικής damped τάσης, αντίστοιχα. Μοντέλα με πολλαπλασιαστική τάση παρουσιάζουν αστάθεια [Gardner, 2006] και συναντάμε σπάνια έως ποτέ χρονοσειρές με τέτοια ποιοτικά χαρακτηριστικά, γι αυτό συνήθως μας ενδιαφέρουν τα μοντέλα που περιέχουν αθροιστική ή αθροιστική damped τάση και βρίσκονται στις πρώτες 3 γραμμές του πίνακα στο σχήμα 1.

2.4.5 Η μέθοδος ETS

Όλες οι μέθοδοι που προαναφέρθηκαν, παράγουν σημειακές προβλέψεις, δηλαδή ένα ρητό αριθμό για κάθε ζητούμενη μελλοντική στιγμή h . Οι Ord, Koehler, and Snyder [1997] και οι Snyder, Koehler, Hyndman, and Ord [2002] έδειξαν πως όλες οι μέθοδοι εκθετικής εξομάλυνσης αποτελούν βέλτιστες σημειακές προβλέψεις από στατιστικά μοντέλα που μπορούν να παράξουν ολόκληρη την κατανομή των προβλέψεων. Τα μοντέλα αυτά καλούνται *innovation state space models* Anderson and Moore [2012], Aoki [1987] και αποτελούν στοχαστικές διαδικασίες παραγωγής δεδομένων που μπορούν να αναπαράγουν κατανομές πιθανοτήτων. Επεκτείνουν της μεθόδους εκθετικής εξομάλυνσης, υπό την έννοια ότι παράγουν τις ίδιες σημειακές προβλέψεις, αλλά με την πληροφορία της κατανομής των προβλέψεων μπορούν να εξαχθούν διαστήματα πρόβλεψης.

Ο Hyndman, Koehler, Ord, and Snyder [2008] παρουσίασαν τη μέθοδο *Exponential Smoothing State Space Model* ή αλλιώς ETS (*Error-Trend-Seasonality*) που επεκτείνει τις μεθόδους εκθετικής εξομάλυνσης με μια συνιστώσα σφάλματος ε_t που προστίθεται στις προϋπάρχουσες εξισώσεις επιπέδου, τάσης και εποχιακότητας. Η ε_t μπορεί να σχετίζεται

με προσθετικό ή πολλαπλασιαστικό τρόπο με τις υπόλοιπες συνιστώσες. Συνεπώς, η ETS περιλαμβάνει $2 \cdot 15 = 30$ μεθόδους πρόβλεψης. Η μέθοδος γραμμικής τάσης του Holt για παράδειγμα, αν υποθέσουμε προσθετικά, κανονικά και ανεξάρτητα κατανομημένα σφάλματα:

$$\varepsilon_t = y_t - \ell_{t-1} - b_{t-1} \sim \mathcal{N}(0, \sigma^2) \quad (20)$$

επεκτείνεται στην επιλογή ETS(A,A,N) (additive error, additive trend, no seasonality):

$$\begin{aligned} \hat{y}_t &= \ell_{t-1} + b_{t-1} + \varepsilon_t \\ \ell_t &= \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t \\ b_t &= b_{t-1} + \beta\varepsilon_t \end{aligned} \quad (21)$$

όπου $\beta = \alpha\beta^*$. Ενώ αντίστοιχα, υποθέτοντας πολλαπλασιαστικά σφάλματα:

$$\varepsilon_t = \frac{y_t - (\ell_{t-1} - b_{t-1})}{\ell_{t-1} - b_{t-1}} \sim \mathcal{N}(0, \sigma^2) \quad (22)$$

προκύπτει το μοντέλο ETS(M,A,N) (multiplicative error, additive trend, no seasonality):

$$\begin{aligned} \hat{y}_t &= (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t) \\ \ell_t &= (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t) \\ b_t &= b_{t-1} + \alpha\beta^*(\ell_{t-1} + b_{t-1})\varepsilon_t \end{aligned} \quad (23)$$

Αντίστοιχα, επεκτείνονται και οι υπόλοιπες μέθοδοι του πίνακα 1, ώστε να περιέχουν συνιστώσα προσθετικού ή πολλαπλασιαστικού σφάλματος ε_t . Στον πίνακα 2 παρουσιάζονται τα μοντέλα εκθετικής εξομάλυνσης προσθετικού σφάλματος της ETS. Τα μοντέλα πολλαπλασιαστικής τάσης παραλείπονται για τους λόγους που εξηγήσαμε.

| Τάση | Εποχιακότητα | | |
|----------------------|--|---|--|
| | Χωρίς | Προσθετική | Πολλαπλασιαστική |
| Χωρίς | $y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ | $y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$ | $y_t = \ell_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/\ell_{t-1}$ |
| Προσθετική | $y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ | $y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$ | $y_t = (\ell_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + b_{t-1})$ |
| Προσθετική Damped | $y_t = \ell_{t-1} + \phi b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ | $y_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$ | $y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + \phi b_{t-1})$ |

Πίνακας 2: Μοντέλα με Αθροιστικό Σφάλμα ε_t

Υποθέτοντας τώρα, πολλαπλασιαστική συνιστώσα σφάλματος για τα μοντέλα της εκθετικής εξομάλυνσης στον πίνακα 1, προκύπτουν τα μοντέλα πολλαπλασιαστικού

σφάλματος της ETS, που παρουσιάζονται στον πίνακα 3. Η παραγωγή διαστημάτων πρόβλεψης που επιτυγχάνει η ETS, αποκτάται μέσω προσομοίωσης διαφορετικών μελλοντικών μονοπατιών (σημειακών προβλέψεων). Η επιλογή του καλύτερου μοντέλου εκθετικής εξομάλυνσης για πρόβλεψη -όπως θα δούμε στην επόμενη ενότητα- γίνεται με αυτόματο και οικονομικό τρόπο.

| Τάση | Εποχιακότητα | | |
|----------------------|---|---|--|
| | Χωρίς | Προσθετική | Πολλαπλασιαστική |
| Χωρίς | $y_t = \ell_{t-1}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$ | $y_t = (\ell_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \alpha(\ell_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + s_{t-m})\varepsilon_t$ | $y_t = \ell_{t-1}s_{t-m}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$ |
| Προσθετική | $y_t = (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ | $y_t = (\ell_{t-1} + b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ | $y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$ |
| Προσθετική Damped | $y_t = (\ell_{t-1} + \varphi b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \varphi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \varphi b_{t-1} + \beta(\ell_{t-1} + \varphi b_{t-1})\varepsilon_t$ | $y_t = (\ell_{t-1} + \varphi b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \varphi b_{t-1} + \alpha(\ell_{t-1} + \varphi b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = \varphi b_{t-1} + \beta(\ell_{t-1} + \varphi b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + \varphi b_{t-1} + s_{t-m})\varepsilon_t$ | $y_t = (\ell_{t-1} + \varphi b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \varphi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \varphi b_{t-1} + \beta(\ell_{t-1} + \varphi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$ |

Πίνακας 3: Μοντέλα με Πολλαπλασιαστικό Σφάλμα ε_t

2.4.6 Επιλογή Μοντέλου

Για την εξαγωγή των μοντέλων μέσω εκθετικής εξομάλυνσης απαιτείται επομένως, ο προσδιορισμός των συντελεστών εξομάλυνσης $\alpha, \beta, \gamma, \varphi$ καθώς και των αρχικών επιπέδων ℓ_0, b_0 και $s_0 \dots s_{m-1}$. Ένα σημαντικό πλεονέκτημα της ETS είναι πως η επιλογή του κατάλληλου μοντέλου μπορεί να γίνει με αυτόματο τρόπο, μέσω της ελαχιστοποίησης κριτηρίων πληροφορίας (*information criteria*) όπως τα **AIC**, **AIC_C** και **BIC**. Στο πλαίσιο της ETS, το κριτήριο πληροφορίας του Akaike ορίζεται ως εξής:

$$\mathbf{AIC} = -2 \log L + 2k \quad (24)$$

όπου L η πιθανοφάνεια του μοντέλου (εξίσωση 25) και k ο συνολικός αριθμός παραμέτρων και αρχικών καταστάσεων που έχουν εκτιμηθεί (συμπεριλαμβανομένης και της διακύμανσης των υπολοίπων/*residuals*).

$$L = n \log \left(\sum_{t=1}^n \frac{e_t^2}{kx_{t-1}^2} \right) + 2 \sum_{t=1}^n \log |kx_{t-1}| \quad (25)$$

Για να δοθεί μεγαλύτερο βάρος στην πολυπλοκότητα του μοντέλου εκθετικής εξομάλυνσης χρησιμοποιείται το "διορθωμένο" κριτήριο πληροφορίας **AIC_c**:

$$\mathbf{AIC}_c = \mathbf{AIC} + 2 \frac{k(k+1)}{T-k-1} \quad (26)$$

καθώς και το Μπαεζιανό Κριτήριο Πληροφορίας **BIC** (Schwarz, 1978) που δίνει κι αυτό μεγαλύτερο βάρος στην πολυπλοκότητα του μοντέλου ώστε να αποφευχθεί φαινόμενο υπερ-προσαρμογές (*over-fitting*) εκφράζεται μαθηματικά ως εξής:

$$\mathbf{BIC} = \mathbf{AIC} + k(\log T - 2) \quad (27)$$

Οι συντελεστές εξομάλυνσης λαμβάνουν τυπικά τιμές στο διάστημα $(0, 1)$, όμως για λόγους ευστάθειας των μοντέλων συνήθως λαμβάνουν τις παρακάτω τιμές:

$$0.1 \leq \alpha, \beta, \gamma \leq 0.9 \text{ και } \beta \leq \varphi \leq 1 \quad (28)$$

Τρεις από τους συνδυασμούς Σφάλματος-Τάσης-Εποχιακότητας μπορούν να οδηγήσουν σε αριθμητικές αστάθειες. Πιο συγκεκριμένα, τα μοντέλα που μπορεί να προκαλέσουν τέτοιες δυσκολίες είναι τα ETS(A,N,M), ETS(A,A,M) και ETS(A, Ad, M), λόγω διαίρεσεων με τιμές πιθανά κοντά στο μηδέν στις εξισώσεις κατάστασης (*state equations*) της ETS. Για το λόγο αυτό τα παραπάνω μοντέλα δεν περιλαμβάνονται στα εξεταζόμενα μοντέλα, όταν η επιλογή γίνεται με αυτόματο τρόπο.

Επιπλέον, τα μοντέλα με πολλαπλασιαστικά σφάλματα είναι χρήσιμα όταν τα δεδομένα είναι αυστηρά θετικά, αλλά γίνονται αριθμητικά ασταθή όταν στα δεδομένα περιέχονται μηδενικές ή αρνητικές παρατηρήσεις. Συνεπώς, κατά την αυτόματη επιλογή μοντέλου εκθετικής εξομάλυνσης, αν ανιχνευθούν μηδενικές ή αρνητικές παρατηρήσεις, τα πιθανά μοντέλα προς επιλογή είναι μόνο τα 6 πλήρως αθροιστικά μοντέλα.

2.5 Η μέθοδος *Theta*

Η μέθοδος *Theta* αναπτύχθηκε από τους Assimakopoulos and Nikolopoulos [2000] και αποτελεί μια μονοδιάστατη μέθοδο πρόβλεψης, που βασίζεται στη μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς με τη βοήθεια της παραμέτρου θ (*theta*) η οποία εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των παρατηρήσεων. Με αυτόν τον τρόπο, η αρχική χρονοσειρά αποσυντίθεται σε τουλάχιστον δύο καινούργιες χρονοσειρές (γραμμές *Theta*), οι οποίες προεκτείνονται στο μέλλον σαν ξεχωριστές χρονοσειρές. Η τελική πρόβλεψη προκύπτει ως συνδυασμός των επιμέρους προβλέψεων, που στη γενική περίπτωση είναι ένας σταθμισμένος μέσος.

2.5.1 Γραμμές *Theta*

Οι γραμμές *Theta* που παράγονται, έχουν την ιδιότητα να διατηρούν τη μέση τιμή και την κλίση της αρχικής χρονοσειράς, αλλά όχι τις τοπικές καμπυλότητες και τη διακύμανση. Για διάφορες τιμές της παραμέτρου θ , μπορεί να γίνει καλύτερη προσέγγιση της μακροπρόθεσμης ή βραχυπρόθεσμης συμπεριφοράς των δεδομένων. Όσο μικρότερη η τιμή της, τόσο μεγαλύτερος ο βαθμός μείωσης των καμπυλοτήτων. Σχηματικά, έχουμε τις εξής περιπτώσεις:

- $\theta > 1$: τονίζονται περισσότερο τα βραχυπρόθεσμα χαρακτηριστικά της χρονοσειράς.
- $\theta < 1$: τονίζονται περισσότερο τα μακροπρόθεσμα χαρακτηριστικά της χρονοσειράς.
- $\theta = 0$: η χρονοσειρά ταυτίζεται με την ευθεία της απλής γραμμικής παλινδρόμησης.
- $\theta = -1$: η χρονοσειρά αντιστοιχεί με τη συμμετρική της αρχικής χρονοσειράς ως προς την ευθεία της γραμμικής παλινδρόμησης.

2.5.2 Κλασσική Μέθοδος *Theta*

Το μοντέλο *Theta* που αποσυνθέτει τη χρονοσειρά σε δύο γραμμές με παραμέτρους $\theta = 0$ και $\theta = 2$ ονομάζεται «κλασσική μέθοδος *Theta*». Έγινε γνωστή στο διαγωνισμό προβλέψεων M3 που περιελάμβανε προβλέψεις για 3003 χρονοσειρές διαφόρων εφαρμογών. Η μέθοδος *Theta* ξεχώρισε στο διαγωνισμό ως η ακριβέστερη μέθοδος, υπεραπέδωσε συνθετότερων μεθόδων και είχε εντυπωσιακά αποτελέσματα σε μηνιαία και μικροοικονομικά δεδομένα. Τα βήματα της κλασσικής μεθόδου *Theta* είναι τα εξής:

0. **Έλεγχος Εποχιακότητας** Πρώτο βήμα για πρόβλεψη με κλασσική *Theta* αποτελεί ο έλεγχος για στατιστικά σημαντική εποχιακή συμπεριφορά της χρονοσειράς.
1. **Αποεποχιακοποίηση** Το δεύτερο βήμα αφορά στην αποεποχιακοποίηση της χρονοσειράς μέσω κλασσικής μεθόδου πολλαπλασιαστικής αποσύνθεσης, και είναι προαιρετικό εφόσον έχει προκύψει εποχιακή συμπεριφορά από το Βήμα 0.
2. **Αποσύνθεση** Η χρονοσειρά αποσυντίθεται σε δύο γραμμές *Theta*, την ευθεία με παράμετρο $\theta=0$ (ευθεία γραμμικής παλινδρόμησης) και με παράμετρο $\theta=2$.
3. **Πρόβλεψη** Οι ευθεία γραμμικής παλινδρόμησης ($\theta=0$) προεκτείνεται κατά τα γνωστά, ενώ η δεύτερη γραμμή ($\theta=2$) προεκτείνεται μέσω απλής εκθετικής εξομάλυνσης.
4. **Συνδυασμός** Οι επιμέρους προβλέψεις των 2 γραμμών συνδυάζονται μέσω αριθμητικού μέσου.
5. **Εποχιακοποίηση** Εφόσον προέκυψε αποεποχιακοποίηση στο βήμα 1, οι προβλέψεις επανεποχιακοποιούνται με βάση τους δείκτες εποχιακότητας.

Ο μαθηματικός τύπος που περιγράφει τη κλασσική μέθοδο *Theta* είναι ο εξής:

$$\mathbf{y}_t^\theta = \theta \cdot \mathbf{y}_{00}, \quad \text{με } \mathbf{y}_{00} = \mathbf{y}_t - 2\mathbf{y}_{t+1} + \mathbf{y}_{t+2} \quad (29)$$

Για τον υπολογισμό των ευθειών *Theta* χωρίς τη χρήση διαφορών δεύτερης τάξης, μπορεί να χρησιμοποιηθεί η σχέση 30, που προτάθηκε από τον Νικολόπουλο (2008) και όπως φαίνεται, οι γραμμές *Theta* προκύπτουν ως γραμμικοί συνδυασμοί των αρχικών παρατηρήσεων και της ευθείας γραμμικής παλινδρόμησης:

$$\mathbf{y}_t^\theta = \theta \cdot \mathbf{y}_t + (1 - \theta)(\alpha + \beta t) \quad (30)$$

όπου α και β είναι οι συντελεστές της ευθείας γραμμικής παλινδρόμησης που περιγράφονται από τις γνωστές σχέσεις:

$$\beta = \frac{\sum_{t=1}^n (\mathbf{y}_t - \bar{\mathbf{y}})(t - \bar{t})}{\sum_{t=1}^n (t - \bar{t})^2} \quad \text{και} \quad \alpha = \bar{\mathbf{y}} - \beta \bar{t} \quad (31)$$

Η μέθοδος Theta έχει αποτελέσει αντικείμενο ενεργούς έρευνας [Nikolopoulos et al., 2012, Petropoulos and Nikolopoulos, 2013, E. and B., 2017, Spiliotis et al., 2020, Fiorucci et al., 2016, Nikolopoulos, 2003, Petropoulos and Nikolopoulos, 2017, Nikolopoulos and Thomakos, 2020] από τη στιγμή που παρουσιάστηκε στον M3 και έπειτα. Οι Hyndman and Billah [2003] παρομοίασαν τις προβλέψεις της κλασσικής μεθόδου Theta, με αυτές της εκθετικής εξομάλυνσης με τάση.

2.6 Μοντέλα ARIMA

Τα μοντέλα *ARIMA* (autoregressive integrated moving averages) αποτελούν ακόμη μια δημοφιλή προσέγγιση στην πρόβλεψη χρονοσειρών, ίσως τη δημοφιλέστερη μαζί με την εκθετική εξομάλυνση. Μελετήθηκαν από τους Box and Jenkins [1970] και σε αντίθεση με τα μοντέλα εκθετικής εξομάλυνσης που συλλαμβάνουν ποιοτικά χαρακτηριστικά με χρονική εξάρτηση όπως την τάση και την εποχιακότητα μιας χρονοσειράς, τα μοντέλα *ARIMA* επιδιώκουν να αποτυπώσουν την αυτοσυσχέτιση των παρατηρήσεων.

Τα ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου είναι στοχαστικά μαθηματικά μοντέλα που περιγράφουν, αναλύουν και προβλέπουν τη διαχρονική εξέλιξη κάποιου φυσικού μεγέθους που εξαρτάται από μη ντετερμινιστικούς παράγοντες. Για να εφαρμοστούν σε μια χρονοσειρά, πρέπει εκείνη να είναι στάσιμη και διακριτή. Τα μοντέλα *ARIMA* δύνανται να εκφραστούν σαν ένας γραμμικός συνδυασμός του τυχαίου παράγοντα (τυχαίο σφάλμα ή σφάλμα πρόβλεψης), των τιμών του μεγέθους σε προηγούμενες χρονικές στιγμές και κάποιων άλλων στοχαστικών παραγόντων. Βέβαια σε πραγματικά δεδομένα είναι ιδιαίτερα δύσκολο να εντοπιστούν αυτοί οι τρεις παράγοντες, αλλά μπορούν να προσεγγιστούν αποτελεσματικά.

Για μια τυχαία χρονοσειρά, τα μοντέλα *ARIMA* βασίζονται στην παραδοχή της αλληλεξάρτησης μεταξύ των τιμών που λαμβάνει η χρονοσειρά για τις διάφορες χρονικές στιγμές. Εξάλλου, η πλειοψηφία των φυσικών μεγεθών δε δίνει τη δυνατότητα πλήρους επίγνωσης των παραγόντων που επηρεάζουν την εξέλιξη τους στο χρόνο, οπότε είναι ιδιαίτερα δύσκολη η περιγραφή τους με ένα ντετερμινιστικό μοντέλο. Από την άλλη μεριά, τα μη ντετερμινιστικά μοντέλα έχουν τη δυνατότητα να περιγράφουν την εξέλιξη ενός μεγέθους, υπολογίζοντας την πιθανότητα η τιμή του να βρίσκεται σε κάποιο διάστημα. Τα μοντέλα *ARIMA* μπορούν επίσης να μοντελοποιήσουν ένα μεγάλο εύρος εποχιακών δεδομένων.

Ένα μοντέλο *ARIMA* αποτελείται από ένα μοντέλο *AR*, ένα μοντέλο *MA* και ένα μοντέλο διαφόρισης που εξασφαλίζει τη στασιμότητα της χρονοσειράς. Στις παρακάτω ενότητες περιγράφουμε τα επιμέρους και το συνολικό μοντέλο:

2.6.1 Μοντέλα AR

Τα αυτοπαλινδρομούμενα μοντέλα υποθέτουν την ύπαρξη γραμμικής σχέσης ανάμεσα σε κάθε παρατήρηση της χρονοσειράς με τις προηγούμενες της. Ένα τέτοιο μοντέλο με τάξη p , μπορεί να γραφτεί ως εξής:

$$\mathbf{y}_t = c + \varphi_1 \mathbf{y}_{t-1} + \varphi_2 \mathbf{y}_{t-2} + \dots + \varphi_p \mathbf{y}_{t-p} + \varepsilon_t \quad (32)$$

όπου $\varphi_1, \varphi_2, \dots, \varphi_p$ είναι οι συντελεστές αυτοπαλινδρόμησης του μοντέλου AR με υστέρηση $i = 1, \dots, p$, ε_t λευκός θόρυβος και c σταθερά για την οποία ισχύει $c = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$. Χρησιμοποιώντας τον τελεστή ολίσθησης (*backward notation*: $B y_t = y_{t-1}$) η σχέση 32 μετασχηματίζεται στην:

$$(1 - \varphi_1 B_1 - \varphi_2 B_2 - \dots - \varphi_p B_p) \bar{y}_t = \varepsilon_t \quad (33)$$

με $\bar{y}_t = y_t - \mu$. Μπορεί κανείς να δει ένα μοντέλο AR ως μια πολλαπλή παλινδρόμηση μεταξύ των παρατηρήσεων της ίδιας χρονοσειράς. Θα αναφερόμαστε σε ένα αυτοπαλινδρομούμενο μοντέλο ως ένα $AR(p)$ μοντέλο.

2.6.2 Μοντέλα MA

Τα μοντέλα κινητού μέσου όρου θεωρούν γραμμικές σχέσεις ανάμεσα στην παρατήρηση της χρονοσειράς που εξετάζεται και στα σφάλματα που εμφάνισε το μοντέλο σε προηγούμενες περιόδους. Η μαθηματική εξίσωση ενός τέτοιου μοντέλου τάξης q μπορεί να γραφτεί ως εξής:

$$y_t = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (34)$$

όπου θ_i οι συντελεστές μερικής αυτοσυσχέτισης του μοντέλου MA με υστέρηση $i = 1, \dots, q$ και $c = \mu$. Η σχέση 34 δεν αποτελεί μια συμβατική σχέση παλινδρόμησης μιας και τα σφάλματα δεν αποτελούν τμήμα των παρατηρήσεων. Επίσης δεν σχετίζεται με τους κινητούς μέσους όρους που είδαμε ως τεχνική εξομάλυνσης και εξαγωγής της τασης-κύκλου μιας χρονοσειράς. Χρησιμοποιώντας τον τελετή ολίσθησης, η σχέση μετατρέπεται στην εξής:

$$\bar{y}_t = (1 - \theta_1 B_1 - \theta_2 B_2 - \dots - \theta_q B_q) \cdot \varepsilon_t \quad (35)$$

Θα αναφερόμαστε σε ένα μοντέλο κινητών μέσων όρων ως ένα $MA(q)$ μοντέλο. Στην επόμενη παράγραφο παρουσιάζουμε τα μοντέλα ARIMA που αποτελούν σύνθεση των παραπάνω μοντέλων με την προσθήκη ενός μοντέλου διαφορίσης.

2.6.3 Μοντέλα ARIMA

Τα μοντέλα ARIMA συνδυάζουν όπως είπαμε, ένα αυτοπαλινδρομικό μοντέλο τάξης p , ένα μοντέλο κινητών μέσων όρων τάξης q και ένα μοντέλο διαφορίσης τάξης d .

Στον πίνακα 4 φαίνονται ορισμένες ειδικές περιπτώσεις μοντέλων ARIMA. Τα μοντέλα ARIMA μπορούν να γίνουν αρκετά πολύπλοκα, οδηγώντας σε υπερεπροσαρμογή. Θα

| Μοντέλο | Περιγραφή |
|----------------------------|---------------------------|
| ARIMA(0,0,0) χωρίς σταθερά | Λευκός Θόρυβος |
| ARIMA(0,1,0) χωρίς σταθερά | Τυχαίος Περίπατος |
| ARIMA(0,1,0) με σταθερά | Τυχαίος Περίπατος με Τάση |
| ARIMA(p,0,0) | Αυτοπαλινδρόμηση |
| ARIMA(0,0,q) | Κινητοί Μέσοι Όροι |

Πίνακας 4: Ειδικές Περιπτώσεις Μοντέλων ARIMA

δούμε παρακάτω τα κριτήρια επιλογής μοντέλων ARIMA για πρόβλεψη. Το ολοκληρωμένο

αυτοπαλινδρομικό μοντέλο κινητών μέσων όρων, μέσω χρήσης του τελεστή ολίσθησης και των σχέσεων 33 και 35 αναπαρίσταται μέσω της μαθηματικής σχέσης:

$$(1 - \varphi_1 \mathbf{B}_1 - \dots - \varphi_p \mathbf{B}_p)(1 - \mathbf{B})^n(1 - \mathbf{B}^m)^N \mathbf{y}_t = c + (1 - \theta_1 \mathbf{B}_1 - \dots - \theta_p \mathbf{B}_q) \cdot \varepsilon_t \quad (36)$$

Η επιλογή μοντέλου *ARIMA*, ανάγεται στον προσδιορισμό των παραμέτρων p, q των τριών επιμέρους μοντέλων. Η εύρεση των κατάλληλων p και q μπορεί να γίνει αυτόματα μέσω του κριτηρίου πληροφορίας *AIC* (Akaike 1973) στη σχέση 37:

$$\mathbf{AIC} = -2 \log(\mathbf{L}) + 2(p + q + k + 1) \quad (37)$$

όπου η παράμετρος k παίρνει τιμές 0 και 1 όταν η σταθερά c είναι μηδενική και διάφορη του μηδενός αντίστοιχα και \mathbf{L} είναι η προσδοκώμενη πιθανοφάνεια ταύτισης του μοντέλου με τα δεδομένα που εκφράζεται μαθηματικά στην εξίσωση 38:

$$\mathbf{L} = \prod_{t=1}^T \sqrt{\frac{1}{2\pi\sigma_t^2}} \cdot e^{-\sum_{t=1}^T \frac{(\mathbf{X}_t - \mathbf{F}(\mathbf{X}_t))^2}{2\sigma_t^2}} \Leftrightarrow \quad (38)$$

$$-2 \log \mathbf{L} = n \cdot \left[\log(2\pi) + 1 + \log \left(\frac{RSS}{n} \right) \right]$$

όπου $\mathbf{F}(\mathbf{X}_t)$ η προβλεπόμενη τιμή από το μοντέλο, n ο αριθμός των παρατηρήσεων, e^{\dots} το σφάλμα πρόβλεψης, σ^2 η διακύμανση των σφαλμάτων του μοντέλου και *RSS* (*Round Sum of Squares*) το άθροισμα των τετραγωνικών σφαλμάτων πρόβλεψης του μοντέλου.

Για να δοθεί μεγαλύτερο βάρος στην πολυπλοκότητα του μοντέλου *ARIMA* χρησιμοποιείται το "διορθωμένο" κριτήριο πληροφορίας *AIC_c*:

$$\mathbf{AIC}_c = \mathbf{AIC} + 2 \frac{(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2} \quad (39)$$

καθώς και το Μπαεζιανό Κριτήριο Πληροφορίας *BIC* (Schwarz, 1978) που δίνει κι αυτό μεγαλύτερο βάρος στην πολυπλοκότητα του μοντέλου ώστε να αποφευχθεί φαινόμενο υπερ-προσαρμογές (*over-fitting*) εκφράζεται μαθηματικά ως εξής:

$$\mathbf{BIC} = \mathbf{AIC} + (\log T - 2)(p + q + k + 1) \quad (40)$$

Καλά μοντέλα προκύπτουν ελαχιστοποιώντας οποιοδήποτε από τα τρία παραπάνω κριτήρια. Συχνότερη είναι η χρήση του *AIC_c*. Για την επιλογή των παραμέτρων p, q πρέπει να έχει προηγηθεί μέσω κάποιου ελέγχου στασιμότητας, η εύρεση του κατάλληλου βαθμού πρώτων διαφορών d που θα εφαρμοστούν στις αρχικές παρατηρήσεις καθώς και διαφόριση της χρονοσειράς αν δεν είναι στάσιμη. Έλεγχος στασιμότητας και εφαρμογή κατάλληλου μοντέλου διαφόρισης παρέχεται αυτόματα κατά την πρόβλεψη με *ARIMA* στα περισσότερα εργαλεία πρόβλεψης. Η παρουσίαση τεχνικών για έλεγχο στασιμότητας χρονοσειρών (τεστ *ADF*, *KPSS*) ξεφεύγει των ορίων της εργασίας μας. Μια πλήρης ανάλυση για τα μοντέλα *ARIMA* μπορεί κανείς να βρει στην 5η και νεότερη έκδοση του βιβλίου *Time series analysis: Forecasting and control* Tunnicliffe Wilson [2016].

2.7 Μέθοδοι Πρόβλεψης Διακοπτόμενης Ζήτησης

Κατά τη λεγόμενη διακοπτόμενη ζήτηση, παρουσιάζονται θετικές παρατηρήσεις ανάμεσα σε μηδενικά διαστήματα. Οι επικρατούσες τεχνικές πρόβλεψης διακοπτόμενης ζήτησης είναι οι μέθοδοι *Croston*, *SBA* και *ADIDA*. Βασικά στοιχεία τους παρουσιάζονται στις επόμενες παραγράφους.

2.7.1 Μέθοδος Croston

Η ιδέα για πρόβλεψη διακοπτόμενης ζήτησης με Croston (Croston 1972) είναι ο διαχωρισμός της εκάστοτε χρονοσειράς σε δύο επιμέρους χρονοσειρές όπου:

- η μια αποτελείται από τα χρονικά διαστήματα μεταξύ των μη μηδενικών ζητήσεων (intervals)
- η άλλη αποτελείται από τις τιμές των ανεξάρτητων μη μηδενικών ζητήσεων (demands).

Οι δύο χρονοσειρές προεκτείνονται ανεξάρτητα, ενώ η πρόβλεψη μέσω της μεθόδου Croston επιτυγχάνεται μέσω του υπολογισμού του πηλίκου των δύο ανεξάρτητων προβλέψεων, δηλαδή:

$$F_{Croston} = \frac{F_{demands}}{F_{intervals}}$$

2.7.2 Μέθοδος SBA

Οι Syntetos και Boylan το 2001 διαπίστωσαν ότι η μέθοδος Croston συχνά είναι θετικά προκατειλημμένη, δηλαδή παρουσιάζει αισιόδοξη τάση στα αποτελέσματα των προβλέψεων της σε σχέση με τις πραγματικές μελλοντικές τιμές αρκετών χρονοσειρών. Συνέδεσαν επίσης το επίπεδο της αισιόδοξης τάσης της μεθόδου με την τιμή της παραμέτρου εξομάλυνσης α που χρησιμοποιείται για την προέκταση των δύο αποσυντεθειμένων χρονοσειρών. Μέγιστη προκατάληψη παρατηρείται όταν το α λάβει τη μεγαλύτερη τιμή, δηλαδή $\alpha=1$. Η μέθοδος *SBA* αποτελεί μια τροποποίηση της μεθόδου Croston που έχει ως στόχο της εξάλειψη της μεροληπτικής (αισιόδοξης) τάσης της Croston και προκύπτει ως εξής:

$$F_{SBA} = \left(1 - \frac{\alpha}{2}\right) \frac{F_{demands}}{F_{intervals}}$$

2.7.3 Μέθοδος ADIDA

Μια ακόμη προσέγγιση για την αποφυγή της "ασυνέχειας" των παρατηρήσεων, είναι η συνάνθρωση των παρατηρήσεων σε χρονικές περιόδους μικρότερης συχνότητας (για παράδειγμα συνάνθρωση ημερήσιων παρατηρήσεων σε μηνιαίες). Με την επιλογή κατάλληλου επιπέδου συνάνθρωσης μπορεί να μειωθεί έντονα έως να απαλειφθεί το φαινόμενο της διακοπτότητας. Έχοντας μια "συνεχή" χρονοσειρά, μπορεί κανείς πλέον να

εφαρμόσει οποιαδήποτε από τις μεθόδους που προσφέρονται για μη διακοπτόμενες χρονοσειρές. Η πρόβλεψη μετά πρέπει να διασπαστεί κατάλληλα για να προκύψουν οι τιμές των αποσυναθροισμένων χρονικών περιόδων (να επιστρέψουμε λόγω χάρη από μηνιαία σε ημερήσια δεδομένα). Η παραπάνω ιδέα οδήγησε στη δημιουργία της μεθόδου *ADIDA* (*Aggregate-Disaggregate Intermittent Demand Approach*) από τους Nikolopoulos et al το 2010 της οποίας η μεθοδολογία σχηματικά είναι η εξής:

1. Συνάθροιση των δεδομένων με κατάλληλο επίπεδο συνάθροισης
2. Πρόβλεψη της συναθροισμένης χρονοσειράς
3. Διαχωρισμός των προβλέψεων που παρήχθησαν στο προηγούμενο βήμα

Για το διαχωρισμό των προβλέψεων από τη συναθροισμένη προς την αρχική χρονοσειρά, χρησιμοποιούνται ορισμένες εμπειρικές μέθοδοι, που παρουσιάζονται παρακάτω:

- *Equal Weights*: Αναφέρεται στον απλό ισοβαρή διαχωρισμό. Ενδείκνυται για χρονοσειρές με μεγάλη τυχαιότητα και χωρίς εποχιακή συμπεριφορά.
- *Previous Weights*: Αναφέρεται στην εφαρμογή των βαρών που έχουν οι m προηγούμενες παρατηρήσεις, όπου m ισούται με το επίπεδο συνάθροισης.
- *Average Weights*: Αναφέρεται στον υπολογισμό των μέσων βαρών που υπολογίζονται αν χωρίσουμε τις παρατηρήσεις σε k ομάδες m παρατηρήσεων η καθεμία, όπου $k \times m$ ισούται με το σύνολο των διαθέσιμων παρατηρήσεων και m ισούται με το επίπεδο συνάθροισης. Ενδείκνυται σε περιπτώσεις όπου η συνιστώσα της εποχιακότητας είναι έντονη.

Η μη επικαλυπτόμενη συνάθροιση δεδομένων μοιάζει να είναι μια πολλά υποσχόμενη προσέγγιση για χρονοσειρές διακοπτόμενης ζήτησης, καθώς οι προβλέψεις σε υψηλότερα επίπεδα συνάθροισης περιόδων είναι γενικά πιο ακριβείς και έχουν μικρότερες διακυμάνσεις σε σχέση με εκείνες μικρότερων επιπέδων συνάθροισης.

Ατυχείς Προβλέψεις #4

«Η ηχογραφημένη μουσική θα καταστρέψει
όλη τη μουσική δεξιοτεχνία.»

συνθέτης *John Philip Sousa*, 1906

3 Δείκτες Ακριβείας

Κεντρική ανάγκη για την παραγωγή καλών προβλέψεων είναι η δυνατότητα σύγκρισης των διαφόρων μοντέλων σε σχέση με την ακρίβεια ή/και την αμεροληψία τους, τόσο *in-sample* όσο και *out-of-sample* (όταν γίνουν διαθέσιμες οι μελλοντικές παρατηρήσεις).

Στον πίνακα 5 παρουσιάζεται η τυπική διάταξη ιστορικών δεδομένων και μοντέλου συν προβλέψεων μιας χρονοσειράς, καθώς και ο διαχωρισμός σε *in-sample* και *out-sample* διαστήματα. Τα μέτρα ακριβείας εφαρμόζονται τόσο ανάμεσα στις ιστορικές παρατηρήσεις και το μοντέλο (*in-sample*), όσο και ανάμεσα στις προβλέψεις και τις μελλοντικές τιμές όταν γίνουν διαθέσιμες (*out-of-sample*). Σημειώνεται πως η επιλογή του σωστού μέτρου ακριβείας εξαρτάται από την εφαρμογή και συχνά οι κατατάξεις απόδοσης των μοντέλων μπορούν να μεταβάλλονται με την εναλλαγή μέτρου. Στο παρόν κεφάλαιο παρουσιάζουμε ορισμένους στατιστικούς δείκτες αξιολόγησης προβλέψεων.

| Περίοδος | Δεδομένα | Πρόβλεψη | |
|----------|-----------|-----------|----------------------|
| 1 | Y_1 | F_1 | <i>in sample</i> |
| 2 | Y_2 | F_2 | |
| 3 | Y_3 | F_3 | |
| ... | ... | ... | |
| $n-1$ | Y_{n-1} | F_{n-1} | |
| n | Y_n | F_n | |
| $n+1$ | | F_{n+1} | <i>out of sample</i> |
| $n+2$ | | F_{n+2} | |
| ... | | ... | |
| $n+h$ | | F_{n+h} | |

Πίνακας 5: Δεδομένα & Προβλέψεις Χρονοσειράς

3.1 Σφάλματα Εξαρτώμενα από την Κλίμακα

Στην πιο απλή του μορφή, το σφάλμα της πρόβλεψης F_t για μια στιγμή t ορίζεται ως η διαφορά της από την αντίστοιχη παρατήρηση: $e_t = Y_t - F_t$. Τα e_t , $|e_t|$, e_t^2 (η εν γένει e_t^n με $n \in \mathbb{R}^+$) αποτελούν παραλλαγές μέτρησης του σφάλματος πρόβλεψης για χρονική στιγμή t , με κοινό χαρακτηριστικό τους την εξάρτηση από τη στάθμη της εκάστοτε χρονοσειράς. Τα παραπάνω σφάλματα, εκφράζονται στις μονάδες μέτρησης που εκφράζονται και οι παρατηρήσεις (όταν αναφερόμαστε σε προσημασμένη ή απρόσημη διαφορά) ή κάποιο μετασχηματισμό των μονάδων (όταν υψώνουμε τη διαφορά παρατήρησης-μοντέλου σε κάποια δύναμη). Έτσι, τα σφάλματα έχουν φυσική σημασία που προκύπτει από τη μονάδα μέτρησης της εκάστοτε χρονοσειράς.

Από την εξαγωγή μέσω όρων (ή άλλων συνδυασμών) των παραπάνω παραλλαγών για το σφάλμα κάθε χρονικής στιγμής, προκύπτουν τα εξαρτώμενα από την κλίμακα (*scale-dependent*) σφάλματα μοντέλου ή πρόβλεψης ως προς τη χρονοσειρά ενδιαφέροντος. Τα σφάλματα που εξαρτώνται από τη στάθμη της χρονοσειράς, είναι χρήσιμα γιατί εκφράζονται στις μονάδες της εκάστοτε χρονοσειράς (ή έναν μετασχηματισμό τους). Ωστόσο, βασικό τους μειονέκτημα είναι η αδυναμία σύγκρισης μοντέλων ή προβλέψεων για χρονοσειρές με άλλη κλίμακα. Παρακάτω αναλύονται ορισμένα από τα βασικότερα (*scale-dependent*) σφάλματα.

3.1.1 Μέσο Σφάλμα

Το Μέσο Σφάλμα (*Mean Error*) υπολογίζεται από τον απλό προσημασμένο μέσο όρο των σφαλμάτων και εκφράζει ένα μέτρο συστηματικότητας του σφάλματος. Περιγράφεται από τον παρακάτω τύπο:

$$ME = \frac{1}{n} \sum_{t=1}^n (Y_t - F_t) \quad (41)$$

Όσο η τιμή αυτού είναι κοντά στο μηδέν, τόσο τα σφάλματα είναι τυχαία και όχι συστηματικά. Αν ο δείκτης παίρνει θετικές τιμές, δηλώνει συστηματική απαισιοδοξία στις προβλέψεις, ενώ αρνητικές τιμές δηλώνουν συστηματική αισιοδοξία. Συχνά ο δείκτης αναφέρεται και ως *bias* (προκατάληψη/μεροληψία) του μοντέλου.

3.1.2 Μέσο Απόλυτο Σφάλμα

Το Μέσο Απόλυτο Σφάλμα (*Mean Absolute Error*) αποτελεί ένα μέτρο της ακρίβειας της πρόβλεψης επί των πραγματικών τιμών, που διατηρεί τις μονάδες μέτρησης της χρονοσειράς. Εκφράζει τη μέση αστοχία της πρόβλεψης, χωρίς να δίνεται σημασία στην κατεύθυνση της πρόβλεψης. Προκύπτει ως εξής:

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - F_t| \quad (42)$$

Όσο μεγαλύτερη είναι η τιμή του δείκτη, τόσο μικρότερη προκύπτει η ακρίβεια της μεθόδου που εφαρμόστηκε. Η εξάρτηση του *MAE* από την κλίμακα της χρονοσειράς δεν επιτρέπει τη χρησιμοποίησή του, για να συγκριθούν προβλέψεις μεταξύ χρονοσειρών με διαφορετικές στάθμες.

3.1.3 Μέσο Τετραγωνικό Σφάλμα

Το Μέσο Τετραγωνικό Σφάλμα (*Mean Squared Error*), σε σύγκριση με τα *ME* και *MAE*, δίνει αρκετά μεγαλύτερο βάρος στα μεγάλα σφάλματα. Ο υπολογισμός του -όπως προκύπτει από την 43- περιλαμβάνει τον αριθμητικό μέσο των τετραγώνων των διαφορών μεταξύ των προβλέψεων και των δεδομένων. Περιγράφεται από τον παρακάτω τύπο:

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2 \quad (43)$$

Η βελτιστοποίηση ως προς *MSE*, βελτιστοποιεί ταυτόχρονα και την ευκλείδεια απόσταση μοντέλου και χρονοσειράς, λόγω της «1-1» ιδιότητας της $f(x) = \sqrt{x}$. Χρησιμοποιείται ευρέως για τον υπολογισμό των βέλτιστων παραμέτρων εξομάλυνσης και την προσαρμογή μοντέλων πρόβλεψης.

3.1.4 Ρίζα Μέσου Τετραγωνικού Σφάλματος

Η Ρίζα Μέσου Τετραγωνικού Σφάλματος (*Root Mean Squared Error*) υπολογίζεται άμεσα ως \sqrt{MSE} . Αποτελεί την ευκλείδεια απόσταση του μοντέλου από τη χρονοσειρά. Προκύπτει ως εξής:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2} \quad (44)$$

Έχει τις ίδιες ιδιότητες με *MSE*, αλλά είναι εκπεφρασμένο στις πραγματικές μονάδες της χρονοσειράς και κατά συνέπεια να έχουν φυσική σημασία οι τιμές που παίρνει.

3.2 Ποσοστιαία Σφάλματα

Όπως έχει επισημανθεί [1993], σφάλματα που εξαρτώνται από την κλίμακα, δε δίνουν πληροφορία για την αναλογία του σφάλματος προς το επίπεδο της στάθμης των παρατηρήσεων. Για την εξαγωγή αυτής της πληροφορίας, είναι απαραίτητος ο υπολογισμός των σφαλμάτων πρόβλεψης σε ποσοστιαία μορφή. Αυτό είναι ιδιαίτερα χρήσιμο όταν θέλουμε για παράδειγμα να συγκρίνουμε την ακρίβεια μιας μεθόδου πρόβλεψης που έχει εφαρμοσθεί σε περισσότερες από μία χρονοσειρές διαφορετικού επιπέδου μέσης τιμής ή όταν οι πραγματικές τιμές είναι ιδιαίτερα υψηλές. Παρακάτω αναλύονται δύο από τα βασικά ποσοστιαία σφάλματα, τα *MAPE* και *sMAPE*.

3.2.1 Μέσο Απόλυτο Ποσοστιαίο Σφάλμα

Το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (*Mean Absolute Percentage Error*) είναι εκπεφρασμένο σε εκατοστιαίες μονάδες και λαμβάνει τιμές μεγαλύτερες ή ίσες του μηδενός, με τις μικρότερες τιμές να υποδηλώνουν και καλύτερη απόδοση της μεθόδου πρόβλεψης. Αποτελεί δηλαδή μία ποσοστιαία έκφραση του απόλυτου σφάλματος, ως προς τη στάθμη της παρατήρησης. Υπολογίζεται ως εξής:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right| \cdot 100 \quad (45)$$

Μειονέκτημα του *MAPE* αποτελεί το ότι οδηγεί σε απροσδιοριστίες για μηδενικές τιμές και σε ακραίες τιμές για τιμές μικρότερες της μονάδας.

3.2.2 Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα

Το Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (*Symmetric Mean Absolute Percentage Error*) [Scott Armstrong and Schnaars, 1986, Flores, 1986] αποτελεί μια παραλλαγή του μέσου απόλυτου ποσοστιαίου σφάλματος και υπολογίζεται από τον τύπο 46:

$$sMAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{\frac{Y_t + F_t}{2}} \right| \cdot 100 \quad (46)$$

Σε αντίθεση με το μέσο απόλυτο σφάλμα, το απόλυτο του σφάλματος δε διαιρείται απλώς με την πραγματική τιμή, αλλά με το ημίθροισμα της πραγματικής τιμής και της πρόβλεψης. Με αυτή την μετατροπή, ο δείκτης αυτός αποκτά πάνω όριο και πλέον μπορεί να πάρει τιμές στο διάστημα $[0\%, 200\%]$. Με αυτό τον τρόπο, ξεπερνιούνται τα προβλήματα απροσδιοριστίας και ακραίων τιμών που προκύπτουν κατά τη χρήση του *MAPE*.

3.3 Σχετικά Σφάλματα

Τόσο τα απόλυτα όσο και τα ποσοστιαία σφάλματα πρόβλεψης, μπορούν να κανονικοποιηθούν ως προς το σφάλμα μιας μεθόδου αναφοράς. Έτσι, προκύπτουν τα σχετικά σφάλματα που στη γενική περίπτωση έχουν αυτή τη μορφή:

$$RelError = \frac{Error_{fmethod}}{Error_{benchmark}}$$

Σκοπός των σχετικών σφαλμάτων είναι να εξακριβωθεί, αν η συγκρινόμενη μέθοδος, που είναι συνήθως πιο πολύπλοκη από τη μέθοδο *benchmark*, οδηγεί σε βελτίωση της προβλεπτικής ικανότητας. Αν δηλαδή αξίζει να θυσιάσουμε υπολογιστικό κόστος προκειμένου να γίνουμε πιο ακριβείς. Για λόγους απλότητας, και όταν δεν υπάρχει προφανής απάντηση, ως μέθοδος αναφοράς τίθεται η απλοϊκή μέθοδος (*naive*) ή κάποια άλλη που χαρακτηρίζεται από σχετική απλότητα.

3.3.1 Μέσο Απόλυτο Κανονικοποιημένο Σφάλμα

Προκειμένου να αντιμετωπιστούν οι περιπτώσεις απροσδιοριστίας των δεικτών *MAPE* και *sMAPE* αλλά και να δοθεί η ίδια βαρύτητα στα μικρά και τα μεγάλα σφάλματα (σε αντίθεση με τους δείκτες *MSE* και *RMSE*), οι Hyndman and Koehler [2006] πρότειναν το Μέσο Απόλυτο Κανονικοποιημένο Σφάλμα (*Mean Absolute Scaled Error*) για αξιολόγηση της ακριβείας πρόβλεψης. Ο στατιστικός αυτός δείκτης προκύπτει ως εξής:

$$MASE = \frac{\frac{1}{h} \sum_{t=n+1}^{n+h} |Y_t - F_t|}{\frac{1}{n-1} \sum_{t=2}^n |Y_t - Y_{t-1}|} \quad (47)$$

Αποτελεί υποπερίπτωση σχετικού σφάλματος κατά την οποία κανονικοποιούμε με τη μέση τιμή των διαφορών πρώτου βαθμού της χρονοσειράς, ή με άλλα λόγια με τις προβλέψεις της *naive* μεθόδου. Η σημασία του παραπάνω μέτρου έχει επαληθευτεί κι από άλλες ερευνητικές προσπάθειες [Franses, 2016].

3.3.2 Ρίζα Μέσου Τετραγωνικού Κανονικοποιημένου Σφάλματος

Η Ρίζα Μέσου Τετραγωνικού Κανονικοποιημένου Σφάλματος (*Root Mean Squared Scaled Error*) έχει παρόμοια χαρακτηριστικά με τον *MASE*, με βασική διαφορά το αυξημένο βάρος

που δίνεται στα μεγαλύτερα σφάλματα λόγω των τετράγωνων των διαφορών πρόβλεψης και παρατήρησης. Περιγράφεται από τον παρακάτω τύπο:

$$RMSSE = \sqrt{\frac{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2}{\frac{1}{n-1} \sum_{t=2}^n (Y_t - Y_{t-1})^2}} \quad (48)$$

Ο *RMMSE*, όπως και ο *MASE* είναι ένα μέτρο με ορισμένα χρήσιμα χαρακτηριστικά. Πιο συγκεκριμένα, είναι ανεξάρτητο της κλίμακας (ή στάθμης) των δεδομένων, δεν ορίζεται μόνο όταν οι προβλέψεις της *Naive* έχουν μηδενικό σφάλμα, αντιμετωπίζει με συμμετρικό τρόπο τα αρνητικά και θετικά σφάλματα, και τέλος έχει καθορισμένη μέση τιμή και πεπερασμένη απόκλιση.

3.3.3 Theil's U-Statistic

Ο δείκτης *U Theil* [1966], συνδυάζει χαρακτηριστικά σχετικών σφαλμάτων, δίνοντας παράλληλα μεγαλύτερο βάρος σε μεγαλύτερα σφάλματα. Ο στατιστικός αυτός δείκτης επιτρέπει τη συγκριτική συσχέτιση της ακρίβειας των προβλέψεων με τις προβλέψεις της μεθόδου *naive* και ταυτόχρονα τετραγωνίζει τα σφάλματα, ούτως ώστε να δίνεται μεγαλύτερο βάρος στις μεγάλες αποκλίσεις. Προκύπτει από τον τύπο:

$$U = \sqrt{\frac{\sum_{t=2}^n \left| \frac{Y_t - F_t}{Y_{t-1}} \right|^2}{\sum_{t=2}^n \left| \frac{Y_t - Y_{t-1}}{Y_{t-1}} \right|^2}} \quad (49)$$

Τιμές μικρότερες της μονάδας δηλώνουν υπεραπόδοση των προβλέψεων της εξεταζόμενης μεθόδου επί της *naive*, ενώ τιμές μεγαλύτερες της μονάδας δηλώνουν υποαπόδοση.

Η επιλογή του κατάλληλου μέτρου για την αξιολόγηση μοντέλων και προβλέψεων, εξαρτάται τόσο από την εφαρμογή και τα ιδιαίτερα χαρακτηριστικά της/των εκάστοτε χρονοσειρών, όσο και από το στόχο και το περιεχόμενο της αξιολόγησης. Τα σύγχρονα εργαλεία πρόβλεψης, παρέχουν βιβλιοθήκες που περιέχουν τους περισσότερους από τους αναφερθέντες δείκτες αξιολόγησης προβλέψεων. Συνεπώς, χρήσιμη είναι η εξαγωγή των και αντιπαραβολή των παραπάνω δεικτών, λαμβάνοντας από τον καθένα τις ειδικές πληροφορίες που μπορεί να αποτυπώσει για την ακρίβεια των προβλέψεων. Στο πλαίσιο της παρούσας εργασίας, θα είναι συχνή η ανάγκη σύγκρισης ακρίβειας προβλέψεων πάνω σε ομάδες χρονοσειρών με διαφορετικές στάθμες και χαρακτηριστικά. Οι προβλέψεις στην εργασία θα συγκριθούν με διάφορα μέτρα ακρίβειας, αλλά η παρουσίαση των αποτελεσμάτων θα γίνει με τη χρήση του *RMSSE*, που παρουσιάστηκε παραπάνω. Περισσότερα για την αξιολόγηση των προβλέψεων στην εργασία, θα αναλυθούν στην ενότητα των προδιαγραφών της μελέτης περίπτωσης (5.4).

Ατυχείς Προβλέψεις #5

«Οι υπολογιστές στο μέλλον ίσως
ζυγίζουν λιγότερο από 1,5 τόνο.»

περιοδικό *Popular Mechanics*, 1949

4 Ιεραρχική Πρόβλεψη

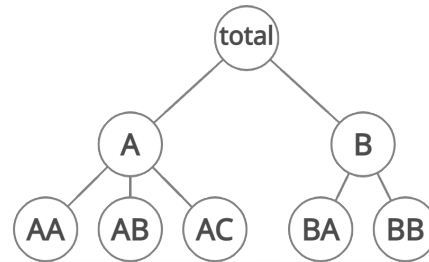
Οι μέθοδοι πρόβλεψης που αναπτύξαμε στο προηγούμενο κεφάλαιο, μπορούν να εφαρμοστούν για την πρόβλεψη της κάθε χρονοσειράς ξεχωριστά. Σε πολλές εφαρμογές, προκύπτουν συλλογές χρονοσειρών που συνδέονται μέσω περιορισμών συνάθροισης, που ονομάζονται *ιεραρχικές χρονοσειρές*. Η ετήσια άφιξη τουριστών στην Ελλάδα, για παράδειγμα, μπορεί να αποσυναθροιστεί ανά περιφέρεια και πύλη εισόδου. Επιπλέον μπορεί να αποσυναθροιστεί ανά είδη και υποκατηγορίες ειδών τουρισμού. Όταν ζητείται η πρόβλεψη τέτοιων συλλογών χρονοσειρών, κατά κανόνα επιθυμούνται προβλέψεις που είναι ευθυγραμμισμένες με τα επίπεδα συνάθροισης. Θέλουμε δηλαδή οι προβλέψεις να υπακούουν στην ιεραρχική δομή των παρατηρήσεων. Η ανεξάρτητη πρόβλεψη των χρονοσειρών με κάποια από τις μεθόδους πρόβλεψης, θα παράξει προβλέψεις που δε θα είναι κατά κανόνα ευθυγραμμισμένες με τη φυσική ιεραρχία που ορίζεται στις παρατηρήσεις. Οι μέθοδοι ιεραρχικής πρόβλεψης, αποτελούν μια συλλογή μεθοδολογιών για την παραγωγή προβλέψεων που υπακούουν στην έμφυτη ιεραρχία των παρατηρήσεων.

4.1 Ιεραρχικές & Ομαδοποιημένες Χρονοσειρές

Θεωρούμε την ιεραρχική δομή του σχήματος 1. Ορίζουμε την πλήρως συναθροισμένη χρονοσειρά "total" ως επίπεδο 0, το πρώτο επίπεδο αποσυναθροίσεως ως επίπεδο 1 και ούτω καθεξής, έως ότου φτάσουμε στο πλέον αποσυναθροισμένο επίπεδο που στην περίπτωση μας είναι το επίπεδο 2 με τις bottom-level χρονοσειρές AA, AB, AC, BA και BB.

Ως \mathbf{b}_t ορίζουμε το διάνυσμα n διαστάσεων που περιέχει όλες τις παρατηρήσεις των bottom-level χρονοσειρών τη χρονική στιγμή t . Έτσι, η συνάθροιση των bottom-level χρονοσειρών μπορεί να προκύψει από τη μητρική αναπαράσταση:

$$\mathbf{y}_t = \mathbf{S} \cdot \mathbf{b}_t \quad (50)$$



Σχήμα 1: Ιεραρχική Χρονοσειρά 2 Επιπέδων

όπου \mathbf{S} είναι ο $m \times n$ πίνακας συνάθροισης, που καθορίζει πως οι bottom-level παρατηρήσεις θα συναθροιστούν στα ανώτερα επίπεδα της ιεραρχίας. Στο παράδειγμα ιεραρχικής χρονοσειράς του σχήματος 1 έχουμε 5 μη συναθροισμένες χρονοσειρές και συνολικά 8 στην ιεραρχία, συνεπώς $n = 5$ και $m = 8$. Ορίζοντας ως y_{X_t} τις παρατηρήσεις της χρονοσειράς X τη χρονική στιγμή t , τα διανύσματα \mathbf{y}_t και \mathbf{b}_t είναι:

$$\mathbf{y}_t = [y_t \ y_{A_t} \ y_{B_t} \ y_{AA_t} \ y_{AB_t} \ y_{AC_t} \ y_{BA_t} \ y_{BB_t}]' \quad (51)$$

$$\mathbf{b}_t = [y_{AA_t} \ y_{AB_t} \ y_{AC_t} \ y_{BA_t} \ y_{BB_t}]' \quad (52)$$

και ο πίνακας \mathbf{S} :

$$\mathbf{S} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ & & \mathbf{1}_5 & & \end{bmatrix} \quad (53)$$

Κάθε περιορισμός συνάθροισης αναπαρίσταται από μια γραμμή του πίνακα \mathbf{S} , ο οποίος μπορεί να διαμεριστεί στα επίπεδα συνάθροισης της ιεραρχίας. Τα επίπεδα 0,1 και 2 της ιεραρχικής δομής του σχήματος 1 απεικονίζονται στην 1η, τις 2-3 και τις 4-8 γραμμές του πίνακα \mathbf{S} στη σχέση 53. Με αυτόν το συμβολισμό μπορούμε να αναπαριστούμε οποιαδήποτε μεγαλύτερη συλλογή χρονοσειρών οργανώνεται σε επίπεδα συνάθροισης.

4.2 Μέθοδοι Ιεραρχικής Πρόβλεψης

Έστω $\hat{\mathbf{y}}_T(h)$ το διάνυσμα των ανεξάρτητων προβλέψεων όλων των χρονοσειρών της ιεραρχίας με ορίζοντα πρόβλεψης h , σε παρατηρήσεις μέχρι τη χρονική στιγμή T και ταξινομημένων όπως στο διάνυσμα \mathbf{y}_t . Όλες οι μέθοδοι που συνδυάζουν γραμμικά τις $\hat{\mathbf{y}}_T(h)$ για να παράξουν τις ευθυγραμμισμένες προβλέψεις $\tilde{\mathbf{y}}_T(h)$, μπορούν να αναπαρασταθούν μέσω της σχέσης:

$$\tilde{\mathbf{y}}_T(h) = \mathbf{S}\mathbf{P}\hat{\mathbf{y}}_T(h) \quad (54)$$

όπου \mathbf{S} είναι ο πίνακας συνάθροισης της εξίσωσης 50 και ο \mathbf{P} ένας $n \times m$ πίνακας, η μορφή του οποίου αντανακλά την προσέγγιση ιεραρχικής πρόβλεψης που εφαρμόζουμε για την παραγωγή ευθυγραμμισμένων προβλέψεων. Αυτή η σημειογραφία προτάθηκε το 2007 από τους Hyndman et al και όπως θα δούμε παρακάτω, μπορεί να εκφράσει όλες τις μεθόδους γραμμικού συνδυασμού των ανεξάρτητων προβλέψεων σε μητρική μορφή, με τις single-level προσεγγίσεις να αποτελούν ειδικές περιπτώσεις της εξίσωσης 54.

4.2.1 Bottom-Up Μέθοδος

Διαχρονικά, η πλέον δημοφιλής προσέγγιση ιεραρχικής πρόβλεψης είναι η bottom-up μέθοδος. Αποτελεί μια απλή μέθοδο που σε αρκετές περιπτώσεις υπεραποδίδει πιο σύνθετων προσεγγίσεων [Makridakis et al., 2020c]. Στην bottom-up μοντελοποιούνται οι μη συναθροισμένες χρονοσειρές και Για να την εκφράσουμε μέσω της εξίσωσης 54 γράφουμε:

$$\mathbf{P} = [\mathbf{0}_{m \times n} \mid \mathbf{1}_{n \times n}] \quad (55)$$

όπου ο $\mathbf{0}_{m \times n}$ είναι ο $m \times n$ μηδενικός πίνακας. Ο πίνακας \mathbf{P} της εξίσωσης 55 απομονώνει τις προβλέψεις των bottom-level χρονοσειρών, οι οποίες στη συνέχεια συναθροίζονται από τον πίνακα \mathbf{S} για να παραχθούν οι ευθυγραμμισμένες προβλέψεις όλης της ιεραρχίας. Η μοντελοποίηση των μη συναθροισμένων χρονοσειρών διατηρεί πληροφορία που χάνεται λόγω συνάθροισης σε χρονοσειρές ανώτερων επιπέδων. Ωστόσο, ο θόρυβος που συνήθως περιέχουν οι bottom-level χρονοσειρές καθιστά δύσκολη τη μοντελοποίησή τους.

4.2.2 Top-Down Μέθοδοι

Ακόμα μια ευρέως εφαρμοζόμενη προσέγγιση ιεραρχικής πρόβλεψης είναι οι top-down μέθοδοι. Στην top-down πρόβλεψη, προβλέπουμε μόνο τη χρονοσειρά "total" του επιπέδου 0 και στη συνέχεια αποσυναθροίζουμε την πρόβλεψη στα χαμηλότερα επίπεδα, βάσει ποσοστών επί των παρατηρήσεων ή των προβλέψεων. Για να εκφράσουμε την top-down προσέγγιση μέσω της 54 γράφουμε τον \mathbf{P} ως:

$$\mathbf{P} = [\mathbf{p} \mid \mathbf{0}_{n \times (m-1)}] \quad (56)$$

όπου $\mathbf{p} = [p_1, p_2, \dots, p_n]$ είναι το διάνυσμα με τα ποσοστά των μη συναθροισμένων χρονοσειρών. Συνεπώς, εδώ ο πίνακας \mathbf{P} κατανέμει τις προβλέψεις της χρονοσειράς "total" στις bottom-level χρονοσειρές σύμφωνα με το \mathbf{p} . Η επιλογή του διανύσματος \mathbf{p} μπορεί να γίνει με τις παρακάτω παραλλαγές.

Ιστορικά Ποσοστά Αυτή η top-down προσέγγιση κατανέμει την πρόβλεψη του επιπέδου 0 βάσει των ιστορικών ποσοστών που έχουν οι παρατηρήσεις στις χρονοσειρές.

Μέσα Ιστορικά Ποσοστά Σε αυτή την παραλλαγή της εξαγωγής ιστορικών ποσοστών, κάθε ποσοστό p_i του διανύσματος \mathbf{p} που ορίσαμε στην 56 αναπαριστά τον αριθμητικό μέσο των ιστορικών ποσοστών της μη συναθροισμένης χρονοσειράς \mathbf{b}_{i_t} προς την πλήρως συναθροισμένη \mathbf{y}_t για τις n χρονικές περιόδους των παρατηρήσεων. Δηλαδή:

$$p_i = \frac{1}{n} \sum_{t=1}^n \frac{\mathbf{b}_{i_t}}{\mathbf{y}_t} \quad i = 1, 2, \dots, n \quad (57)$$

Το διάνυσμα \mathbf{p} λέμε ότι αναπαριστά τα μέσα ιστορικά ποσοστά (*average historical proportions*) των χρονοσειρών στην ιεραρχία.

Ποσοστά Ιστορικών Μέσων Εδώ, κάθε p_i εκφράζει το ποσοστό της μέσης ιστορικής στάθμης-αξίας της μη συναθροισμένης χρονοσειράς \mathbf{b}_{i_t} προς την στάθμη της πλήρως συναθροισμένης \mathbf{y}_t για τις n χρονικές περιόδους των παρατηρήσεων. Δηλαδή:

$$p_i = \sum_{t=1}^n \frac{\mathbf{b}_{i_t}}{n} / \sum_{t=1}^n \frac{\mathbf{y}_t}{n} \quad i = 1, 2, \dots, n \quad (58)$$

Το διάνυσμα \mathbf{p} θα αναπαριστά τα ποσοστά των ιστορικών μέσων (*proportions of historical averages*) των χρονοσειρών στην ιεραρχία. Σημειώνουμε πως το διάνυσμα \mathbf{p} και στις 2 προσεγγίσεις περιέχει τα ποσοστά των μη συναθροισμένων χρονοσειρών, τα οποία πολλαπλασιαζόμενα με τον πίνακα \mathbf{S} παράγουν τα ποσοστά όλων των επιπέδων συνάθροισης.

Ποσοστά Επί των Προβλέψεων Σε αυτή την προσέγγιση top-down πρόβλεψης που προτάθηκε από τους Athanasopoulos et al. [2009], προβλέπονται ανεξάρτητα όλες οι χρονοσειρές στην ιεραρχία και στη συνέχεια εξάγουμε τα ποσοστά των προβλέψεων που έχουν οι χρονοσειρές απόγονοι ως προς τις χρονοσειρές γονείς. Ορίζουμε ως $\hat{\mathbf{y}}_{j,n}^{(l)}(h)$ την πρόβλεψη με ορίζοντα h της χρονοσειράς που βρίσκεται l επίπεδα πάνω από τη χρονοσειρά i , και $\hat{\mathbf{s}}_{i,n}^{(l+1)}(h)$ το άθροισμα των προβλέψεων όσων χρονοσειρών είναι άμεσοι απόγονοι του i . Τότε κάθε ποσοστό p_i μπορεί να γραφεί ως:

$$p_i = \prod_{l=0}^{K-1} \frac{\hat{\mathbf{y}}_{i,n}^{(l)}(h)}{\hat{\mathbf{s}}_{i,n}^{(l+1)}(h)} \quad i = 1, 2, \dots, n \quad (59)$$

με K να είναι το πλήθος των επιπέδων στην ιεραρχία. Εναλλακτικές μέθοδοι ιεραρχικής πρόβλεψης, όπως middle-out προσεγγίσεις μπορούν να προκύψουν με επιλογή κατάλληλου πίνακα \mathbf{P} .

4.2.3 Μέθοδοι Γραμμικού Συνδυασμού

Σε αντίθεση με τις παραπάνω προσεγγίσεις, μέθοδοι γραμμικού συνδυασμού των ανεξάρτητων προβλέψεων, όπως η οικογένεια μεθόδων που αναπτύχθηκε αρχικά από τους Hyndman, Ahmed, Athanasopoulos, and Shang [2011] και επαναδιατυπώθηκε από τους Wickramasuriya, Athanasopoulos, and Hyndman [2019], προβλέπει ανεξάρτητα όλες τις χρονοσειρές στην ιεραρχία και προσαρμόζει *ad hoc* τις προβλέψεις με τους περιορισμούς συνάθροισης, μέσω ενός μοντέλου παλινδρόμησης.

Υποθέτοντας αμερόληπτες ανεξάρτητες προβλέψεις $\hat{\mathbf{y}}_t(h)$, έδειξαν πως οι εναρμονισμένες προβλέψεις $\tilde{\mathbf{y}}_t(h)$ παραμένουν αμερόληπτες αν ο πίνακας \mathbf{P} ικανοποιεί τη σχέση:

$$\mathbf{SPS} = \mathbf{S} \Leftrightarrow \mathbf{PS} = \mathbf{1}_n \quad (60)$$

Για κάθε τέτοιο \mathbf{P} , έδειξαν πως ο πίνακας διακύμανσης των σφαλμάτων των εναρμονισμένων προβλέψεων $\tilde{\mathbf{y}}_t(h)$ αναπαρίσταται μέσω της:

$$\text{Var}[\mathbf{y}_{t+h} - \tilde{\mathbf{y}}_t(h)] = \mathbf{SPW}_h \mathbf{P}' \mathbf{S}' \quad (61)$$

όπου \mathbf{W}_h είναι ο πίνακας διακύμανσης - συνδιακύμανσης των ανεξάρτητων προβλέψεων η εκτίμηση του οποίου απαιτείται. Ο πίνακας \mathbf{P} που ικανοποιεί την 60 και ελαχιστοποιεί το $\text{tr}(\mathbf{SPW}_h \mathbf{P}' \mathbf{S}')$ της εξίσωσης 61, που αποτελεί το ζητούμενο για βέλτιστες αμερόληπτες προβλέψεις, έδειξαν ότι είναι ο:

$$\mathbf{P} = (\mathbf{S}' \mathbf{W}_h^\dagger \mathbf{S})^{-1} \mathbf{S}' \mathbf{W}_h^\dagger \quad (62)$$

με \mathbf{W}_h^\dagger τον γενικευμένο αντίστροφο του \mathbf{W}_h . Άρα, μέσω των σχέσεων 54 και 62 οι βέλτιστες αμερόληπτες εναρμονισμένες προβλέψεις $\tilde{\mathbf{y}}_t(h)$ θα είναι:

$$\tilde{\mathbf{y}}_t(h) = \mathbf{S}(\mathbf{S}' \mathbf{W}_h^\dagger \mathbf{S})^{-1} \mathbf{S}' \mathbf{W}_h^\dagger \hat{\mathbf{y}}_n(h) \quad (63)$$

Η μέθοδος αυτή καλείται *Minimum Trace* εναρμόνιση. Για την εφαρμογή της απαιτείται η εκτίμηση του πίνακα \mathbf{W}_h που υλοποιείται με τους παρακάτω εκτιμητές:

1. *OLS* Για τον εκτιμητή απλών ελαχίστων τετραγώνων θέτουμε:

$$\mathbf{W}_h = k_h \mathbf{1}, \forall h, \text{ και } k_h > 0 \quad (64)$$

Αποτελεί την πιο απλή προσέγγιση, αγνοώντας τόσο την κλίμακα όσο και τις σχέσεις μεταξύ των χρονοσειρών. Είναι βέλτιστος αν οι ανεξάρτητες προβλέψεις είναι ασυσχέτιστες και παρουσιάζουν σταθερή διακύμανση k_h , υποθέσεις που συναντάμε αρκετά σπάνια στις εφαρμογές ιεραρχικών χρονοσειρών.

2. *WLS_v* Για τον εκτιμητή σταθμισμένων ελαχίστων τετραγώνων, με βάρη που προκύπτουν από τη διακύμανση των σειρών (*WLS* με *variance scaling*) θέτουμε:

$$\mathbf{W}_h = k_h \cdot \text{diag}(\hat{\mathbf{W}}_1) \quad \forall h, \text{ με } k_h > 0 \quad (65)$$

όπου ο

$$\mathbf{W}_1 = \frac{1}{T} \sum_{t=1}^T [y_{t+1} - \hat{\mathbf{y}}_t(1)] \cdot [\mathbf{y}_{t+1} - \hat{\mathbf{y}}_t(1)]' \quad (66)$$

είναι ο αμερόληπτος εκτιμητής των in-sample σφαλμάτων και των σφαλμάτων των ανεξάρτητων προβλέψεων για $h = 1$. Αυτή η προσέγγιση λαμβάνει υπ' όψη την κλίμακα των χρονοσειρών στην ιεραρχία αλλά αγνοεί τις μεταξύ τους σχέσεις.

3. *WLS_s* Για τον εκτιμητή σταθμισμένων ελαχίστων τετραγώνων που προτάθηκε από τους Athanasopoulos et al για χρονικές ιεραρχίες, θέτουμε:

$$\mathbf{W}_h = k_h \cdot \text{diag}(\mathbf{S}_{m \times n} \mathbf{1}_{n \times 1}) \quad \forall h, \text{ με } k_h > 0 \quad (67)$$

Αυτή η προσέγγιση υποθέτει σταθερή διακύμανση k_h για τα σφάλματα στις bottom-level χρονοσειρές και συσχέτιση μόνο μεταξύ των κόμβων με σχέσεις προγόνων - απογόνων. Έτσι, λέμε ότι υλοποιεί δομική κλιμάκωση (*structural scaling*). Και αυτή η προσέγγιση δεν λαμβάνει υπ' όψη τις σχέσεις ανάμεσα στις χρονοσειρές της ιεραρχίας.

4. *MinT_{shrink}* Αυτός ο εκτιμητής συρρικνώνει τα μη διαγώνια στοιχεία του $\hat{\mathbf{W}}_1$, υποθέτοντας πάλι σταθερή διακύμανση σφαλμάτων k_h , εκτιμώντας το \mathbf{W}_h μέσω της:

$$\mathbf{W}_h = k_h \cdot [\lambda_D \cdot \text{diag}(\hat{\mathbf{W}}_1) + (1 - \lambda_D) \hat{\mathbf{W}}_1] \quad (68)$$

όπου ο λ_D είναι η παράμετρος έντασης της συρρίκνωσης και δίνεται από τη σχέση:

$$\lambda_D = \sum_{i \neq j} \hat{\text{Var}}(\hat{r}_{ij}) / \sum_{i \neq j} \hat{r}_{ij}^2 \quad (69)$$

με το \hat{r}_{ij} να εκφράζει το ij -οστό στοιχείο του πίνακα συσχέτισης των in-sample παρατηρήσεων και για $h = 1$. Βασικά πλεονεκτήματα της μεθόδου είναι ότι λαμβάνει υπ' όψη τις σχέσεις μεταξύ των χρονοσειρών στην ιεραρχία και ότι οι εκτιμήσεις που επιστρέφει μπορούν να χρησιμοποιηθούν για κατασκευή διαστημάτων εμπιστοσύνης της πρόβλεψης.

Ατυχείς Προβλέψεις #6

«Το άλογο ήρθε για να μείνει, το αυτοκίνητο
είναι ένα μιχλιμπίδι, μια φαντασιοπληξία.»

πρόεδρος *Michigan Savings Bank*, 1903
προσπαθώντας να πείσει το δικηγόρο του *Henry
Ford* να μην επενδύσει στη *Ford Motor Company*.

5 Μελέτη Περίπτωσης

Όπως είδαμε, η παραγωγή εναρμονισμένων προβλέψεων μέσω κάποιας μεθόδου γραμμικού συνδυασμού (όπως η *Minimum Trace* συνάνθροιση) μπορεί υπό συνθήκες να επιδρά θετικά στην ακρίβεια των προβλέψεων μιας ομαδοποιημένης χρονοσειράς σε σχέση με τις ανεξάρτητες προβλέψεις των επιπέδων στις ιεραρχίες της. Θυμίζουμε πως βασικός στόχος των μεθόδων ιεραρχικής πρόβλεψης είναι η παραγωγή προβλέψεων που υπαχούν στην ιεραρχική δομή των παρατηρήσεων. Ωστόσο, κατά την εφαρμογή μεθόδων γραμμικού συνδυασμού, η μεταφορά πληροφορίας μεταξύ των επιπέδων συνάνθροισης ενισχύει με εξωγενή πληροφορία την ανεξάρτητη πρόβλεψη της κάθε χρονοσειράς. Αυτή η εξωγενής πληροφορία που θα προκύπτει από ευθυγράμμιση της πρόβλεψης μιας χρονοσειράς με διαφορετικές ιεραρχίες, καθώς και ο τρόπος που οι ευθυγραμμίσεις με διαφορετικές ιεραρχίες θα επιδρούν στην ακρίβεια αλλά και την αμεροληψία των προβλέψεων θα αποτελέσει κομμάτι της παρούσας εργασίας και του πειράματος που θα εκτελέσουμε.

Οι ομαδοποιημένες χρονοσειρές, δηλαδή οι συλλογές χρονοσειρών που οργανώνονται σε μία ή περισσότερες ιεραρχικές δομές συνάνθροισης, ποικίλλουν σε μέγεθος ανάλογα με την εφαρμογή. Οι τριμηνιαίες αφίξεις τουριστών στην Ελλάδα για παράδειγμα μπορούν να αποσυναθροιστούν ανά περιφέρεια/πόλη/πύλη εισόδου (1η ιεραρχία) καθώς και σε κατηγορίες/υποκατηγορίες τουρισμού/πύλες εισόδου (2η ιεραρχία). Τα επίπεδα συνάνθροισης που μπορούν να προκύψουν, θα είναι όλοι οι δυνατοί συνδυασμοί ζευγών επιπέδων, δηλαδή το εξωτερικό γινόμενο των επιπέδων των δύο ιεραρχιών της συλλογής. Παραδείγματα συνδυασμών επιπέδων μπορούν να είναι: τριμηνιαίες αφίξεις στην περιφέρεια Θεσσαλονίκης για τουρισμό υγείας, τρίμηνες αφίξεις πανελλήνια για ιαματικό τουρισμό (υποκατηγορία τουρισμού υγείας) κλπ. Από συλλογές σαν κι αυτή που περιέχουν εκατοντάδες μη συναθροισμένες χρονοσειρές, προκύπτουν ιεραρχίες της τάξης των εκατοντάδων έως λίγων χιλιάδων χρονοσειρών. Η ανεξάρτητη πρόβλεψη όλων των χρονοσειρών σε μια τέτοια ομαδοποιημένη χρονοσειρά παραμένει υπολογιστικά προσβάσιμη.

Αντίθετα, για μια επιχείρηση που δραστηριοποιείται στις λιανικές πωλήσεις (*retail*) και μπορεί να εμπορεύεται δεκάδες χιλιάδες προϊόντα, οργανωμένα σε πολλαπλές ιεραρχίες που αφορούν αποσυναθροίσεις των πωλήσεων τόσο ανά κατηγορίες και υποκατηγορίες προϊόντων όσο και ανά γεωγραφικά τμήματα και σημεία πώλησης, η συλλογή ομαδικών χρονοσειρών που θα περιλαμβάνει όλα τα επίπεδα συνάνθροισης θα είναι της τάξης των εκατομμυρίων χρονοσειρών. Αντίστοιχα μεγέθη θα έχουν και ομαδοποιημένες χρονοσειρές οικονομικών δεικτών ή κλιματικών μεγεθών.

Για την παραγωγή εναρμονισμένων -με κάποια ιεραρχία- προβλέψεων μέσω των μεθόδων γραμμικού συνδυασμού απαιτείται η ανεξάρτητη πρόβλεψη κάθε χρονοσειράς που περιέχεται στην ιεραρχία. Όπως είδαμε στην παραπάνω παράγραφο, υπάρχουν συλλογές με εκατομμύρια χρονοσειρές που προκύπτουν σε εφαρμογές πρόβλεψης. Η ανεξάρτητη πρόβλεψη όλων αυτών των χρονοσειρών είναι συχνά υπολογιστικά ασύμφορη, ιδιαίτερα όταν ο στόχος της εξαγωγής πληροφορίας από την ιεραρχία των παρατηρήσεων δεν είναι η παραγωγή εναρμονισμένων προβλέψεων σε κάθε πιθανό επίπεδο συνάνθροισης, αλλά η ενίσχυση τμήματος των ανεξάρτητων προβλέψεων. Για παράδειγμα, αν ζητηθεί η πρόβλεψη των πωλήσεων δύο καταστημάτων μιας επιχείρησης με πολλά καταστήματα, η εξαντλητική

συμπερίληψη των χρονοσειρών στα διάφορα επίπεδα της ιεραρχίας θα απαιτήσει μεγάλο όγκο υπολογιστικών πόρων χωρίς εγγύηση πως κάθε επίπεδο συνάθροισης θα συνεισφέρει ουσιαστικά στη βελτίωση των προβλέψεων του εξεταζόμενου συνόλου χρονοσειρών.

5.1 Στόχοι

Μέχρι στιγμής δεν έχει προταθεί κάποιος αλγόριθμος για εύρεση ιεράρχησης δεδομένων με τη βέλτιστη προβλεπτική ικανότητα. Η συνεισφορά κάθε επιπέδου συνάθροισης τόσο στη βελτίωση των προβλέψεων όσο και στην αύξηση των απαιτούμενων πόρων δεν είναι εκ των προτέρων γνωστή και εξαρτάται σε μεγάλο βαθμό από τις ιδιαιτερότητες των δεδομένων της κάθε εφαρμογής. Ερωτήματα όπως: ποια επίπεδα "αξίζει" να συμπεριληφθούν στην ιεραρχία, με ποια σειρά, ποιο θα είναι το βέλτιστο ύψος του ιεραρχικού δέντρου ή ποια είναι η σχέση βελτίωσης στην ακρίβεια και κόστους πρόβλεψης συνήθως επιλύονται με *ad hoc* τεχνικές που δεν επιδέχονται γενίκευσης πέρα από δεδομένα παρόμοιας μορφής και πεδίου αναφοράς.

Πάνω σε αυτό το κενό και σε πιθανές προτάσεις εύρεσης των κατάλληλων ιεραρχιών για ενίσχυση της πρόβλεψης πωλήσεων θα επικεντρωθεί η προσπάθεια στο υπόλοιπο της εργασίας. Θεωρώντας ως σταθερές της μελέτης τόσο τη μοντελοποίηση και πρόβλεψη βάσης των χρονοσειρών, όσο και τη μέθοδο εναρμόνισης με τη εκάστοτε ιεραρχία, στόχος είναι η ανάδειξη της πληροφορίας που μπορεί να εξαχθεί από τη συμπερίληψη των γραμμικών περιορισμών που ορίζονται από την ιεραρχική δομή συνάθροισης των παρατηρήσεων. Στόχος του πειράματος που θα διεξαχθεί, είναι να απαντηθεί το ερώτημα: «Δεδομένης μιας ομαδοποιημένης χρονοσειράς, ποια είναι η βέλτιστη σε απόδοση³ ιεραρχία, με την οποία μπορούμε να εναρμονίσουμε τις ανεξάρτητες προβλέψεις που μας ενδιαφέρουν;» Με τον όρο απόδοση συμπεριλαμβάνεται η ακρίβεια αλλά και το κόστος πρόβλεψης. Η μελέτη περίπτωσης θα κινηθεί πάνω σε δύο βασικούς άξονες:

- Θα συγκριθούν προβλέψεις που έχουν ευθυγραμμιστεί με 48 διαφορετικές ιεραρχίες και προβλέπουν 12 επίπεδα συνάθροισης της συλλογής 42.840 χρονοσειρών πωλήσεων του συνόλου δεδομένων που τέθηκε υπό πρόβλεψη στο διαγωνισμό *M5*. Στη συνέχεια θα εξαχθούν μετρήσεις για την απόδοση των ιεραρχιών και θα παρουσιαστούν τα αποτελέσματα.
- Μέσω της αυτοματοποίησης της παραπάνω διαδικασίας σε κώδικα, θα σχημαγραφηθεί μια μεθοδολογία που θα επιλέγει τη βέλτιστη -σε συνδυασμό ακρίβειας και πόρων πρόβλεψης- ιεραρχία για ευθυγράμμιση των ενδιαφερόμενων προβλέψεων.

Πριν παρουσιαστεί η μεθοδολογία και τα αποτελέσματά της, στην επόμενη παράγραφο (5.2) γίνεται μια σύντομη επισκόπηση του συνόλου δεδομένων που θα εξετάσουμε στο πείραμά μας.

³Ως βέλτιστη σε απόδοση εννοείται σε ακρίβεια και κόστος πρόβλεψης, πάνω σε ένα ή περισσότερα εξεταζόμενα επίπεδα συνάθροισης

5.2 Σύνολο Δεδομένων

Όσο τα δεδομένα αυξάνονται σε όγκο και τα διαφορετικά επίπεδα συνάθροισης πληθαίνουν σε χρονοσειρές, ο κατάλληλος συνδυασμός επιπέδων για την παραγωγή μιας ιεραρχίας με βάση την οποία θα εναρμονιστεί η πρόβλεψη βάσης μιας ή περισσότερων χρονοσειρών, με στόχο τη βελτίωση της ακρίβειας των προβλέψεων και ταυτόχρονα την αποφυγή περιττού κόστους πρόβλεψης (όπως παράλειψη επιπέδων με πολλές χρονοσειρές που δε συνεισφέρουν στην ακρίβεια της πρόβλεψης) δεν είναι προφανής.

Ορισμένοι τομείς στους οποίους προκύπτει η ανάγκη για προβλέψεις που εναρμονίζονται με κάποια ιεραρχία επιπέδων συνάθροισης, αφορούν μικρότερες συλλογές χρονοσειρών και κατά συνέπεια η συμπερίληψη όλων των επιπέδων συνάθροισης είναι υπολογιστικά προσβάσιμη. Ωστόσο, η επιλογή κατάλληλης ιεραρχίας γίνεται κρισιμότερη, όταν οι χρονοσειρές που περιέχονται σε μια συλλογή ιεραρχικών ή ομαδοποιημένων χρονοσειρών ξεπερνούν τις λίγες χιλιάδες, καθιστώντας το κόστος της πρόβλεψης απαιτητικότερο. Χαρακτηριστικό παράδειγμα εφαρμογής πρόβλεψης στην οποία εμφανίζονται πολυπληθείς συλλογές ομαδοποιημένων χρονοσειρών είναι οι πωλήσεις λιανικής. Σε τέτοιες εφαρμογές, από τον αποφασίζοντα διατίθενται συλλογές ομαδοποιημένων χρονοσειρών που αφορούν πωλήσεις χιλιάδων μονάδων που οργανώνονται σε πολλαπλές ιεραρχίες, τόσο με γεωγραφικά κριτήρια όσο και ανά κατηγορίες και υποκατηγορίες προϊόντων.

Για να μπορέσει να μελετηθεί σε βάθος το παραπάνω φαινόμενο, καθώς και να εξαχθούν ακριβή και ευσταθή αποτελέσματα από τη σύγκριση προβλέψεων που έχουν εναρμονιστεί με διαφορετικές ιεραρχίες, χρειάζεται η αξιοποίηση ενός συνόλου δεδομένων που θα είναι επαρκώς μεγάλο και οργανωμένο σε επίπεδα συνάθροισης για να αποτελέσει τη μελέτη περίπτωσης της εργασίας. Ένα τέτοιο σύνολο δεδομένων τέθηκε δημόσια υπό πρόβλεψη στο διαγωνισμό $M5^4$ και περιλαμβάνει δεδομένα πωλήσεων της Walmart σε μια μεγάλη γεωγραφική περιφέρεια που οργανώνονται σε 12 επίπεδα γεωγραφικής και ανά προϊόν συνάθροισης. Στο υπόλοιπο της παρούσας παραγράφου, παρουσιάζονται τα μεταδεδομένα του συνόλου δεδομένων και διεξάγεται μια σύντομη επεξηγηματική ανάλυση των δεδομένων της μελέτης περίπτωσης.

Τα δεδομένα προέρχονται από το διαγωνισμό $M5$ στον οποίο τέθηκαν δημόσια υπό πρόβλεψη. Το διαγωνισμό φιλοξένησε η δημοφιλής πλατφόρμα [kaggle](#) ενώ προσφέρθηκαν 50.000\$, ως χρηματικό έπαθλο για τις ακριβέστερες υποβολές προβλέψεων. Συγ-

⁴Οι διαγωνισμοί M που οργανώνονται από τον [Σπύρο Μακρινιάκη](#), αποτελούν μια σειρά ανοιχτών διαγωνισμών για αξιολόγηση και σύγκριση της ακρίβειας των διαφορετικών μεθόδων πρόβλεψης χρονοσειρών. Σκοπεύουν στην βαθύτερη κατανόηση και εξέλιξη της μεθοδολογίας της πρόβλεψης μέσω της σύγκρισης της απόδοσης διαφορετικών μεθόδων πάνω σε καλώς ορισμένα προβλήματα του πραγματικού κόσμου. Οι [Makridakis, Hibon, and Moser \[1979\]](#) έκαναν μία πρώτη προσπάθεια να εκτιμήσουν την ακρίβεια των μεθόδων χρησιμοποιώντας 111 χρονοσειρές μακροοικονομικών δεδομένων. Από εκείνη την πρώτη προσπάθεια έχουν μεσολαβήσει πέντε διαγωνισμοί [1979, 1993, 2000, 2018b, 2020c] με συνεχώς αυξανόμενο ενδιαφέρον τόσο ανάμεσα στην ερευνητική κοινότητα όσο και σε όσους εφαρμόζουν στην πράξη τεχνικές προβλέψεων. Στον πρώτο διαγωνισμό συμμετείχαν μόνο οι [Makridakis](#) και [Hibon](#) ενώ στον πέμπτο διαγωνισμό συμμετείχαν 5507 ομάδες με συνολικά 7092 συμμετέχοντες από 101 χώρες. Περισσότερα στοιχεία για τους διαγωνισμούς M , τα συμπεράσματα που προέκυψαν από τα αποτελέσματά τους, καθώς και την επίδραση που έχουν στον τομέα της πρόβλεψης θα αναλυθούν στην παράγραφο [Α](#) του παραρτήματος.

κέντρωσε μεγάλο ενδιαφέρον, ενώ αποτέλεσε τον πρώτο διαγωνισμό M με την πλέον ξεκάθαρη υπεροχή των μεθόδων μηχανικής μάθησης έναντι των στατιστικών μεθόδων.

Παραχωρήθηκαν από τη Walmart, εταιρεία που δραστηριοποιείται εντός και εκτός ΗΠΑ στις λιανικές πωλήσεις (retail) και αποτελεί την εταιρεία με τα

υψηλότερα έσοδα παγκοσμίως. Αφορούν πωλήσεις μονάδων (unit sales) διαφόρων προϊόντων σε καταστήματα των ΗΠΑ, τα οποία οργανώνονται σε επίπεδα συνάθροισης με βάση τις κατηγορίες των προϊόντων και γεωγραφικά κριτήρια. Πιο συγκεκριμένα, το σύνολο δεδομένων περιλαμβάνει πωλήσεις που οργανώνονται σε:

- 3 κατηγορίες (Hobbies, Foods, Household)
- 7 υποκατηγορίες - τμήματα (departments) προϊόντων (Foods 1, ..., Foods 3, Hobbies 1, Hobbies 2, Household 1, Household 2).

Τα προϊόντα των παραπάνω κατηγοριών, οργανώνονται γεωγραφικά σε:

- 3 πολιτείες των ΗΠΑ (California, Texas, Wisconsin)
- 10 καταστήματα (Cal 1, ..., Cal 4, Tex 1, ..., Tex 3, Wis 1, ..., Wis 3).

Η παραπάνω οργάνωση των προϊόντων σε μια γεωγραφική ιεραρχία και μια ιεραρχία κατηγοριών προϊόντων, επιτρέπει την αναφορά σε 12 επίπεδα συναθροίσεως.

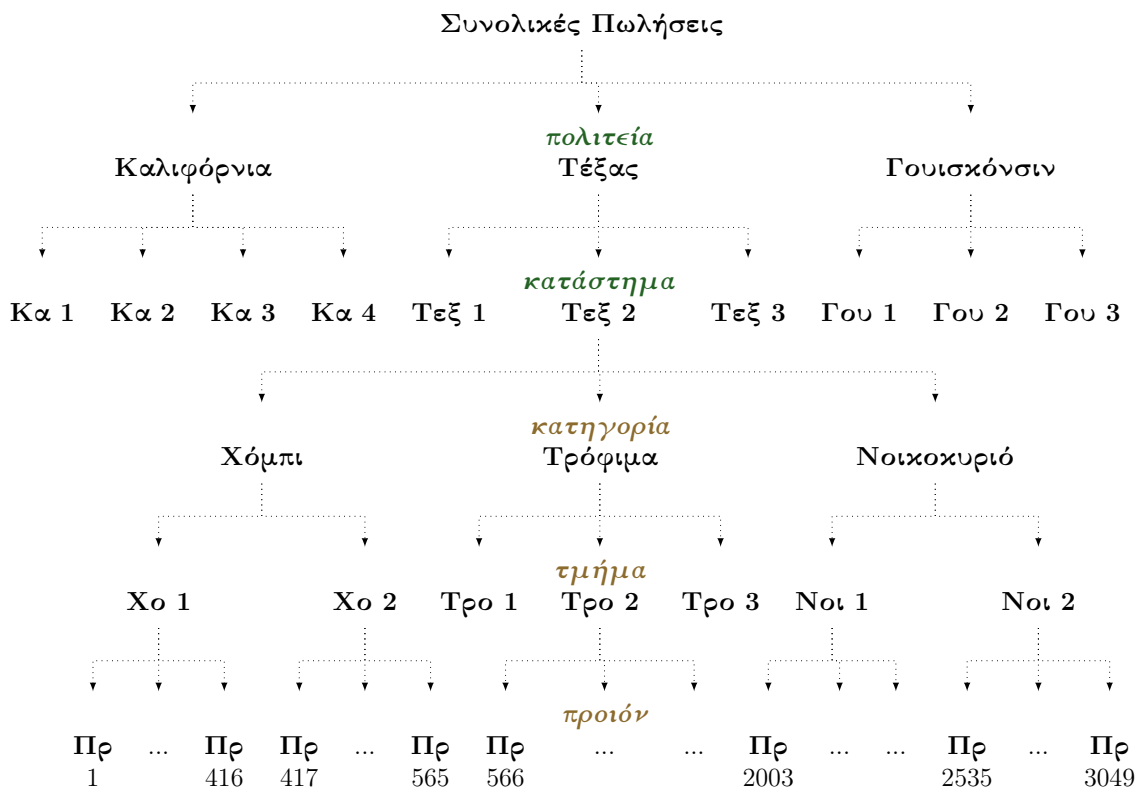
| Μονάδα Μέτρησης | Συχνότητα Παρατηρήσεων | Ημερομηνίες Παρατηρήσεων | Διάστημα Παρατηρήσεων | Πηγή |
|---|-----------------------------|--------------------------|--------------------------|--------------|
| πλήθος πωλήσεων ($\in \mathbb{Z}^{0+}$) | ημερήσια δεδομένα (yearday) | 29/1/2011 έως 19/6/2016 | 1969 ημέρες ή 5,4 χρόνια | Walmart Inc. |
| Πλήθος Προϊόντων | Πλήθος Ιεραρχικών | Πλήθος Επιπέδων | Πλήθος Χρονοσειρών | |
| 3049 | 2 | 12 | 42840 | |

Πίνακας 6: Ταυτότητα Συνόλου Δεδομένων Διαγωνισμού $M5$

| Επίπεδο Συνάθροισης | id | Πλήθος Σειρών | Περιγραφή Επιπέδου |
|---------------------|----|---------------|-----------------------------------|
| Σύνολο | 1 | 1 | Συνολικές Πωλήσεις Μονάδων (Π.Μ.) |
| Πολιτεία | 2 | 3 | Π.Μ. ανά Πολιτεία |
| Κατάστημα | 3 | 10 | Π.Μ. ανά Κατάστημα |
| Κατηγορία | 4 | 3 | Π.Μ. ανά Κατηγορία Προϊόντων |
| Τμήμα | 5 | 7 | Π.Μ. ανά Τμήμα (Υποκατηγορία) |
| Πολιτεία/Κατηγορία | 6 | 9 | Π.Μ. ανά Πολιτεία & Κατηγορία |
| Πολιτεία/Τμήμα | 7 | 21 | Π.Μ. ανά Πολιτεία & Τμήμα |
| Κατάστημα/Κατηγορία | 8 | 30 | Π.Μ. ανά Κατάστημα & Κατηγορία |
| Κατάστημα/Τμήμα | 9 | 70 | Π.Μ. ανά Κατάστημα & Τμήμα |
| Προϊόν | 10 | 3049 | Π.Μ. ανά Προϊόν |
| Προϊόν/Πολιτεία | 11 | 9147 | Π.Μ. ανά Προϊόν & Πολιτεία |
| Προϊόν/Κατάστημα | 12 | 30490 | Π.Μ. ανά Προϊόν & Κατάστημα |
| Συνολικές Σειρές: | | 42840 | |

Πίνακας 7: Περιγραφή Επιπέδων | $M5$ Dataset

Το αναγνωριστικό, το πλήθος χρονοσειρών και περιγραφή για καθένα από τα επίπεδα συναθροίσεως στο σύνολο δεδομένων του $M5$ παρουσιάζονται στον Πίνακα 7, ενώ η



Σχήμα 2: Ομαδοποιημένη Χρονοσειρά Συνόλου Δεδομένων M5

ιεραρχία της ομαδοποιημένης χρονοσειράς (με μερική διάταξη ανάμεσα στα επίπεδα των δύο ιεραρχικών χρονοσειρών) παρατίθεται στο Σχήμα 2. Μαζί με τις παρατηρήσεις πωλήσεων μονάδων προϊόντων, παραχωρήθηκαν και επεξηγηματικές/εξωγενείς μεταβλητές (*explanatory variables*) όπως οι τιμές των προϊόντων, περίοδοι εκπτώσεων και διακοπών, ειδικά γεγονότα (*special events*). Το σύνολο δεδομένων (όπως παρέχεται μέσα από την επίσημη σελίδα των διαγωνισμών στο github: [Mcompetitions](#)) απαρτίζεται από τα παρακάτω csv αρχεία:

1. Ένα αρχείο *sales_train_validation.csv* που περιέχει τα δεδομένα των πωλήσεων για τα 30490 προϊόντα (3049 προϊόντα σε 10 καταστήματα το καθένα). Το αρχείο έχει μια στήλη για καθεμιά από τις 1941 ημέρες στο διάστημα μεταξύ 29/1/2011 και 22/5/2016, χωρίς την περίοδο των 28 ημερών για το validation (δηλαδή ως τις 19/6/2016). Περιέχει επίσης στήλες για κατηγορία, υποκατηγορία, προϊόν, πολιτεία και κατάστημα κάθε μονάδας.
2. Ένα αρχείο *sell_prices.csv* που περιέχει τα αναγνωριστικά καταστήματος και προϊόντος για τις 30490 μονάδες, μαζί με μια στήλη με τον εβδομαδιαίο μέσο όρο της τιμής πώλησης για καθένα από αυτά.
3. Ένα αρχείο *calendar.csv* που για καθεμιά ημερομηνία πωλήσεων στο πρώτο αρχείο,

περιέχει χαρακτηριστικά όπως, μέρα της εβδομάδας, μήνας, έτος. Περιέχει επίσης πληροφορίες για (μέχρι 2) ειδικά γεγονότα ή αργίες κάθε ημέρας (όνομα και περιγραφή). Τέλος, περιέχει 3 flags για το αν υπάρχει περίοδος εκπτώσεων μέσω SNAP (εκπτώτικα κουπόνια που παρέχονται από την ομοσπονδιακή κυβέρνηση με σε άτομα με οικονομικά προβλήματα) σε καθεμιά από τις 3 πολιτείες του συνόλου δεδομένων.

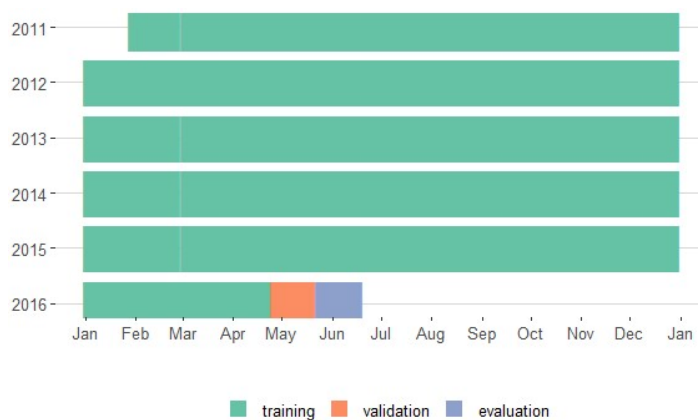
Τα παραπάνω αποτελούν βασικές μεταπληροφορίες - ταυτότητα του συνόλου δεδομένων του διαγωνισμού *M5*. Στην επόμενη παράγραφο παρουσιάζεται μια σύντομη επισκόπηση του περιεχομένου του *M5 Dataset*.

5.3 Επεξηγηματική Ανάλυση Δεδομένων

Το περιεχόμενο του συνόλου δεδομένων του *M5* παρουσιάζει ορισμένα ενδιαφέροντα χαρακτηριστικά, μερικά από τα οποία είναι απαραίτητη γνώση για όποιον κληθεί να προβλέψει το dataset. Η ισχυρή εποχιακότητα πολλαπλών περιόδων, ο υψηλός θόρυβος (πλειοψηφικές οι μηδενικές παρατηρήσεις) ή η μη κοινή χρονική έναρξη και λήξη των παρατηρήσεων για τα διαφορετικά προϊόντα είναι ορισμένα από αυτά. Στις παρακάτω μικρές ενότητες, παρουσιάζεται το περιεχόμενο του συνόλου και επεξηγούνται τα ιδιαίτερα χαρακτηριστικά του.

5.3.1 Διάρκεια Παρατηρήσεων

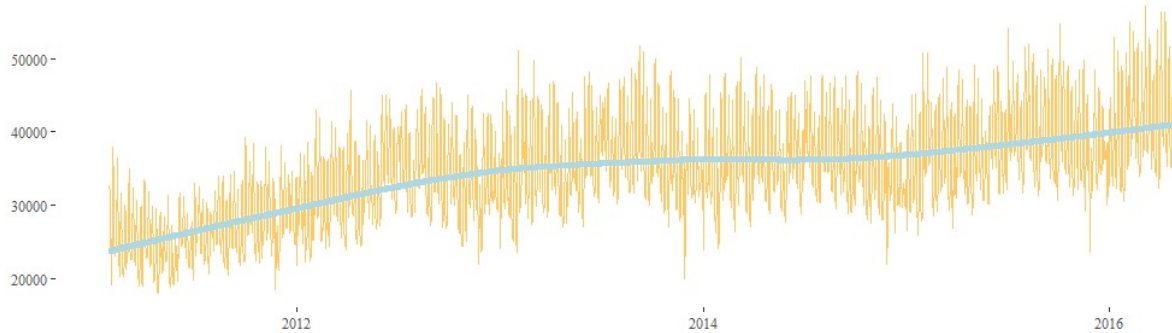
Οι παρατηρήσεις είναι ημερήσιες και η διάρκειά τους είναι από τις 29 Ιανουαρίου 2011 έως τις 19 Ιουνίου 2016. Δηλαδή καλύπτουν ένα χρονικό διάστημα 1969 ημερών ή περίπου 5,4 χρόνια (5 χρόνια, 4 μήνες και 23 ημέρες). Όπως φαίνεται και στο σχήμα 3 οι πρώτες 1913 ημέρες παραχωρήθηκαν -στο πλαίσιο του *M5*- για την περίοδο της εκπαίδευσης (training set), οι επόμενες 28 ημέρες (1914 έως 1941) από τις 25 Απριλίου 2016 έως τις 22 Μαΐου 2016 χρησιμοποιήθηκαν ως validation set και οι τελευταίες 28 ημέρες για την αξιολόγηση των υποβολών (test set). Η διάρκεια των παρατηρήσεων είναι αρκετά μεγάλη και κρίνεται επαρκής για την εκπαίδευση τόσο των στατιστικών μεθόδων, όσο και μεθόδων μηχανικής μάθησης και νευρωνικών δικτύων. Κατά τη διεξαγωγή του πειράματός μας θα χρησιμοποιήσουμε υποσύνολα των πρώτων 1941 ημερών για εκπαίδευση (training set + validation set) και θα συγκρίνουμε τα αποτελέσματα πάνω στο test set, όπως ορίζονται στο σχήμα 3. Στόχος αυτής της επιλογής είναι τα αποτελέσματα της παρούσας δουλειάς να μπορούν να συγκριθούν με τις υποβολές του διαγωνισμού *M5*.



Σχήμα 3: Διάρκεια & Περίοδοι | *M5 Dataset*

5.3.2 Συνολικές Πωλήσεις

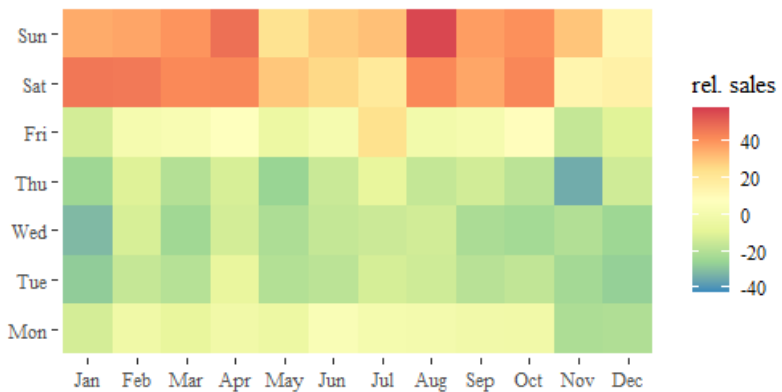
Η χρονοσειρά των συνολικών πωλήσεων στα 10 καταστήματα της Walmart στο παραπάνω διάστημα, είναι το πλέον συναθροισμένο επίπεδο συνάθροισης (επίπεδο 1), αποτελεί την κορυφή των δύο ιεραρχιών στην ομαδική χρονοσειρά μας και παρουσιάζεται στο σχήμα 4 με πορτοκαλί χρώμα. Φανερή γίνεται η ύπαρξη πολλαπλών περιόδων επο-



Σχήμα 4: Συνολικές Πωλήσεις (πορτοκαλί) & Smoothing Fit (μπλε) | M5 Dataset

χιακότητας (εβδομαδιαία, μηνιαία) που φαίνονται καλύτερα στο σχήμα 5. Σε εβδομαδιαίο επίπεδο παρατηρείται σαφής άνοδος πωλήσεων που ξεκινάει μεσοβδόμαδα (Πέμπτη) και κορυφώνεται τα σαββατοκύριακα. Με μέτρο τις πωλήσεις της Δευτέρας (δείκτης εποχιακότητας 1,0) παρατηρούμε κάμψη και στασιμότητα τις ημέρες Τρίτη, Τετάρτη, Πέμπτη (δ.ε. 0,9) και έντονη άνοδο μέχρι το Σάββατο (δ.ε. 1,2) που παραμένει και την Κυριακή. Τα στοιχεία αυτά αφορούν τη χρονοσειρά των συνολικών πωλήσεων (επίπεδο 0). Σε επόμενη ενότητα θα δούμε ποια είναι η εποχιακή συμπεριφορά και ορισμένων από τα χαμηλότερα επίπεδα συνάθροισης της ομαδοποιημένης χρονοσειράς.

Πέρα από την ισχυρή εποχιακότητα, παρατηρείται και μια ανοδική τάση στις παρατηρήσεις. Διακρίνουμε καθαρότερα τη συμπεριφορά αυτή μέσα από μια smoothing fit καμπύλη που φιλτράρει τις διακυμάνσεις της αρχικής χρονοσειράς και φαίνεται στο σχήμα 4 των συνολικών πωλήσεων με χρώμα μπλε. Πέρα από την αύξηση στη ζήτηση,



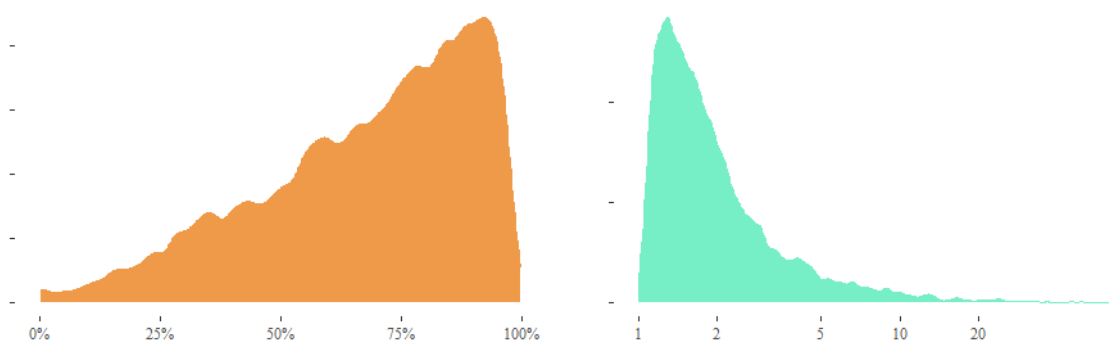
Σχήμα 5: Heatmap Εβδομαδιαίας & Μηνιαίας Εποχιακότητας Συνολικών Πωλήσεων (Επίπεδο 1) | M5 Dataset

μια κρυμμένη μεταβλητή που σχετίζεται με την ανοδική τάση, είναι και η "είσοδος" νέων προϊόντων καθ' όλη τη διάρκεια μέτρησης των πωλήσεων. Τέλος, μπορούμε να διακρίνουμε ορισμένες μύτες-κορυφές (*spikes*) που οφείλονται σε ειδικά γεγονότα (πχ Τελικοί

NBA, Cinco de Mayo, Γιορτή της Μητέρας κλπ), αργίες (Καθολικό/Ορθόδοξο Πάσχα, Πρωτοχρονιά κλπ) των ΗΠΑ ή εκπτώτικες περιόδους (κουπόνια προγράμματος SNAP).

5.3.3 Τιμές των Παρατηρήσεων

Μεγάλο ποσοστό των παρατηρήσεων στις 42840 χρονοσειρές του συνόλου είναι μηδενικές. Για πρώτη φορά στους διαγωνισμούς M επιλέγεται ένα σύνολο χρονοσειρών με τόσο έντονη διακοπτότητα. Στο σχήμα 6 παρουσιάζονται οι πυκνότητες πιθανότητας των τιμών που συναντάμε στο σύνολο δεδομένων. Στο σχήμα 6a παραθέτουμε την πυκνότητα πιθανότητας των ποσοστών των μηδενικών παρατηρήσεων για καθεμιά από τις 42840 χρονοσειρές του συνόλου δεδομένων. Μπορεί κανείς να παρατηρήσει πως η πλειοψηφία των χρονοσειρών του συνόλου έχει πάνω από 50% μηδενικές παρατηρήσεις. Στο Σχήμα



(a) Π.Π. Ποσοστών των Μηδενικών Παρατηρήσεων - Γραμμική Κλίμακα (b) Π.Π. Τιμών των Μη Μηδενικών Παρατηρήσεων - Λογαριθμική Κλίμακα

Σχήμα 6: Πυκνότητες Πιθανοτήτων Τιμών Παρατηρήσεων - 42840 σειρές | M5 Dataset

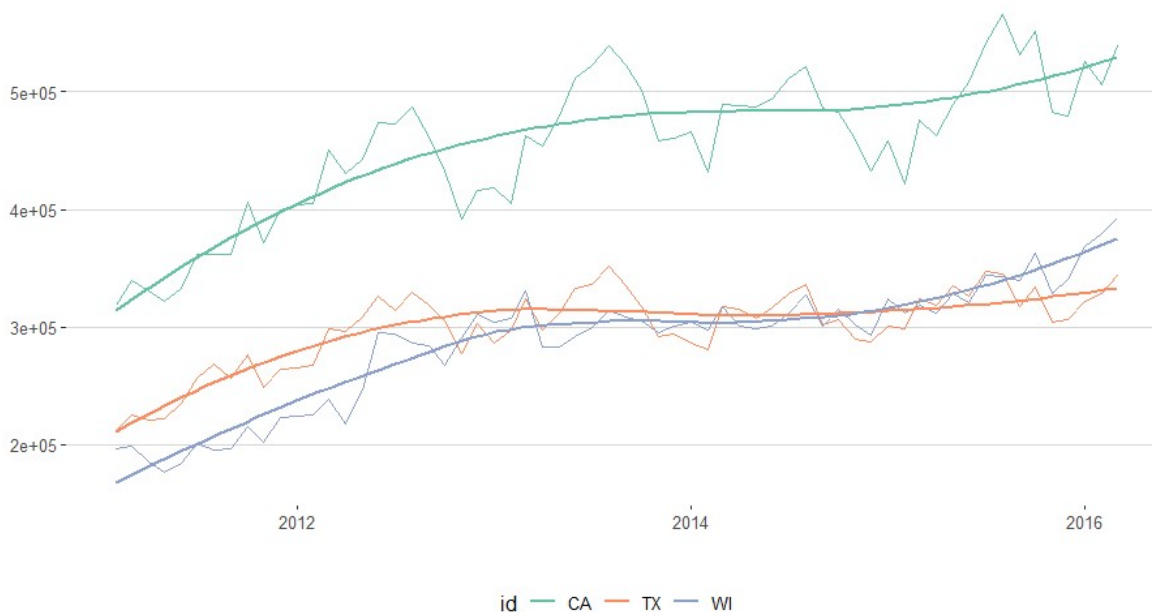
6b παρουσιάζεται η πυκνότητα πιθανότητας των μη μηδενικών τιμών των παρατηρήσεων, αποτυπωμένες σε λογαριθμική κλίμακα, μιας και η συντριπτική πλειοψηφία των μη μηδενικών παρατηρήσεων είναι κάτω από 20, ενώ οι μέγιστες ημερήσιες πωλήσεις μπορεί να ξεπερνούν τις 70 για μικρό αριθμό προϊόντων. Παρατηρούμε ότι η μεγάλη πλειοψηφία των μη μηδενικών παρατηρήσεων που αφορούν τις ημερήσιες πωλήσεις μονάδων είναι 1 και 2.

Από την παραπάνω ανάλυση, γίνεται κατανοητό πως για την πλειοψηφία των παρατηρήσεων, οι παρατηρούμενες τιμές είναι 0, 1 ή 2. Αυτό σημαίνει πως όποιος κληθεί να προβλέψει το σύνολο δεδομένων θα έρθει αντιμέτωπος με έντονο θόρυβο και διακοπτόμενη ζήτηση ειδικά όσο κινείται προς πιο αποσυναθροισμένα επίπεδα και ειδικά για τα πλέον αποσυναθροισμένα επίπεδα 10, 11, 12.

Για τις ανάγκες της επισκόπησης που επιχειρούμε σε αυτή την παράγραφο, και για να ξεπεράσουμε το παραπάνω πρόβλημα του θορύβου, θα μετασχηματίσουμε τα δεδομένα μας, αθροίζοντας τις ημερήσιες πωλήσεις σε μηνιαίες. Έτσι θα εξερευνήσουμε πιο ομαλές χρονοσειρές (με λιγότερες παρατηρήσεις για το ίδιο χρονικό διάστημα) ώστε να δώσουμε έμφαση στα μοτίβα και τα ποιοτικά στοιχεία των χρονοσειρών χωρίς περιττό θόρυβο.

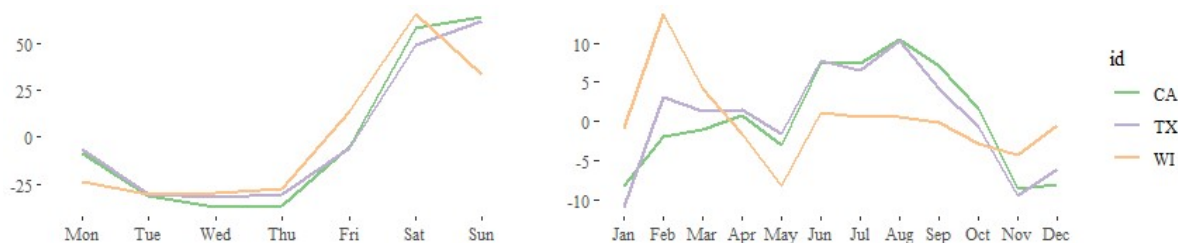
5.3.4 Πωλήσεις ανά Πολιτεία

Στο σχήμα 7 φαίνονται οι πωλήσεις μονάδων, συναθροισμένων στις τρεις πολιτείες: Καλιφόρνια, Τέξας, Γουισκόνσιν. Παρατηρούμε πως η Καλιφόρνια έχει τις υψηλότερες πωλήσεις ανάμεσα στις πολιτείες. Μια κρυμμένη μεταβλητή γι αυτή την υπεροχή της Καλιφόρνια είναι πως στο σύνολο δεδομένων περιέχονται 4 καταστήματα σε αυτή την πολιτεία σε αντίθεση με τις άλλες 2 για τις οποίες έχουμε δεδομένα από 3 καταστήματα στην καθεμιά. Το Γουισκόνσιν προσέγγιζε σταδιακά τις πωλήσεις στο Τέξας, ξεπερνών-



Σχήμα 7: Μηνιαίες Πωλήσεις Μονάδων ανά Πολιτεία & Smoothing Fit | *M5* Dataset

τας το επακολούθως στους τελευταίους μήνες των παρατηρήσεων. Παρατηρούνται επίσης έντονες «πτώσεις» στη στάθμη της Καλιφόρνια εντός των ετών 2013, 2014 και 2015 που γίνονται φανερά σε μικρότερη έκταση και στις υπόλοιπες πολιτείες. Αυτές οι πτώσεις, παρ' ότι δε συμβαίνουν κάθε έτος (πχ το 2012) πιθανώς αποτελούν ένδειξη ετήσιας εποχιακότητας. Στο σχήμα 8 φαίνεται η εβδομαδιαία και η μηνιαία εποχιακότητα ανά πο-



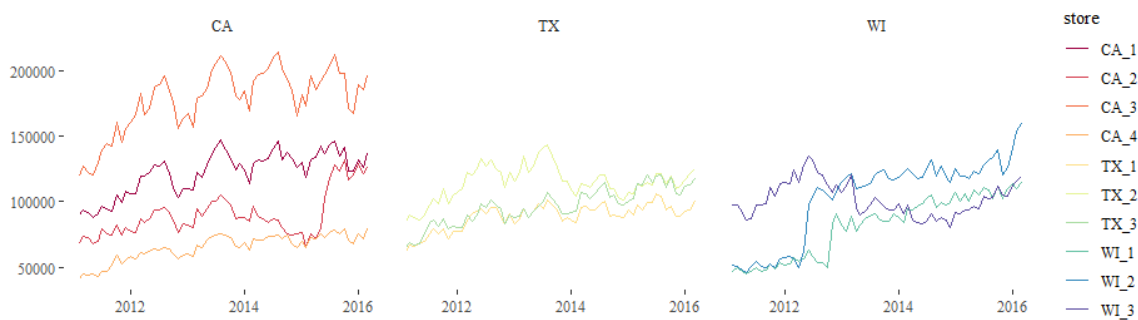
Σχήμα 8: Εβδομαδιαία & Μηνιαία Εποχιακότητα Πωλήσεων ανά Πολιτεία | *M5* Dataset

λιτεία. Παρατηρούμε πως οι τρεις πολιτείες συμπεριφέρονται πανομοιότυπα ως προς τις σχετικές πωλήσεις των αντίστοιχων ημερών της εβδομάδας ή μηνών του χρόνου, με το Γουισκόνσιν να διαφοροποιείται στη μηνιαία εποχιακότητα σε σχέση με τη συνολική εικόνα

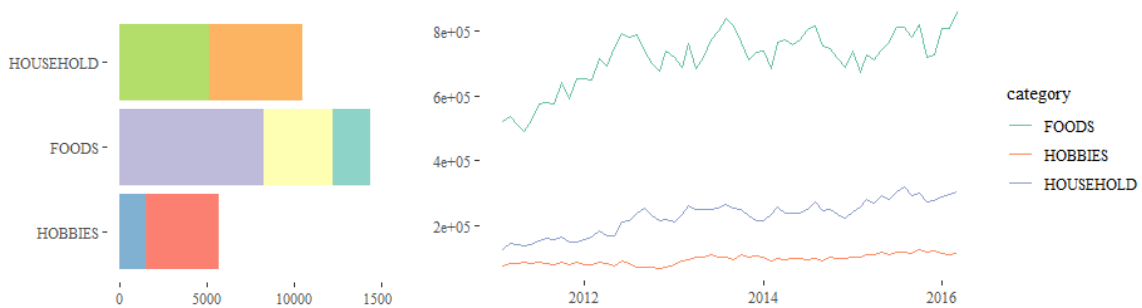
των καταστημάτων στις άλλες δύο πολιτείες.

5.3.5 Πωλήσεις ανά Κατάστημα & Κατηγορία Προϊόντος

Σε αυτή την ενότητα, αρχικά παρουσιάζουμε τις πωλήσεις μονάδων, συναθροισμένων ανά κατάστημα. Προκύπτουν έτσι οι 10 χρονοσειρές του σχήματος 9a, διαχωρισμένες σε 3 όψεις (*facets*), μια για κάθε πολιτεία. Παρατηρούμε πως τα καταστήματα στην Καλιφόρνια είναι καθαρά διαχωρισμένα από άποψη όγκου πωλήσεων. Το κατάστημα 4 έχει πωλήσεις μονάδων σχεδόν όσο τα υπόλοιπα τρία καταστήματα μαζί, ενώ ενδιαφέρουσα είναι η πορεία του καταστήματος 2 που αρχικά παρατηρούνται καθοδικές πωλήσεις προς το επίπεδο του καταστήματος 4, ενώ το τελευταίο έτος των παρατηρήσεων εμφανίζει ισχυρή επανάκαμψη και φτάνει στο επίπεδο του καταστήματος 1. Οι πωλήσεις των καταστημάτων στο Τέξας



(a) Πωλήσεις Μονάδων ανά Κατάστημα



(b) Πλήθος Προϊόντων ανά Τμήμα

(c) Πωλήσεις Μονάδων ανά Κατηγορία Προϊόντων

Σχήμα 9: Πωλήσεις ανά Κατάστημα & Κατηγορία | M5 Dataset

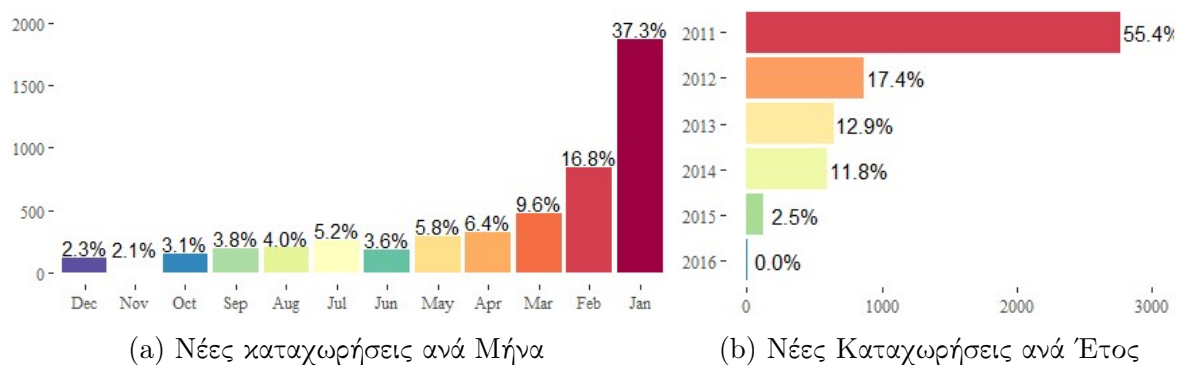
εμφανίζουν πανομοιότυπη συμπεριφορά, με το Τέξας 3 να φτάνει από το επίπεδο του Τέξας 1 στο επίπεδο του Τέξας 2 κατά τη διάρκεια των παρατηρήσεων. Στο Γουισκόνσιν παρατηρούνται δύο έντονες αλλαγές επιπέδου στα καταστήματα 1 και 2 το 2012, ενώ το κατάστημα 3 εμφανίζει σταδιακή πτώση στάθμης και επανάκαμψη.

Παρουσιάζονται επίσης πωλήσεις μονάδων ανά κατηγορία προϊόντων, καθώς και το πλήθος των προϊόντων που περιέχει κάθε τμήμα και κατηγορία στα σχήματα 9c και 9b

αντίστοιχα. Η κατηγορία Τροφίμων (*Foods*) περιέχει τα περισσότερα προϊόντα. Το τμήμα *Foods 1* για παράδειγμα, έχει πιο πολλές ετικέτες από ό,τι η κατηγορία των Χόμπι. Δεύτερη σε πλήθος προϊόντων έρχεται η κατηγορία *Household*. Στις χρονοσειρές των πωλήσεων ανά κατηγορία (σχήμα 9c) παρατηρούμε ότι η απόσταση μεταξύ των χρονοσειρών των κατηγοριών *Foods* και *Household* είναι μεγαλύτερη από αυτή του πλήθους των προϊόντων στην κάθε κατηγορία. Αντίστοιχα βλέπουμε πως η απόσταση μεταξύ των κατηγοριών *Hobbies* και *Household* είναι μικρότερη από τη διαφορά πλήθους προϊόντων στις 2 κατηγορίες. Αυτό καταδεικνύει πως πωλούνται περισσότερες μονάδες ανά προϊόν στην κατηγορία *Foods*, ακόλουθα στην κατηγορία *Household* και τέλος στην κατηγορία *Hobbies*.

5.3.6 Κατανομές Αφτηρίας & Λήξης Παρατηρήσεων

Ακόμη ένα ιδιαίτερο χαρακτηριστικό που παρουσιάζει το σύνολο δεδομένων του *M5* είναι πως παρά την ονομαστική διάρκεια των 1969 ημερών, για αρκετά προϊόντα η διάρκεια των παρατηρήσεων είναι μικρότερη της ονομαστικής. Αυτό συμβαίνει διότι έχουμε είτε μεταγενέστερες εισόδους είτε προγενέστερες εξόδους ή εσωτερικά (των 1969 ημερών) διαστήματα που δεν προσφέρεται το εν λόγω προϊόν.



Σχήμα 10: Ποσοστά Νέων Καταχωρήσεων ανά Μήνα & Έτος | *M5* Dataset

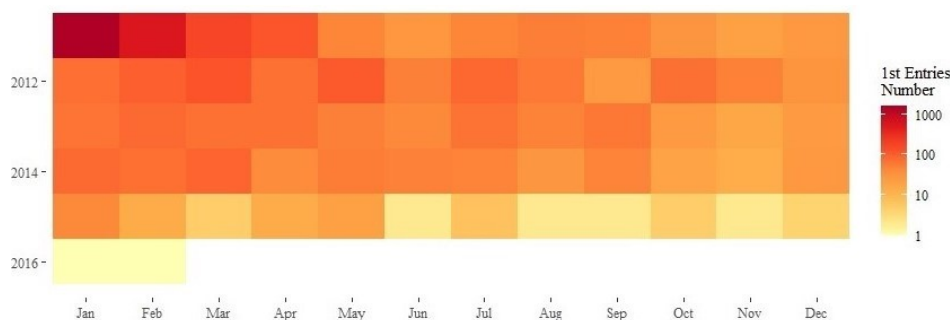
Σχετικές Συχνότητες Αφτηρίας Παρατηρήσεων ανά Έτος & Μήνα

Ένας απλός τρόπος για να αποτυπώσουμε στατιστικά αυτό το φαινόμενο είναι με τα ραβδογράμματα του σχήματος 10. Στο σχήμα 10b βλέπουμε πως το πρώτο έτος (2011) διαθέσιμες για πώληση ήταν κάτι παραπάνω από τις μισές μονάδες (μονάδα), δηλαδή $[3049 \cdot 55,4\%] = 1690$ προϊόντα. Αυτό εξηγεί και κομμάτι της ανοδικής τάσης που ανιχνεύσαμε στις συνολικές πωλήσεις. Στο σχήμα 10a φαίνεται και η διασπορά των πρώτων μη μηδενικών καταχωρήσεων ανά μήνα. Στους Ιανουάριους των 5,4 ετών του συνόλου εισήχθη μόνο το 37,3% των προϊόντων ή $[3049 \cdot 37,3\%] = 1134$ προϊόντα.

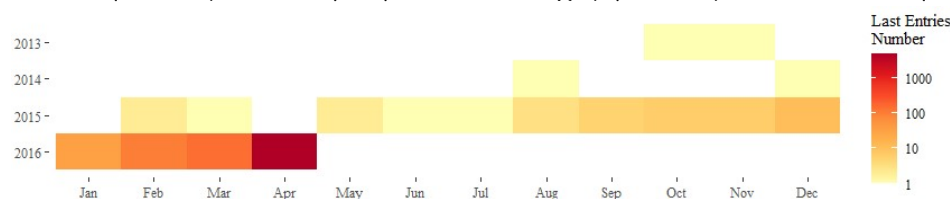
Heatmaps Πρώτων και Τελευταίων Μη Μηδενικών Παρατηρήσεων

Ένας ακόμη τρόπος επισκόπησης της διασποράς των πρώτων και τελευταίων μη μηδενικών πωλήσεων για τις μονάδες του συνόλου δεδομένων είναι μέσω heatmap. Στο heatmap

του σχήματος 11a μπορούμε να δούμε την πορεία των νέων καταχωρήσεων προϊόντων για καθένα από τους 62 μήνες του training set. Αντίστοιχα όπως είπαμε, παρουσιάζονται και πρόωρες αποχωρήσεις προϊόντων. Στο heatmap του σχήματος 11b βλέπουμε πως ήδη από το τρίτο έτος παρατηρήσεων (2013) έχουμε τελευταίες μη μηδενικές καταχωρήσεις σε ορισμένες μονάδες. Το 2014 έχουμε παρόμοιο πλήθος τελευταίων μη μηδενικών καταχωρήσεων, ενώ το 2015 έχουμε αύξηση του πλήθους αποχωρήσεων μονάδων που κινείται ανοδικά τους πρώτους μήνες του 2016 και κορυφώνεται στο τέλος της διάρκειας του dataset. Η ιδιαιτερότητα αυτή των κενών στις παρατηρήσεις του συνόλου δεδομένων,



(a) Heatmap Πλήθους Πρώτων Μη Μηδενικών Καταχωρήσεων Προϊόντων ανά Μήνα & Έτος



(b) Heatmap Πλήθους Τελευταίων Μη Μηδενικών Καταχωρήσεων Προϊόντων ανά Μήνα & Έτος

Σχήμα 11: Heatmaps Νέων Καταχωρήσεων ανά Μήνα & Έτος | M5 Dataset

πρέπει να αντιμετωπιστεί με κατάλληλο τρόπο κατά την προεπεξεργασία των δεδομένων ούτως ώστε να καταλήξουμε σε προβλέψεις με νόημα.

Έχοντας ολοκληρώσει την επεξηγηματική ανάλυση του συνόλου δεδομένων που θα αποτελέσει τη μελέτη περίπτωσης της παρούσας εργασίας, στην επόμενη παράγραφο παρουσιάζουμε τις προδιαγραφές του πειράματος καθώς και την πειραματική διάταξη.

5.4 Προδιαγραφές & Πειραματική Διάταξη

Το πείραμά μας θα επικεντρωθεί στη σύγκριση προβλέψεων που θα αφορούν τα 12 επίπεδα του συνόλου δεδομένων του M5. Οι προβλέψεις αυτές θα έχουν ευθυγραμμιστεί με όσες διαφορετικές ιεραρχίες μπορούν να προκύψουν συνδυάζοντας επίπεδα συνάθροισης από τις 2 ιεραρχίες στην ομαδοποιημένη χρονοσειρά του συνόλου δεδομένων. Οι συγκρίσεις θα αφορούν τόσο την ακρίβεια των παραγόμενων προβλέψεων όσο και το κόστος τους. Στόχο, όπως είπαμε, θα έχουμε την εξαγωγή της βέλτιστης ιεραρχίας (ή βέλτιστων ιεραρχιών) για πρόβλεψη που θα προκύπτει από τη σύγκριση των μετρήσεων που θα κάνουμε πάνω σε συνδυασμούς κριτηρίων ακρίβειας-κόστους που θα θέσουμε.

Στις παρακάτω ενότητες παρουσιάζονται οι επιλογές που κάναμε για το πείραμα και η πειραματική μεθοδολογία.

5.4.1 Επιλογή Ιεραρχιών & Εξεταζόμενα Επίπεδα Συνάθροισης

Πρώτο βήμα για την κατάστρωση του πειράματος είναι η επιλογή ενός συνόλου δεδομένων που επιτρέπει -μέσω της οργάνωσής του σε επίπεδα συνάθροισης- την αναφορά σε αρκετές διαφορετικές ιεραρχίες επιπέδων. Το σύνολο δεδομένων του *M5* -όπως είδαμε και στην παραπάνω παράγραφο- περιέχει για κάθε χρονοσειρά πωλήσεων μιας μονάδας προϊόντος, πληροφορία για το κατάστημα (*store id*), την πολιτεία (*state id*), το προϊόν (*item id*), το τμήμα (*dept id*) και την κατηγορία (*cat id*) της μονάδας. Με άλλα λόγια, τα δεδομένα είναι οργανωμένα σε δύο ιεραρχίες που αποτελούνται από τα προαναφερθέντα επίπεδα συνάθροισης. Για τις ανάγκες του πειράματος, δημιουργήσαμε μια λίστα με όλες

Επιλεγμένες Ιεραρχίες Συνόλου Δεδομένων *M5*

| Αναγνωριστικό Ιεραρχίας (Κλειδί Συνάθροισης) | Υψος | Πλήθος | |
|--|-------------------------------|--------|----|
| total/state/id | total/id | 2 | 1 |
| total/store/id | total/cat/id | 3 | 5 |
| total/dept/id | total/item/id | | |
| total/state/store/id | total/state/cat/id | 4 | 14 |
| total/state/dept/id | total/state/item/id | | |
| total/store/cat/id | total/store/dept/id | | |
| total/cat/state/id | total/cat/store/id | | |
| total/dept/state/id | total/dept/store/id | | |
| total/cat/dept/id | total/cat/item/id | | |
| total/dept/item/id | total/item/state/id | | |
| total/state/store/cat/id | total/state/store/dept/id | 5 | 18 |
| total/state/cat/dept/id | total/state/cat/item/id | | |
| total/state/dept/store/id | total/state/dept/item/id | | |
| total/store/cat/dept/id | total/cat/dept/state/id | | |
| total/cat/dept/store/id | total/cat/dept/item/id | | |
| total/cat/state/store/id | total/cat/state/dept/id | | |
| total/cat/state/item/id | total/cat/store/dept/id | | |
| total/dept/state/store/id | total/cat/item/state/id | | |
| total/dept/state/item/id | total/dept/item/state/id | | |
| total/state/store/cat/dept/id | total/state/cat/store/dept/id | 6 | 10 |
| total/state/cat/dept/store/id | total/state/cat/dept/item/id | | |
| total/cat/dept/state/store/id | total/cat/dept/state/item/id | | |
| total/cat/dept/item/state/id | total/cat/state/dept/store/id | | |
| total/cat/state/store/dept/id | total/cat/state/dept/item/id | | |
| ομαδοποιημένη X/Σ: (total/cat/dept/item/id)*(total/state/store/id) | | | 48 |

Πίνακας 8: Λίστα Ιεραρχιών Συνόλου Δεδομένων *M5*

τις (υπο)ιεραρχίες που μπορούν να προκύψουν από την ομαδοποιημένη χρονοσειρά του συνόλου δεδομένων. Στον πίνακα 8 παρουσιάζεται η εξαντλητική απαρίθμηση των ιεραρχιών που προκύπτουν από συνδυασμό των επιπέδων του συνόλου δεδομένων. Η εξαγωγή αυτής της λίστας αποτελεί ένα καλώς ορισμένο πρόβλημα συνδυαστικής μιας και ψάχνουμε

το σύνολο των διατάξεων όλων των υποσυνόλων των αναφερθέντων επιπέδων υπό τους περιορισμούς:

1. να έχουν ως τελευταίο επίπεδο τις πλήρως αποσυναθροισμένες χρονοσειρές (*id*), οι οποίες περιέχουν πωλήσεις προϊόντος ανά κατάσταση (*item/store*)⁵
2. όταν εμφανίζονται επίπεδα που ανήκουν σε μία από τις δύο ξεχωριστές ιεραρχίες του συνόλου δεδομένων (*total/cat/dept/item/id* και *total/state/store/id*) στην υπό εξέταση ιεραρχία, να ακολουθούν τη φυσική τους διάταξη (πχ τα επίπεδα *state* και *store* να εμφανίζονται μόνο με αυτή τη σειρά).

Μέσω ευθυγράμμισης των ανεξάρτητων προβλέψεων με τις παραπάνω ιεραρχίες, θα παράγουμε προβλέψεις για καθένα από τα 12 επίπεδα του συνόλου δεδομένων του *M5*. Τα 12 επίπεδα θα αποτελέσουν, δηλαδή, τα επίπεδα συνάθροισης στα οποία θα συγκριθούν οι προβλέψεις μας. Στοιχεία τους παραθέτουμε στον πίνακα 9.

Έχοντας πλέον στη διάθεσή μας ένα ικανοποιητικό μέγεθος 48 ιεραρχιών προς εξέταση για 12 εξεταζόμενα επίπεδα συνάθροισης, μπορούμε να προχωρήσουμε στην παρουσίαση των προδιαγραφών των ανεξάρτητων και ιεραρχικών προβλέψεων που θα εξάγουμε για τις ανάγκες του πειράματός μας.

| Επίπεδο Συνάθροισης | <i>id</i> | Πλήθος Σειρών |
|------------------------|-----------|------------------|
| total | 1 | 1 |
| state | 2 | 3 |
| store | 3 | 10 |
| category | 4 | 3 |
| department | 5 | 7 |
| state/category | 6 | 9 |
| state/department | 7 | 21 |
| store/category | 8 | 30 |
| store/department | 9 | 70 |
| item | 10 | 3049 |
| item/state | 11 | 9147 |
| item/store | 12 | 30490 |
| Συνολικές Σειρές: | | 42840 |

Πίνακας 9: Εξεταζόμενα Επίπεδα

5.4.2 Προδιαγραφές Προβλέψεων

Για να μελετήσουμε στοχευμένα τη σχέση ανάμεσα στην ευθυγράμμιση των ανεξάρτητων προβλέψεων με διαφορετικές ιεραρχίες και την απόδοση (ακρίβεια και κόστος) των παραγόμενων προβλέψεων, θα χρησιμοποιήσουμε ως σταθερές του πειράματός μας τόσο την επιλογή μοντέλου για ανεξάρτητη πρόβλεψη των χρονοσειρών σε κάθε επίπεδο, όσο και για τη μέθοδο ιεραρχικής πρόβλεψης με την οποία θα ευθυγραμμίσουμε τις ανεξάρτητες προβλέψεις ώστε να υπακούν στους περιορισμούς που θα τίθενται από κάθε εξεταζόμενη ιεραρχία. Στις επόμενες δύο ενότητες παρουσιάζουμε τις μεθόδους ανεξάρτητης και ιεραρχικής πρόβλεψης αντίστοιχα, με τις οποίες θα προβλέψουμε στο πείραμά μας.

Προβλέψεις Βάσης Για την πρόβλεψη βάσης ή αλλιώς ανεξάρτητη πρόβλεψη (δηλαδή την πρόβλεψη μιας χρονοσειράς χωρίς εναρμόνιση με την ιεραρχία στην οποία ανήκει) θα χρησιμοποιήσουμε εκθετική εξομάλυνση. Όπως είδαμε και στην

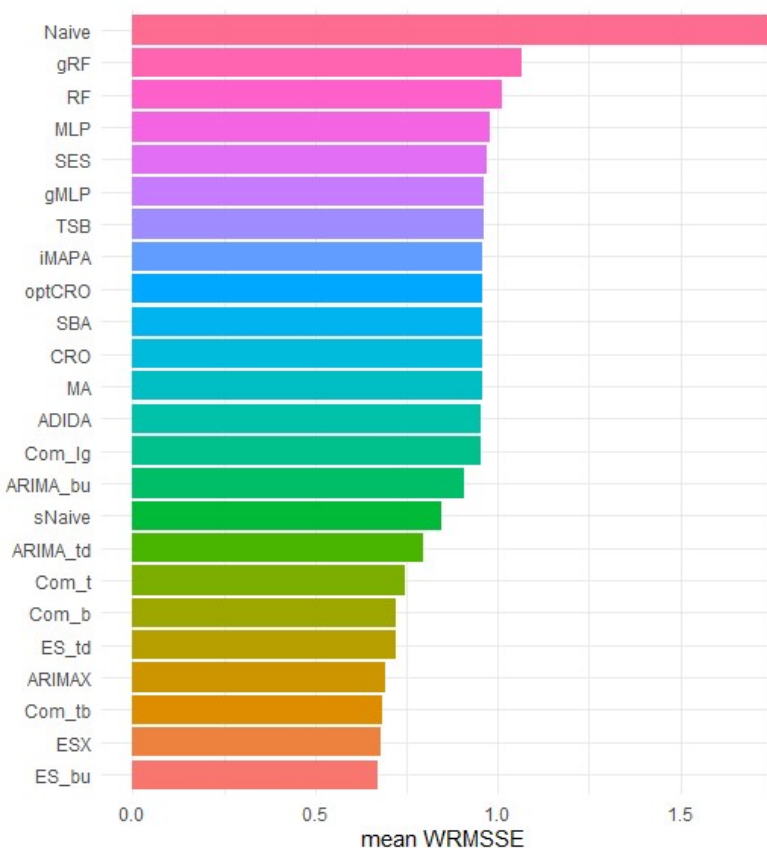
⁵Στη γενική περίπτωση, μια ιεραρχία μπορεί να μην καταλήγει στο πλέον αποσυναθροισμένο επίπεδο (bottom level). Ωστόσο,

παράγραφο 2.4 οι μέθοδοι εκθετικής εξομάλυνσης αποτελούν μια αρκετά δημοφιλή οικογένεια στατιστικών μεθόδων πρόβλεψης, ιδιαίτερα σε επιχειρησιακές προβλέψεις και προβλέψεις πωλήσεων λιανικής. Ένα από τα χαρακτηριστικά ευρήματα του διαγωνισμού *M5* ήταν πως ο συνδυασμός ανεξάρτητης πρόβλεψης των μη συναθροισμένων χρονοσειρών με εκθετική εξομάλυνση και bottom up συνάθροιση υπεραπέδωσε όλων των υπόλοιπων μεθόδων που χρησιμοποιήθηκαν ως benchmarks για την ακρίβεια των προβλέψεων του συνόλου δεδομένων του διαγωνισμού *M5*.

Στο σχήμα 12 φαίνεται η καλύτερη απόδοση της εκθετικής εξομάλυνσης με bottom up συνάθροιση (ES_{bu}) επί των υπόλοιπων 23 benchmarks, σε ό,τι αφορά στο μέσο σφάλμα (mean WRMSSE) των προβλέψεων στα 12 διαφορετικά επίπεδα συνάθροισης, όπως διατέθηκαν δημόσια μέσα από τη σελίδα του διαγωνισμού στο github ([Mcompetitions](#)).

Πέρα από την υψηλότερη ακρίβεια της εκθετικής εξομάλυνσης με bottom up συνάθροιση επί των υπόλοιπων μεθόδων που αποτέλεσαν τα benchmarks του διαγωνισμού, η μέθοδος πήγε πολύ

καλά και συγκρινόμενη με τις υποβολές των διαγωνιζόμενων ομάδων. Πιο συγκεκριμένα, μόνο το 7,5% των υποβληθέντων προβλέψεων ή αλλιώς οι προβλέψεις 415 από τις 5507 ομάδες υπεραπέδωσαν σε ακρίβεια της ES_{bu} [Makridakis, Spiliotis, and Assimakopoulos, 2020c]. Όπως είδαμε στην παράγραφο 2.4, η πρόβλεψη με εκθετική εξομάλυνση περιλαμβάνει μια σειρά από μεθόδους που εφαρμόζονται σε χρονοσειρές με διαφορετικά ποιοτικά χαρακτηριστικά, όπως εποχιακότητα (προσθετική ή πολλαπλασιαστική) και τάση (προσθετική, προσθετική damped, πολλαπλασιαστική) και η επιλογή του κάθε μοντέλου γίνεται με αυτόματο τρόπο. Έτσι, η εκθετική εξομάλυνση καλύπτει μια ευρεία γκάμα χρονοσειρών και ιδιαίτερα χρονοσειρές πωλήσεων λιανικής. Συνεπώς η εκθετική εξομάλυνση αποτελεί μια αξιοπρεπή επιλογή ως σταθερά για την ανεξάρτητη πρόβλεψη όλων των χρονοσειρών στο πείραμά μας.



Σχήμα 12: Μέση Απόδοση Benchmarks | *M5* Dataset

Ιεραρχικές Προβλέψεις Για την ευθυγράμμιση/εναρμόνιση (*reconciliation*) των ανεξάρτητων προβλέψεων με κάθε μια από τις 48 εξεταζόμενες ιεραρχίες του πειράματός μας, θα χρησιμοποιήσουμε *Minimum Trace* εναρμόνιση. Όπως είδαμε και στο κεφάλαιο 4 η *Minimum Trace* δέχεται ως είσοδο ένα σύνολο ανεξάρτητων προβλέψεων ολόκληρης της ιεραρχίας και στη συνέχεια εναρμονίζει τις προβλέψεις με την ιεραρχία, ελαχιστοποιώντας το σφάλμα που προκύπτει από την απόκλιση των προβλέψεων με τους περιορισμούς συνάνθροισης που ορίζουν οι παρατηρήσεις. Έτσι αξιοποιείται η πληροφορία από όλα τα επίπεδα της ιεραρχίας, σε αντίθεση με τις κλασσικές μεθόδους ιεραρχικής πρόβλεψης όπως οι *bottom up*, *top down* και *middle out* (single level προσεγγίσεις) κατά τις οποίες προβλέπονται οι χρονοσειρές ενός μόνο επιπέδου συνάνθροισης και στη συνέχεια παράγονται προβλέψεις για τα υπόλοιπα επίπεδα βάσει της φυσικής ιεραρχίας των δεδομένων.

Ωστόσο, δεν είναι εκ των προτέρων γνωστό αν όλα τα επίπεδα περιέχουν υψηλή προβλεπτική αξία, η οποία θα βελτιώσει την ακρίβεια μέσω της εξωγενούς πληροφορίας που θα μεταφερθεί από την εναρμόνιση με *Minimum Trace*. Για να μελετηθεί το ποια επίπεδα συνάνθροισης και με ποια διάταξη προσφέρουν βέλτιστες προβλέψεις ανάμεσα στις πιθανές ιεραρχίες επιπέδων, η *Minimum Trace* στην εργασία θα τροφοδοτείται κάθε φορά με μια διαφορετική ιεραρχία προς εξέταση. Άρα το σφάλμα που προκύπτει από τη μη συμφιλίωση των προβλέψεων βάσης με την ιεραρχία, και αποτελεί την ποσότητα που ελαχιστοποιεί η *Minimum Trace*, θα περιλαμβάνει κάθε φορά διαφορετικά επίπεδα συνάνθροισης και με άλλη διάταξη. Συνεπώς κύριος στόχος τ

5.4.3 Προδιαγραφές Σύγκρισης Απόδοσης

Έχοντας αναλύσει πόσες και ποιες ιεραρχίες θα εξετάσουμε, καθώς και τις μεθόδους ανεξάρτητων και ιεραρχικών προβλέψεων που θα χρησιμοποιήσουμε, παρουσιάζουμε με ποιον τρόπο θα συγκρίνουμε την απόδοση των διαφορετικών ιεραρχιών στο πείραμά μας.

Μέτρα Ακρίβειας Στο πλαίσιο της ανάλυσής μας χρησιμοποιήσαμε μέτρων ακριβείας για τη σύγκριση των αποτελεσμάτων. Ωστόσο, η παρουσίαση των αποτελεσμάτων θα γίνει πάνω στη σύγκριση των ριζών των κανονικοποιημένων μέσω των τετραγωνικών σφαλμάτων μεταξύ των προβλέψεων και των μελλοντικών τιμών, δηλαδή στο μέτρο *Root Mean Squared Scaled Error*:

$$RMSSE := \sqrt{\frac{\frac{1}{h} \sum_{t=n+1}^{n+h} (\mathbf{y}_t - \hat{\mathbf{y}}_t)^2}{n}}{\frac{1}{n-1} \sum_{t=2} (\mathbf{y}_t - \mathbf{y}_{t-1})^2}} \quad (70)$$

όπου \mathbf{y}_t είναι η μελλοντική τιμή τη στιγμή t , $\hat{\mathbf{y}}_t$ οι προβλέψεις για τη στιγμή t , n είναι το μήκος του συνόλου εκπαίδευσης και h ο ορίζοντας πρόβλεψης. Ο *RMMSE* είναι ένα μέτρο με ορισμένα χαρακτηριστικά, χρήσιμα στο πείραμά μας. Πιο συγκεκριμένα:

- είναι ανεξάρτητο της κλίμακας (ή στάθμης) των δεδομένων

- δεν ορίζεται μόνο όταν οι προβλέψεις της *Naive* έχουν μηδενικό σφάλμα
- αντιμετωπίζει με συμμετρικό τρόπο τα αρνητικά και θετικά σφάλματα
- έχει καθορισμένη μέση τιμή και πεπερασμένη απόκλιση

Το μέτρο ακριβείας του διαγωνισμού *M5* ήταν το *Weighted RMSSE*, δηλαδή ένας RMMSE σταθμισμένος με ένα βάρος w_i που εκφράζει την ανθροιστική αξία της κάθε χρονοσειράς σε δολάρια (τιμή πώλησης \times μονάδες πώλησης τις τελευταίες 28 ημέρες των ιστορικών παρατηρήσεων):

$$WRMSSE := \sum_{i=1}^{42840} w_i \cdot RMSSE_i \quad (71)$$

Ο *WRMSSE* είναι μέτρο κατάλληλο για βελτιστοποίηση κερδών, ωστόσο η δική μας εργασία δεν έχει τέτοιο στόχο. Παρά το γεγονός ότι η μέτρηση με *WRMSSE* θα μας επέτρεπε την απευθείας σύγκριση με τις προβλέψεις που υποβλήθηκαν στον *M5*, θα επιλέξουμε να παίρνουμε ίσα βάρη, και συνεπώς απλούς μέσους όρους των RMMSE των προβλέψεων για να εξάγουμε ένα μέτρο ακριβείας για τα διάφορα επίπεδα συνάνθροισης που αποτελούνται από 1 έως 30490 χρονοσειρές.

Για να γίνει πιο καθαρή η επίδραση που θα έχει ο γραμμικός συνδυασμός επιπέδων συνάνθροισης στην ακρίβεια των προβλέψεων, σε σχέση με την *bottom up*, ως την πιο απλή μέθοδο ιεραρχικής πρόβλεψης, πέρα από τον *RMSSE*, οι προβλέψεις -έστω *fc*- κάθε εξεταζόμενης ιεραρχίας θα συγκριθούν και ως προς το σχετικό με τη *bottom up* σφάλμα πρόβλεψης:

$$RelRMSSE := \frac{RMSSE_{fc}}{RMSSE_{bu}} \quad (72)$$

Όπου $RMSSE_{fc}$ είναι το σφάλμα πρόβλεψης της εκάστοτε εξεταζόμενης μεθόδου και $RMSSE_{bu}$ το σφάλμα της μεθόδου *bottom up*.

Μέτρα Κόστους Πέρα από την ακρίβεια των προβλέψεων, μια ακόμη βασική παράμετρος για την ολοκληρωμένη μελέτη της απόδοσης καθεμιάς εκ των 48 ιεραρχιών του πειράματός μας είναι και η σύγκριση του κόστους πρόβλεψης. Το κόστος αυτό μπορεί να μετρηθεί με διάφορους τρόπους. Στο πείραμά μας για τον καθορισμό του κόστους πρόβλεψης κάθε ιεραρχίας, θα μετρήσουμε:

1. το πλήθος των χρονοσειρών της κάθε ιεραρχίας
2. το χρόνο πρόβλεψης για κάθε ιεραρχία
3. το μέγεθος των παραγόμενων αντικειμένων στη μνήμη

Σημειώνεται πως οι μετρήσεις χρόνου και μνήμης ως δείκτες κόστους πρόβλεψης, είναι μετρήσεις στενά δεμένες με το μηχάνημα που εκτελεί τον κώδικα. Ωστόσο, το πλήθος των χρονοσειρών σε κάθε ιεραρχία αποτελεί έναν σκληρό αριθμό (*hard number*), δηλαδή ανεξάρτητο της όποιας τεχνολογίας εφαρμόζεται για την παραγωγή προβλέψεων που

επιδέχεται γενικότερης χρήσης στη μελέτη μας και θα αποτελέσει τον πιο αξιόπιστο δείκτη κόστους πρόβλεψης.

Στις παραπάνω ενότητες παρουσιάστηκαν οι προδιαγραφές και τα επιμέρους στοιχεία του της πειραματικής διάταξης. Πιο συγκεκριμένα μελετήθηκε το:

- ποια δεδομένα θα εξεταστούν
- ποιες ιεραρχίες του συνόλου δεδομένων θα συγκριθούν και πάνω σε ποια επίπεδα
- ποιες μέθοδοι θα χρησιμοποιηθούν για ανεξάρτητη και ιεραρχική πρόβλεψη
- με ποια μέτρα θα συγκριθούν ακρίβεια και κόστος πρόβλεψης

Στην επόμενη παράγραφο παρουσιάζεται η μεθοδολογία που θα εφαρμοστεί για την εξαγωγή των μετρήσεων και θα αποτελέσει σύνθεση των παραπάνω στοιχείων.

5.5 Πειραματική Διαδικασία

Έχοντας δει τα επιμέρους βήματα της μεθοδολογίας που θα συγκρίνει ιεραρχίες δεδομένων ως προς την προβλεπτική τους ικανότητα, στην παρούσα παράγραφο αναπτύσσεται η πειραματική διαδικασία που θα ακολουθηθεί. Η διαδικασία/αλγόριθμος από τον οποίο θα προκύψει η κατάταξη των εξεταζόμενων ιεραρχιών ως προς την απόδοσή τους στην πρόβλεψη των 12 επιπέδων συνάντρωσης του συνόλου δεδομένων, παρουσιάζεται στις παρακάτω ενότητες, ενώ στην επόμενη παράγραφο (5.6) παρουσιάζεται η ανάπτυξη της παρούσας μεθοδολογίας σε ένα πρόγραμμα σε γλώσσα R που θα αυτοματοποιεί τη διαδικασία παραγωγής και σύγκρισης των προβλέψεων, που θα έχουν ευθυγραμμιστεί με καθεμία εκ των 48 εξεταζόμενων ιεραρχιών του συνόλου.

5.5.1 Προεπεξεργασία Δεδομένων

Πρώτο βήμα της μεθοδολογίας είναι η προεπεξεργασία των δεδομένων και η προσαρμογή τους στις ανάγκες του πειράματος. Η προεπεξεργασία ξεκινά με την τροποποίηση της μορφής των δεδομένων, ώστε να εξάγονται τελικά οι μη συναθροισμένες (*bottom-level*) χρονοσειρές του συνόλου δεδομένων, μαζί με την απαραίτητη πληροφορία για τον τρόπο που μπορούν να οργανωθούν ώστε να προκύψουν τα υπόλοιπα επίπεδα συνάντρωσης που περιέχονται στη συλλογή των ομαδοποιημένων χρονοσειρών της μελέτης περίπτωσης. Το βήμα αυτό είναι προαπαιτούμενο για οποιαδήποτε ανάλυση σαν κι αυτή της εργασίας, και η υλοποίησή του είναι στενά δεμένη τόσο με τα ίδια τα δεδομένα όσο και με την τεχνολογία με την οποία κανείς προεπεξεργάζεται, μοντελοποιεί και προβλέπει δεδομένα. Η παραπάνω διαδικασία δε θα αναλυθεί περαιτέρω καθώς ξεπερνά το ενδιαφέρον της παρούσας εργασίας. Δύο όμως μεταγενέστερα βήματα της προεπεξεργασίας, άπτονται του ενδιαφέροντος της παρουσίασης, καθώς επεμβαίνουν δυνητικά στην ακρίβεια των προβλέψεων. Τα βήματα αυτά αφορούν τόσο τη διάρκεια των ιστορικών παρατηρήσεων με τις οποίες θα εκπαιδευτούν τα μοντέλα εκθετικής εξομάλυνσης, όσο και το πλήθος των μονάδων που

θα συμπεριλάβουμε στα τρεξίματα. Η επεξεργασία της διάρκειας των παρατηρήσεων και το φιλτράρισμα των χρονοσειρών παρουσιάζονται στις δύο επόμενες ενότητες. Το σύνολο δεδομένων στη μελέτη περίπτωσης των πωλήσεων της Walmart που μελετά η παρούσα εργασία, όπως είδαμε στην παράγραφο 5.2, έχει:

- *Υψηλή Διακοπτότητα (High Intermittency)*
Όπως βλέπουμε και στο παράδειγμα πωλήσεων του προϊόντος Hobbies_2_010 στο κατάστημα Wi_1 στο σχήμα 13, αλλά αντιλαμβανόμαστε και από τις κατανομές του σχήματος 6 στην επεξηγηματική ανάλυση του συνόλου δεδομένων, στις χρονοσειρές των χαμηλών επιπέδων συνάθροισης (με κύρια τα επίπεδα 10,11,12, δηλαδή προϊόν, προϊόν ανά πολιτεία, προϊόν ανά κατάστημα) προκύπτουν διαστήματα μηδενικών πωλήσεων -ταθερά ή μη- ανάμεσα σε θετικές τιμές πωλήσεων, οι οποίες πλειοψηφικά είναι 1,2 ή 3.
- *Ασυγχρόνιστες Ενάρξεις & Λήξεις Παρατηρήσεων*
Όπως έγινε φανερό στα heatmaps του σχήματος 11 αλλά και τα ιστογράμματα του σχήματος 10, οι πρώτες και τελευταίες παρατηρήσεις πωλήσεων, των διαφορετικών μονάδων στα 10 καταστήματα, είναι ημερολογιακά μη κοινές μεταξύ τους.



Σχήμα 13: Παράδειγμα Διακοπτόμενης Ζήτησης | *id μονάδας*: Hobbies/2/010/Wi/1

Επεξεργασία Διάρκειας Παρατηρήσεων Για στατιστικές μεθόδους πρόβλεψης όπως η εκθετική εξομάλυνση που θα χρησιμοποιήσουμε στο πείραμά μας, σε συνδυασμό με τη φύση των χρονοσειρών που θα προβλέψουμε, δεν είναι αναγκαία τόσο μεγάλη διάρκεια παρατηρήσεων. Οι 280 εβδομάδες των ιστορικών παρατηρήσεων καλύπτουν τις ανάγκες για συνθετότερες στατιστικές μεθόδους ή μεθόδους μηχανικής μάθησης, στις οποίες μια διάρκεια χιλιάδων ημερησίων παρατηρήσεων μιας ποσότητας, μπορεί να μεταβάλει σημαντικά την ακρίβεια των προβλέψεων. Ωστόσο για τις ανάγκες της παρούσας πειραματικής διαδικασίας, λίγοι μήνες δεδομένων κρίνονται επαρκείς για την εκπαίδευση των μοντέλων με εκθετική εξομάλυνση. Ακόμη και αν υποθεθεί πως παραπάνω εβδομάδες παρατηρήσεων θα επηρέαζαν σε στατιστικά σημαντικό βαθμό την ακρίβεια των προβλέψεων, τα σφάλματα των προβλέψεων βάσης με εκθετική εξομάλυνση θα είναι κοινά για κάθε εξεταζόμενη ιεραρχία και θα αποτελέσουν σταθερά του πειράματος.

Για τους παραπάνω λόγους, επιλέχθηκε να μοντελοποιηθούν οι πωλήσεις για ένα διάστημα 13 εβδομάδων (91 ημέρες ή ≈ 3 μήνες). Πιο συγκεκριμένα, θα αξιοποιηθούν οι 9 τελευταίες εβδομάδες του training set, μαζί με τις 4 εβδομάδες του evaluation set, όπως παραχωρήθηκαν στο διαγωνισμό M5. Οι προβλέψεις θα αξιολογηθούν πάνω στο test set. Με τον περιορισμό αυτό στη διάρκεια των παρατηρήσεων επιτυγχάνονται τα εξής:

- Επιταχύνεται σημαντικά η διεξαγωγή του κάθε πειράματος.

- Αντιμετωπίζονται σε μεγάλο βαθμό οι μη κοινές ενάρξεις πωλήσεων μονάδων.
- Η αξιολόγηση της ακρίβειας των προβλέψεων θα μπορεί να συγκριθεί με τις υποβολές των ομάδων του διαγωνισμού M5.

Φιλτράρισμα Χρονοσειρών με Μηδενικές Πωλήσεις Για την εφαρμογή της μεθόδου συνάντησης *Minimum Trace*, όπως και της προγενέστερης *Optimal Combination*, δεν επιτρέπεται η χρήση σταθερών εκτιμήσεων και προβλέψεων σε όλο τον ορίζοντα πρόβλεψης [Hyndman et al., 2011, Wickramasuriya et al., 2019]. Συνεπώς, χρειάζεται κατάλληλος χειρισμός για τις πωλήσεις μονάδων στο σύνολο δεδομένων τα προϊόντα που -για κάποιον από τους δύο λόγους που αναφέραμε- έχουν σχεδόν αποκλειστικά μηδενικές τιμές, οι οποίες θα συνεπάγονται μηδενικές προβλέψεις.

Όπως ελέχθη, με την επεξεργασία της διάρκειας των παρατηρήσεων, αντιμετωπίζονται σε μεγάλο βαθμό οι μη κοινές ενάρξεις πωλήσεων αλλά όχι εντελώς. Το ζήτημα αυτό αντιμετωπίζεται με διάφορες τακτικές. Μια επιλογή μπορεί να είναι η προσθήκη ενός μικρού θορύβου (της τάξης του 10% της ελάχιστης απόστασης παρατηρήσεων) ώστε να εγγυώνται μη σταθερές προβλέψεις. Εναλλακτικά, μπορεί να επιλεγεί η προσθήκη ενός φίλτρου που θα απομονώνει τις χρονοσειρές που περιέχουν ένα αρκετά μικρό αριθμό μη μηδενικών πωλήσεων⁶. Επιλέχθηκε η δεύτερη τακτική, συνεπώς ένα ακόμη βήμα της προεπεξεργασίας των δεδομένων είναι το φιλτράρισμα (*filtering*) των χρονοσειρών του συνόλου δεδομένων με -σχεδόν αποκλειστικά- μηδενικές παρατηρήσεις.

Για τους παραπάνω λόγους, αποφασίστηκε να φιλτραριστούν όσες χρονοσειρές έχουν μη μηδενικές παρατηρήσεις, που δεν ξεπερνούν ένα κατώφλι της τάξης των λίγων μονάδων, παραδείγματος χάρη 8 μη μηδενικές πωλήσεις στις 91 ημέρες των παρατηρήσεων. Το φιλτράρισμα θα προηγείται της δειγματοληψίας, ώστε να εξασφαλίζεται πως στα δείγματα θα περιέχονται μονάδες με στατιστικά σημαντικό αριθμό -στο πλαίσιο της εργασίας- μη μηδενικών πωλήσεων.

5.5.2 Δειγματοληψία Χρονοσειρών

Ο αριθμός των προϊόντων στο σύνολο δεδομένων του M5 είναι όπως είδαμε 3049 και ο αριθμός όλων των χρονοσειρών 42840. Μετά την προεπεξεργασία των δεδομένων, η πειραματική διαδικασία περιλαμβάνει την εξαγωγή ενός συνόλου δειγμάτων του συνόλου δεδομένων, πάνω στο οποίο θα εξαχθούν οι προβλέψεις που θα έχουν εναρμονιστεί με τις εξεταζόμενες ιεραρχίες. Πιο συγκεκριμένα, θα εξαχεται ένα πλήθος τυχαίων δειγμάτων που περιέχουν ένα σταθερό αριθμό προϊόντων ανά υποκατηγορία προϊόντων. Για παράδειγμα, μπορούν να εξαχθούν 30 δείγματα με 50 προϊόντα ανά υποκατηγορία το καθένα. Ο αριθμός των 50 προϊόντων ανά υποκατηγορία, ρυθμίζει και το συντελεστή δειγματοληψίας

⁶Η αιτιολόγηση της χρήσης του φίλτρου χρονοσειρών -σχεδόν- μηδενικών πωλήσεων, είναι πως χρονοσειρές που περιέχουν λίγες μη μηδενικές πωλήσεις (π.χ. σε ποσοστό μικρότερο του 10% των παρατηρήσεων), με τις μη μηδενικές τιμές τους να είναι πλειοψηφικά 1,2 ή 3 όπως προκύπτει και από το σχήμα 6, θα οδηγήσουν σε οριακά μηδενικές προβλέψεις και συνεπώς αφενός δε θα εμποδίζουν την εφαρμογή της *Minimum Trace* και αφετέρου δε θα επηρεάζουν σε σημαντικό βαθμό την ακρίβεια των μετρήσεων.

(*downsampling factor*) αφού δίνει ένα δείγμα με $50 \times 7 \times 10 = 3500$ χρονοσειρές (50 προϊόντα σε 7 υποκατηγορίες σε 10 καταστήματα). Ορισμένοι από τους λόγους που επιλέχθηκε η δειγματοληψία του συνόλου δεδομένων για τη διαδικασία αξιολόγησης των εξεταζόμενων ιεραρχιών, είναι:

1. Επειδή οι χρόνοι τρεξίματος του προγράμματος που αναπτύχθηκε για τη σύγκριση των ιεραρχιών είναι της τάξης των ημερών, τα μικρότερα samples επιτρέπουν ταχύτερη και πιο στοχευμένη διεξαγωγή κάθε τρεξίματος (πχ επιλογή διαφορετικών ομάδων προϊόντων ή γεωγραφικών τμημάτων κλπ).
2. Τα συμπεράσματα για την απόδοση των προβλέψεων που ευθυγραμμίζονται με διαφορετικές ιεραρχίες, έχει αξία να εξεταστεί αν γίνονται φανερά και σε μικρότερο πλήθος προϊόντων, και αν ναι σε ποιο ποσοστό μεγέθους επί του συνόλου δεδομένων γίνονται καθαρότερα.
3. Το σύνολο δεδομένων του *M5* αποτελεί ένα στιγμιότυπο πωλήσεων της Walmart σε μια μεγάλη γεωγραφική περιφέρεια. Αυτό σημαίνει πως η σύγκριση των 48 εξεταζόμενων ιεραρχιών, πάνω στην πρόβλεψη -αποκλειστικά- όλων των προϊόντων του συνόλου δεδομένων, θα προσφέρει συμπεράσματα που θα είναι λιγότερο γενικά από ό,τι η σύγκριση των ιεραρχιών σε πολλά δείγματα του συνόλου δεδομένων.⁷
4. Αν στόχο αποτελεί η εύρεση της βέλτιστης ιεραρχίας για εναρμόνιση, πιθανό πρώτο βήμα θα ήταν η αξιολόγηση προβλέψεων ιεραρχιών πάνω σε μικρότερα samples του εκάστοτε συνόλου δεδομένων.

Έχοντας αναλύσει την προεπεξεργασία και δειγματοληψία των δεδομένων, στην επόμενη παράγραφο παρουσιάζεται η δημιουργία των 48 εξεταζόμενων ιεραρχιών.

5.5.3 Κατασκευή Ιεραρχιών & Μέθοδοι Ιεραρχικής Πρόβλεψης

Έχοντας ολοκληρώσει την προεπεξεργασία των δεδομένων, μπορούμε να προχωρήσουμε στην κατασκευή των 48 ιεραρχιών που παρουσιάστηκαν στην παράγραφο 5.4.1. Όπως ελέχθη και στην εισαγωγή της παραγράφου 5.2, το σύνολο δεδομένων περιέχει 30490 bottom level χρονοσειρές, δηλαδή τις πωλήσεις κάθε προϊόντος σε κάθε κατάστημα του dataset.

Όπως φαίνεται και στο σχήμα 14, μαζί με τις παρατηρήσεις των πωλήσεων, στο αρχείο εισόδου παρέχεται και η πληροφορία για την οργάνωση των μονάδων σε επίπεδα συνάνθρωισης. Η πληροφορία αυτή παρέχεται σε μορφή αναγνωριστικών για το προϊόν

⁷Για περαιτέρω γενίκευση των αποτελεσμάτων και απαλοιφή της επίδρασης των calendar effects στις παρατηρήσεις, θα μπορούσαν ενδεχομένως να συγκριθούν μοντέλα που έχουν εξαχθεί και από άλλα χρονικά διαστήματα, πέρα από τις τελευταίες 13 εβδομάδες που επιλέχθηκαν στην εργασία. Στο πλαίσιο της παρούσας εργασίας δεν επιλέχθηκε μια τέτοια στρατηγική, ωστόσο αποτελεί χρήσιμη μελλοντική προέκταση.

| id | item_id | dept_id | cat_id | store_id | state_id | d_1 | d_2 | d_3 |
|-------------------------------|---------------|-----------|---------|----------|----------|-----|-----|-----|
| HOBBIES_1_001_CA_1_validation | HOBBIES_1_001 | HOBBIES_1 | HOBBIES | CA_1 | CA | 0 | 0 | 0 |
| HOBBIES_1_002_CA_1_validation | HOBBIES_1_002 | HOBBIES_1 | HOBBIES | CA_1 | CA | 0 | 0 | 0 |
| HOBBIES_1_003_CA_1_validation | HOBBIES_1_003 | HOBBIES_1 | HOBBIES | CA_1 | CA | 0 | 0 | 0 |
| HOBBIES_1_004_CA_1_validation | HOBBIES_1_004 | HOBBIES_1 | HOBBIES | CA_1 | CA | 0 | 0 | 0 |
| HOBBIES_1_005_CA_1_validation | HOBBIES_1_005 | HOBBIES_1 | HOBBIES | CA_1 | CA | 0 | 0 | 0 |

Σχήμα 14: Πρώτες 10 Γραμμές & Στήλες | *sales_train_validation.csv*

(*item_id*), το τμήμα-υποκατηγορία (*dept_id*) και την κατηγορία (*cat_id*) καθώς και για την πολιτεία (*state_id*) και το κατάστημα πώλησης (*store_id*).

Η συνάνθροιση στα 12 επιθυμητά επίπεδα του συνόλου δεδομένων που παρουσιάστηκαν στον πίνακα 9, μπορεί να υλοποιηθεί με χρήση διαφόρων τεχνικών, που σχετίζονται με διαφορετικές βιβλιοθήκες/πακέτα της γλώσσας R. Στην παρούσα εργασία, για συνάνθροιση των αποσυναθροισμένων χρονοσειρών, χρησιμοποιήθηκε η συνάρτηση `aggregate_key()` που περιλαμβάνεται στο πακέτο `fable` των O'Hara-Wild, Hyndman, and Wang. Η `aggregate_key()` δέχεται ως όρισμα ένα αλφαριθμητικό που αποτελεί το κλειδί συνάνθροισης. Συνδυάζει δηλαδή, επίπεδα (που εμφανίζονται ως στήλες στο αρχείο των δεδομένων, όπως στο σχήμα 14) με χρήση των συμβόλων "/" και "*" για να ορίσει ιεράρχηση ή ομαδοποίηση επιπέδων αντίστοιχα. Ένα παράδειγμα κλειδιού είναι το `(cat/dept/item/id)*(state/store/id)`. Με τον τρόπο αυτό θα κατασκευαστούν οι 48 ιεραρχίες που περιγράφηκαν στην παράγραφο 5.4.1.

Όπως αναλύθηκε και στην παράγραφο 5.4.2, οι σημειακές προβλέψεις βάσης θα εξαχθούν με εκθετική εξομάλυνση, που είναι υλοποιημένη ως *ETS* από τους O'Hara-Wild et al. στο `fable`. Στη συνέχεια η εναρμόνιση των προβλέψεων βάσης κάθε εξεταζόμενου δείγματος θα εναρμονίζεται με την εκάστοτε ιεραρχία με τη μέθοδο *Minimum Trace*. Η *Minimum Trace* εναρμονίζει τις προβλέψεις βάσης με όλα τα επίπεδα συνάνθροισης που μπορούν να προκύψουν από τη δήλωση των μη συναθροισμένων χρονοσειρών. Για να αξιολογηθούν διαφορετικές εναρμονίσεις, στη *Minimum Trace* θα τροφοδοτούνται δεδομένα που έχουν οργανωθεί με διαφορετικό τρόπο σε επίπεδα συνάνθροισης.

Πέρα από τη σύγκριση των προβλέψεων που έχουν εναρμονιστεί με καθεμιά εκ των εξεταζόμενων ιεραρχιών που περιγράφηκαν στην παράγραφο, θα συμπεριληφθούν και οι μέθοδοι ιεραρχικής πρόβλεψης *bottom up* και *top down*, ως μέθοδοι σημεία αναφοράς (*benchmarks*). Στόχος της παρούσας εργασίας, πέρα από τη σύγκριση των διαφορετικών ιεραρχιών μεταξύ τους, θα είναι και η απάντηση στο αν και κατά πόσο, προβλέψεις που έχουν εναρμονιστεί με *Minimum Trace*, βελτιώνουν την ακρίβεια πρόβλεψης σε σχέση με τις μεθόδους που μοντελοποιούν ένα μοναδικό επίπεδο συνάνθροισης (*single-level methods*), δηλαδή των *bottom up* και *top down*.

5.5.4 Εξαγωγή Μετρήσεων Απόδοσης

Το πρόγραμμα που αναπτύχθηκε στο πλαίσιο της παρούσας εργασίας, παράγει κατά τον τερματισμό του ένα συγκεντρωτικό αρχείο με πληροφορίες για τις προβλέψεις των διαφορετικών ιεραρχιών πάνω στα διαφορετικά επίπεδα συνάθροισης. Πιο συγκεκριμένα, αποτελεί έναν πίνακα με στήλες που περιλαμβάνουν αναγνωριστικά για την ιεραρχία (*key_name*, *key_code*) και το επίπεδο που προβλέπεται (*level*), μια σειρά από μέτρα ακρίβειας που αφορούν στην ακρίβεια των προβλέψεων⁸, το πλήθος των χρονοσειρών (*series_n*) σε κάθε ιεραρχία και το ύψος (*tree_height*) κάθε ιεραρχίας, τον χρόνο εκτέλεσης για τη μοντελοποίηση και πρόβλεψη κάθε ιεραρχίας (*exec_time*) καθώς και το μέγεθος σε MB (*mem_size*) που απαιτείται για τη μοντελοποίηση κάθε ιεραρχίας. Πέντε σειρές από το παραγόμενο, κατά την εκτέλεση του προγράμματος, αρχείο, φαίνεται στο σχήμα 15.

| key_name | key_code | RMSSE | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 | exec_time | series_n | tree_height | mem_size | level |
|------------------------|----------|--------|----------|----------|----------|---------|--------|--------|--------|-----------|----------|-------------|----------|-------|
| bottom up | 51 | 0.7200 | -93.6663 | 220.3541 | 181.9180 | -4.8155 | 7.9283 | 0.7609 | 0.4547 | 21.7 | 1530 | 2 | 103.0 | 1 |
| cat/dept/id | 13 | 0.5427 | -19.4366 | 166.0930 | 139.5384 | -1.3029 | 5.8951 | 0.5836 | 0.3860 | 20.0 | 1540 | 4 | 103.9 | 1 |
| cat/dept/item/id | 31 | 0.5490 | -24.3074 | 168.0169 | 142.0917 | -1.4821 | 6.0041 | 0.5943 | 0.3787 | 25.2 | 1715 | 5 | 116.7 | 1 |
| cat/dept/item/state/id | 46 | 0.5484 | -28.1050 | 167.8345 | 140.7432 | -1.6667 | 5.9674 | 0.5886 | 0.3824 | 40.7 | 2221 | 6 | 151.6 | 1 |
| cat/dept/state/id | 29 | 0.5446 | -7.9610 | 166.6641 | 141.5820 | -0.7491 | 5.9338 | 0.5922 | 0.3736 | 21.2 | 1561 | 5 | 105.6 | 1 |

Σχήμα 15: 5 Πρώτες Σειρές του Εξαγόμενου Αρχείου Μετρήσεων

5.6 Ανάπτυξη σε Κώδικα

Η μεθοδολογία που αναπτύχθηκε στις παραπάνω παραγράφους θα υλοποιηθεί σε ένα πρόγραμμα σε γλώσσα R. Η προεπεξεργασία και ο χειρισμός των συλλογών ιεραρχικών ή ομαδοποιημένων χρονοσειρών θα γίνει μέσα από τις βιβλιοθήκες `tidyr` και `dplyr`, ενώ η πρόβλεψη, η εναρμόνιση και ο εξαγωγή της ακριβείας και κόστους πρόβλεψης θα γίνει με χρήση της βιβλιοθήκης `fable`. Το `fable` αποτελεί μια νέα βιβλιοθήκη για ιεραρχική πρόβλεψη στην R, που έχει αναπτυχθεί από τους O'Hara-Wild, Hyndman, and Wang [2021]. Το πακέτο αυτό αποτελεί τη νέα πρόταση των συγγραφέων για ιεραρχική πρόβλεψη στην R, που έρχεται να αντικαταστήσει τα πλέον ώριμα, δημοφιλή πακέτα: `hts` [Rob J. Hyndman and Wickramasuriya, 2021] και `forecast` [Hyndman and Khandakar, 2008]. Το πακέτο παρέχει βελτιωμένο προγραμματιστικό API για την κατασκευή, μοντελοποίηση και εναρμόνιση ιεραρχικών ή ομαδοποιημένων χρονοσειρών, ωστόσο οι χρόνοι και τα μεγέθη σε μνήμη που απαιτούνται για τη μοντελοποίηση είναι σημαντικά αυξημένοι σε σχέση με τα κόστη των `hts` και `forecast`.

Μέσω της ανάπτυξης του παραπάνω προγράμματος, αυτοματοποιείται η κατασκευή, μοντελοποίηση, εναρμόνιση, πρόβλεψη και σύγκριση των εξεταζόμενων ιεραρχιών.

⁸Η παρουσίαση της εργασίας θα επικεντρωθεί στη σύγκριση του RMSSE ως μέτρο ακρίβειας πρόβλεψης, όπως αναφέρθηκε και στην παράγραφο 5.4.3. Οι υπόλοιποι στατιστικοί δείκτες συμπεριλήφθηκαν για λόγους πληρότητας της ανάλυσης.

Οι απαραίτητες πληροφορίες για την ακρίβεια και το κόστος των -εναρμονισμένων με διαφορετικές ιεραρχίες- προβλέψεων ενός ή περισσότερων επιπέδων συνάθροισης, μπορούν να εξαχθούν με αυτόματο τρόπο, και πάνω σε διαφορετικά δείγματα του συνόλου δεδομένων του *M5*. Το πρόγραμμα, πέρα από την ελευθερία στην επιλογή μεθόδου μοντελοποίησης κάθε χρονοσειράς, αλλά και μεθόδου εναρμόνισης/ιεραρχικής πρόβλεψης, παρέχει ελευθερία επιλογής των εξεταζόμενων δειγμάτων, όπως και ελευθερία στην επιλογή τόσο του υποσυνόλου των 48 διαφορετικών ιεραρχιών (παράγραφος 5.4.1) που θα συγκριθούν, όσο και στο υποσύνολο των 12 επιπέδων συνάθροισης του συνόλου δεδομένων, πάνω στα οποία θα γίνει η σύγκριση. Στόχος αυτής της σχεδιαστικής επιλογής στην ανάπτυξη του κώδικα είναι η διατήρηση αυξημένων βαθμών ελευθερίας που θα επιτρέψει την εκτέλεση διαφορετικών σεναρίων που εξαρτώνται από τις παραπάνω τρεις παραμέτρους.

Αλγόριθμος 1 Σχεδιάγραμμα Αλγορίθμου Επιλογής Βέλτιστης Ιεραρχίας

Require: data (*historical data with inherent aggregation constraints*)

Ensure: $\forall l \in \text{levels } hts_{opt}^l := \min_{error, cost} \{fc_{hts}\}$

```

1: ts ← preprocess(data)                                ▷ format, duration, filtering of data
2: samples ← {si ← sample(ts,rand(i),extra_constraints)}    ▷ set of samples
3: keys ← {ki ← combine(aggregation levels)}                ▷ set of aggregation keys
4: for s ∈ samples do
5:   for k ∈ keys do
6:     hts ← aggregation(s,k)                             ▷ aggregate sample data
7:     fitbase ← ETS(hts)                                  ▷ model time series
8:     fitMinT ← MinT(fitbase)                             ▷ reconcile models with hts constraints
9:     fc ← forecast(fitMinT, h)    ▷ produce coherent forecasts for given horizon h
10:    a ← accuracy(fc, hts)                                ▷ measure accuracy
11:    c ← cost(hts, fitbase, fitMinT)                       ▷ and cost
12:    infos ← infos + <k,a,c>                             ▷ save hts info for sample s
13:   end for
14:   info ← info + infos                                  ▷ save sample s info
15: end for
16: htsoptl ← rank(info) ▷ select optimal hts per level by ranking info of selected samples

```

Στον αλγόριθμο 1 περιγράφεται σε ψευδοκώδικα ο σκελετός του αλγορίθμου που αναπτύχθηκε για την εύρεση βέλτιστης ιεραρχίας προς πρόβλεψη. Απαιτούμενο του αλγορίθμου είναι η τροφοδότηση με δεδομένα που οργανώνονται σε επίπεδα συνάθροισης, ενώ το αποτέλεσμα είναι η επιλογή μιας ή περισσότερων ιεραρχιών ως βέλτιστη παράμετρο για τις προβλέψεις κάθε επιπέδου (hts_{opt}^l). Όπως φαίνεται, πρώτο βήμα είναι η προεπεξεργασία των δεδομένων (*data*) που περιλαμβάνει την αλλαγή της μορφής των δεδομένων σε αυτή της χρονοσειράς, μεταβολή της διάρκειας (έναρξη-λήξη) των ιστορικών παρατηρήσεων και πέρασμα ορισμένων φίλτρων για απομόνωση χρονοσειρών. Επόμενο βήμα είναι η δημιουργία του συνόλου δειγμάτων των δεδομένων (*samples*), καθώς

και του συνόλου των ιεραρχιών που θα εξεταστούν, μέσω των αναγνωριστικών κάθε αποδεκτής ιεραρχίας επιπέδων (*keys*). Στη συνέχεια, για κάθε συνδυασμό δείγματος s και εξεταζόμενης ιεραρχίας επιπέδων συνάθροισης με κλειδί k , δημιουργείται η αντίστοιχη συλλογή ιεραρχικών χρονοσειρών, μοντελοποιούνται με ETS (fit_{base}), εναρμονίζονται με MinT (fit_{MinT}), προβλέπονται για ορίζοντα πρόβλεψης h και εξάγονται πληροφορίες για την ακρίβεια (a) και το κόστος (c) πρόβλεψης των επιλεγμένων επιπέδων συνάθροισης. Η εύρεση της ή των βέλτιστων ιεραρχιών για πρόβλεψη, γίνεται μέσω της σύγκρισης των προβλέψεων των εξεταζόμενων ιεραρχιών πάνω στο σύνολο όλων των δειγμάτων.

Ατυχείς Προβλέψεις #7

«Η αλήθεια είναι πως καμιά διαδικτυακή βάση δεδομένων
δε θα αντικαταστήσει την καθημερινή εφημερίδα.»

Clifford Stoll, «The Internet? Bah!» 1995

6 Αποτελέσματα

Στο κεφάλαιο αυτό, παρουσιάζονται τα αποτελέσματα αξιολόγησης της προβλεπτικής αξίας των ιεραρχιών που προκύπτουν στη μελέτη περίπτωσης του συνόλου δεδομένων του δι-αγωνισμού *M5*. Ως προβλεπτική αξία, εννοείται ένας συνυπολογισμός της βελτίωσης ακρίβειας και του κόστους της μοντελοποίησης. Η πειραματική διάταξη που ορίστηκε στην παράγραφο 5.4, αντιμετωπίζει ως σταθερές του πειράματος τόσο τη μέθοδο ανεξάρτητης μοντελοποίησης των χρονοσειρών στα επίπεδα των εξεταζόμενων ιεραρχιών (εκθετική εξομάλυνση - *ETS*), όσο και τη μέθοδο εναρμόνισης των ανεξάρτητων προβλέψεων με τους περιορισμούς συνάθροισης σε κάθε εξεταζόμενη ιεραρχία (*Minimum Trace*)⁹. Με τον τρόπο αυτό, ο μόνος παράγοντας που καθορίζει την ακρίβεια και το κόστος των προβλέψεων θα είναι τα χαρακτηριστικά των εξεταζόμενων ιεραρχιών.

Η παραπάνω θεώρηση δίνει ρόλο υποκειμένου στις εξεταζόμενες ιεραρχίες, που υπονοεί ενεργητική κλίση. Έτσι, στο υπόλοιπο της παρούσας εργασίας, μια ιεραρχία θα μπορεί να «προβλέπει» ένα ή περισσότερα επίπεδα συνάθροισης, να «κοστίζει» περισσότερο ή λιγότερο από μία άλλη και γενικότερα να «αποδίδει» με τρόπο που θα επιτρέπει την ποσοτική σύγκρισή της με άλλες και τελικά την κατάταξη της. Στις παραγράφους που ακολουθούν:

1. παρουσιάζονται τα αποτελέσματα από τις εκτελέσεις του script (5.6) για επιλεγμένα samples του συνόλου δεδομένων του *M5* (5.2) και γίνεται μια ανάλυση ακρίβειας/κόστους πρόβλεψης ανάμεσα στις εξεταζόμενες ιεραρχίες.
2. αναλύονται τα ευρήματα που προκύπτουν από τη σύγκριση απόδοσης των προβλέψεων που έχουν εναρμονιστεί με τις εξεταζόμενες ιεραρχίες, τόσο μεταξύ τους όσο και με τα benchmarks των bottom up και top down μεθόδων.
3. επισημαίνονται τα συμπεράσματα που εξήχθησαν από την ανάλυση της μελέτης περίπτωσης και προτείνονται ιδέες για μελλοντικές προεκτάσεις πάνω στην ανάπτυξη μεθοδολογιών για ενίσχυση πρόβλεψης επιπέδων συνάθροισης, με αυτόματη επιλογή βέλτιστης σε απόδοση ιεραρχίας για εναρμόνιση προβλέψεων.

Τα δεδομένα που παράγονται κατά την εκτέλεση του προγράμματος, επιτρέπουν τη διεξαγωγή μιας ανάλυσης ακρίβειας-κόστους ανάμεσα στις προβλέψεις που έχουν εξαχθεί με καθεμιά από τις 50 επιλογές ιεραρχικής πρόβλεψης (*Minimum Trace* με 48 διαφορετικές ιεραρχίες, bottom up, top down με μέσα ιστορικά ποσοστά), συγκρινόμενες πάνω στα 12 επίπεδα συνάθροισης που ορίζονται στο σύνολο δεδομένων. Η αξιολόγηση των 50 επιλογών θα γίνει πάνω σε ολόκληρη την ιεραρχία, αλλά και πάνω σε υποσύνολα των 12 επιπέδων συνάθροισης και θα αφορά τόσο την ακρίβεια όσο και το κόστος πρόβλεψης σε κάθε ομάδα αξιολόγησης. Πιο συγκεκριμένα, η παρουσίαση και ανάλυση των αποτελεσμάτων θα αναφερθεί στη σύγκριση των 50 επιλογών πρόβλεψης πάνω:

⁹Πέρα από τη *Minimum Trace* εναρμόνιση για τις 48 διαφορετικές ιεραρχίες, θα προστεθούν και οι single-level μέθοδοι ιεραρχικής πρόβλεψης top down και bottom up ως μέθοδοι σημεία-αναφοράς(benchmarks).

- Στη συνολική απόδοση της πρόβλεψης όλων των (12) επιπέδων συνάθροισης. Το μέτρο ακριβείας προβλέψεων που επιλέχθηκε για την παρουσίαση (*RMSSE*) επιτρέπει το συνδυασμό των σφαλμάτων χρονοσειρών με διαφορετικές στάθμες, συνεπώς μπορεί να εξαχθεί η μέση ακρίβεια των προβλέψεων ανάμεσα σε επίπεδα όπως για παράδειγμα οι συνολικές πωλήσεις (*level 1*) και οι πωλήσεις μονάδων ανά προϊόν και κατάσταση (*level 12*).
- Στην απόδοση της πρόβλεψης επιπέδων υψηλής συνάθροισης: συνολικές πωλήσεις (επίπεδο 1), πωλήσεις ανά πολιτεία (επίπεδο 2), κατάσταση (επίπεδο 3), κατηγορία προϊόντων (επίπεδο 4) και τμήμα (ή υποκατηγορία προϊόντων) (επίπεδο 5).
- Στην απόδοση της πρόβλεψης επιπέδων μεσαίας συνάθροισης: πωλήσεις ανά πολιτεία και κατηγορία (επίπεδο 6), ανά πολιτεία και υποκατηγορία (επίπεδο 7), ανά κατάσταση και κατηγορία (επίπεδο 8) και ανά κατάσταση και τμήμα (επίπεδο 9).
- Στην απόδοση της πρόβλεψης επιπέδων χαμηλής συνάθροισης: πωλήσεις ανά προϊόν (επίπεδο 10), ανά προϊόν και πολιτεία (επίπεδο 11) και ανά προϊόν και κατάσταση (επίπεδο 12) που αποτελεί το επίπεδο των μη συναθροισμένων χρονοσειρών (*bottom level*).

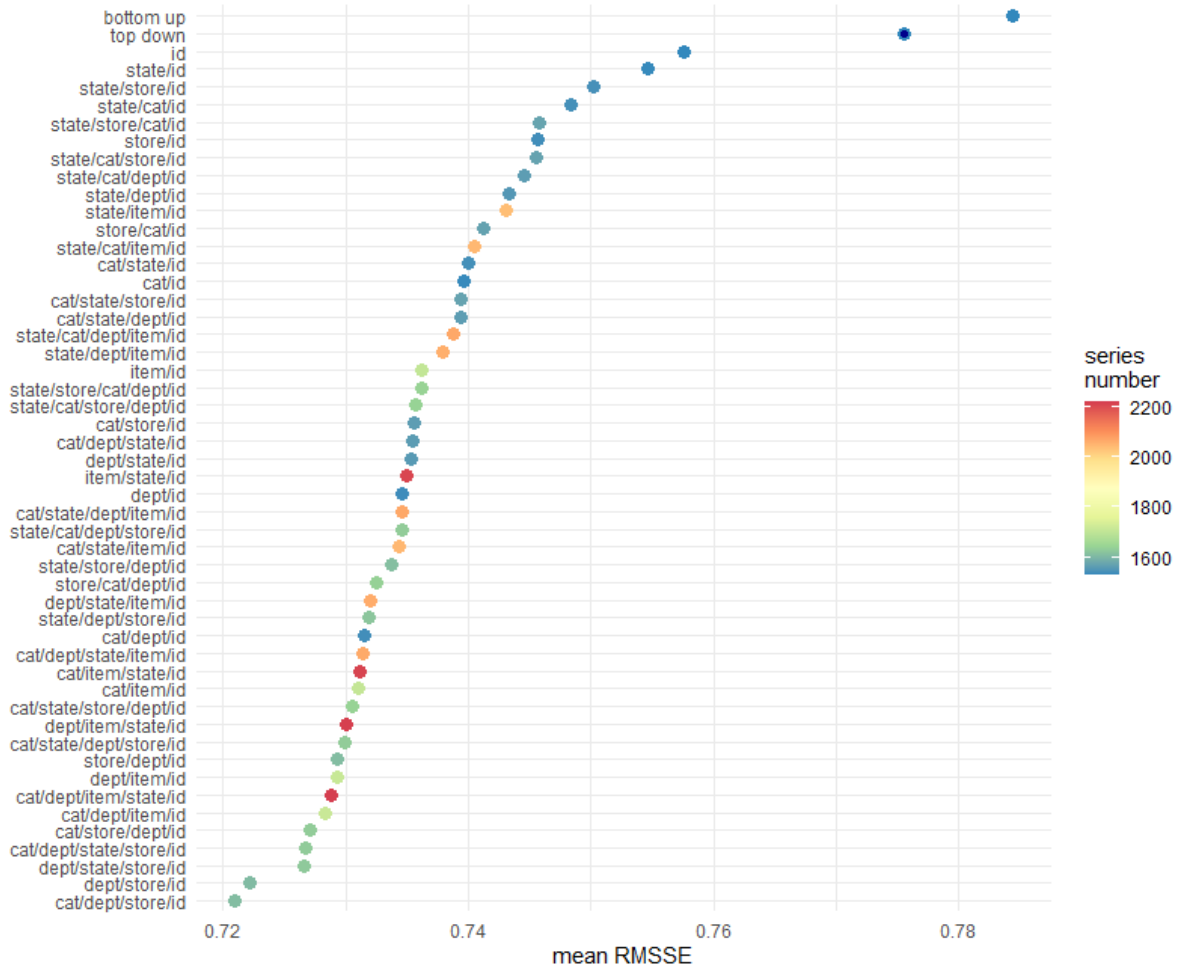
Πέρα από τα παραπάνω, παρουσιάζεται και μια συσταδοποίηση ακριβείας και κόστους πρόβλεψης ιεραρχιών, με χρήση μηχανικής μάθησης (*k-nearest neighbours*). Από μια τέτοια διαδικασία, μπορούν να συγκρίνονται ιεραρχίες, χωρίς να είναι απαραίτητη η εκ των προτέρων πληροφορία για κάθε ιεραρχία, όπως το ποια επίπεδα συνάθροισης περιέχει και πόσες χρονοσειρές υπάρχουν στο καθένα από αυτά. Τα δεδομένα πάνω στα οποία θα διεξαχθεί η ανάλυση, προκύπτουν κατά την εκτέλεση του προγράμματος που αυτοματοποιεί τη μεθοδολογία που αναλύθηκε στην παράγραφο 5.5, για τα επιλεγμένα δείγματα (*samples*) του συνόλου δεδομένων.

6.1 Μέση Απόδοση σε όλα τα Επίπεδα Συνάθροισης

Σε αυτή την παράγραφο παρουσιάζονται οι συγκρίσεις της συνολικής απόδοσης (μέσης ακριβείας και κόστους) των εξεταζόμενων ιεραρχιών πάνω σε όλα τα επίπεδα του συνόλου δεδομένων, όπως προκύπτουν από τις εκτελέσεις του προγράμματος (παράγραφος 5.6) για επιλεγμένα δείγματα του συνόλου δεδομένων. Το μέτρο ακριβείας που επιλέχθηκε (παράγραφος 5.4.3) για τη σύγκριση των ιεραρχιών, δηλαδή ο *RMSSE*, επιτρέπει το συνδυασμό σφαλμάτων που προέρχονται από επίπεδα με διαφορετικές στάθμες. Έτσι, μπορούν να εξαχθούν αριθμητικοί μέσοι των σφαλμάτων μεταξύ των ακριβειών στα επιμέρους επίπεδα συνάθροισης (*mean RMSSE*), ώστε να προκύπτει μια μέση ακρίβεια πρόβλεψης κάθε ιεραρχίας στο σύνολο των επιπέδων του συνόλου δεδομένων, από την οποία θα εξαχθούν οι επιθυμητές κατατάξεις. Στις επόμενες δύο παραγράφους παρουσιάζονται οι κατατάξεις των ιεραρχιών ως προς την ακρίβεια πρόβλεψης στα 12 επίπεδα συνάθροισης, αλλά και το κόστος για τη μοντελοποίηση των χρονοσειρών που περιέχει η καθιεμία.

6.1.1 Κατάταξη ως προς την Ακρίβεια Πρόβλεψης

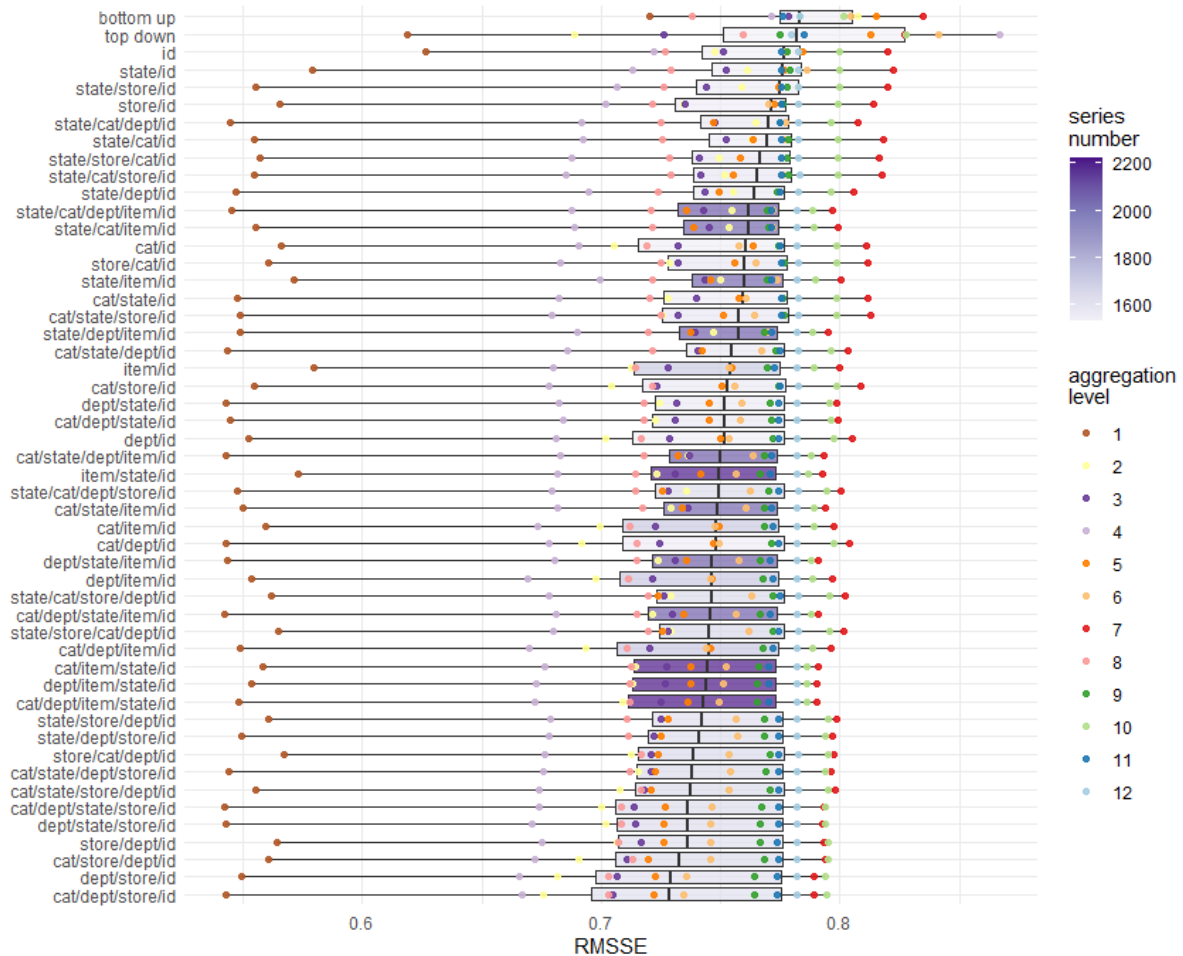
Στο σχήμα 16 παρουσιάζεται η κατάταξη των 50 εξεταζόμενων επιλογών πρόβλεψης ως προς τη μέση ακρίβεια πρόβλεψης ($mean\ RMSSE$) σε όλα τα επίπεδα συνάθροισης και στο σχήμα 17 παρουσιάζονται τα boxplots της ακρίβειας πρόβλεψης κάθε εξεταζόμενης μεθόδου στα 12 επίπεδα, με την ίδια σειρά κατάταξης. Η χρωματική κλίμακα αναφέρεται στο πλήθος των χρονοσειρών που περιέχει κάθε εξεταζόμενη επιλογή πρόβλεψης και στα δύο σχήματα, ενώ στο σχήμα 17 χρωματίζονται διαφορετικά οι προβλέψεις για κάθε επίπεδο συνάθροισης. Η κατάταξη ως προς τη μέση ακρίβεια πρόβλεψης των 50 εξεταζόμενων



Σχήμα 16: Κατάταξη των 50 Επιλογών Πρόβλεψης ως προς τη Μέση Ακρίβεια
 Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών - 25 Προϊόντα Ανά Υποκατηγορία

μεθόδων που προκύπτει (σχήμα 16), αναλύεται στο σχήμα 17 στις επιμέρους προβλέψεις των εξεταζόμενων μεθόδων στα 12 επίπεδα συνάθροισης του συνόλου δεδομένων του $M5$. Σε αυτή την εικόνα των αποτελεσμάτων γίνεται αντιληπτό, πως οι κατατάξεις ακρίβειας των ιεραρχιών για τις επιμέρους προβλέψεις των 12 επιπέδων συνάθροισης, παρουσιάζει διαφορές σε σχέση με την κατάταξη της μέσης ακρίβειας των προβλέψεων των εξεταζόμενων ιεραρχιών σε όλα τα επίπεδα, όπως προκύπτει από το μέσο RMSSE. Οι προβλέψεις δηλαδή

του κάθε επιπέδου, που φαίνονται στο σχήμα ως ομόχρωμα σημεία, δεν κατατάσσονται με τη σειρά των ιεραρχιών, όπως ορίζεται στον άξονα y. Αυτό σημαίνει πως χρειάζεται μια επιμέρους ανάλυση των προβλέψεων σε κάθε επίπεδο ή ομάδα επιπέδων για εύρεση βέλτιστης ιεραρχίας. Σημαίνει επίσης, πως αν ως μέση ακρίβεια πρόβλεψης οριζόταν ένας σταθμισμένος μέσος (αντί απλού) που να πριμοδοτεί ορισμένα επίπεδα ενδιαφέροντος ή αναλογικά με τη στάθμη, οι κατατάξεις ως προς την ακρίβεια θα παρουσίαζαν πιθανές μεταβολές. Σημειώνεται εδώ πως στο σχήμα 17, όπως αναφέρεται και στον τίτλο του,



Σχήμα 17: Ανάλυση της Κατάταξης του Σχ. 16 στις Ακρίβειες Πρόβλεψης των Ιεραρχιών πάνω στα 12 Επίπεδα Συνάθροισης

Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών - Boxplots χωρίς Outliers

οι προβλέψεις του επιπέδου 1, ήταν εκτός του $1,5 \cdot IQR^{10}$. Αποφασίστηκε να συμπεριληφθούν, καθώς ο αποκλεισμός τους από τη μέση τιμή του boxplot επηρέαζε την κατάταξη. Με την επιλογή που έγινε, οι κατατάξεις των σχημάτων 16 και 17 πλέον ταυτίζονται.

¹⁰ *Inter-Quartile Range (IQR)* ενός δείγματος είναι η απόσταση ανάμεσα στο 25ο και το 75ο εκατοστημόριο των δεδομένων. Τα boxplot κατά σύμβαση περιλαμβάνουν στο «κουτί» όσες παρατηρήσεις ανήκουν στο *IQR*, ενώ στην ευθεία του boxplot ανήκει το $1,5 \cdot IQR$

Η παραπάνω μεταβολή της σύμβασης στα boxplots που απεικονίζονται στο σχήμα 17, αναφέρεται για αποφυγή σύγχυσης από τη μεριά του αναγνώστη.

Στον πίνακα 10 παρουσιάζονται τα σφάλματα και το κόστος για την ιεραρχία με την καλύτερη μέση ακρίβεια, που ήταν η «cat/dept/store/id» με μέσο RMSSE 0.721, την ιεραρχία με το υψηλότερο κόστος πρόβλεψης, που ήταν η «cat/dept/item/state/id» με 2221 χρονοσειρές, καθώς και τα 2 benchmarks: bottom up, top down.

| Μέθοδος \ Απόδοση | Σφάλμα | | Κόστος | | | Επίπεδα προς Αξιολόγηση |
|------------------------|-------------------|----------------------|---------------|------------|----------------|-------------------------|
| | <i>mean RMSSE</i> | <i>Relmean RMSSE</i> | Πλήθος Σειρών | Μνήμη [MB] | Χρόνος [λεπτά] | |
| cat/dept/store/id | 0.721 | 0.920 | 1610 | 109.3 | 22.3 | 1 έως 12 |
| cat/dept/item/state/id | 0.729 | 0.930 | 2221 | 151.6 | 40.7 | 1 έως 12 |
| bottom up | 0.784 | 1 | 1531 | 103 | 21.7 | 1 έως 12 |
| top down | 0.776 | 0.990 | 1 | | | 1 έως 12 |

Πίνακας 10: Απόδοση Επιλεγμένων Μεθόδων σε όλα τα Επίπεδα Συνάθροισης

Γίνεται φανερό αφενός η βελτίωση της ακρίβειας που πετυχαίνει η *Minimum Trace* σε σχέση με τα benchmarks, όσο και το γεγονός ότι οι 38% παραπάνω χρονοσειρές της *cat/dept/item/state/id* δεν οδήγησαν σε βελτίωση της μέσης ακρίβειας πρόβλεψης, σε σχέση με την *cat/dept/store/id*. Όπως φαίνεται και στα διαγράμματα, η δεύτερη σε ακρίβεια ιεραρχία *dept/store/id* είχε μικρή διαφορά από την *cat/dept/store/id*. Το δείγμα, του οποίου τα αποτελέσματα φαίνονται στα σχήματα 16 και 17, περιέχει 1530 bottom-level χρονοσειρές, οργανωμένες σε 25 προϊόντα ανά υποκατηγορία, και έχουν φιλτραριστεί 220 χρονοσειρές που περιείχαν λιγότερες από 8 μη μηδενικές παρατηρήσεις.

- Μια πρώτη παρατήρηση σε σχέση με την κατάταξη του σχήματος 16 είναι πως η *Minimum Trace* μέθοδος εναρμόνισης προβλέπει με μεγαλύτερη ακρίβεια -για καθεμιά εκ των 48 ιεραρχιών- σε σχέση με τις single-level μεθόδους bottom up και top down. Το παραπάνω αποτέλεσμα εντοπίστηκε στις μετρήσεις των περισσότερων από τα εξεταζόμενα δείγματα. Αντίστοιχη εικόνα φαίνεται και στην κατάταξη του σχήματος 29¹¹ όπου μόνο η πιο απλή ιεραρχία (total/id) δεν ξεπερνά την top down με μέσα ιστορικά ποσοστά.
- Ακόμη μια παρατήρηση είναι πως οι ιεραρχίες που περιέχουν ορισμένα επίπεδα γεωγραφικής και ανά προϊόν οργάνωσης, τείνουν να προβλέπουν με καλύτερη μέση ακρίβεια το σύνολο των 12 επιπέδων συνάθροισης, πάνω στα οποία γίνεται η σύγκριση των επιλογών πρόβλεψης. Όπως θα αναλυθεί παρακάτω, η πιθανότητα κάποιων επιπέδων συνάθροισης να περιέχονται στις ιεραρχίες που καταλαμβάνουν τις πρώτες θέσεις στην κατάταξη της μέσης ακρίβειας προβλέψεων, είναι αυξημένη σε σχέση με τα υπόλοιπα επίπεδα.

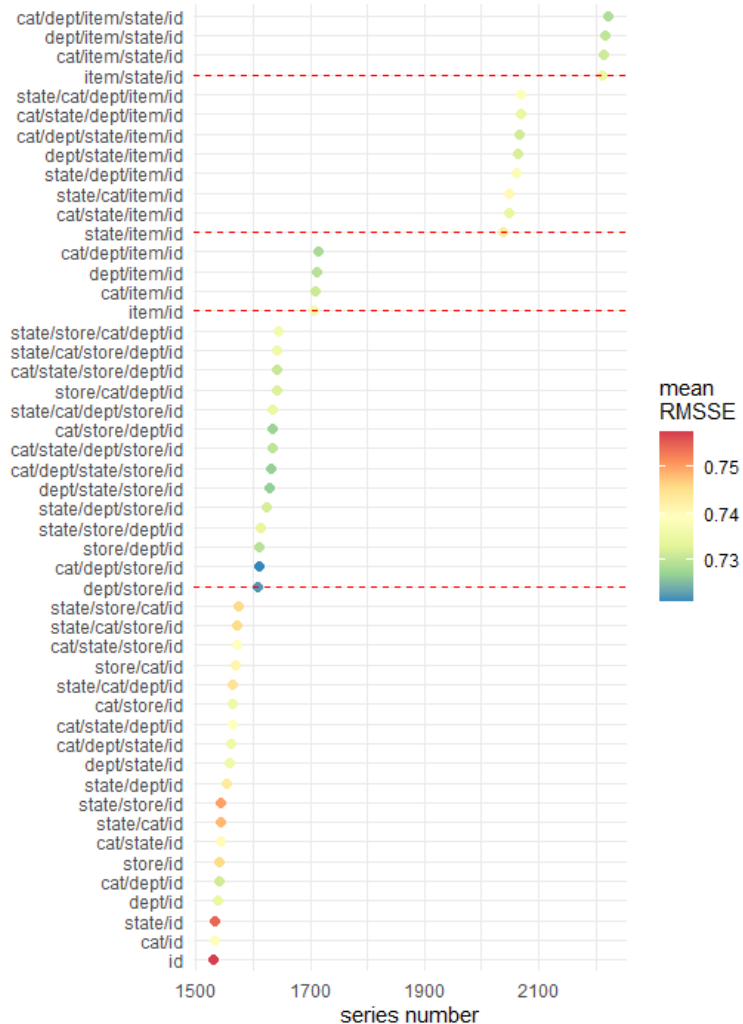
¹¹Στην παράγραφο Β του παραρτήματος της εργασίας, φαίνονται οι κατατάξεις των ιεραρχιών ως προς τη μέση ακρίβεια σε μερικά ακόμη δείγματα του συνόλου δεδομένων.

6.1.2 Κατάταξη ως προς το Κόστος Πρόβλεψης

Στην παραπάνω ενότητα, οι 50 εξεταζόμενες μέθοδοι, κατατάχθηκαν ως προς την ακρίβεια πρόβλεψης μεσοσταθμικά σε όλα τα επίπεδα του συνόλου δεδομένων. Σε αυτή την ενότητα, θα παρουσιαστεί η κατάταξή τους ως προς το κόστος πρόβλεψης. Το εξεταζόμενο δείγμα των δεδομένων περιέχει 1530 bottom-level χρονοσειρές, οργανωμένες σε 25 προϊόντα ανά υποκατηγορία, και έχουν φιλτραριστεί 220 χρονοσειρές που περιείχαν λιγότερες από 8 μη μηδενικές παρατηρήσεις.

Για την εξαγωγή των ιεραρχικών προβλέψεων, η bottom up μέθοδος θα χρειαστεί τις 1530 bottom-level χρονοσειρές. Η top down μέθοδος (με μέσα ιστορικά ποσοστά) θα χρειαστεί την πλήρως συναθροισμένη χρονοσειρά των συνολικών πωλήσεων και τα ιστορικά ποσοστά κάθε επιπέδου, όπως αυτά προκύπτουν από τα δεδομένα. Για τις 48 εξεταζόμενες ιεραρχίες, η εναρμόνιση με *Minimum Trace* προαπαιτεί τη μοντελοποίηση των χρονοσειρών σε κάθε επίπεδο συνάθροισης που περιέχεται στην ιεραρχία, ο αριθμός των οποίων θα κυμαίνεται στο διάστημα 1531 έως 2221 χρονοσειρών. Αυτό σημαίνει πως για το εξεταζόμενο δείγμα, προκύπτουν ιεραρχίες με έως και 45% περισσότερες χρονοσειρές, από αυτές στο χαμηλότερο επίπεδο. Σε ολόκληρο το σύνολο δεδομένων, οι bottom-level χρονοσειρές

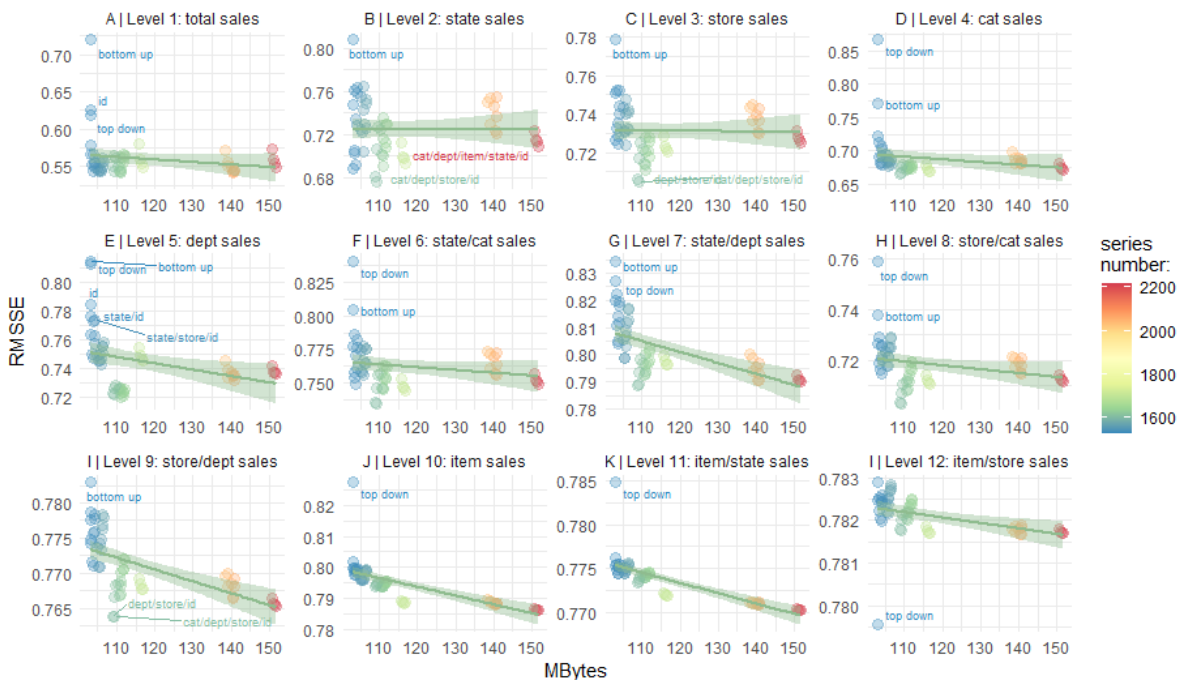
είναι 34090 και εκείνες σε όλα τα επίπεδα 42840, συνεπώς μιλάμε για μια αύξηση μέχρι και 20%. Αυτή η αύξηση των χρονοσειρών, θα οδηγήσει σε αύξηση των απαιτούμενων πόρων



Σχήμα 18: Κατάταξη των 48 Ιεραρχιών ως προς το Πλήθος των Περιεχομένων Χρονοσειρών

Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών
Χρωματική Κλίμακα για τη Μέση Ακρίβεια

ακρίβειας (*mean RMSSE*) και πλήθους χρονοσειρών¹². Στο σχήμα 19 βλέπουμε το βαθμό στον οποίο σχετίζεται η μέση ακρίβεια πρόβλεψης κάθε μεθόδου, με το πλήθος των περιεχόμενων χρονοσειρών σε κάθε ιεραρχία. Οι 48 εξεταζόμενες ιεραρχίες, μαζί με τα 2 benchmarks προέβλεψαν με ακρίβεια που κυμαίνεται στο διάστημα [0.721, 0.784]. Η ευθεία γραμμικής παλινδρόμησης που φαίνεται στο σχήμα με πράσινο χρώμα, έχει αρνητική κλίση, άρα υποδεικνύει μια σχέση μείωσης σφάλματος πρόβλεψης ως προς την αύξηση των χρονοσειρών που περιέχει κάθε ιεραρχία. Ωστόσο, όπως γίνεται φανερό, οι ακριβέστερες προβλέψεις δεν προκύπτουν από τις ιεραρχίες που περιέχουν τις περισσότερες χρονοσειρές.

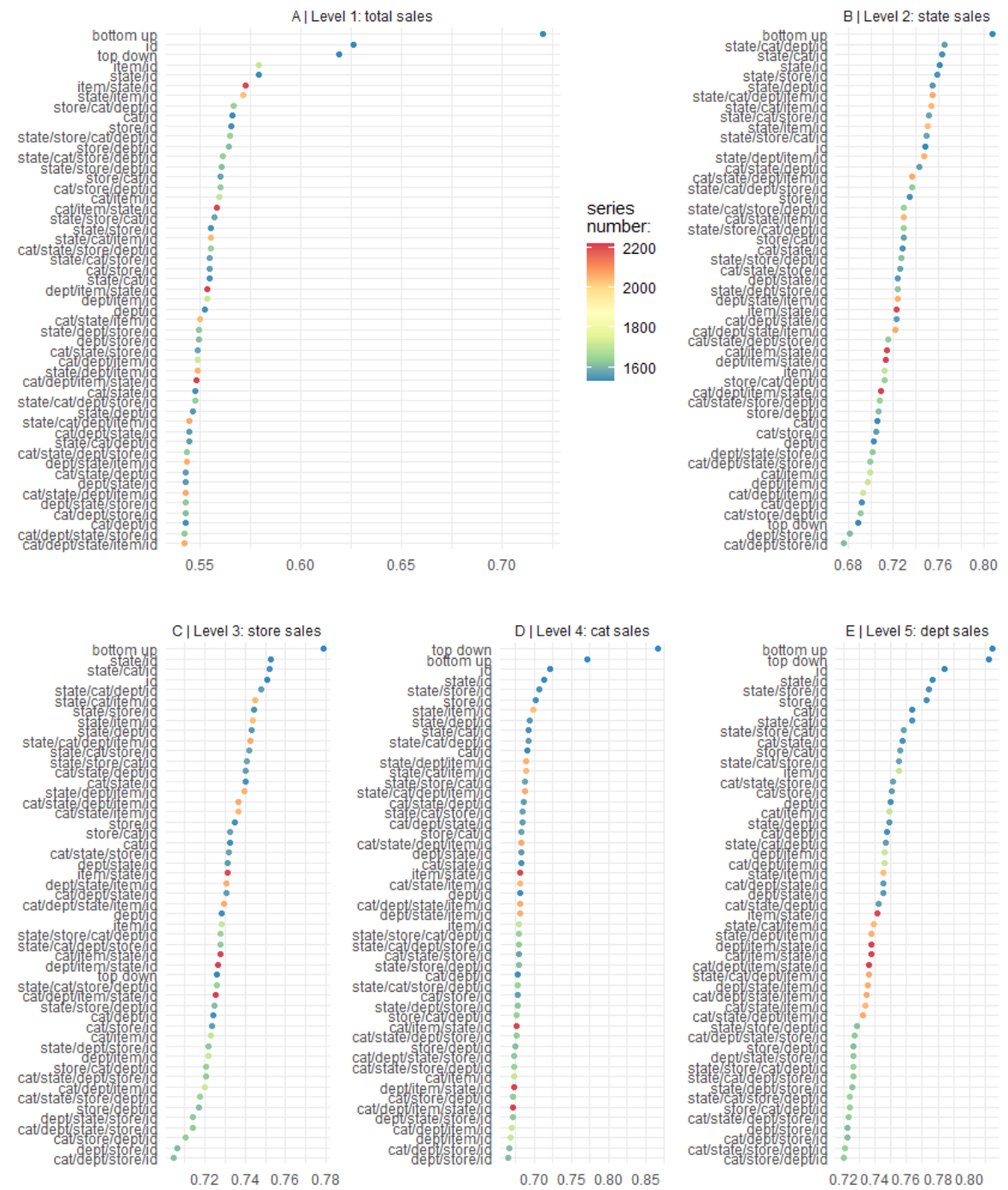


Σχήμα 20: Συσχέτιση Μέσης Ακρίβειας & Κόστους σε Μνήμη - Ανάλυση στα 12 Επίπεδα Συνάθροισης

Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών - Χρωματική Κλίμακα για το Πλήθος των Περιεχόμενων Χρονοσειρών κάθε Ιεραρχίας

Παρόμοια εικόνα παρουσιάζεται και όταν αναλυθεί η σχέση μέσης ακρίβειας και κόστους στα 12 επίπεδα συνάθροισης. Όπως φαίνεται και στο σχήμα 20, η αύξηση του κόστους πρόβλεψης (στην προκειμένη περίπτωση κόστος σε μνήμη) δείχνει να βελτιώνει μεσοσταθμικά την ακρίβεια στα περισσότερα επίπεδα, όπως δηλώνει η αρνητική κλίση της ευθείας παλινδρόμησης. Ωστόσο, η συστάδα ιεραρχιών που περιέχει την ιεραρχία με την καλύτερη μέση ακρίβεια, χρειάζεται μεσοσταθμικά το 73% της μνήμης που απαιτεί η πολυπληθέστερη σε χρονοσειρές ιεραρχία.

¹²Το πλήθος των χρονοσειρών που θα χρειαστεί να προβλεφθούν σε κάθε ιεραρχία, όπως θα δούμε και στην παράγραφο 6.3, σχετίζεται γραμμικά με τις μετρήσεις που έγιναν στους χρόνους αλλά και στην απαιτούμενη μνήμη.



Σχήμα 21: Κατάταξη των 48 Ιεραρχιών και των Benchmarks ως προς τη Μέση Ακρίβεια των Προβλέψεων στα 5 Υψηλότερα Επίπεδα Συνάθροισης
 Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών - Χρωματική Κλίμακα για το Πλήθος των Περιεχόμενων Χρονοσειρών σε κάθε Ιεραρχία

6.2 Μέση Απόδοση σε Ομάδες Επιπέδων Συνάθροισης

Στην παρακάτω παράγραφο παρουσιάζεται η μέση απόδοση των εξεταζόμενων μεθόδων, πάνω σε ομάδες επιπέδων υψηλής, μέσης και χαμηλής συνάθροισης.

6.2.1 Επίπεδα Υψηλής Συνάθροισης

Τα επίπεδα συνάθροισης που βρίσκονται ψηλά στην ιεραρχία, περιέχουν λίγες και ομαλές χρονοσειρές με υψηλή στάθμη, που αποτυπώνουν δηλαδή υψηλή αξία για τον αποφασίζοντα στο πλαίσιο των πωλήσεων λιανικής. Στη μελέτη περίπτωσης, ως υψηλά επίπεδα ορίστηκαν οι συνολικές πωλήσεις (επίπεδο 1), πωλήσεις ανά πολιτεία (επίπεδο 2), κατάσταση (επίπεδο 3), κατηγορία προϊόντων (επίπεδο 4) και τμήμα (ή υποκατηγορία προϊόντων) (επίπεδο 5). Τα επίπεδα αυτά περιέχουν από 1 έως 9 χρονοσειρές με υψηλή στάθμη και αισθητά μειωμένο θόρυβο σε σχέση με τα χαμηλά επίπεδα του συνόλου δεδομένων.

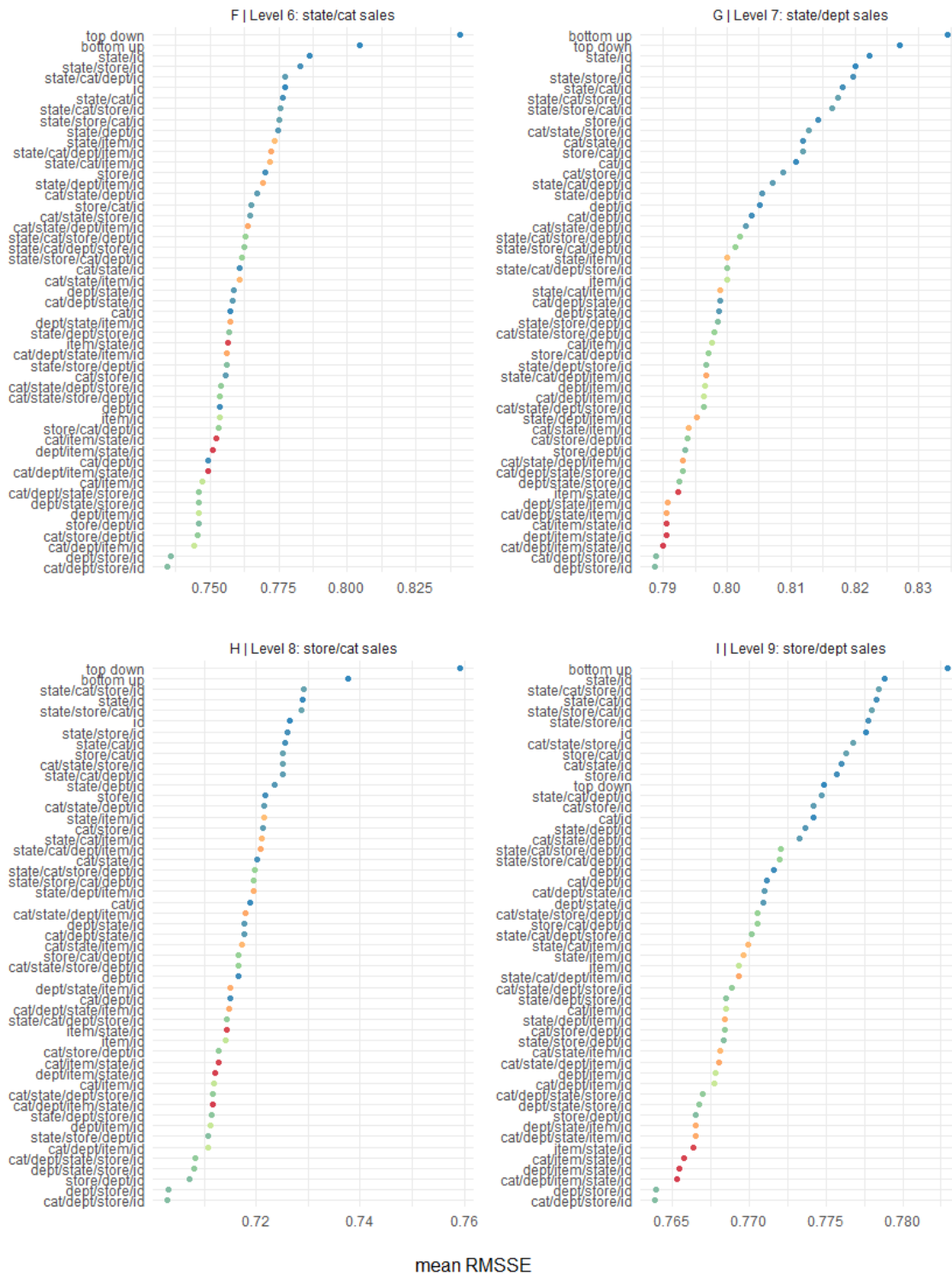
Οι προβλέψεις αυτών των επιπέδων είναι συνήθως κρισιμότερες από ό,τι χαμηλότερων επιπέδων. Οι προβλέψεις ανά πολιτεία ή ανά κατηγορία προϊόντων δρομολογούν αντίστοιχες δαπάνες και προγραμματισμό. Συνεπώς, το σφάλμα των προβλέψεων θα μεταφραστεί σε υψηλότερο κόστος από ό,τι το σφάλμα σε χαμηλότερα επίπεδα. Για το λόγο αυτό, η πρόβλεψη ενός υψηλού επιπέδου συνάθροισης όπως οι πωλήσεις της κατηγορίας Τροφίμων ή της πολιτείας της Καλιφόρνια, συνδέεται συνήθως με κάποια διοικητική θέση/τμήμα της επιχείρησης.

Στο σχήμα 21 παρουσιάζονται οι κατατάξεις των εξεταζόμενων ιεραρχιών, πάνω στα 5 υψηλότερα επίπεδα συνάθροισης, όπως ορίστηκαν παραπάνω. Σταθερή παραμένει η βελτιωμένη ακρίβεια των ιεραρχιών με την *Minimum Trace* σε σχέση με τα benchmarks, όπως παρουσιάστηκε στο σχήμα 16 πάνω σε όλα τα επίπεδα συνάθροισης. Η ιεραρχία *cat/dept/store/item/id* προέβλεψε με υψηλότερη ακρίβεια το επίπεδο 1, ενώ οι ιεραρχίες *cat/dept/store/id* και *dept/store/id* -που είχαν και την καλύτερη μέση ακρίβεια στα 12 επίπεδα- προέβλεψαν καλύτερα τα επίπεδα 2 έως 5.

| Μέθοδος \ Απόδοση | Σφάλμα | | Κόστος | | | Επίπεδα προς Αξιολόγηση |
|------------------------|-------------------|----------------------|---------------|------------|----------------|-------------------------|
| | <i>mean RMSSE</i> | <i>Relmean RMSSE</i> | Πλήθος Σειρών | Μνήμη [MB] | Χρόνος [λεπτά] | |
| cat/dept/store/id | 0.662 | 0.851 | 1610 | 109.3 | 22.3 | 1 έως 5 |
| cat/dept/item/state/id | 0.678 | 0.871 | 2221 | 151.6 | 40.7 | 1 έως 5 |
| bottom up | 0.778 | 1 | 1530 | 103 | 21.7 | 1 έως 5 |
| top down | 0.742 | 0.954 | 1 | | | 1 έως 5 |

Πίνακας 11: Απόδοση Επιλεγμένων Μεθόδων στα Επίπεδα Συνάθροισης 1-5

Στον πίνακα 11 παρουσιάζεται η ακρίβεια και το κόστος πρόβλεψης για την ακριβέστερη ιεραρχία στα 5 υψηλότερα επίπεδα *cat/dept/store/id*, την πιο κοστοβόρα *cat/dept/item/state/id* καθώς και τα 2 benchmarks. Παρατηρείται συνέπεια στο ότι η ακριβέστερη ιεραρχία περιέχει αισθητά λιγότερες χρονοσειρές από την πολυπληθέστερη ιεραρχία, η οποία προβλέπει με καλή ακρίβεια, αλλά βρίσκεται εκτός των κορυφαίων σε ακρίβεια.



Σχήμα 22: Κατάταξη των 48 Ιεραρχιών και των Benchmarks ως προς τη Μέση Ακρίβεια των Προβλέψεων στα 4 Μεσαία Επίπεδα Συνάθροισης

6.2.2 Επίπεδα Μέσης Συνάθροισης

Τα 4 αυτά επίπεδα περιέχουν από 6 έως 70 χρονοσειρές που παραμένουν ομαλές λόγω της υψηλής συνάθροισης που διατηρούν. Περιλαμβάνουν χρονοσειρές για τις πωλήσεις ανά πολιτεία και κατηγορία (επίπεδο 6), ανά πολιτεία και υποκατηγορία (επίπεδο 7), ανά κατάσταση και κατηγορία (επίπεδο 8) και ανά κατάσταση και υποκατηγορία (επίπεδο 9).

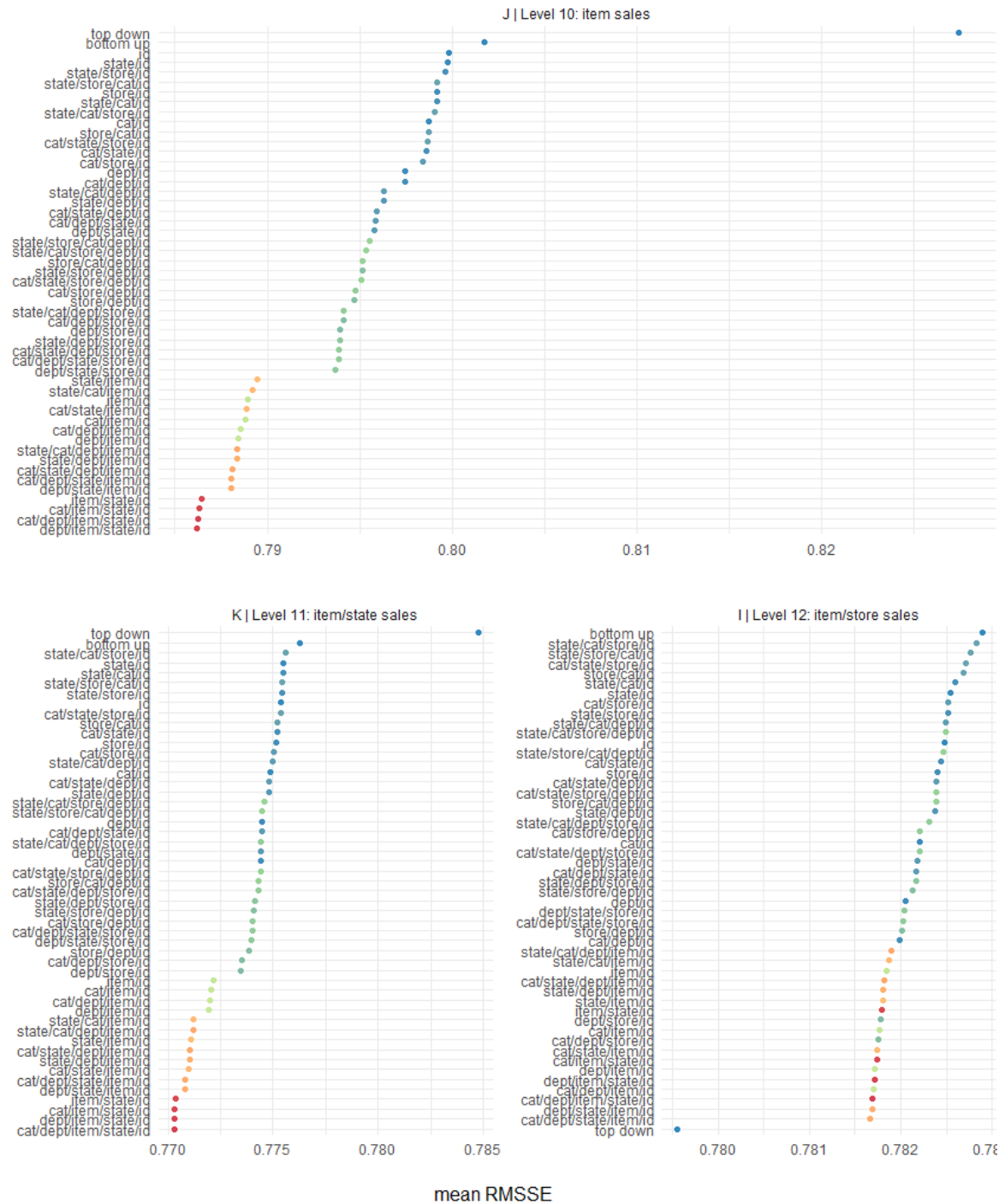
Και σε αυτή την ομάδα επιπέδων συνάθροισης, όπως και στα υψηλά επίπεδα, είναι χαρακτηριστική η υπεροχή ως προς την ακρίβεια, ιεραρχιών που περιέχουν τα επίπεδα των πωλήσεων ανά υποκατηγορία προϊόντων και ανά κατάσταση. Ταυτόχρονα, όταν αυτές οι ιεραρχίες δεν περιλαμβάνουν το επίπεδο 10, δηλαδή τις χρονοσειρές πωλήσεων κάθε προϊόντος σε όλα τα καταστήματα (*item*), τότε περιέχουν αισθητά λιγότερες χρονοσειρές με αποτέλεσμα να μειώνεται σημαντικά το κόστος πρόβλεψης που περιλαμβάνει τη μοντελοποίηση των χρονοσειρών και την εναρμόνιση με την εκάστοτε ιεραρχία.

| Μέθοδος \ Απόδοση | Σφάλμα | | Κόστος | | | Επίπεδα προς Αξιολόγηση |
|------------------------|-------------------|----------------------|---------------|------------|----------------|-------------------------|
| | <i>mean RMSSE</i> | <i>Relmean RMSSE</i> | Πλήθος Σειρών | Μνήμη [MB] | Χρόνος [λεπτά] | |
| cat/dept/store/id | 0.742 | 0.937 | 1610 | 109.3 | 22.3 | 6 έως 9 |
| cat/dept/item/state/id | 0.750 | 0.947 | 2221 | 151.6 | 40.7 | 6 έως 9 |
| bottom up | 0.792 | 1 | 1530 | 103 | 21.7 | 6 έως 9 |
| top down | 0.809 | 1.021 | 1 | | | 6 έως 9 |

Πίνακας 12: Απόδοση Επιλεγμένων Μεθόδων στα Μεσαία Επίπεδα 6-9

Στον πίνακα 12 φαίνεται η μέση απόδοση της ιεραρχίας με την υψηλότερη ακρίβεια στα μεσαία επίπεδα, ανάμεσα στις 48 εξεταζόμενες ιεραρχίες, που είναι η *cat/dept/store/id*. Όπως προαναφέρθηκε, ιεραρχίες όπως οι *dept/store/id* αλλά και άλλες παρήγαγαν προβλέψεις με μέση ακρίβεια που διαφέρει στα όρια του στατιστικού λάθους με την πρώτη. Συνεπώς οι βέλτιστες προβλέψεις μπορούν να προκύπτουν από μία ή περισσότερες ιεραρχίες, που έχουν όμως παρόμοια χαρακτηριστικά. Μαζί με την απόδοση της *cat/dept/store/id*, στον πίνακα 12 παρουσιάζονται οι αποδόσεις της ιεραρχίας με το υψηλότερο κόστος μοντελοποίησης (*cat/dept/item/state/id*), καθώς και των benchmarks bottom up και top down.

Οι προβλέψεις που έχουν εναρμονιστεί με μία από τις εξεταζόμενες ιεραρχίες, παραμένουν ακριβέστερες από τα 2 benchmarks και στις προβλέψεις των μεσαίων επιπέδων του συνόλου δεδομένων. Όπως φαίνεται και στο σχήμα 22, οι μέθοδοι bottom up και top down εναλλάσσονται στις τελευταίες θέσεις της κατάταξης, πέρα από το επίπεδο 9, δηλαδή τις πωλήσεις ανά κατάσταση και υποκατηγορία προϊόντων, στο οποίο η top down ξεπέρασε σε ακρίβεια τις ιεραρχίες που περιέχουν αποκλειστικά τα επίπεδα των κατηγοριών, των πολιτειών και των καταστημάτων. Παρά τη σχετική κατάταξη των μεθόδων, ήδη από το επίπεδο 9 γίνεται φανερή η μείωση των διακυμάνσεων των ακριβειών, που εντείνεται στα επίπεδα χαμηλά στην ιεραρχία. Οι προβλέψεις των 50 επιλογών, βρίσκονται στο διάστημα [0.76, 0.79] ως προς τον RMSSE.



Σχήμα 23: Κατάταξη των 48 Ιεραρχιών και των Benchmarks ως προς τη Μέση Ακρίβεια των Προβλέψεων στα 3 Χαμηλότερα Επίπεδα Συνάθροισης
 Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών - Χρωματική Κλίμακα για το Πλήθος των Περιεχομένων Χρονοσειρών σε κάθε Ιεραρχία

6.2.3 Επίπεδα Χαμηλής Συνάθροισης

Τελευταία ομάδα συνόλων πάνω στην οποία θα συγκριθούν οι προβλέψεις που έχουν εναρμονιστεί με τις 48 εξεταζόμενες ιεραρχίες, είναι τα επίπεδα που βρίσκονται χαμηλά στην ιεραρχία. Τα μη συναθροισμένα ή χαμηλής συνάθροισης επίπεδα, αποτελούν συνήθως τη μεγάλη πλειοψηφία των χρονοσειρών στις συλλογές ομαδοποιημένων χρονοσειρών. Παρουσιάζουν υψηλό θόρυβο, και στην περίπτωση του εξεταζόμενου συνόλου υψηλή διακοπτότητα (*high intermitency*). Τα χαμηλά επίπεδα περιλαμβάνουν πωλήσεις ανά προϊόν (επίπεδο 10), ανά προϊόν και πολιτεία (επίπεδο 11) και ανά προϊόν και κατάσταση (επίπεδο 12) που αποτελεί το bottom level επίπεδο με τις μη συναθροισμένες χρονοσειρές.

Σταθερή παρατήρηση που προέκυψε από τα αποτελέσματα εκτελέσεων διαφορετικών δειγμάτων του συνόλου δεδομένων, όπως φαίνεται και στο σχήμα 23, είναι πως οι αποκλίσεις στην ακρίβεια των προβλέψεων για τα χαμηλά επίπεδα συνάθροισης που έχουν εναρμονιστεί με τις 48 εξεταζόμενες ιεραρχίες είναι αισθητά χαμηλότερες από τις αποκλίσεις σε υψηλότερα επίπεδα. Στον πίνακα 13 παρουσιάζονται -κατά σύμβαση με τα

| Μέθοδος \ Απόδοση | Σφάλμα | | Κόστος | | | Επίπεδα προς Αξιολόγηση |
|------------------------|-------------------|----------------------|---------------|------------|----------------|-------------------------|
| | <i>mean RMSSE</i> | <i>Relmean RMSSE</i> | Πλήθος Σειρών | Μνήμη [MB] | Χρόνος [λεπτά] | |
| dept/item/state/id | 0.779 | 0.989 | 2218 | 151.3 | 22.3 | 10 έως 12 |
| cat/dept/item/state/id | 0.780 | 0.991 | 2221 | 151.6 | 40.7 | 10 έως 12 |
| dept/store/id | 0.783 | 0.995 | 1607 | 109 | 22.7 | 10 έως 12 |
| bottom up | 0.787 | 1 | 1530 | 103 | 21.7 | 10 έως 12 |
| top down | 0.797 | 1.013 | 1 | | | 10 έως 12 |

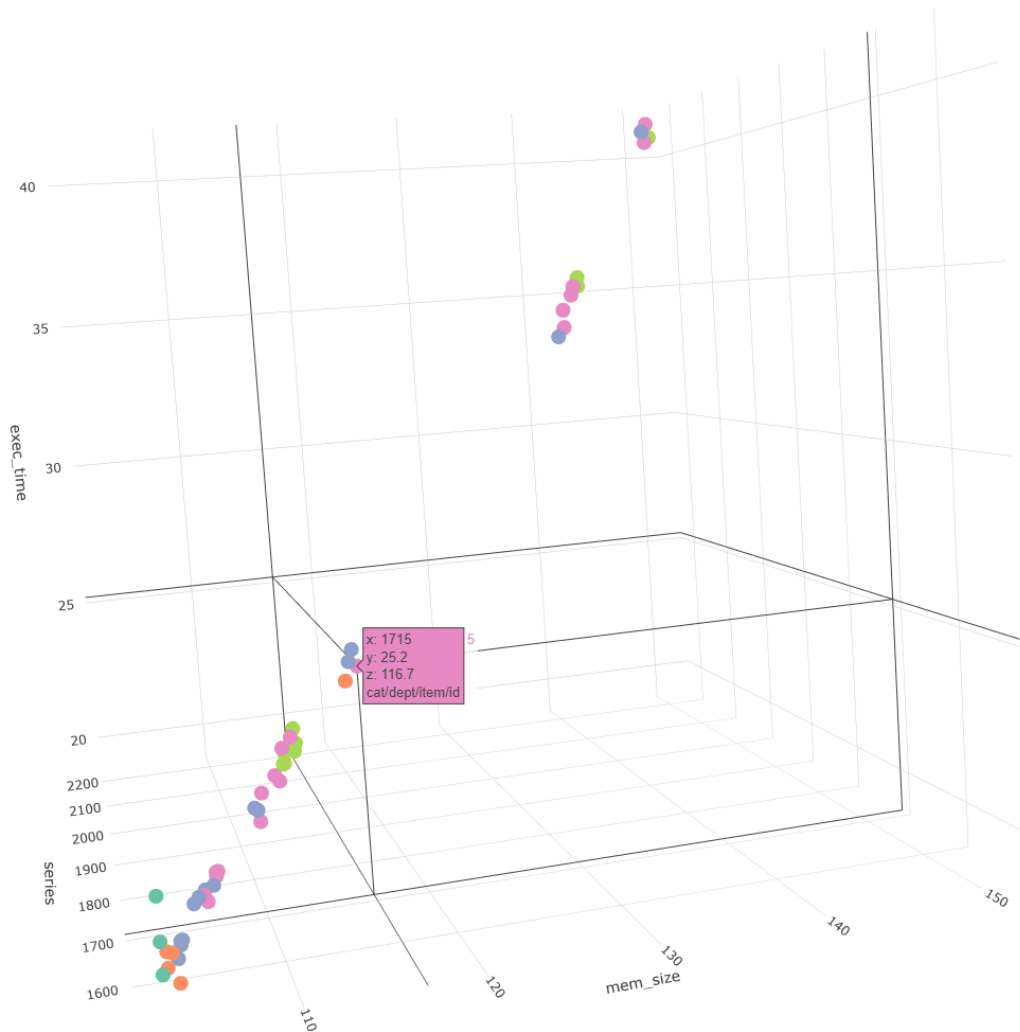
Πίνακας 13: Απόδοση Επιλεγμένων Μεθόδων στα Χαμηλά Συνάθροισης 10-12

υψηλά και μεσαία επίπεδα συνάθροισης που εξετάστηκαν παραπάνω- η μέση απόδοση της ακριβέστερης ιεραρχίας, της πιο κοστοβόρας ιεραρχίας καθώς και των 2 single level μεθόδων που ορίστηκαν ως benchmarks. Παρατηρείται πάλι υψηλότερη ακρίβεια προβλέψεων των χαμηλών επιπέδων για εναρμονισμένες με *Minimum Trace* προβλέψεις από ό,τι με bottom up ή top down. Επιπλέον, οι ιεραρχίες που περιέχουν το επίπεδο 10 (item) αποτελούν τις ιεραρχίες που παράγουν τις βέλτιστες προβλέψεις στα χαμηλά επίπεδα. Υπενθυμίζεται πως η συμπερίληψη του επιπέδου 10, δεν έδωσε τις καλύτερες προβλέψεις τόσο μεσοσταθμικά όσο και στα υψηλά και μεσαία επίπεδα της συνάθροισης.

Ωστόσο, η παραπάνω ανάλυση έχει σχετικά μικρή αξία, μιας και οι αποκλίσεις στις ακρίβειες είναι της τάξης του 2ου δεκαδικού ψηφίου του RMSSE στο επίπεδο 10 (item) και 3ου δεκαδικού ψηφίου στα επίπεδα 11 (item/state) και 12 (item/store). Οι οριακές ισοπαλίες στις ακρίβειες πρόβλεψης των χαμηλών επιπέδων συνάθροισης, επιτρέπουν την απαλοιφή του από τη στάθμιση για επιλογή βέλτιστης ιεραρχίας. Οι ιεραρχίες που απέδωσαν καλύτερα στα υψηλά και μεσαία επίπεδα συνάθροισης για παράδειγμα, μπορούν να αποτελέσουν σύμφωνα με τα δεδομένα, αξιόπιστες επιλογές και για την πρόβλεψη και των χαμηλών επιπέδων στην ιεραρχία. Για το λόγο αυτό, στον πίνακα 13 περιλαμβάνεται και η απόδοση της ιεραρχίας dept/store/id.

6.3 Αλληλοσυσχέτιση Μέτρων Κόστους Πρόβλεψης

Κατά την εφαρμογή της μεθοδολογίας, εξήχθησαν διαφορετικά μέτρα αξιολόγησης του κόστους για προβλέψεις εναρμονισμένες με καθεμιά εκ των εξεταζόμενων ιεραρχιών. Όπως είδαμε αναλυτικότερα στην παράγραφο 5.4.3, τα αρχεία που παράγονται κατά τον τερματισμό του προγράμματος, περιέχουν πληροφορία για το χρόνο που χρειάστηκε η ανεξάρτητη μοντελοποίηση όλων των χρονοσειρών κάθε εξεταζόμενης ιεραρχίας (*exec_time*) σε λεπτά της ώρας, το μέγεθος του μεγαλύτερου αντικειμένου που παράγεται κατά την πρόβλεψη με κάθε εξεταζόμενη ιεραρχία, δηλαδή τις απαιτήσεις μνήμης του υπολογιστικού συστήματος της πρόβλεψης (*mem_size*) σε MB, καθώς και το πλήθος των χρονοσειρών που περιέχονται σε κάθε εξεταζόμενη ιεραρχία (*series_n*).

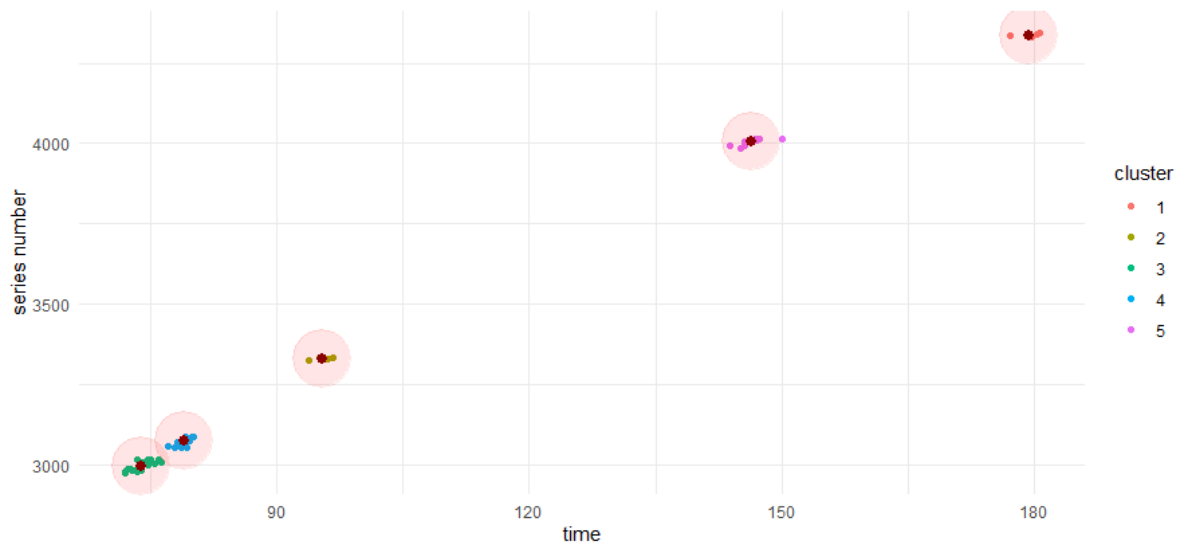


Σχήμα 24: Αλληλοσυσχέτιση των Μέτρων Κόστους Πρόβλεψης στη Μελέτη Περίπτωσης
Όπως φαίνεται και στο σχήμα 24, υπάρχει γραμμική συσχέτιση ανάμεσα στις τρεις ξε-

χωριστές μετρήσεις κόστους πρόβλεψης. Αυτό σημαίνει πως η διάφορα ανάμεσα στο πλήθος των περιεχόμενων χρονοσειρών σε δύο ιεραρχίες, μεταφράζεται σε γραμμική μεταβολή τόσο στον χρόνο όσο και στη μνήμη που απαιτεί η πρόβλεψη της ιεραρχίας. Συνεπώς, η μελέτη κόστους μπορεί να περιοριστεί αποκλειστικά στο πλήθος των περιεχόμενων χρονοσειρών σε κάθε ιεραρχία.

6.4 Συστάδες Κόστους Πρόβλεψης

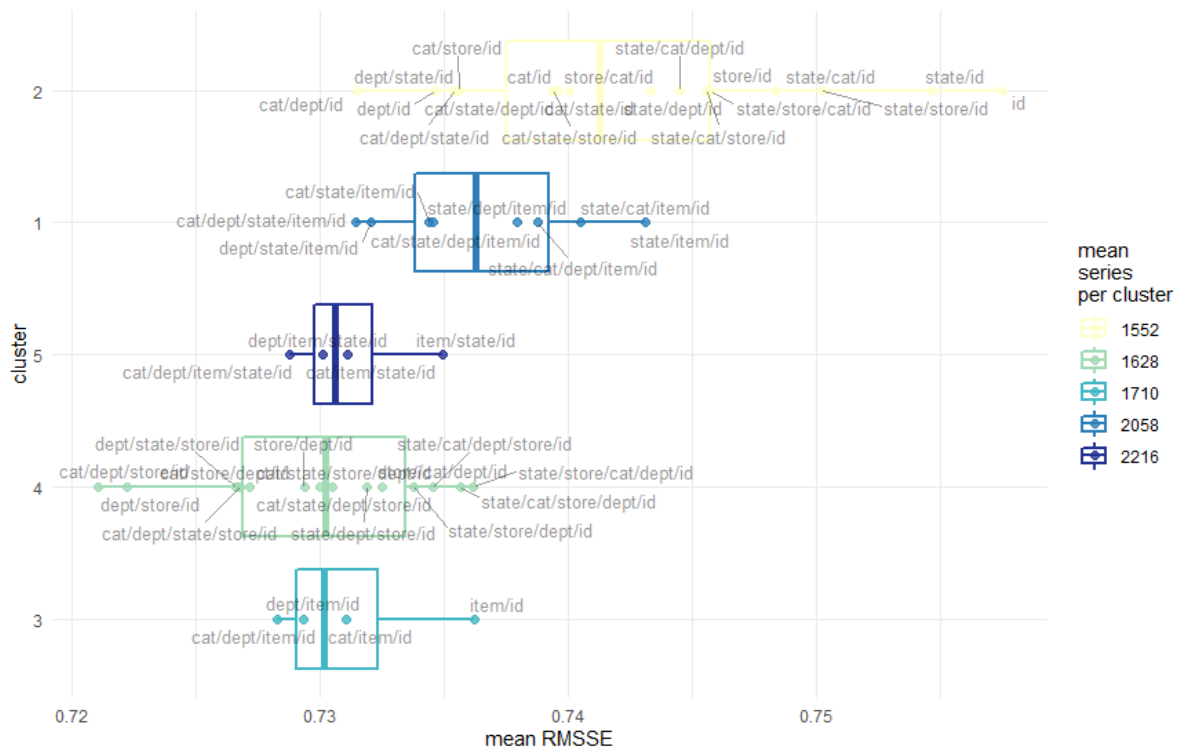
Ένα στοιχείο που παρατηρήθηκε με συνέπεια κατά την ανάλυση των αποτελεσμάτων διαφορετικών δειγμάτων, είναι η ομαδοποίηση ή συσταδοποίηση (*clustering*) των ιεραρχιών τόσο ως προς την ακρίβεια πρόβλεψης, όσο και ως προς το πλήθος των περιεχόμενων χρονοσειρών. Το φαινόμενο αυτό, που έγινε φανερό και στα σχήματα 18, 19, 20 και 24, μπορεί να εξηγηθεί από τη γνώση ενδογενών πληροφοριών για την κάθε ιεραρχία, όπως το ποια επίπεδα περιέχει, με ποια σειρά και πόσες χρονοσειρές περιέχει το καθένα. Ωστόσο, η παραπάνω συσταδοποίηση μπορεί να προκύψει και χωρίς γνώση των παραπάνω, με χρήση μηχανικής μάθησης και συγκεκριμένα της μεθόδου των κοντινότερων γειτόνων.



Σχήμα 25: Συσταδοποίηση Ιεραρχιών ανά Κόστος Πρόβλεψης | k -Κοντινότεροι Γείτονες Χρωματική Κλίμακα ανά Συστάδα και Αποτύπωση Κέντρων Κόστους Δείγμα 1531 Μη Συναθροισμένων Χρονοσειρών

Όπως φαίνεται και στο σχήμα 25, η συσταδοποίηση των ιεραρχιών σε ομάδες με παρόμοιο κόστος πρόβλεψης μέσω εφαρμογής της μεθόδου κοντινότερων γειτόνων, καταλήγει στην ύπαρξη 5 συστάδων-clusters πολυπλοκότητας. Σημειώνεται πως η παρακάτω ομαδοποίηση προέκυψε χωρίς γνώση για το ποιες είναι οι εξεταζόμενες ιεραρχίες. Πλέον, υπάρχει η δυνατότητα να εξαχθούν μέσες ακρίβειες και κόστη πρόβλεψης για καθένα από τα διαφορετικά clusters πολυπλοκότητας. Τα αποτελέσματα αυτής της ανάλυσης παρουσιάζον-

ται στο σχήμα 26. Τα clusters έχουν ταξινομηθεί βάσει της μέσης ακρίβειας πρόβλεψης των περιεχόμενων χρονοσειρών, που προκύπτει από τα boxplots. Με χρωματική κλίμακα αναφέρεται το μέσο πλήθος χρονοσειρών που περιέχονται ανά cluster. Και από αυτή τη σκοπιά, παραμένει το γεγονός ότι το cluster με την υψηλότερη πολυπλοκότητα δεν παράγει τις καλύτερες μέσες προβλέψεις. Αντίθετα, οι ιεραρχίες στα clusters με γαλάζιο και πράσινο χρώμα (δηλαδή μέσο πλήθος χρονοσειρών 1710 και 1628 αντίστοιχα) παρουσιάζουν οριακά καλύτερη ακρίβεια σε σχέση με την ακρίβεια του cluster στο οποίο κάθε ιεραρχία περιέχει μεσοσταθμικά 2216 χρονοσειρές.



Σχήμα 26: Μέση Απόδοση ανά Cluster Κόστους Πρόβλεψης των Ιεραρχιών

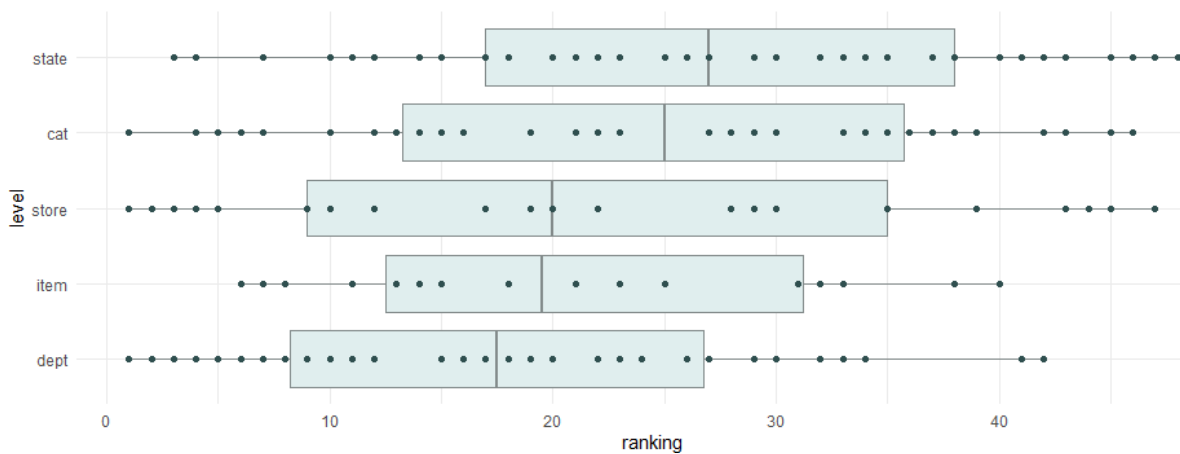
Στο σχήμα 26, πέρα από τα boxplots, φαίνονται και οι μέσες ακρίβειες των ιεραρχιών που περιέχονται στα clusters. Γίνεται φανερό, πως καθοριστικός παράγοντας αύξησης του πλήθους των περιεχόμενων χρονοσειρών σε κάθε ιεραρχία είναι η ύπαρξη και η θέση του επιπέδου των πωλήσεων ανά προϊόν (επίπεδο 10 - *item*). Τα clusters που περιέχουν τις ιεραρχίες με τις χαμηλότερες μέσες περιεχόμενες χρονοσειρές (κίτρινο και πράσινο χρώμα), περιέχουν ιεραρχίες χωρίς το επίπεδο *item*. Το cluster με γαλάζιο χρώμα, περιέχει το επίπεδο *item*, αλλά αποκλειστικά στο προτελευταίο επίπεδο (πάνω από το *id*) και με τα υπόλοιπα επίπεδα να προέρχονται αποκλειστικά από την ιεραρχία των κατηγοριών (και όχι τη γεωγραφική), που παράγει μικρότερο παράγοντα διακλάδωσης στο δέντρο των ιεραρχιών του cluster. Το cluster με σκούρο γαλάζιο περιέχει ιεραρχίες με το *item* αποκλειστικά στο προτελευταίο επίπεδο, ενώ το cluster με σκούρο μπλε χρώμα περιέχει το επίπεδο των

προϊόντων 2 επίπεδα πάνω από το *id*, με το επίπεδο *state* να βρίσκεται ενδιάμεσα.

Από τα παραπάνω, γίνεται φανερό πως κύριος παράγοντας αύξησης του κόστους πρόβλεψης των εξεταζόμενων ιεραρχιών στη μελέτη περίπτωσης που εξετάζεται στην εργασία, είναι η συμπερίληψη ή όχι, καθώς και η θέση του επιπέδου των πωλήσεων ανά προϊόν. Ωστόσο, το συγκεκριμένο επίπεδο, φαίνεται να μην έχει μετρήσιμη προβλεπτική αξία, ιδιαίτερα στα επίπεδα μεσαίας και υψηλής συνάθροισης (όπως παρουσιάστηκε και στο σχήμα 20). Στο πλαίσιο της παρούσας ανάλυσης, η συμπερίληψη του επιπέδου των πωλήσεων ανά προϊόν κρίνεται μη αναγκαία, καθώς όσες ιεραρχίες περιείχαν το συγκεκριμένο επίπεδο δεν παρήγαγαν βελτιωμένες προβλέψεις σε σχέση με πιο ολιγοπληθείς ιεραρχίες -τόσο μεσοσταθμικά όσο και ανά επίπεδο- ενώ η μοντελοποίησή τους απαιτούσε υψηλούς πόρους.

6.5 Κατατάξεις Ακρίβειας Επιπέδων Συνάθροισης

Ακόμη μια σκοπιά από την οποία μπορούν να αναλυθούν τα αποτελέσματα της μελέτης περίπτωσης, είναι η προβλεπτική αξία που προσφέρει η συμπερίληψη καθενός από τα πέντε επίπεδα των ξεχωριστών ιεραρχιών της συλλογής του συνόλου δεδομένων. Στο σχήμα 27 παρουσιάζεται η μέση κατάταξη κάθε επιπέδου συνάθροισης, όπως προκύπτει από τις εμφανίσεις του στις κατατάξεις των 48 ιεραρχιών και 2 μεθόδων αναφοράς που παρουσιάστηκαν στο σχήμα 16. Η κάθετη θεώρηση των σημείων του σχήματος 27

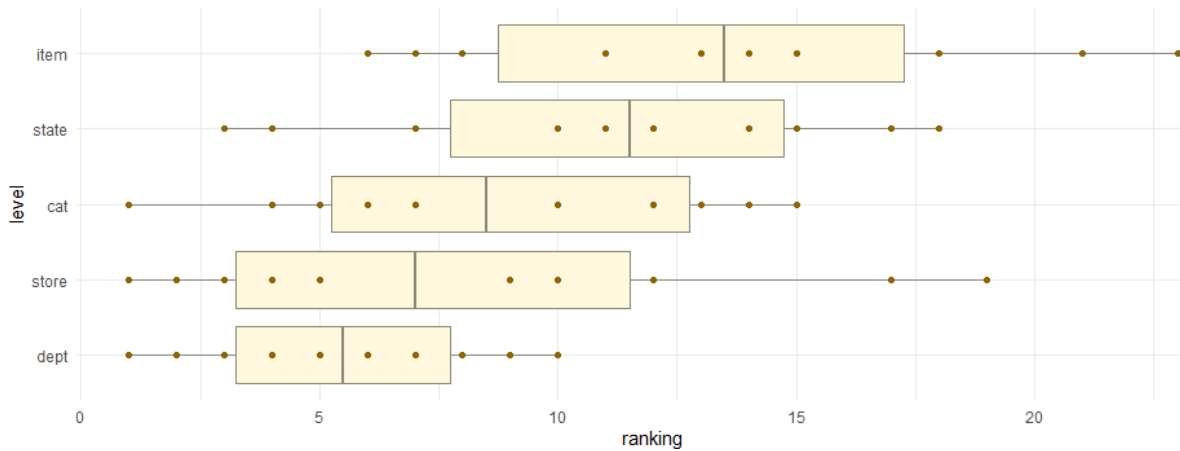


Σχήμα 27: Μέση Κατάταξη Εμφάνισης Επιπέδων στις 50 Εξεταζόμενες Μεθόδους

αποκαλύπτει τις κατατάξεις των ιεραρχιών. Για παράδειγμα, πρώτη σε ακρίβεια ιεραρχία ήταν η *total/cat/dept/store/id* που αποτελεί την πρώτη από αριστερά κάθετη τριάδα¹³. Η κατάταξη του σχήματος δείχνει μεγαλύτερη πυκνότητα εμφάνισης των επιπέδων *dept* και *store* στις πρώτες θέσεις της κατάταξης ακρίβειας. Ωστόσο η τελική κατάταξη τοποθετεί το επίπεδο *item* δεύτερο ως προς τη μέση κατάταξη ακρίβειας των ιεραρχιών στις οποίες εμφανίζεται. Η κατάταξη του σχήματος 27 αντικρούει στη μέχρι τώρα εξαγόμενη εικόνα

¹³Η *total/cat/dept/store/id* περιέχει στην πραγματικότητα πέντε επίπεδα συνάθροισης. Από την παρούσα μελέτη αφέρθηκαν εκτός τα επίπεδα *total* και *id*, διότι περιέχονται σε κάθε ιεραρχία εξ υποθέσεως.

της χαμηλής προβλεπτικής ικανότητας του επιπέδου των προϊόντων. Αυτό όμως που αξιολογείται στο σχήμα 27, είναι η μέση κατάταξη σε όλες τις -50- εξεταζόμενες μεθόδους. Αν ζητούμενο είναι η προβλεπτική ικανότητα που ενδεχομένως περιέχει το κάθε επίπεδο συνάθροισης, τότε μια καταλληλότερη εικόνα των αποτελεσμάτων είναι η μέση κατάταξη των επιπέδων, ως προς τις ακριβέστερες και όχι όλες τις ιεραρχίες στις οποίες εμφανίστηκε. Μια τέτοια εικόνα δίνεται στο σχήμα 28, όπου εμφανίζεται η μέση κατάταξη των πέντε



Σχήμα 28: Μέση Κατάταξη Επιπέδων στις 10 Ακριβέστερες Εμφανίσεις

επιπέδων συνάθροισης στις 10 πρώτες εμφανίσεις στην κατάταξη ακρίβειας των 50 εξεταζόμενων μεθόδων. Στο σχήμα 28 γίνεται πλέον φανερή η συμβολή των επιπέδων dept και store σε υψηλές κατατάξεις των ιεραρχιών στις οποίες περιέχονται, ως προς την ακρίβεια πρόβλεψης. Η παραπάνω θεώρηση, ενδεχομένως «πριμοδοτεί» επίπεδα συνάθροισης που εμφανίστηκαν σε ιεραρχίες με υψηλή ακρίβεια, αλλά δεν αποτέλεσαν τα ίδια ενισχυτικό παράγοντα. Ένα τέτοιο παράδειγμα είναι η φαινομενική «πρωτιά» του επιπέδου cat, ενώ τα επίπεδα dept και store προσέφεραν από μόνα τους εξίσου ακριβείς προβλέψεις.

Στον πίνακα 14 παρουσιάζονται οι μέσες κατατάξεις ακρίβειας των εξεταζόμενων επιπέδων συνάθροισης, τόσο ανάμεσα σε όλες τις μεθόδους πρόβλεψης, όσο και ανάμεσα στις 10

| Κατατάξεις \ Επίπεδο | Επίπεδο | | | | |
|---|---------|-------|-------|-------------|-------|
| | state | store | cat | dept | item |
| Μέση Κατάταξη στις 50 Μεθόδους Πρόβλεψης | 27.1 | 22.1 | 24.3 | 18.2 | 20.9 |
| Μέση Κατάταξη στις 10 Ακριβέστερες Εμφανίσεις | 11.1 | 8.2 | 8.7 | 5.5 | 13.6 |
| Απόλυτες Συχνότητες Εμφάνισης | 33 | 20 | 29 | 29 | 16 |
| Σχετικές Συχνότητες Εμφάνισης | 68.8% | 41.7% | 60.4% | 60.4% | 33.3% |

Πίνακας 14: Μέση Κατάταξη Ακρίβειας & Συχνότητες Εμφάνισης Επιπέδων Συνάθροισης στις Εξεταζόμενες Ιεραρχίες

ακριβέστερες εμφανίσεις των επιπέδων στις κατατάξεις της παραγράφου 6.1. Φανερή είναι η πρωτιά του επιπέδου των τμημάτων προϊόντων και στις δύο περιπτώσεις. Τα επίπεδα

store, cat, state διατηρούν αυτή τη σειρά κατάταξης, ενώ το επίπεδο πωλήσεων ανά προϊόν «πέφτει» στην 5η θέση ανάμεσα στις δύο κατατάξεις.

Στον πίνακα, φαίνονται και οι απόλυτες και σχετικές συχνότητες εμφάνισης των επιπέδων συνάθροισης στην ιεραρχία. Το επίπεδο πωλήσεων ανά προϊόν εμφανίζεται λιγότερες φορές, αφενός γιατί αποτελεί το χαμηλότερο επίπεδο της ιεραρχίας των κατηγοριών -και άρα οι πιθανές διατάξεις του στις εκάστοτε ιεραρχίες είναι περιορισμένες, και αφετέρου διότι η εξ υποθέσεως συμπερίληψη του επιπέδου των πωλήσεων μονάδων/μη συναθροισμένων χρονοσειρών (επίπεδο 12, *item/store*) οδηγούσε στον αποκλεισμό ορισμένων ιεραρχήσεων που περιείχαν τα δύο παραπάνω επίπεδα (ιεραρχίες της μορφής *total/*/item/*/store*/id* ή *total/*/store*/item*/id*) λόγω της περιττής επανάληψης των εν λόγω επιπέδων. Αυτός είναι και λόγος της μειωμένης συχνότητας εμφάνισης του επιπέδου των πωλήσεων ανά κατάσταση (*store*).

Ατυχείς Προβλέψεις #8

«Ο ερχομός της ασύρματης εποχής θα κάνει τον πόλεμο αδύνατο, αφού θα τον κάνει γελοίο.»

Guglielmo Marconi
εφευρέτης του τηλεγράφου, 1912

7 Συμπεράσματα

Από τα αποτελέσματα της παραγράφου 6, προέκυψαν ορισμένα συμπεράσματα καθώς και νέα ερωτήματα για μελλοντική έρευνα, τα οποία συνοψίζονται στο παρακάτω κεφάλαιο.

7.1 Ευρήματα

Από τις μετρήσεις της απόδοσης των προβλέψεων που είχαν εναρμονιστεί με τις εξεταζόμενες ιεραρχίες, πάνω στα 12 επίπεδα συνάθροισης του συνόλου δεδομένων για τη μελέτη περίπτωσης των πωλήσεων της Walmart, προέκυψαν ορισμένα ενδιαφέροντα ευρήματα για την προβλεπτική αξία που μπορεί να έχει μια ιεραρχία, πάνω στις προβλέψεις ενός ή περισσότερων επιπέδων. Αναλυτικότερα:

- Η εναρμόνιση με επίπεδα που περιέχουν περισσότερες χρονοσειρές δεν οδηγεί απαραίτητα σε καλύτερες προβλέψεις για ένα ή περισσότερα επίπεδα συνάθροισης.
- Η προβλεπτική αξία των επιπέδων σε μια συλλογή ιεραρχικών ή ομαδοποιημένων χρονοσειρών, δεν είναι ισοδύναμη για κάθε εξεταζόμενο επίπεδο ή ομάδα επιπέδων συνάθροισης.
- Τα επίπεδα που βρίσκονται στο μέσο της ή των ξεχωριστών ιεραρχιών κάθε συλλογής ιεραρχικών ή ομαδοποιημένων χρονοσειρών, παρέχουν υψηλή προβλεπτική αξία τόσο μεσοσταθμικά όσο και ανά επίπεδο συνάθροιση. Ταυτόχρονα, η προσθήκη επιπέδων μεσαίας συνάθροισης στην εκάστοτε ιεραρχία, διατηρεί σε χαμηλά επίπεδα το κόστος πρόβλεψης.
- Οι προβλέψεις που συνδυάζουν γραμμικά, μοντέλα από περισσότερα επίπεδα, προέβλεψαν με μεγαλύτερη ακρίβεια σε σχέση με τις μεθόδους bottom up και top down στη συντριπτική πλειοψηφία των δειγμάτων του συνόλου δεδομένων. Ωστόσο, ο κανόνας αυτός δεν είναι καθολικός, αφού μπορεί να προκύψουν και δείγματα στα οποία τα benchmarks υπεραποδίδουν¹⁴.
- Κατά την κατάταξη των ιεραρχιών ως προς την ακρίβεια αλλά και το κόστος πρόβλεψης, προκύπτουν clusters ιεραρχιών που εμφανίζουν παρόμοια ακρίβεια και χρόνους, μεσοσταθμικά αλλά και ανά επίπεδο. Συνεπώς η βέλτιστη οργάνωση επιπέδων για την ενίσχυση των προβλέψεων, μπορεί να περιλαμβάνει ένα σύνολο ιεραρχιών με κοινά χαρακτηριστικά, και όχι απαραίτητα μια μοναδική ιεραρχία.
- Κύριος παράγοντας εξάρτησης του μεγέθους κάθε ιεραρχίας, είναι το πλήθος των χρονοσειρών στα επίπεδα που περιλαμβάνει. Στη μελέτη περίπτωσης, το επίπεδο συνάθροισης που χαρακτηρίζει σε μεγάλο βαθμό το μέγεθος της συλλογής -και κατ'επέκταση το κόστος της μοντελοποίησης- είναι αυτό των πωλήσεων ανά προϊόν (*item*). Τέτοια επίπεδα, μπορούν να ελέγχονται ως προς την προβλεπτική τους

¹⁴Περισσότερα ως προς τα δείγματα-εξαιρέσεις αναλύονται στην παράγραφο Β του παραρτήματος.

ικανότητα και αν δεν προσφέρουν μετρήσιμη πληροφορία μπορούν να μη συμπεριλαμβάνονται στην οργάνωση που θα επιλεγεί για εναρμόνιση, όπως έγινε στην παρούσα εργασία.

- Ακόμη ένας παράγοντας εξάρτησης του μεγέθους της ιεραρχίας είναι η θέση στην οποία θα καταλαμβάνουν τα πολυπληθή επίπεδα στην ιεραρχία. Όσο ψηλότερα τοποθετείται ένα επίπεδο όπως το (*item*), τόσο υψηλότερος γίνεται και ο παράγοντας διακλάδωσης (*branching factor*) της εκάστοτε ιεραρχίας και άρα το συνολικό πλήθος των περιεχόμενων χρονοσειρών. Παράλληλα η τοποθέτηση ενός πολυπληθούς επιπέδου ψηλότερα στην ιεραρχία δεν έδειξε να βελτιώνει τις προβλέψεις, τόσο μεσοσταθμικά όσο και ανά επίπεδο συνάθροισης.

7.2 Μελλοντικές Προεκτάσεις

Σε μια εποχή στην οποία τα δεδομένα πληθαίνουν και συνδέονται με αυξανόμενους ρυθμούς, η δυνατότητα ενίσχυσης των προβλέψεων μιας χρονοσειράς με μοντέλα που αποτυπώνουν συναθροίσεις και διαμερίσεις της, μπορεί να αποτελέσει ένα χρήσιμο εργαλείο για καλύτερες προβλέψεις και πιο αποτελεσματική υποστήριξη λήψης απόφασης. Στη συλλογή ομαδοποιημένων χρονοσειρών του συνόλου δεδομένων του διαγωνισμού *M5* για παράδειγμα, θα μπορούσε ένας αναλυτής να προσθέσει τις χρονοσειρές πωλήσεων του ανταγωνισμού της Walmart και η κορυφή της ιεραρχίας να είναι πλέον οι πωλήσεις όλων των επιχειρήσεων με *super* και *mega markets* στις ΗΠΑ. Αυτό βέβαια προϋποθέτει τη μεταβολή της θεώρησης των ιεραρχικών ή ομαδοποιημένων χρονοσειρών, από οντότητες που παρέχει *ad hoc* ο αποφασίζων, σε παραμέτρους βελτίωσης των ίδιων των προβλέψεων. Πάνω σε αυτό το στόχο, η παρούσα εργασία εξήγαγε ορισμένα ευρήματα σε σχέση με το ποιες θεωρούνται καλές ιεραρχίες για ενίσχυση της πρόβλεψης πωλήσεων λιανικής σε διάφορα επίπεδα συνάθροισης.

Από τα αποτελέσματα της εργασίας προκύπτουν νέα ερωτήματα και υποθέσεις προς εξέταση. Για αρχή, αναγκαία είναι η έρευνα περισσότερων συνόλων δεδομένων, τόσο εντός των πωλήσεων λιανικής που αποτελούν χαρακτηριστικά παραδείγματα μεγάλων ομαδοποιημένων χρονοσειρών, όσο και σε άλλες εφαρμογές πρόβλεψης στις οποίες προκύπτουν συλλογές χρονοσειρών που οργανώνονται σε επίπεδα συνάθροισης. Το συμπέρασμα που προκύπτει από την παρούσα εργασία είναι πως τα επίπεδα που βρίσκονται στο μέσο της ή των ξεχωριστών ιεραρχιών κάθε συλλογής ιεραρχικών ή ομαδοποιημένων χρονοσειρών, παρέχουν υψηλή προβλεπτική αξία τόσο μεσοσταθμικά όσο και ανά επίπεδο συνάθροισης. Ωστόσο το σύνολο δεδομένων, οργανώνεται σε επίπεδα που ορισμένα περιέχουν το πολύ μερικές δεκάδες χρονοσειρές ενώ τα υπόλοιπα περιέχουν τουλάχιστον χιλιάδες. Η εμπειρική επαλήθευση του συμπεράσματος της παρούσας εργασίας είναι απαραίτητη και σε σύνολα δεδομένων που περιέχουν διαφορετικές οργανώσεις, με πλήρη χρονοσειρών ανά επίπεδο συνάθροισης που εμφανίζουν διαφορετικές κατανομές πλήθους.

Οι ιεραρχίες που είχαν εναρμονιστεί με *Minimum Trace*, παρήγαγαν στη συντριπτική τους πλειοψηφία, ακριβέστερες προβλέψεις από τις μεθόδους *bottom up* και *top down*. Ωστόσο, υπήρξαν και δείγματα στα οποία οι *single level* μέθοδοι απέδωσαν καλύτερα από

τις εναρμονισμένες με *Minimum Trace* προβλέψεις, κυρίως λόγω της κακής απόδοσης των τελευταίων. Το αποτέλεσμα αυτό δείχνει πως ακόμη και σε δείγματα του ίδιου συνόλου δεδομένων μπορούν να προκύπτουν μεταβολές στις κατατάξεις ακρίβειας ανάμεσα στις εξεταζόμενες μεθόδους. Ένα ερώτημα που τίθεται προς μελέτη, είναι το ποια είναι τα χαρακτηριστικά των δειγμάτων στα οποία οι single level προσεγγίσεις προβλέπουν ακριβέστερα από τις μεθόδους γραμμικού συνδυασμού επιπέδων συνάθροισης.

Τέλος, μια σημαντική συνεισφορά στον τομέα της ιεραρχικής πρόβλεψης θα ήταν η ανάπτυξη μια γενικής μεθοδολογίας, που θα επιλέγει με αυτόματο τρόπο και για δεδομένα με διαφορετικά χαρακτηριστικά, τη βέλτιστη οργάνωση σε επίπεδα συνάθροισης για εναρμόνιση των προβλέψεων με κριτήριο τη βελτιστοποίηση της ακρίβειας και του κόστους της πρόβλεψης ενός ή περισσότερων επιπέδων συνάθροισης. Τόσο η επιλογή της βέλτιστης οργάνωσης για ενίσχυση της πρόβλεψης, όσο και η εύρεση των δεδομένων που αποτελούν την βέλτιστη οργάνωση, θα μπορούσε να ενσωματώνεται σε μια αυτόματη μέθοδο η οποία εν τέλει θα προέβλεπε αξιοποιώντας την πληροφορία που εξήχθη με ενδογενή κριτήρια.

Ατυχείς Προβλέψεις #9

«Ένας πύραυλος δεν θα μπορέσει ποτέ,
να ξεπεράσει την ατμόσφαιρα της Γης.»

άρθρο στις *New York Times*, 1920.
αποσύρθηκε «ελαφρά τη καρδία» κατά
την προσσελήγωση του *Apollo 11* το 1969.

Βιβλιογραφία

- B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Dover Books on Electrical Engineering. Dover Publications, 2012. ISBN 9780486136899. URL <https://books.google.gr/books?id=iYMqLQp49UMC>.
- Masanao Aoki. Studies of economic interdependence by state space modeling of time series: Us-japan example. *Annales d'Économie et de Statistique*, pages 225–252, 1987. ISSN 0769489X, 22726497. URL <http://www.jstor.org/stable/20075655>.
- Mahsa Ashouri, Rob J Hyndman, and Galit Shmueli. Fast forecast reconciliation using linear models. *Journal of Computational and Graphical Statistics*, 0(0):1–20, 2021. doi: 10.1080/10618600.2021.1939038. URL <https://doi.org/10.1080/10618600.2021.1939038>.
- V. Assimakopoulos and K. Nikolopoulos. The theta model: a decomposition approach to forecasting. *International Journal of Forecasting*, 16(4):521–530, 2000. ISSN 0169-2070. doi: [https://doi.org/10.1016/S0169-2070\(00\)00066-2](https://doi.org/10.1016/S0169-2070(00)00066-2). URL <https://www.sciencedirect.com/science/article/pii/S0169207000000662>. The M3-Competition.
- G. Athanasopoulos, P. Gamakumara, A. Panagiotelis, R.J. Hyndman, and M. Affan. *Macroeconomic Forecasting in the Era of Big Data. Advanced Studies in Theoretical and Applied Econometrics*, volume 52, pages 7–14. Springer, 2020. doi: https://doi.org/10.1007/978-3-030-31150-6_21. URL https://link.springer.com/chapter/10.1007%2F978-3-030-31150-6_21.
- George Athanasopoulos, Roman A. Ahmed, and Rob J. Hyndman. Hierarchical forecasts for australian domestic tourism. *International Journal of Forecasting*, 25(1):146–166, 2009. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2008.07.004>. URL <https://www.sciencedirect.com/science/article/pii/S0169207008000691>.
- George Athanasopoulos, Rob J. Hyndman, Nikolaos Kourentzes, and Fotios Petropoulos. Forecasting with temporal hierarchies. *European Journal of Operational Research*, 262(1):60–74, 2017. ISSN 0377-2217. doi: <https://doi.org/10.1016/j.ejor.2017.02.046>. URL <https://www.sciencedirect.com/science/article/pii/S0377221717301911>.
- G.E.P. Box and G.M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day series in time series analysis and digital processing. Holden-Day and Wisconsin Madison Uni. Dept. of Statistics., 1970. ISBN 9780816210947. URL <https://books.google.gr/books?id=5BVfnXaq03oC>.
- R.G. Brown. *Statistical Forecasting for Inventory Control*. McGraw-Hill, 1959. URL <https://books.google.gr/books?id=oKI8AAAAIAAJ>.

- R.G. Brown. Smoothing, forecasting and prediction of discrete time series. *Journal of the Royal Statistical Society. Series A (General)*, 127(2):292–293, 1960. ISSN 00359238. URL <http://www.jstor.org/stable/2344012>.
- A.D. Little Brown RG. Exponential smoothing for predicting demand. *Philip Morris Records; Master Settlement Agreement*, 1956. URL <https://www.industrydocuments.ucsf.edu/docs/jzlc0130>.
- Margaux Brégère and Malo Huard. Online hierarchical forecasting for power consumption data. *International Journal of Forecasting*, 2021. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2021.05.011>. URL <https://www.sciencedirect.com/science/article/pii/S0169207021000947>.
- John F. Muth Charles C. Holt, Franco Modigliani and Herbert A. Simon. Planning production, inventories, and work force. *American Journal of Agricultural Economics*, 44(2):650–652, 1960. doi: <https://doi.org/10.2307/1235878>. URL <https://onlinelibrary.wiley.com/doi/abs/10.2307/1235878>.
- Jose A. Fiorucci, Tiago R. Pellegrini, Francisco Louzada, Fotios Petropoulos, and Anne B. Koehler. Models for optimising the theta method and their relationship to state space models. *International Journal of Forecasting*, 32(4):1151–1161, 2016. doi: 10.1016/j.ijforecast.2016. URL <https://ideas.repec.org/a/eee/intfor/v32y2016i4p1151-1161.html>.
- Benito E Flores. A pragmatic view of accuracy measurement in forecasting. *Omega*, 14(2):93–98, 1986. ISSN 0305-0483. doi: [https://doi.org/10.1016/0305-0483\(86\)90013-7](https://doi.org/10.1016/0305-0483(86)90013-7). URL <https://www.sciencedirect.com/science/article/pii/0305048386900137>.
- Philip Hans Franses. A note on the mean absolute scaled error. *International Journal of Forecasting*, 32(1):20–22, 2016. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2015.03.008>. URL <https://www.sciencedirect.com/science/article/pii/S0169207015000448>.
- Everette S. Gardner. Exponential smoothing: The state of the art—part ii. *International Journal of Forecasting*, 22(4):637–666, 2006. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2006.03.005>. URL <https://www.sciencedirect.com/science/article/pii/S0169207006000392>.
- Everette S. Gardner and Ed. McKenzie. Note—seasonal exponential smoothing with damped trends. *Management Science*, 35(3):372–376, 1989. doi: 10.1287/mnsc.35.3.372. URL <https://doi.org/10.1287/mnsc.35.3.372>.
- Everette S. Gardner Jr. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1):1–28, 1985. doi: <https://doi.org/10.1002/for.3980040103>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.3980040103>.

- Tom Steel George Athanasopoulos and Don Weatherburn. Forecasting prison numbers: a grouped time series approach. *Monash University*, 2019.
- Clive Granger and P. Newbold. Spurious regressions in econometrics. *Journal of Econometrics*, 2(2):111–120, 1974. URL <https://EconPapers.repec.org/RePEc:eee:econom:v:2:y:1974:i:2:p:111-120>.
- Ross Hollyman, Fotios Petropoulos, and Michael E. Tipping. Understanding forecast reconciliation. *European Journal of Operational Research*, 294(1):149–160, 2021. ISSN 0377-2217. doi: <https://doi.org/10.1016/j.ejor.2021.01.017>. URL <https://www.sciencedirect.com/science/article/pii/S0377221721000199>.
- Charles C. Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1):5–10, 1957. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2003.09.015>. URL <https://www.sciencedirect.com/science/article/pii/S0169207003001134>. 2004 reprint of the 1957 report to the Office of Naval Research (ONR 52).
- G. Hyndman, R.J. & Athanasopoulos. Forecasting: principles and practice, 3rd edition, 2021. URL <https://otexts.com/fpp3/>.
- Rob Hyndman, Anne Koehler, Keith Ord, and Ralph Snyder. *Forecasting with exponential smoothing. The state space approach*. Springer, 01 2008. doi: 10.1007/978-3-540-71918-2.
- Rob J. Hyndman. A brief history of forecasting competitions. *International Journal of Forecasting*, 36(1):7–14, 2020. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2019.03.015>. URL <https://www.sciencedirect.com/science/article/pii/S016920701930086X>. M4 Competition.
- Rob J. Hyndman and Baki Billah. Unmasking the theta method. *International Journal of Forecasting*, 19(2):287–290, 2003. ISSN 0169-2070. doi: [https://doi.org/10.1016/S0169-2070\(01\)00143-1](https://doi.org/10.1016/S0169-2070(01)00143-1). URL <https://www.sciencedirect.com/science/article/pii/S0169207001001431>.
- Rob J. Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(3):1–22, 2008. doi: 10.18637/jss.v027.i03. URL <https://www.jstatsoft.org/index.php/jss/article/view/v027i03>.
- Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2006.03.001>. URL <https://www.sciencedirect.com/science/article/pii/S0169207006000239>.

- Rob J Hyndman, Anne B Koehler, Ralph D Snyder, and Simone Grose. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3):439–454, 2002. ISSN 0169-2070. doi: [https://doi.org/10.1016/S0169-2070\(01\)00110-8](https://doi.org/10.1016/S0169-2070(01)00110-8). URL <https://www.sciencedirect.com/science/article/pii/S0169207001001108>.
- Rob J. Hyndman, Roman A. Ahmed, George Athanasopoulos, and Han Lin Shang. Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis*, 55(9):2579–2589, 2011. ISSN 0167-9473. doi: <https://doi.org/10.1016/j.csda.2011.03.006>. URL <https://www.sciencedirect.com/science/article/pii/S0167947311000971>.
- Rob J Hyndman, Alan Lee, and Earo Wang. Fast computation of reconciled forecasts for hierarchical and grouped time series. Monash Econometrics and Business Statistics Working Papers 17/14, Monash University, Department of Econometrics and Business Statistics, 2014. URL <https://ideas.repec.org/p/msh/ebswps/2014-17.html>.
- Jooyoung Jeon, Anastasios Panagiotelis, and Fotios Petropoulos. Probabilistic forecast reconciliation with applications to wind power and electric load. *European Journal of Operational Research*, 279(2):364–379, 2019. ISSN 0377-2217. doi: <https://doi.org/10.1016/j.ejor.2019.05.020>. URL <https://www.sciencedirect.com/science/article/pii/S0377221719304242>.
- Alex J. Koning, Philip Hans Franses, Michèle Hibon, and H.O. Stekler. The m3 competition: Statistical tests of the results. *International Journal of Forecasting*, 21(3):397–409, 2005. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2004.10.003>. URL <https://www.sciencedirect.com/science/article/pii/S0169207004000810>.
- S. Makridakis, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1(2):111–153, 1982. doi: <https://doi.org/10.1002/for.3980010202>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.3980010202>.
- Spyros Makridakis. Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*, 9(4):527–529, 1993. ISSN 0169-2070. doi: [https://doi.org/10.1016/0169-2070\(93\)90079-3](https://doi.org/10.1016/0169-2070(93)90079-3). URL <https://www.sciencedirect.com/science/article/pii/0169207093900793>.
- Spyros Makridakis and Michèle Hibon. The m3-competition: results, conclusions and implications. *International Journal of Forecasting*, 16(4):451–476, 2000. ISSN 0169-2070. doi: [https://doi.org/10.1016/S0169-2070\(00\)00057-1](https://doi.org/10.1016/S0169-2070(00)00057-1). URL <https://www.sciencedirect.com/science/article/pii/S0169207000000571>. The M3-Competition.

- Spyros Makridakis, Michele Hibon, and Claus Moser. Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society. Series A (General)*, 142(2):97–145, 1979. ISSN 00359238. URL <http://www.jstor.org/stable/2345077>.
- Spyros Makridakis, Chris Chatfield, Michèle Hibon, Michael Lawrence, Terence Mills, Keith Ord, and LeRoy F. Simmons. The m2-competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, 9(1):5–22, 1993. ISSN 0169-2070. doi: [https://doi.org/10.1016/0169-2070\(93\)90044-N](https://doi.org/10.1016/0169-2070(93)90044-N). URL <https://www.sciencedirect.com/science/article/pii/016920709390044N>.
- Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3):1–26, 03 2018a. doi: 10.1371/journal.pone.0194889. URL <https://doi.org/10.1371/journal.pone.0194889>.
- Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. The m4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34(4):802–808, 2018b. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2018.06.001>. URL <https://www.sciencedirect.com/science/article/pii/S0169207018300785>.
- Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. The m4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1):54–74, 2020a. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2019.04.014>. URL <https://www.sciencedirect.com/science/article/pii/S0169207019301128>. M4 Competition.
- Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. Predicting/hypothesizing the findings of the m4 competition. *International Journal of Forecasting*, 36(1):29–36, 2020b. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2019.02.012>. URL <https://www.sciencedirect.com/science/article/pii/S0169207019301098>. M4 Competition.
- Spyros Makridakis, Evangelos Spiliotis, and Vassilis Assimakopoulos. The m5 accuracy competition: Results, findings and conclusions. M5, 10 2020c. URL https://www.researchgate.net/publication/344487258_The_M5_Accuracy_competition_Results_findings_and_conclusions.
- Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. The m5 competition: Background, organization, and implementation. *International Journal of Forecasting*, 2021. ISSN 0169-2070. doi: <https://doi.org/10.1016/j.ijforecast.2021.07.007>. URL <https://www.sciencedirect.com/science/article/pii/S0169207021001187>.
- Paolo Mancuso, Veronica Piccialli, and Antonio M. Sudoso. A machine learning approach for forecasting hierarchical time series. *Expert Systems with Applications*, 182:115102,

2021. ISSN 0957-4174. doi: <https://doi.org/10.1016/j.eswa.2021.115102>. URL <https://www.sciencedirect.com/science/article/pii/S0957417421005431>.
- Konstantinos Nikolopoulos and Dimitrios D. Thomakos. *Forecasting with the Theta Method*, pages 1–5. American Cancer Society, 2020. ISBN 9781118445112. doi: <https://doi.org/10.1002/9781118445112.stat08270>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat08270>.
- Konstantinos Nikolopoulos, Vassilios Assimakopoulos, Nikolaos Bougioukos, Akrivi Litsa, and Fotios Petropoulos. The theta model: An essential forecasting tool for supply chain planning. In Gary Lee, editor, *Advances in Automation and Robotics, Vol. 2*, pages 431–437, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. ISBN 978-3-642-25646-2.
- V. Nikolopoulos, K. & Assimakopoulos. Theta intelligent forecasting information system. *Industrial Management & Data Systems*, 103(9):711–726, 1 2003. ISSN 0263-5577. doi: 10.1108/02635570310506133. URL <https://doi.org/10.1108/02635570310506133>.
- José Manuel Oliveira and Patrícia Ramos. Assessing the performance of hierarchical forecasting methods on the retail sector. *Entropy*, 21(4), 2019. ISSN 1099-4300. doi: 10.3390/e21040436. URL <https://www.mdpi.com/1099-4300/21/4/436>.
- J. K. Ord, A. B. Koehler, and R. D. Snyder. Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association*, 92(440):1621–1629, 1997. doi: 10.1080/01621459.1997.10473684. URL <https://doi.org/10.1080/01621459.1997.10473684>.
- Mitchell O’Hara-Wild, Rob Hyndman, and Earo Wang. *fable: Forecasting Models for Tidy Time Series*, 2021. URL <https://CRAN.Rproject.org/package=fable>.
- Xiongfeng Pan, Jing Zhang, Changyu Li, Xianyou Pan, and Jinbo Song. Analysis of china’s regional wind power generation efficiency and its influencing factors. *Energy & Environment*, 30(2):254–271, 2019. doi: 10.1177/0958305X18788820. URL <https://doi.org/10.1177/0958305X18788820>.
- C. Carl Pegels. Exponential forecasting: Some new variations. *Management Science*, 15 (5):311–315, 1969. ISSN 00251909, 15265501. URL <http://www.jstor.org/stable/2628137>.
- Clint L.P. Pennings and Jan van Dalen. Integrated hierarchical forecasting. *European Journal of Operational Research*, 263(2):412–418, 2017. ISSN 0377-2217. doi: <https://doi.org/10.1016/j.ejor.2017.04.047>. URL <https://www.sciencedirect.com/science/article/pii/S0377221717303946>.
- Fotios Petropoulos and Konstantinos Nikolopoulos. Optimizing theta model for monthly data. In *ICAART 2013 - Proceedings of the 5th International Conference on Agents*

- and Artificial Intelligence*, ICAART 2013 - Proceedings of the 5th International Conference on Agents and Artificial Intelligence, pages 190–195, May 2013. ISBN 9789898565389. 5th International Conference on Agents and Artificial Intelligence, ICAART 2013 ; Conference date: 15-02-2013 Through 18-02-2013.
- Fotios Petropoulos and Konstantinos Nikolopoulos. The Theta Method. *Foresight: The International Journal of Applied Forecasting*, pages 11–17, Summer 2017. URL <https://ideas.repec.org/a/for/ijafaa/y2017i46p11-17.html>.
- Earo Wang Rob J. Hyndman, Alan Lee and Shanika Wickramasuriya. *hts: Hierarchical and Grouped Time Series*. CRAN, 2021. URL <https://pkg.earo.me/hts/>.
- J. Scott Armstrong and Steven P. Schnaars. Long-range forecasting: From crystal ball to computer, 2nd ed. (wiley, new york, 1985) [uk pound]22.95 (paper), pp. 689. *International Journal of Forecasting*, 2(3):387–390, 1986. URL <https://EconPapers.repec.org/RePEc:eee:intfor:v:2:y:1986:i:3:p:387-390>.
- Han Lin Shang and Rob J. Hyndman. Grouped functional time series forecasting: An application to age-specific mortality rates. *Journal of Computational and Graphical Statistics*, 26(2):330–343, 2017. doi: 10.1080/10618600.2016.1237877. URL <https://doi.org/10.1080/10618600.2016.1237877>.
- Ralph D. Snyder, Anne B. Koehler, Rob J. Hyndman, and J. Keith Ord. Exponential Smoothing for Inventory Control: Means and Variances of Lead-Time Demand. Monash Econometrics and Business Statistics Working Papers 3/02, Monash University, Department of Econometrics and Business Statistics, February 2002. URL <https://ideas.repec.org/p/msh/ebswps/2002-3.html>.
- Evangelos Spiliotis, Vassilios Assimakopoulos, and Spyros Makridakis. Generalizing the theta method for automatic forecasting. *European Journal of Operational Research*, 284(2):550–558, 2020. ISSN 0377-2217. doi: <https://doi.org/10.1016/j.ejor.2020.01.007>. URL <https://www.sciencedirect.com/science/article/pii/S0377221720300242>.
- Evangelos Spiliotis, Mahdi Abolghasemi, Rob J. Hyndman, Fotios Petropoulos, and Vassilios Assimakopoulos. Hierarchical forecast reconciliation with machine learning. *Applied Soft Computing*, 112:107756, 2021. ISSN 1568-4946. doi: <https://doi.org/10.1016/j.asoc.2021.107756>. URL <https://www.sciencedirect.com/science/article/pii/S1568494621006773>.
- Souhaib Ben Taieb, James W. Taylor, and Rob J. Hyndman. Hierarchical probabilistic forecasting of electricity demand with smart meter data. *Journal of the American Statistical Association*, 116(533):27–43, 2021. doi: 10.1080/01621459.2020.1736081. URL <https://doi.org/10.1080/01621459.2020.1736081>.

- James W. Taylor. Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19(4):715–725, 2003. ISSN 0169-2070. doi: [https://doi.org/10.1016/S0169-2070\(03\)00003-7](https://doi.org/10.1016/S0169-2070(03)00003-7). URL <https://www.sciencedirect.com/science/article/pii/S0169207003000037>.
- Henri Theil. Applied economic forecasting. *American Journal of Agricultural Economics*, 49(1_Part_I):241–243, 1966. URL https://EconPapers.repec.org/RePEc:oup:ajagec:v:49:y:1967:i:1_part_i:p:241-243.
- Granville Tunnicliffe Wilson. Time series analysis: Forecasting and control, 5th edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. isbn: 978-1-118-67502-1. *Journal of Time Series Analysis*, 37, 03 2016. doi: 10.1111/jtsa.12194.
- Marco Villegas García and Diego Pedregal. Supply chain decision support systems based on a novel hierarchical forecasting approach. *Decision Support Systems*, 114, 08 2018. doi: 10.1016/j.dss.2018.08.003.
- Shanika L. Wickramasuriya, George Athanasopoulos, and Rob J. Hyndman. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 114(526):804–819, 2019. doi: 10.1080/01621459.2018.1448825. URL <https://doi.org/10.1080/01621459.2018.1448825>.
- Peter R. Winters. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3):324–342, 1960. doi: 10.1287/mnsc.6.3.324. URL <https://doi.org/10.1287/mnsc.6.3.324>.
- Gokhan Mert Yagli, Dazhi Yang, and Dipti Srinivasan. Automatic hourly solar forecasting using machine learning models. *Renewable and Sustainable Energy Reviews*, 105(C):487–498, 2019. doi: 10.1016/j.rser.2019.02.00. URL <https://ideas.repec.org/a/eee/rensus/v105y2019icp487-498.html>.
- Dazhi Yang, Jan Kleissl, Christian A. Gueymard, Hugo T.C. Pedro, and Carlos F.M. Coimbra. History and trends in solar irradiance and pv power forecasting: A preliminary assessment and review using text mining. *Solar Energy*, 168:60–101, 2018. ISSN 0038-092X. doi: <https://doi.org/10.1016/j.solener.2017.11.023>. URL <https://www.sciencedirect.com/science/article/pii/S0038092X17310022>. Advances in Solar Resource Assessment and Forecasting.
- Σπηλιώτης Ε. and Ασημακόπουλος Β. *Ανάπτυξη Πλαισίου Αυτοματοποιημένης Προέκτασης Χρονοσειρών Μέσω Της Γενίκευσης Της Μεθόδου Πρόβλεψης Θ*. PhD thesis, Εθνικό Μετσόβιο Πολυτεχνείο, 2017. URL <http://artemis.cslab.ece.ntua.gr:8080/jspui/handle/123456789/9095>.

Ατυχείς Προβλέψεις #10

«Δεν υπάρχει πιθανότητα το iPhone να κατακτήσει σημαντικό μερίδιο της αγοράς. Καμία πιθανότητα.»

*Steve Ballmer
CEO της Microsoft, 2007*

A Διαγωνισμοί M

Οι διαγωνισμοί M που οργανώνονται από τον Σπύρο Μακρινδάκη, αποτελούν μια σειρά ανοιχτών διαγωνισμών για αξιολόγηση και σύγκριση της ακρίβειας των διαφορετικών μεθόδων πρόβλεψης χρονοσειρών. Σκοπεύουν στην βαθύτερη κατανόηση και εξέλιξη της μεθοδολογίας της πρόβλεψης μέσω της σύγκρισης της απόδοσης διαφορετικών μεθόδων πάνω σε καλώς ορισμένα προβλήματα του πραγματικού κόσμου. Οι Granger and Newbold [1974] ήταν οι πρώτοι οι οποίοι τη δεκαετία του 1970 θέλησαν να εκτιμήσουν εμπειρικά την ακρίβεια των μεθόδων πρόβλεψης μέσω της εφαρμογής τους πάνω σε σετ δεδομένων. Λόγω βέβαια του υψηλού υπολογιστικού κόστους για την εποχή, ο αριθμός των χρονοσειρών που χρησιμοποιήθηκε στις μελέτες τους ήταν περιορισμένος.

Οι Makridakis, Hibon, and Moser [1979] έκαναν μία πρώτη προσπάθεια να εκτιμήσουν την ακρίβεια των μεθόδων χρησιμοποιώντας 111 χρονοσειρές μακροοικονομικών δεδομένων. Το βασικότερο συμπέρασμα της μελέτης τους ήταν ότι οι πιο περίπλοκες μέθοδοι δεν παρήγαγαν αναγκαστικά και ακριβέστερες προβλέψεις. Η εν λόγω έρευνα συνάντησε αντιδράσεις από ομάδα στατιστικολόγων, οι οποίοι δεδομένης της μη θεωρητικής τεκμηρίωσης των αποτελεσμάτων από τους Makridakis και Hibon, συνέχιζαν να υποστηρίζουν πως πιο σύνθετα μοντέλα είχαν υψηλότερη προβλεπτική ικανότητα.

Οι διαγωνισμοί M γνωρίζουν συνεχώς αυξανόμενο ενδιαφέρον. Στον πρώτο διαγωνισμό συμμετείχαν μόνο οι Makridakis και Hibon ενώ στον πέμπτο διαγωνισμό συμμετείχαν 5507 ομάδες με συνολικά 7092 συμμετέχοντες από 101 χώρες. Μια ιστορική παρουσίαση των διαγωνισμών πρόβλεψης, με ενδιαφέροντα χαρακτηριστικά και συμπεράσματα και για τους διαγωνισμούς M δημοσιεύτηκε και από τον Hyndman [2020].

A.1 Διαγωνισμός M1

Ο πρώτος διαγωνισμός M διεξήχθη το 1982 [Makridakis et al., 1982] και περιελάμβανε προβλέψεις από 24 μοντέλα και παραλλαγές τους για καθεμιά από 1001 χρονοσειρές πάνω σε μικροοικονομικά, μακροοικονομικά, βιομηχανικά και δημογραφικά μεγέθη. Η εξαγωγή των προβλέψεων έγινε από τους Makridakis και Hibon και είχε ως στόχο την υποστήριξη της ιδέας, οι μέθοδοι πρόβλεψης -πέρα από τις μαθηματικές τους ιδιότητες και προδιαγραφές- να συγκρίνονται πάνω στην πρόβλεψη πραγματικών δεδομένων. Βασικό εμπειρικό συμπέρασμα του M1 ήταν πως περιπλοκότερες μέθοδοι δεν παρήγαγαν αναγκαστικά καλύτερες προβλέψεις από πιο απλές μεθόδους. Πιο αναλυτικά:

1. Στη γενική περίπτωση οι πιο πολύπλοκες μέθοδοι δεν μπορούν να εγγυηθούν και καλύτερες προβλέψεις.
2. Με την αύξηση του ορίζοντα πρόβλεψης αυξάνονται τα σφάλματα πρόβλεψης
3. Η κατάταξη των μεθόδων που συμμετέχουν στο διαγωνισμό μεταβάλλεται ανάλογα με το μέτρο απόδοσης που χρησιμοποιείται για τη σύγκριση.
4. Ο συνδυασμός απλών μεθόδων αποτελεί συνήθως καλύτερη επιλογή από την πρόβλεψη με τις επιμέρους μεθόδους.

Τα δεδομένα που τέθηκαν υπό πρόβλεψη στον Μ1, μπορούν να ανακτηθούν [εδώ](#).

Α.2 Διαγωνισμός Μ2

Ο δεύτερος διαγωνισμός Μ διεξήχθη στο διάστημα 1993-1995 Makridakis et al. [1993], με συμμετοχή περισσότερων ερευνητών για την εξαγωγή προβλέψεων. Στόχος του διαγωνισμού υπήρξε η περαιτέρω ενίσχυση των αποτελεσμάτων του Μ1. Καινοτομία του διαγωνισμού ήταν πως πέρα από τις 6 μακροοικονομικές χρονοσειρές τεσσάρων εταιρειών, οι ερευνητές είχαν πρόσβαση και σε εξωγενή πληροφορία μέσω της επικοινωνίας τους με μέλη των επιχειρήσεων ή της συλλογής στοιχείων που αφορούσαν την οικονομία, την αγορά κλπ.

Τα εμπειρικά συμπεράσματα που προέκυψαν ήταν παρόμοια με αυτά του Μ1, με τις πιο απλές μεθόδους να σημειώνουν καλύτερη απόδοση συγκριτικά με τις πιο σύνθετες και τον συνδυασμό αυτών να αποδεικνύεται καλύτερη επιλογή από την εφαρμογή των επιμέρους μεθόδων μεμονωμένα. Τέλος, η ακρίβεια των διαφόρων μεθόδων πρόβλεψης φάνηκε να επηρεάζεται αρνητικά από το μέγεθος του ορίζοντα πρόβλεψης ενώ, ανάλογα με το κριτήριο αξιολόγησης που χρησιμοποιόταν, άλλαζε η κατάταξη των μεθόδων που συμμετείχαν στο διαγωνισμό.

Τα δεδομένα που τέθηκαν υπό πρόβλεψη στον Μ2, μπορούν να ανακτηθούν [εδώ](#).

Α.3 Διαγωνισμός Μ3

Ο διαγωνισμός Μ3 διεξήχθη το 2000 [Makridakis and Hibon, 2000, Koning et al., 2005]. Αποτέλεσε το μεγαλύτερο διαγωνισμό μέχρι τη διεξαγωγή του και τα συμπεράσματά του αξιοποιήθηκαν από πληθώρα ερευνητών. Χορηγός υπήρξε το *International Journal of Forecasting* και συμμετείχαν 25 ομάδες, εκ των οποίων οι 19 ήταν ακαδημαϊκές και οι 6 εξειδικευμένα πακέτα λογισμικού επιχειρηματικών προβλέψεων. Βασικά χαρακτηριστικά του διαγωνισμού ήταν αφενός το σαφώς μεγαλύτερο πλήθος χρονοσειρών διαφόρων κατηγοριών (3003 χρονοσειρές στα: μικροοικονομικά, μακροοικονομικά, βιομηχανία, χρηματοοικονομικά, δημογραφικά και λοιπά) και αφετέρου η συμμετοχή μη στατιστικών μεθόδων πρόβλεψης, όπως νευρωνικά δίκτυα (*neural networks*) και έξυπνα συστήματα (*expert systems*). Συνοπτικά, από τις στατιστικές μεθόδους που συμμετείχαν αναφέρουμε τις μεθόδους εκθετικής εξομάλυνσης και τα μοντέλα ARIMA, από τα νευρωνικά δίκτυα τις τεχνικές *RBF* και από τα έξυπνα συστήματα τα *ForecastPro*, *SmartForecasts*, *Autocast* και *Autobox*. Με αυτό τον τρόπο, τα αποτελέσματα των προηγούμενων διαγωνισμών Μ θα δοκιμάζονταν σε μεγαλύτερο όγκο δεδομένων και για περισσότερες και πιο εξειδικευμένες μεθόδους. Σύμφωνα με τον Rob J. Hyndman, *Editor-in-Chief* του περιοδικού *International Journal of Forecasting* (IJF) και ενός εκ των κορυφαίων ερευνητών της πρόβλεψης: "*The M3 data have continued to be used since 2000 for testing new time series forecasting methods. In fact, unless a proposed forecasting method is competitive against the original M3 participating methods, it is difficult to get published in the IJF.*"

Τα δεδομένα που τέθηκαν υπό πρόβλεψη στον M3, μπορούν να ανακτηθούν [εδώ](#).

A.4 Διαγωνισμός M4

Ο τέταρτος διαγωνισμός Makridakis et al. [2020a] ανακοινώθηκε το Νοέμβριο του 2017. Ξεκίνησε την 1η Ιανουαρίου του 2018 και τελείωσε στις 31 Μαΐου του 2018. Ο M4 αναπαρήγαγε αλλά και επέκτεινε τα συμπεράσματα των προηγούμενων τριών διαγωνισμών, χρησιμοποιώντας ένα ποικιλόμορφο σύνολο δεδομένων για να αξιολογηθούν οι εξεταζόμενες μέθοδοι σε προβλέψεις με διαφορετικά χαρακτηριστικά. Για να προκύψουν ακριβή και αξιόπιστα αποτελέσματα, ο M4 έκανε χρήση 100000 χρονοσειρών από πραγματικές εφαρμογές και περιελάμβανε όλες τις κύριες ποσοτικές μεθόδους πρόβλεψης χρονοσειρών, συμπεριλαμβανομένων και αυτών που βασίζονται στην τεχνητή νοημοσύνη (μηχανική μάθηση), πέρα από τις παραδοσιακές στατιστικές μεθόδους. Στο blog του ο Rob J Hyndman, έγραψε για τον M4: "*The "M" competitions organized by Spyros Makridakis have had an enormous influence on the field of forecasting. They focused attention on what models produced good forecasts, rather than on the mathematical properties of those models. For that, Spyros deserves congratulations for changing the landscape of forecasting research through this series of competitions.*"

Ο διαγωνισμός, πέρα από την παραγωγή σημειακών προβλέψεων, εξέταζε και διαστήματα εμπιστοσύνης πρόβλεψης. Τα βασικά συμπεράσματα που προέκυψαν από την αξιολόγηση των αποτελεσμάτων του διαγωνισμού M4 [Makridakis et al., 2018b], είναι:

1. Ο συνδυασμός μεθόδων υπήρξε η πιο πετυχημένη επιλογή πρόβλεψης. Στις 17 μεθόδους με τις ακριβέστερες προβλέψεις, οι 14 αποτελούσαν συνδυασμού μεθόδων, κυρίως στατιστικών.
2. Τη μεγαλύτερη «έκπληξη» αποτέλεσε μια υβριδική μέθοδος που συνδύαζε στατιστικές μεθόδους και μεθόδους μηχανικής μάθησης. Προέβλεψε με τη μεγαλύτερη ακρίβεια τόσο σημειακές όσο και πιθανοτικές προβλέψεις και κατατέθηκε από τον Slawek Smyl, επιστήμονα δεδομένων στην Uber Technologies. Σύμφωνα με το sMAPE, βελτίωσε κατά 10% την ακρίβεια των προβλέψεων σε σχέση με το benchmark COMB, ενώ στον M3 η ακριβέστερη μέθοδος βελτίωσε κατά 4% τις προβλέψεις σε σχέση με το ίδιο benchmark.
3. Η δεύτερη ακριβέστερη μέθοδος ήταν συνδυασμός 7 στατιστικών μεθόδων και μιας μεθόδου μηχανικής μάθησης, που αφορούσε τον υπολογισμό των βαρών του σταθμισμένου μέσου των 7 στατιστικών μεθόδων που θα αποτελούσε την τελική πρόβλεψη. Κατατέθηκε από κοινού από το Πανεπιστήμιο της A Coruna και το Πανεπιστήμιο Monash στην Αυστραλία.
4. Η πρώτη και δεύτερη ακριβέστερη μέθοδος που περιγράφηκαν παραπάνω, πέτυχαν εντυπωσιακά αποτελέσματα και στον υπολογισμό των διαστημάτων εμπιστοσύνης, εκτιμώντας ορθά το 95%. Αποτελούν τις πρώτες μεθόδους που πέτυχαν τέτοιες

ακρίβειες και δεν υποεκτιμούν την αβεβαιότητα σημαντικά, όπως μέχρι εκείνη την περίοδο.

5. Οι 6 μέθοδοι καθαρής μηχανικής μάθησης παρουσίασαν μέτρια έως κακή απόδοση. Καμία δεν υπεραπέδωσε του benchmark COMB ενώ μόλις μία υπεραπέδωσε της Naive2. Αντίστοιχες μέθοδοι μηχανικής μάθησης με αυτές που αξιολογήθηκαν στον M4, μελετήθηκαν και από τους Makridakis, Spiliotis, and Assimakopoulos [2018a].

Οι Makridakis, Spiliotis, and Assimakopoulos [2020b] δημοσίευσαν τις υποθέσεις που είχαν κάνει για τα επερχόμενα αποτελέσματα του M4, ορισμένες από τις οποίες αποτελούν διαχρονικά χαρακτηριστικά των διαγωνισμών M και άλλες που συμπεριλάμβαναν την ολοένα αυξανόμενη προβλεπτική δύναμη των μεθόδων μηχανικής μάθησης. Στη μεγάλη πλειοψηφία τους επαληθεύτηκαν, όπως προκύπτει κι από τα αποτελέσματα του M4.

Τα δεδομένα που τέθηκαν υπό πρόβλεψη στον M4, μπορούν να ανακτηθούν [εδώ](#).

A.5 Διαγωνισμός M5

Ο πέμπτος διαγωνισμός M5 [Makridakis et al., 2021] ξεκίνησε στις 3 Μαρτίου 2020 και τα αποτελέσματα ανακοινώθηκαν την 1η Ιουλίου του 2020. Αφορούσε πραγματικά δεδομένα πωλήσεων και διεξήχθη στην πλατφόρμα Kaggle. Περιελάμβανε υψηλά οικονομικά βραβεία για τους νικητές, που ανέρχονταν συνολικά σε 100.000\$. Τα δεδομένα παρήχθησαν από την Walmart, την εταιρεία με τα περισσότερα έσοδα στον πλανήτη και αφορούσαν 42000 ιεραρχικές χρονοσειρές ημερήσιων παρατηρήσεων, ξεκινώντας από τις πωλήσεις μονάδων και φτάνοντας μέχρι τη συνολική ζήτηση σε μια μεγάλη γεωγραφική περιφέρεια. Πέρα από τα δεδομένα των πωλήσεων, το σύνολο δεδομένων της Walmart περιελάμβανε πληροφορίες για τις τιμές των προϊόντων, διαφημιστικές/προωθητικές ενέργειες και επίπεδα αποθεμάτων μαζί με την ημέρα στην οποία αναφέρεται κάθε στοιχείο. Ο κύριος στόχος του διαγωνισμού ήταν η συμμετοχή όσων εφαρμόζουν τις προβλέψεις αντί των ερευνητών/ακαδημαϊκών. Ο M5 συγκέντρωσε υψηλό ενδιαφέρον, με τη συμμετοχή περίπου 6000 συμμετεχόντων και ομάδων. Βασικά στοιχεία που προέκυψαν από την αξιολόγηση των αποτελεσμάτων του διαγωνισμού M5 από τους Makridakis, Spiliotis, and Assimakopoulos [2020c] είναι:

1. Η υπεροχή των μεθόδων μηχανικής μάθησης, σε σχέση με τις παραδοσιακές στατιστικές μεθόδους. Για χρόνια, τα εμπειρικά αποτελέσματα έδειχναν πως συνθετότερες μέθοδοι πρόβλεψης δεν οδηγούσαν απαραίτητα σε υψηλότερη ακρίβεια. Ο M4 ήταν ο πρώτος διαγωνισμός στον οποίο 2 μέθοδοι μηχανικής μάθησης ήταν ακριβέστερες από τις στατιστικές. Στο διαγωνισμό M5 -δηλαδή σχεδόν 4 χρόνια μετά- όλες οι κορυφαίες μέθοδοι που υποβλήθηκαν χρησιμοποιούσαν καθαρή μηχανική μάθηση.
2. Η αξία του συνδυασμού προβλέψεων. Ο διαγωνισμός M5 επιβεβαίωσε το συμπέρασμα τόσο των προηγούμενων διαγωνισμών M όσο και ερευνητών της

πρόβλεψης, ότι συνδυασμός μεθόδων -ακόμη και απλών- οδηγεί σε βελτίωση της ακρίβειας.

3. Η αξία του "cross-learning", δηλαδή της αξιοποίησης της πληροφορίας που προκύπτει από το γεγονός ότι οι χρονοσειρές του διαγωνισμού ήταν συσχετισμένες¹⁵
4. Η σημαντικές διαφορές ανάμεσα στις νικητήριες υποβολές και τα benchmarks που χρησιμοποιήθηκαν για την πρόβλεψη πωλήσεων. Ο διαγωνισμός αξιολόγησε 24 benchmarks που περιλαμβάνουν τις κυρίαρχες στατιστικές μεθόδους πρόβλεψης. Οι νικητήριες μέθοδοι παρουσίασαν σημαντική βελτίωση στην ακρίβεια της πρόβλεψης και ήταν μεσοσταθμικά ακριβέστερες κατά περισσότερο από 20%.
5. Η ευεργετική επίδραση των εξωγενών προσαρμογών. Ενώ στο διαγωνισμό M2 οι εξωγενείς προσαρμογές στις προβλέψεις δεν έδειξαν να βελτιώνουν σημαντικά την ακρίβεια πρόβλεψης, στο διαγωνισμό M5 η 2η και 5η σε ακρίβεια μέθοδοι χρησιμοποίησαν τέτοιες προσαρμογές σε μορφή πολλαπλασιαστών, για να ενισχύσουν τις προβλέψεις τους.
6. Η προσθετική αξία κατάλληλων στρατηγικών επικύρωσης (*Cross Validation* ή *CV*). Οι 4 κορυφαίες υποβολές του διαγωνισμού, καθώς και η πλειονότητα των 50 κορυφαίων υποβολών, χρησιμοποίησαν τεχνικές επικύρωσης των προβλέψεων, όπου τουλάχιστον οι 4 τελευταίες ημέρες από τα παράθυρα 28 ημερών των ιστορικών παρατηρήσεων χρησιμοποιήθηκαν για την αξιολόγηση των προβλέψεων, παρέχοντας μια καλή προσέγγιση της post-sample απόδοσης των προβλέψεων.
7. Η σημασία των εξωγενών/επεξηγηματικών μεταβλητών. Οι μέθοδοι πρόβλεψης χρονοσειρών, είναι τυπικά αποτελεσματικές στην αποτύπωση των μοτίβων (στάθμη, τάση, εποχιακότητα) που εμφανίζονται στις ιστορικές παρατηρήσεις. Ωστόσο, μέθοδοι που λαμβάνουν υπόψη αποκλειστικά τις ιστορικές παρατηρήσεις, αποτυγχάνουν να αποτυπώσουν την επίδραση αργιών, ειδικών γεγονότων, εκπτώσεων, τιμών και πιθανά του καιρού. Όλες οι νικητήριες μέθοδοι αξιοποίησαν την παραπάνω πληροφορία που περιεχόταν στο σύνολο δεδομένων. Για παράδειγμα, οι *monsaraida* και άλλες κορυφαίες ομάδες αντιλήφθηκαν και αξιοποίησαν τη θετική επίδραση που έχουν στις προβλέψεις, χαρακτηριστικά που σχετίζονται με τις τιμές πώλησης των προϊόντων. Η αξία των εξωγενών μεταβλητών έγινε αισθητή και κατά τη σύγκριση των 24 benchmarks που αξιολογήθηκαν πάνω στο σύνολο δεδομένων του διαγωνισμού. Πιο συγκεκριμένα,

Τα δεδομένα που τέθηκαν υπό πρόβλεψη στον M5, μπορούν να ανακτηθούν [εδώ](#).

¹⁵Για πρώτη φορά στο διαγωνισμό M5 παραχωρήθηκαν χρονοσειρές που συσχετίζονται μέσω περιορισμών συνάντησης. Στους προηγούμενους διαγωνισμούς οι χρονοσειρές που παρέχονταν ήταν ασυσχέτιστες.

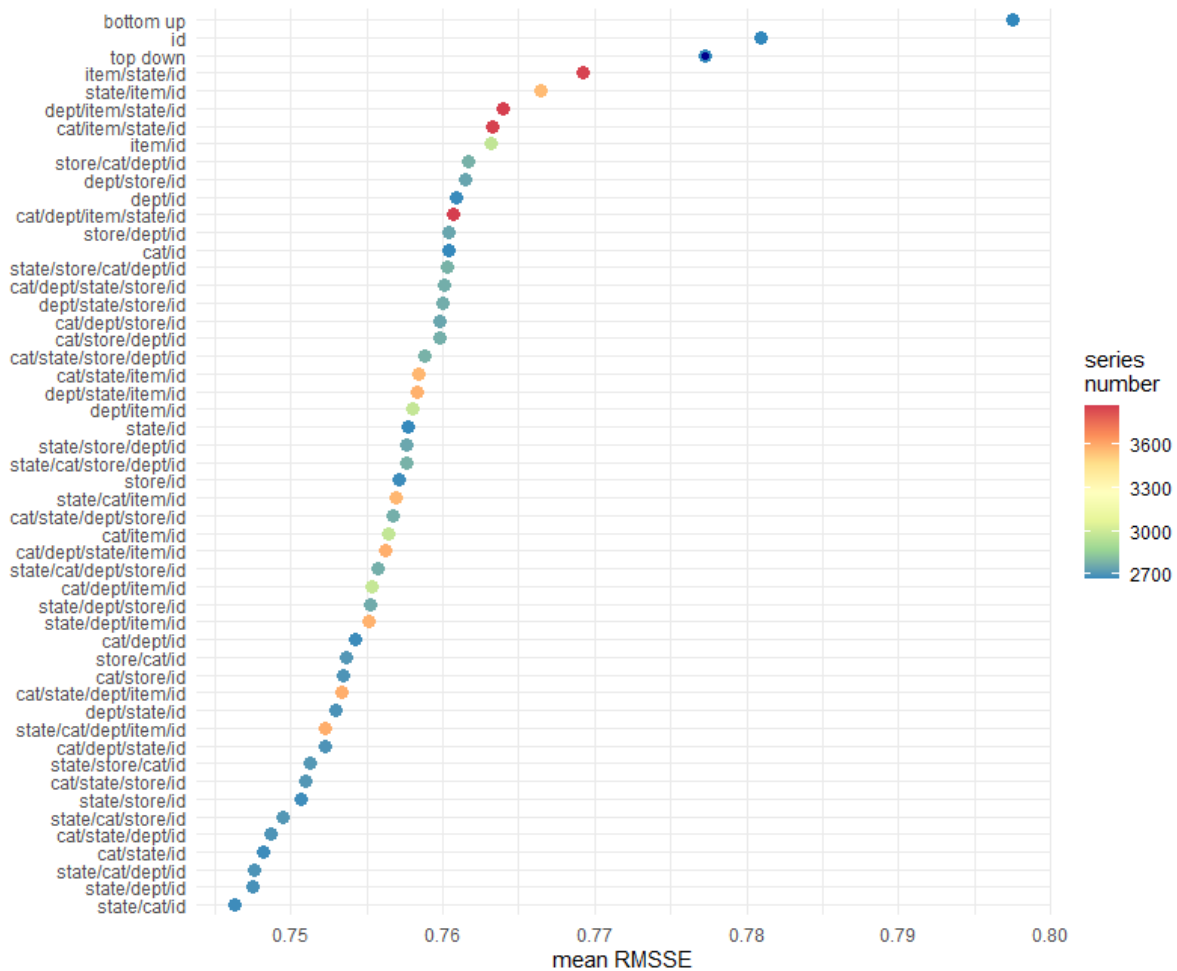
Ατυχείς Προβλέψεις #11

«Η πτήση με μηχανές βαρύτερες από τον αέρα είναι μη πρακτική και ανάξια λόγου, αν όχι εντελώς αδύνατη.»

*Simon Newcomb, Αμερικανο-Καναδός
αστρονόμος και μαθηματικός, 1902*

B Λοιπά Διαγράμματα Μελέτης Περίπτωσης

Όπως φάνηκε, οι προβλέψεις που είχαν εναρμονιστεί μέσω της μεθόδου *Minimum Trace* με κάποια από τις 48 εξεταζόμενες ιεραρχίες, παρουσίαζαν βελτίωση στην ακρίβεια των προβλέψεων τόσο μεσοσταθμικά σε όλα τα επίπεδα, όσο και σε καθένα από τα 12 ξεχωριστά επίπεδα. Ωστόσο, -και λόγω των μικρών διαφορών στην ακρίβεια- εμφανίζονταν διαφορές στη σχετική κατάταξη της ακρίβειας πρόβλεψης των εξεταζόμενων ιεραρχιών. Σε λίγες περιπτώσεις μάλιστα, όπως θα δούμε και παρακάτω τα benchmarks παρήγαγαν ακριβέστερες προβλέψεις από τις εναρμονισμένες με *Minimum Trace* εξεταζόμενες ιεραρχίες.

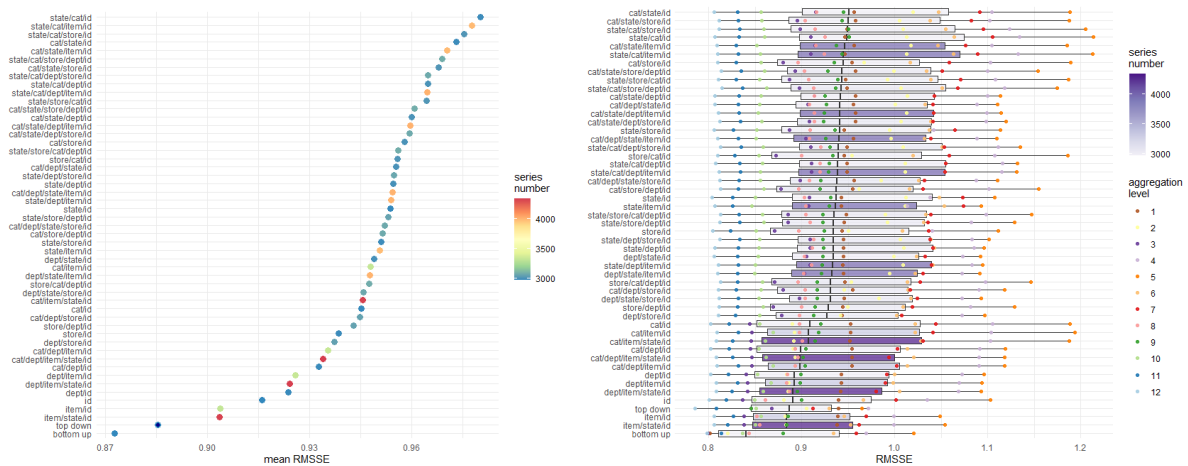


Σχήμα 29: Κατάταξη ως προς τη Μέση Ακρίβεια των 50 Επιλογών Πρόβλεψης
Δείγμα 2672 Μη Συναθροισμένων Χρονοσειρών - 25 προϊόντα ανά υποκατηγορία

Στα σχήματα 29 και 30α που παρουσιάζονται στην παρούσα παράγραφο (B) του παραρτήματος της εργασίας, φαίνονται οι κατατάξεις των ιεραρχιών ως προς τη μέση ακρίβεια σε δύο άλλα δείγματα του συνόλου δεδομένων. Οι ιεραρχίες που προκύπτουν από το

δείγμα που προβλέφθηκε στο σχήμα 29 περιέχουν από 2672 έως 3870 χρονοσειρές, ενώ οι ιεραρχίες του δείγματος στο σχήμα 30a περιέχουν από 2976 έως 4342 χρονοσειρές.

Από την αντιπαράθεση των σχημάτων 16, 29 και 30a, στα οποία απεικονίζονται οι κατατάξεις της μέσης ακρίβειας των εξεταζόμενων ιεραρχιών και των benchmarks για διαφορετικά δείγματα του συνόλου δεδομένων, γίνεται φανερό πως οι κατατάξεις μπορεί να είναι διαφορετικές για διαφορετικά υποσύνολα του εκάστοτε συνόλου, κι εν προκειμένω του συνόλου του M5. Ανακατατάξεις παρατηρούνται τόσο ανάμεσα στις 48 ιεραρχίες με τη *Minimum Trace* -με όλες να αποδίδουν καλύτερα από τα benchmarks: top down και bottom up (όπως ανάμεσα στα σχήματα 16 και 29) όσο και ανάμεσα στην απόδοση των benchmarks σε σχέση με τις εξεταζόμενες εναρμονισμένες προβλέψεις (όπως ανάμεσα στα σχήματα 29 και 30a).



(a) Κατάταξη ως προς τη Μέση Ακρίβεια των 50 Επιλογών Πρόβλεψης (b) Ανάλυση της Κατάταξης του Σχ. 30a στις Ακρίβειες Πρόβλεψης των 12 Επιπέδων

Σχήμα 30: Κατάταξη και Boxplots Δείγματος στο οποίο τα Benchmarks προέβλεψαν με Υψηλότερη Μεσοσταθμική Ακρίβεια από τις Εξεταζόμενες Ιεραρχίες
 Δείγμα 2976 Μη Συναρθρισμένων Χρονοσειρών - 50 προϊόντα ανά υποκατηγορία

Η συμπερίληψη των παραπάνω διαγραμμάτων έχει ως στόχο να αναδείξει τις μεταβολές που προκύπτουν στην κατάταξη της ακρίβειας των προβλέψεων ανά δείγμα, των 50 εξεταζόμενων μεθόδων πρόβλεψης. Όπως αναφέρθηκε στα συμπεράσματα της παρούσας εργασίας (κεφάλαιο 7), στη μεγάλη πλειοψηφία των δειγμάτων προέκυψαν κατατάξεις με κοινά χαρακτηριστικά αλλά και επιμέρους διαφορές και ιδιαιτερότητες. Με βάση αυτά τα αποτελέσματα, η εύρεση μίας ή περισσότερων ιεραρχιών που ενισχύει κατά βέλτιστο τρόπο τις προβλέψεις βάσης μια συλλογής ιεραρχικών ή ομαδοποιημένων χρονοσειρών, θα είναι πιο αξιόπιστη αν προκύπτει από συνδυασμό μετρήσεων και σύγκρισης σε διαφορετικά χρονικά διαστήματα των ιστορικών παρατηρήσεων και πιθανά πάνω σε διαφορετικά δείγματα των δεδομένων.

ΠΡΟΒΛΕΨΕΙΣ ΕΙΔΙΚΩΝ (1964):

- ✘ 1980: Ρομπότ οικιακοί βοηθοί.
- ✘ 1980: Άνθρωποι στον Άρη.
- ✘ 1995: 50 χρόνια επέκτασης ζωής.
- ✘ 1995: Αυτόματες ψηφοφορίες/εκλογές.
- ✘ 1998: Απευθείας εγγραφή στην εγκέφαλο.
- ✘ 1999: Στρατιωτική δύναμη στη Σελήνη.
- ✘ 2000: Επικοινωνία με εξωγήινους.
- ✘ 2015: Τεχνητό κώμα για ταξίδι στο χρόνο.
- ✘ 2020: Εκτροπή πιθήκων για εργασία.

Ατυχείς Προβλέψεις #12

Το 1964 η *RAND Corporation* συντάξε μια αναφορά πάνω σε μακροπρόθεσμες προβλέψεις, ρωτώντας 82 ειδικούς από διάφορα επιστημονικά πεδία να κάνουν μια σειρά από προβλέψεις για το μέλλον της τεχνολογίας. Έξι δεκαετίες μετά, όπως προκύπτει και από τη λίστα που δημοσίευσε η *RAND* (αριστερά), οι ειδικοί απέτυχαν πλήρως να συλλάβουν την ταχύτητα με την οποία θα εξελισσόταν η μέχρι τότε υπάρχουσα τεχνολογία. Η πιο εξόφθαλμη αστοχία είναι ίσως η εκτροπή πιθήκων για χειρωνακτικές και οικοκυρικές εργασίες. Τέτοιοι ισχυρισμοί αναπαράγονται και σε άλλες περιπτώσεις. Το 1967, ο βραβευμένος με *Nobel* Χημικός *Glenn T. Seaborg*, σχολίαζε σε μια ομιλία του στο *Woman's National Democratic Club* της *Washington* των *ΗΠΑ*: «Στον 21ο αιώνα, όσα σπίτια δεν έχουν ρομπότ μέσα στην ντουλάπα τους, θα μπορούν να έχουν έναν «εσωτερικό» πίθηκο για τις δουλειές του σπιτιού».