



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ  
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ  
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΗΣ

**Ανάπτυξη Διαδικτυακής Εφαρμογής Προβλέψεων και  
Αξιολόγησής τους με Αυτόματη Συλλογή και Ανανέωση των  
Χρονοσειρών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΚΑΡΟΖΗΣ Ε. ΘΕΟΔΩΡΟΣ

**Επιβλέπων:** Βασίλειος Ασημακόπουλος

Καθηγητής, Ε.Μ.Π.

**Υπεύθυνη:** Νικολέττα-Ζαμπέτα Λεγάκη

Υποψήφια Διδάκτωρ, Ε.Μ.Π.

Αθήνα, Οκτώβριος 2015





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ  
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ  
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΗΣ

**Ανάπτυξη Διαδικτυακής Εφαρμογής Προβλέψεων και  
Αξιολόγησής τους με Αυτόματη Συλλογή και Ανανέωση των  
Χρονοσειρών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΚΑΡΟΖΗΣ Ε. ΘΕΟΔΩΡΟΣ

**Επιβλέπων:** Βασίλειος Ασημακόπουλος

Καθηγητής, Ε.Μ.Π.

**Υπεύθυνη:** Νικολέττα-Ζαμπέτα Λεγάκη

Υποψήφια Διδάκτωρ, Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την \_\_\_\_\_ 2015

\_\_\_\_\_  
Βασίλειος Ασημακόπουλος

Καθηγητής Ε.Μ.Π

\_\_\_\_\_  
Ιωάννης Ψαράς

Καθηγητής Ε.Μ.Π.

\_\_\_\_\_  
Δημήτριος Ασκούνης

Αναπληρωτής Καθηγητής Ε.Μ.Π

Αθήνα, Οκτώβριος 2015

---

Καρόζης Ε. Θεόδωρος

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

Copyright © 2015 Καρόζης Ε. Θεόδωρος

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

# Περίληψη

Σκοπός της παρούσας διπλωματικής εργασίας είναι ο σχεδιασμός και η υλοποίηση μίας διαδικτυακής εφαρμογής προβλέψεων για συγκεκριμένες κατηγορίες δεδομένων και η συλλογή δεδομένων με live ανατροφοδότηση. Η ταχεία ανάπτυξη της τεχνολογίας σε συνδυασμό με την κατακόρυφη αύξηση του ενδιαφέροντος για την επιστήμη των προβλέψεων, καθιστά τις διαδικτυακές εφαρμογές ένα ενδιαφέρον πεδίο σχεδιασμού και δημιουργίας εφαρμογών προβλέψεων, οι οποίες αποτελούν ένα βασικό εργαλείο στην διαδικασία λήψης των επιχειρησιακών, και όχι μόνο, αποφάσεων.

Αρχικά, αναφέρεται το θεωρητικό υπόβαθρο γύρω από τις προβλέψεις, όπου αναλύεται η έννοια των χρονοσειρών και η διαδικασία επεξεργασίας των χαρακτηριστικών της, το οποίο προηγείται της παραγωγής προβλέψεων. Κατόπιν, παρουσιάζονται αναλυτικά οι μέθοδοι που χρησιμοποιούνται στην εφαρμογή, όπως επίσης και οι δείκτες αξιολόγησης που χρησιμοποιούνται ευρέως για την σύγκριση των μεθόδων.

Στη συνέχεια, παρουσιάζεται ο σχεδιασμός και η υλοποίηση της εφαρμογής. Η εισαγωγή των χρονοσειρών γίνεται από τον χρήστη, όπως επίσης και η επιλογή της παρουσίασης τους σε μορφή πίνακα ή σε διαγραμματική απεικόνιση. Οι χρονοσειρές επιλέγονται μέσα από τις κατηγορίες των αγαθών, των μετοχών, του τουρισμού και της ανεργίας που αντιστοιχούν σε συγκεκριμένα πεδία εφαρμογής των προβλέψεων. Θα πρέπει να τονιστεί πως τα δεδομένα ανανεώνονται συνεχώς με τις πιο πρόσφατες διαθέσιμες παρατηρήσεις. Επίσης δημιουργούνται οι προβλέψεις, για την επιλεγμένη από τον χρήστη χρονοσειρά. Στην κατηγορία των αγαθών υπάρχει μία λειτουργία που επιτρέπει την αυτόματη αξιολόγηση και σύγκριση των μεθόδων πρόβλεψης.

Στο δεύτερο στάδιο της διπλωματικής διερευνήθηκε περαιτέρω η αξιολόγηση και η σύγκριση των μεθόδων πρόβλεψης χρονοσειρών, όπως είναι η naïve, οι μέθοδοι εκθετικής εξομάλυνσης, η απλή γραμμική παλινδρόμηση, η μέθοδος κινητού μέσου όρου και η κλασική μέθοδος Theta, με σκοπό την εξαγωγή συμπερασμάτων σχετικά με την εφαρμογή τους για πρόβλεψη στις συγκεκριμένες χρονοσειρές. Η εξαγωγή και η αξιολόγηση των προβλέψεων αφορά ορισμένα βασικά αγαθά, όπως το αργό πετρέλαιο, το χρυσό, το ασήμι, κλπ.

Τέλος, καταγράφονται τα αποτελέσματα και τα συμπεράσματα τόσο της εφαρμογής, όσο και του πειράματος. Αναφέρονται επίσης οι δυνατότητες και οι πιθανές προεκτάσεις της εφαρμογής με προσθήκη περισσότερων λειτουργιών. Επίσης συμπεριλαμβάνεται εγχειρίδιο χρήσης της υλοποιημένης εφαρμογής.

**Λέξεις κλειδιά:** Τεχνικές προβλέψεων, Μέθοδοι Χρονοσειρών, Διαδικτυακές Εφαρμογές, Εφαρμογή Παραγωγής Προβλέψεων, Αυτόματη Αξιολόγηση Ακρίβειας, Κατάταξη Μεθόδων Πρόβλεψης

# ***Abstract***

The aim of this thesis is the implementation of a forecasting web application for particular datasets and the data collection with live feedback. The rapid development of technology combined with a sharp increase in interest in the science of forecasting makes the web applications an interesting field for designing and creating forecasting applications, which are a key tool in the process of making business decisions.

First of all, the thesis referred to the theoretical background surrounding the field of the forecasting, where is analyzed the notion of time series and the procedure of handling of its characteristics, which precedes the production of forecasts. Then, methods which are used in the application are detailed, as well as the evaluation indicators which are widely used to compare methods.

Then, the author introduce and explain the implementation of the application. The import of time series is done by the user, as well as the choice of the presentation in tabular form or in diagrammatic representation. Time series are selected through the categories of commodities, stocks, tourism and unemployment corresponding to specific fields of forecasting. It should be emphasized that the data is constantly updated with the latest available observations. Also, forecasts are created for the user's selected timeline. In the category of commodities, there is a function that allows the automatic evaluation and comparison of specific forecasting methods.

In the second stage of this thesis, it was investigated the further evaluation and comparison of forecasting methods, such as naïve, methods of exponential smoothing, simple linear regression, moving average method and the Classic Theta Model, in order to extract conclusions regarding their forecast implementation in the specific time series. The extraction and evaluation of forecasts was made for certain commodities, such as Crude Oil, Gold, Silver, etc.

Finally, there was written the results and conclusions of both the application and the experiment. They indicate the possibilities and possible impact of the application by adding more functions. Also, a user manual of the implemented application is included.

**Keywords:** Forecasting Techniques, Time series Methods, Web Applications, Forecasting System, Automated Accuracy Evaluation, Ranking of Forecasting Methods

# Πρόλογος

Η διπλωματική αυτή εργασία εκπονήθηκε στα πλαίσια των ερευνητικών δραστηριοτήτων της Μονάδας Συστημάτων Πρόβλεψης και Στρατηγικής κατά το ακαδημαϊκό έτος 2014 – 2015. Η μονάδα υπάγεται στον Τομέα Ηλεκτρικών Βιομηχανικών Διατάξεων και Συστημάτων Αποφάσεων, της Σχολής Ηλεκτρολόγων Μηχανικών & Μηχανικών Η/Υ, του Εθνικού Μετσόβιου Πολυτεχνείου.

Αρχικά, θα ήθελα να ευχαριστήσω τον Καθηγητή κ. Βασίλη Ασημακόπουλο για τη στήριξη που μου παρείχε στην ολοκλήρωση της διπλωματικής μου εργασίας. Ακόμα, θα ήθελα να ευχαριστήσω τον Καθηγητή κ. Ιωάννη Ψαρρά και τον Αναπληρωτή Καθηγητή κ. Δημήτριο Ασκούνη για την τιμή που μας έκαναν να συμμετάσχουν στην επιτροπή εξέτασης της εργασίας.

Θερμές ευχαριστίες θα ήθελα επίσης να απευθύνω στην Υποψήφια Διδάκτορα κ. Νικολέτα-Ζαμπέτα Λεγάκη, η συμβολή της οποίας υπήρξε καθοριστική για την εκπόνηση της παρούσας διπλωματικής εργασίας. Η βοήθεια και οι συμβουλές της τόσο σε οργανωτικό, όσο και σε επιστημονικό επίπεδο υπήρξαν πολύτιμες.

Τέλος, ευχαριστώ την οικογένεια μου και τους φίλους μου για την συνεχή ηθική συμπαράσταση και υποστήριξή τους.

Θεόδωρος Ε. Καρόζης

Αθήνα, Οκτώβριος 2015





## Περιεχόμενα

1. Ευρεία Περίληψη .....	13
1.1 Εισαγωγή .....	13
1.2 Τεχνικές Προβλέψεων .....	13
1.3 Περιγραφή της Εφαρμογής.....	15
1.4 Σύγκριση Μεθόδων Πρόβλεψης.....	19
1.4.3 Σύγκριση και Επιλογή Μεθόδων .....	23
1.5 Συμπεράσματα-Προοπτικές .....	24
2. Εισαγωγή .....	29
2.1 Χρήση Διαδικτυακών Εφαρμογών .....	29
2.2 Σύνδεση Προβλέψεων-Διαδικτυακών Εφαρμογών .....	30
3. Τεχνικές Προβλέψεων .....	33
3.1 Εισαγωγή .....	33
3.2 Κατηγορίες Μεθόδων Πρόβλεψης.....	35
3.2.1 Ποσοτική πρόβλεψη ( <i>quantitative forecasting</i> ) .....	35
3.2.2 Ποιοτική πρόβλεψη ( <i>qualitative ή judgmental forecasting</i> ) .....	37
3.3 Χρονοσειρές.....	38
3.3.1 Έννοια της Χρονοσειράς .....	38
3.3.2 Ποιοτικά Χαρακτηριστικά της Χρονοσειράς .....	39
3.3.3 Αναπαράσταση Χρονοσειρών.....	40
3.4 Βασικά Βήματα Προβλέψεων.....	42
3.5 Πεδία εφαρμογής των Προβλέψεων .....	44
3.6 Μέθοδοι Πρόβλεψης Συνεχούς Ζήτησης .....	45
3.6.1 Απλοϊκή μέθοδος – <i>Naive</i> .....	45
3.6.2 Μέθοδοι Μέσου Όρου .....	46
3.6.3 Μέθοδοι Εκθετικής Εξομάλυνσης .....	47
3.6.4 Απλή Γραμμική Παλινδρόμηση .....	56
3.6.5 Μοντέλο <i>Theta</i> .....	58
3.7 Δείκτες Αξιολόγησης Προβλέψεων.....	60
4. Περιγραφή της Εφαρμογής .....	67
4.1 Τεχνολογία Λογισμικού.....	67
4.2 Λογισμικό της Εφαρμογής.....	68
4.2.1 Γλώσσα Προγραμματισμού της Εφαρμογής .....	68
4.2.2 Τεχνολογία που χρησιμοποιήθηκε για την άντληση των δεδομένων .....	70
4.3 Υλοποίηση της Εφαρμογής.....	72
4.3.1 Είσοδοι – Έξοδοι Προγράμματος .....	72

4.3.2	<i>Γενική Περιγραφή των Dataset</i> .....	74
4.3.3	<i>Βασικά Πακέτα της Εφαρμογής</i> .....	79
5.	<i>Χρησιμότητα των Δεδομένων(Dataset)</i> .....	83
5.1	<i>Ανάλυση των Αγαθών (Commodities)</i> .....	83
5.1.1	<i>Σημασία της τιμής των αγαθών</i> .....	83
5.1.2	<i>Διαμόρφωση της τιμής των αγαθών</i> .....	85
5.2	<i>Ανάλυση των δεδομένων των τιμών των μετοχών (Stocks)</i> .....	97
5.3	<i>Ανάλυση των Δεδομένων του Τουρισμού (Tourism)</i> .....	99
5.4	<i>Ανάλυση της Ανεργίας (Unemployment)</i> .....	102
5.4.1	<i>Προσδιορισμός της ανεργίας</i> .....	102
5.4.2	<i>Σημασία της καταγραφής της ανεργίας</i> .....	104
6.	<i>Εγχειρίδιο (Manual) Εφαρμογής</i> .....	107
6.1	<i>Είσοδος στην Εφαρμογή και Αρχική Σελίδα</i> .....	107
6.2	<i>Σενάρια Χρήσης</i> .....	109
6.2.1	<i>Επιλογή των Εμπορευμάτων (Commodities)</i> .....	109
6.2.2	<i>Επιλογή των Μετοχών (Stocks)</i> .....	114
6.2.3	<i>Επιλογή του Τουρισμού (Tourism)</i> .....	117
6.2.4	<i>Επιλογή της Ανεργίας (Unemployment)</i> .....	121
7.	<i>Σύγκριση Μεθόδων Πρόβλεψης</i> .....	127
7.1	<i>Γενική Μεθοδολογία</i> .....	127
7.2	<i>Αξιολόγηση Μεθόδων για Πρόβλεψη Πρώτης και Τρίτης Χρονικής Περιόδου</i> .....	129
7.3	<i>Σύγκριση Μεθόδων σύμφωνα με τον δείκτη MAPE</i> .....	166
7.3.1	<i>Προβλέψεις για την επόμενη περίοδο (h=1)</i> .....	167
7.3.2	<i>Προβλέψεις για την τρίτη περίοδο (h=3)</i> .....	168
7.4	<i>Επιλογή Καταλληλότερου Μοντέλου Πρόβλεψης για Κάθε Αγαθό</i> .....	170
7.4.1	<i>Προβλέψεις για την επόμενη περίοδο (h=1)</i> .....	170
7.4.2	<i>Προβλέψεις για την τρίτη περίοδο (h=3)</i> .....	171
7.5	<i>Κατάταξη και Βαθμολόγηση των Μεθόδων Πρόβλεψης Συνολικά στα Αγαθά</i> .....	172
7.5.1	<i>Κατάταξη Μεθόδων Παραγωγής Προβλέψεων για την Επόμενη Περίοδο (h=1)</i> .....	172
7.5.2	<i>Κατάταξη Μεθόδων Παραγωγής Προβλέψεων για την Τρίτη Περίοδο</i> .....	178
8.	<i>Συμπεράσματα-Προοπτικές</i> .....	185
8.1	<i>Συμπεράσματα</i> .....	185
8.2	<i>Προοπτικές</i> .....	190
8.2.1	<i>Προοπτικές Εφαρμογής</i> .....	190
8.2.2	<i>Προοπτικές Ακρίβειας Μεθόδων</i> .....	192

9. Βιβλιογραφία ..... 195



# **1. Ευρεία Περίληψη**

## **1.1 Εισαγωγή**

Οι προβλέψεις αποτελούν ένα αναπόσπαστο κομμάτι της καθημερινότητάς μας. Αυτό αφορά είτε τις επιχειρήσεις, είτε τα κράτη, ή ακόμα και το κάθε άτομο ξεχωριστά. Για αυτό το λόγο το πεδίο των προβλέψεων έχει γίνει αντικείμενο εκτεταμένης έρευνας και ανάπτυξης. Σε αυτό βοηθάει πολύ η ανάπτυξη της τεχνολογίας που παρέχει πλέον στους ερευνητές και στους στατιστικολόγους πολύτιμα στατιστικά εργαλεία που κάνει την υπολογιστική διαδικασία, σαφώς πιο εύκολη. Αυτό γίνεται πιο εμφανές αν αναλογιστεί κανείς πως τα τελευταία είκοσι χρόνια, παρατηρείται έντονη η τάση του σχεδιασμού και της δημιουργίας διάφορων πληροφοριακών συστημάτων για την εφαρμογή μεθόδων πρόβλεψης και την εξαγωγή προβλέψεων.

Φυσικά με την διάδοση του διαδικτύου υπάρχει μία προσπάθεια δημιουργίας διαδικτυακών εφαρμογών οι οποίες σχετίζονται με τον τομέα των προβλέψεων. Εξαιτίας των πλεονεκτημάτων των διαδικτυακών εφαρμογών δημιουργήθηκε η ιδέα της υλοποίησης μίας τέτοιας εφαρμογής χρησιμοποιώντας ένα στατιστικό περιβάλλον λογισμικού, το οποίο ενσωματώνει καινοτόμα χαρακτηριστικά, όπως είναι η αυτόματη σύγκριση των μεθόδων πρόβλεψης και η συνεχής αυτόματη ανανέωση των διαθέσιμων χρονοσειρών. Στη συνέχεια της παρούσας διπλωματικής αναλύεται το θεωρητικό υπόβαθρο που χρησιμοποιήθηκε στην υλοποίηση της εφαρμογής, αλλά και ενός πειράματος σχετικά με την σύγκριση των μεθόδων πρόβλεψης σε συγκεκριμένες χρονοσειρές.

## **1.2 Τεχνικές Προβλέψεων**

Το πρώτο και σημαντικότερο βήμα, για την εξαγωγή προβλέψεων αλλά και γενικότερα για την στατιστική ανάλυση, είναι η συλλογή των απαραίτητων ιστορικών στοιχείων της μεταβλητής που θέλουμε να μελετήσουμε. Τα δεδομένα θα πρέπει σαφώς να είναι επικαιροποιημένα έτσι ώστε η πρόβλεψη να συμβαδίζει με βάση το τωρινό περιβάλλον της μεταβλητής για να είναι όσο το δυνατόν πιο ακριβής η

πρόβλεψη. Το δεύτερο βήμα αφορά στη χρήση της κατάλληλης μεθόδου πρόβλεψης, με σκοπό την εξαγωγή προβλέψεων μεγαλύτερης ακρίβειας. Τελευταίο βήμα, είναι η αξιολόγηση των προβλέψεων που παρήχθησαν χρησιμοποιώντας κατάλληλους στατιστικούς δείκτες.

### **1.2.1 Μέθοδοι Προβλέψεων**

Οι δύο μεγάλες κατηγορίες που εντάσσονται οι τεχνικές προβλέψεων που έχουν αναπτυχθεί μέχρι σήμερα είναι οι εξής:

- *Ποιοτικές προβλέψεις* (qualitative ή judgmental forecasting)
- *Ποσοτικές προβλέψεις* (quantitative forecasting)

Στα πλαίσια της διπλωματικής θα ασχοληθούμε με τις ποσοτικές μεθόδους πρόβλεψης και ειδικότερα με τις μεθόδους των χρονοσειρών. Συγκεκριμένα στην εφαρμογή χρησιμοποιούνται οι εξής οχτώ μέθοδοι πρόβλεψης:

- Απλοϊκή μέθοδος - Naive
- Κινητός μέσος όρος – KMO
- Απλή γραμμική παλινδρόμηση – LRL
- Απλή εκθετική εξομάλυνση – SES
- Εκθετική εξομάλυνση γραμμικής τάσης – Holt
- Εκθετική εξομάλυνση μη γραμμικής τάσης – Damped
- Error-Trend-Seasonal ή Exponential Smoothing method – ETS
- Κλασική μέθοδος Theta

### **1.2.2 Δείκτες Αξιολόγησης Προβλέψεων**

Σε αυτή την κατηγορία απαιτείται μίας πρόβλεψης, που προκύπτει από την εφαρμογή κάποιας μεθόδου πρόβλεψης επί της πραγματικής χρονοσειράς. Η συγκεκριμένη κατηγορία έχει μεγάλη σημασία στον κλάδο των προβλέψεων καθώς αποτελεί ένα τρόπο αξιολόγησης των μεθόδων. Για την διεξαγωγή του πειράματος που αφορά την σύγκριση των μεθόδων πρόβλεψης, χρησιμοποιήθηκαν σαν δείκτες αξιολόγησης τα εξής πέντε σφάλματα:

- Μέσο σφάλμα (Mean Error)

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

- Μέσο απόλυτο σφάλμα (Mean Absolute Error)

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

- Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error)

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

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

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

- Μέσο απόλυτο κανονικοποιημένο σφάλμα (Mean Absolute Scaled Error)

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

## 1.3 Περιγραφή της Εφαρμογής

Με την έκρηξη της τεχνολογίας έχει γίνει πιο έντονη από ποτέ η ανάπτυξη της τεχνολογίας λογισμικού. Η τεχνολογία λογισμικού είναι η βελτίωση της διαδικασίας μεταξύ της ανθρωπότητας και της τεχνολογίας των υλικών (hardware technology). Βέβαια μεγάλο ρόλο της τεχνολογίας λογισμικού παίζουν οι γλώσσες προγραμματισμού και για αυτό το λόγο γίνεται συνεχής προσπάθεια βελτίωσής τους. Η χρησιμοποίηση μίας γλώσσας προγραμματισμού για την υλοποίηση εφαρμογών αποτελεί από μόνο του ένα μεγάλο κομμάτι της τεχνολογίας λογισμικού, το οποίο συνεχώς εξελίσσεται και γιγαντώνεται, αφού καθημερινά δημιουργούνται καινούργιες εφαρμογές.

### 1.3.1 Γλώσσα προγραμματισμού και περιβάλλον υλοποίησης της εφαρμογής

Η εφαρμογή υλοποιήθηκε με την γλώσσα προγραμματισμού *R*. Η *R* είναι μία γλώσσα προγραμματισμού κατάλληλη για στατιστική επεξεργασία τεράστιου όγκου

δεδομένων. Αυτό συμβαίνει λόγω των πάρα πολλών διαθέσιμων πακέτων σχετικά με την στατιστική επεξεργασία των δεδομένων, αλλά και της εφαρμογής ορισμένων μεθόδων πρόβλεψης.

Ένα πακέτο που χρησιμοποιείται είναι το R-shiny, είναι ένα πακέτο της R ανοιχτού κώδικα που παρέχει ένα κομψό και ισχυρό πλαίσιο διαδικτύου (web framework) για τη δημιουργία διαδικτυακών εφαρμογών χρησιμοποιώντας τη γλώσσα προγραμματισμού R. Με βάση αυτό το πακέτο δημιουργήθηκε η εφαρμογή, αφού επιτρέπει στον χρήστη να μετατρέψει τις αναλύσεις του σε διαδραστικές εφαρμογές διαδικτύου. Με αυτό το σκεπτικό μπορεί επομένως η εφαρμογή να δέχεται κάποιες εισόδους, να γίνεται η κατάλληλη επεξεργασία και να βγάζει κάποια αποτελέσματα.

Τα δεδομένα που επεξεργάζεται ο χρήστης αντλούνται μέσω του API ορισμένων μεγάλων εταιρειών, που το παρέχουν ελεύθερα για να μπορεί ο οποιοσδήποτε να αντλεί και να επεξεργάζεται τα δεδομένα που έχουν στις βάσεις τους. Συνήθως αυτό γίνεται με την βοήθεια ενός ήδη υλοποιημένου πακέτου της R.

Τα δεδομένα που είναι υλοποιημένα στην εφαρμογή είναι τα εξής:

- Αγαθά (Commodities)
- Μετοχές (Stocks)
- Τουρισμός (Tourism)
- Ανεργία (Unemployment)

Όπου οι χρονοσειρές είναι είτε ημερήσιες, είτε μηνιαίες, ή και ετήσιες.

### **1.3.2 Είσοδοι-Έξοδοι της Εφαρμογής**

Όλες οι εισοδοί δίνονται από τον χρήστη. Αρχικά η επιλογή του αφορά την κατηγορία των δεδομένων, τα οποία αναφέρθηκαν προηγουμένως, που θέλει να μελετήσει. Στη συνέχεια πρέπει να δώσει το χρονικό ορίζοντα που θέλει, ώστε να αναπαρασταθούν αναλόγως τα δεδομένα. Η αναπαράσταση μπορεί να γίνει είτε σε μορφή πίνακα με την καταγραφή των ιστορικών δεδομένων, είτε σε διαγραμματική παρουσίαση.

Όσον αφορά τη διαγραμματική παρουσίαση ο χρήστης έχει σαν επιλογή τις οχτώ μεθόδους πρόβλεψης που αναφέρθηκαν πιο πάνω και με την ενεργοποίηση της ανάλογης επιλογής εμφανίζεται η προσαρμοσμένη χρονοσειρά προβλέψεων στα ιστορικά δεδομένα, μαζί με την χρονοσειρά των παρατηρήσεων.

Σε μία εφαρμογή προβλέψεων, φυσικά δεν θα μπορούσε να λείπει η λειτουργία, η οποία αφορά την εξαγωγή προβλέψεων για την επόμενη χρονική περίοδο. Οι προβλέψεις αυτές παράγονται με χρήση των οχτώ μεθόδων που αναφέρθηκαν προηγουμένως και αναγράφονται σε μορφή πίνακα. Η πρόβλεψη που γίνεται βρίσκεται σε άμεση σχέση με τον χρονικό ορίζοντα που θέτει ο χρήστης για την μελέτη της χρονοσειράς. Για παράδειγμα, ο χρήστης μπορεί να δηλώσει πως θέλει να μελετήσει την χρονοσειρά από 23/10/1989 μέχρι και 10/11/2009. Τότε οι σημειακές προβλέψεις που θα παραχθούν θα είναι για την επόμενη χρονική περίοδο, δηλαδή



για την 11/11/2009, παρόλο που μπορεί να διατίθενται η πραγματική τιμή και για την συγκεκριμένη ημερομηνία. Αυτό έγινε σε περίπτωση που ο χρήστης θέλει από μόνος του να κρίνει την πρόβλεψη για την συγκεκριμένη ημερομηνία, να μπορεί να το κάνει κρατώντας τις παραχθείσες προβλέψεις, εν συγκρίσει με την πραγματική τιμή, ώστε να μπορέσει να κάνει κα ατομικά αν θέλει μία πρόχειρη αξιολόγηση των μεθόδων πρόβλεψης.

Τώρα ακόμη μία έξοδος της εφαρμογής υλοποιήθηκε έχοντας υπόψιν τα προηγούμενα, δηλαδή την αξιολόγηση και την σύγκριση των μεθόδων με την μελέτη κάποιων σφαλμάτων. Αυτό υλοποιήθηκε για την κατηγορία των αγαθών και με τη βοήθεια της τεχνικής της κυλιόμενης πρόβλεψης (rolling forecasting). Η αξιολόγηση και η κατάταξη των οχτώ μεθόδων έγινε με βάση τα εξής σφάλματα:

- Το μέσο σφάλμα (*Mean Error – ME*)
- Το μέσο απόλυτο σφάλμα (*Mean Absolute Error – MAE*)
- Το συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (*symmetric Mean Absolute Percentage Error – sMAPE*)

Η επιλογή φυσικά του σφάλματος γίνεται από τον χρήστη και κατόπιν η εφαρμογή μετά από κατάλληλη επεξεργασία εμφανίζει σε ένα πίνακα με σειρά κατάταξης από την καλύτερη μέθοδο πρόβλεψης προς την χειρότερη ανάλογα με τα σφάλματα και δίπλα αναγράφεται η τιμή του σφάλματος που έχει επιλεγεί.

Θα πρέπει να τονιστεί πως κάθε φορά που ανανεώνονται τα δεδομένα και γίνονται συνεχώς καινούργια διαθέσιμα, ανανεώνεται ταυτόχρονα η αξιολόγηση των μεθόδων πρόβλεψης. Εκεί εξάλλου έγκειται η καινοτομία της εφαρμογής, δηλαδή στην αυτόματη ανανέωση των δεδομένων με το που γίνουν διαθέσιμα και κατά συνέπεια της αξιολόγησης και της κατάταξης των μεθόδων πρόβλεψης.

### **1.3.3 Περιγραφή των Dataset**

Όπως αναφέρθηκε και προηγουμένως ο χρήστης έχει την δυνατότητα να επιλέξει μέσα από τέσσερις κατηγορίες δεδομένων, τα αγαθά, τις μετοχές, την τουρισμό και την ανεργία.

#### **Αγαθά (Commodities)**

Σε αυτή την κατηγορία υπάρχουν διαθέσιμα δεκαπέντε βασικά αγαθά εκ των οποίων οι παρατηρούμενες τιμές των αγαθών μπορεί να είναι είτε σε ημερήσια βάση, είτε σε μηνιαία ή και τα δύο. Τα αγαθά αυτά είναι τα εξής:

- Αργό πετρέλαιο (Crude Oil)
- Ασήμι (Silver)
- Αλουμίνιο (Aluminum)
- Φυσικό Αέριο (Natural Gas)

- Χαλκός (Copper)
- Λευκόχρυσος ή Πλατίνα (Platinum)
- Ρύζι (Rice)
- Σιτάρι (Wheat)
- Σόγια (Soybean)
- Κακάο (Cocoa)
- Καφές (Coffee)
- Βαμβάκι (Cotton)
- Κάρβουνο (Coal)

### ***Μετοχές (Stocks)***

Οι μετοχές είναι ημερήσιες χρονοσειρές εκ των οποίων ο χρονικός ορίζοντας των παρατηρήσεων είναι αδύνατον να προσδιοριστεί για κάθε μία από τις μετοχές που είναι διαθέσιμες μέσα από την εφαρμογή, διότι ο αριθμός αυτών των μετοχών είναι πραγματικά τεράστιος και φθάνει τις δεκάδες χιλιάδες.

### ***Τουρισμός και Ανεργία (Tourism and Unemployment)***

Οι χρονοσειρές του τουρισμού και της ανεργίας είναι ετήσιες και αφορούν κάθε χώρα. Πιο συγκεκριμένα η χρονοσειρά του τουρισμού εξετάζει τις αφίξεις των τουριστών στη χώρα επιλογής του χρήστη, κατά την διάρκεια ενός έτους. Όσον αφορά τη χρονοσειρά της ανεργίας, εκείνη εξετάζει το ποσοστό της ανεργίας της χώρας επιλογής του χρήστη, στη διάρκεια του έτους.

### ***1.3.4 Βασικά πακέτα της εφαρμογής***

Όπως έχει ήδη αναφερθεί, για την δημιουργία της εφαρμογής χρησιμοποιήθηκαν μερικά πακέτα (packages) χωρίς τα οποία θα ήταν δύσκολη έως αδύνατη η υλοποίηση της εφαρμογής. Τα βασικά πακέτα είναι αυτά που χρησιμοποιήθηκαν για την άντληση των δεδομένων, την διαγραμματική απεικόνισή τους, όπως επίσης και την εφαρμογή των μεθόδων πρόβλεψης πάνω στα επιλεγμένα από το χρήστη δεδομένα. Αυτά τα πακέτα είναι τα εξής:

- Quandl
- Quantmod
- WDI
- ggplot2
- forecast

## 1.4 Σύγκριση Μεθόδων Πρόβλεψης

### 1.4.1 Γενική Μεθοδολογία

Στα οικονομικά, ένα αγαθό θεωρείται ως ένα εμπορεύσιμο προϊόν που παράγεται για να ικανοποιήσει τις επιθυμίες και τις ανάγκες. Σε αυτά μπορεί να περιλαμβάνονται αγαθά και υπηρεσίες. Πρόκειται για μία φυσική ουσία, όπως τα τρόφιμα, τα δημητριακά και τα μέταλλα, τα οποία είναι ανταλλάξιμα με άλλο προϊόν του ίδιου τύπου, και τα οποία οι επενδυτές αγοράζουν ή πωλούν, συνήθως μέσω συμβολαίων μελλοντικής εκπλήρωσης (future contracts). Η τιμή του εμπορεύματος εξαρτάται από την προσφορά και τη ζήτηση.

Προφανώς η σημασία της τιμής των αγαθών είναι τεράστια για οποιοδήποτε άνθρωπο, επιχείρηση ή οργανισμό και ακόμα σε επίπεδο κρατών για την χάραξη της οικονομικής πολιτικής. Αυτό συμβαίνει διότι ορισμένα αγαθά, όπως είναι το αργό πετρέλαιο αποτελεί την πρώτη ύλη για πάρα πολλά πράγματα. Μερικά από αυτά είναι οι μετακινήσεις και η παραγωγή ηλεκτρισμού, τα οποία με τη σειρά τους οδηγούν στη δημιουργία και καλλιέργεια και άλλων αγαθών και υπηρεσιών. Επομένως, οι παράγοντες που διαμορφώνουν τη τιμή ενός αγαθού και η πρόβλεψη της τιμής αυτής είναι απαραίτητη για την οποιαδήποτε χάραξη οικονομικής πολιτικής, είτε πρόκειται για τον οικογενειακό προϋπολογισμό ενός σπιτιού, είτε για τα έσοδα-έξοδα μίας επιχείρησης, ή και ακόμα για την διαμόρφωση του κρατικού προϋπολογισμού.

Στη παρούσα διπλωματική εργασία έγινε μία προσπάθεια συγκέντρωσης ορισμένων χρονοσειρών κάποιων βασικών αγαθών, τα οποία αναφέρθηκαν ήδη, και στη συνέχεια η εφαρμογή των οχτώ μεθόδων πρόβλεψης πάνω σε αυτές τις χρονοσειρές. Η εξαγωγή των προβλέψεων έγινε χρησιμοποιώντας την τεχνική της κυλιόμενης πρόβλεψης (rolling forecasting). Η παραγωγή των προβλέψεων και η ανάλογη μελέτη έγινε για εξαγωγή προβλέψεων της επόμενης περιόδου (one-step ahead) και των επόμενων τριών (three-steps ahead), είτε πρόκειται για ημερήσιες, είτε για μηνιαίες χρονοσειρές.

Η αξιολόγηση των μεθόδων έγινε με βάση πέντε ήδη σφαλμάτων, τα οποία είναι τα εξής:

- Μέσο σφάλμα – *Mean Error*
- Μέσο απόλυτο σφάλμα - *Mean Absolute Error*
- Μέσο απόλυτο ποσοστιαίο σφάλμα - *Mean Absolute Percentage Error*
- Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα - *symmetric Mean Absolute Percentage Error*
- Μέσο απόλυτο κανονικοποιημένο σφάλμα - *Mean Absolute scaled Error*

Οι μέθοδοι που χρησιμοποιήθηκαν για την εξαγωγή των προβλέψεων είναι οι οχτώ μέθοδοι που έχουν ήδη αναφερθεί. Κατόπιν έγινε η σύγκριση των μεθόδων με βάση τα παραπάνω σφάλματα.

Γενικά η διαδικασία που ακολουθήθηκε είναι η εξής:

**Βήμα 1<sup>ο</sup>:** Συγκέντρωση πραγματικών τιμών.

**Βήμα 2<sup>ο</sup>:** Παραγωγή προβλέψεων.

**Βήμα 3<sup>ο</sup>:** Αποθήκευση πραγματικών τιμών και προβλέψεων.

**Βήμα 4<sup>ο</sup>:** Υπολογισμός σφαλμάτων.

**Βήμα 5<sup>ο</sup>:** Αξιολόγηση και σύγκριση των μεθόδων με βάση τα σφάλματα.

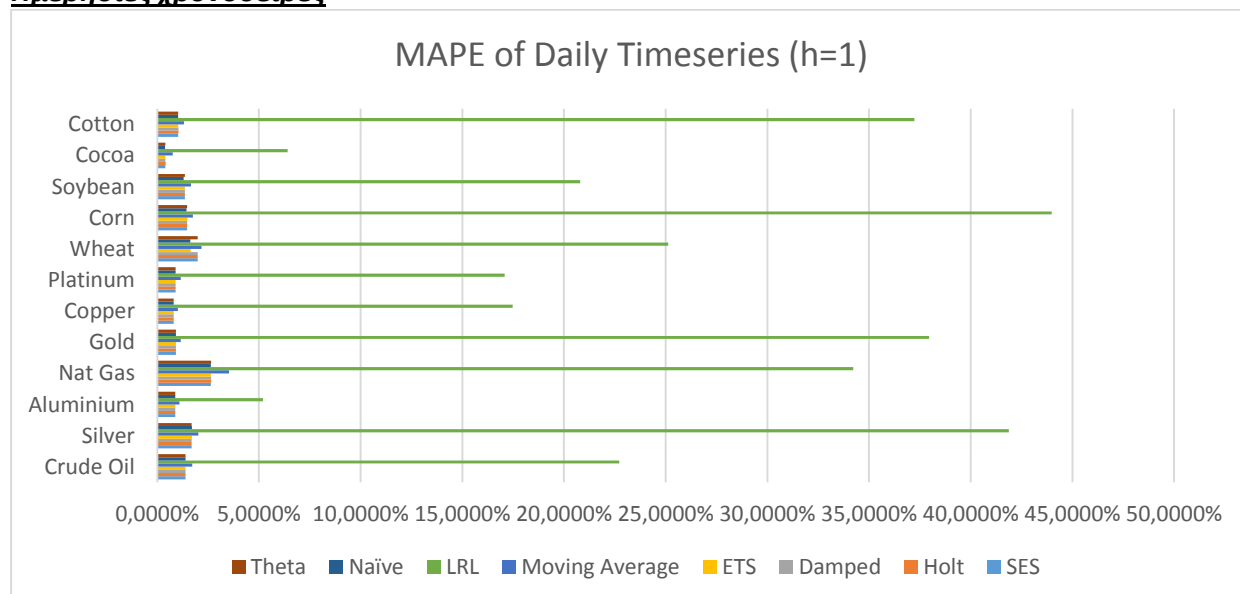
### 1.4.2 Σύγκριση Μεθόδων σύμφωνα με τον δείκτη MAPE

Για τις χρονοσειρές των αγαθών υπολογίστηκαν τα παραπάνω σφάλματα που αναφέρθηκαν στην παραπάνω παράγραφο, και για τις οχτώ μεθόδους πρόβλεψης. Στη συνέχεια συγκρίθηκαν οι μέθοδοι πρόβλεψης με βάση τον δείκτη MAPE, διότι είναι ο πλέον κατάλληλος δείκτης για την σύγκριση των μεθόδων σε διαφορετικές χρονοσειρές. Φυσικά ο διαχωρισμός που έγινε εδώ, έγινε με βάση των ορίζοντα πρόβλεψης των παραγόμενων προβλέψεων, δηλαδή της επόμενης περιόδου ( $h=1$ ) και της τρίτης περιόδου ( $h=3$ ). Επίσης ο δεύτερος διαχωρισμός έγινε με βάση το κριτήριο του χρονικού ορίζοντα των χρονοσειρών, δηλαδή σε ημερήσιες και σε μηνιαίες. Επομένως προέκυψαν τα εξής διαγράμματα με βάση τον παραπάνω διαχωρισμό.

#### 1.4.2.1 Υπολογισμός MAPE για παραγωγή προβλέψεων της επόμενης περιόδου ( $h=1$ )

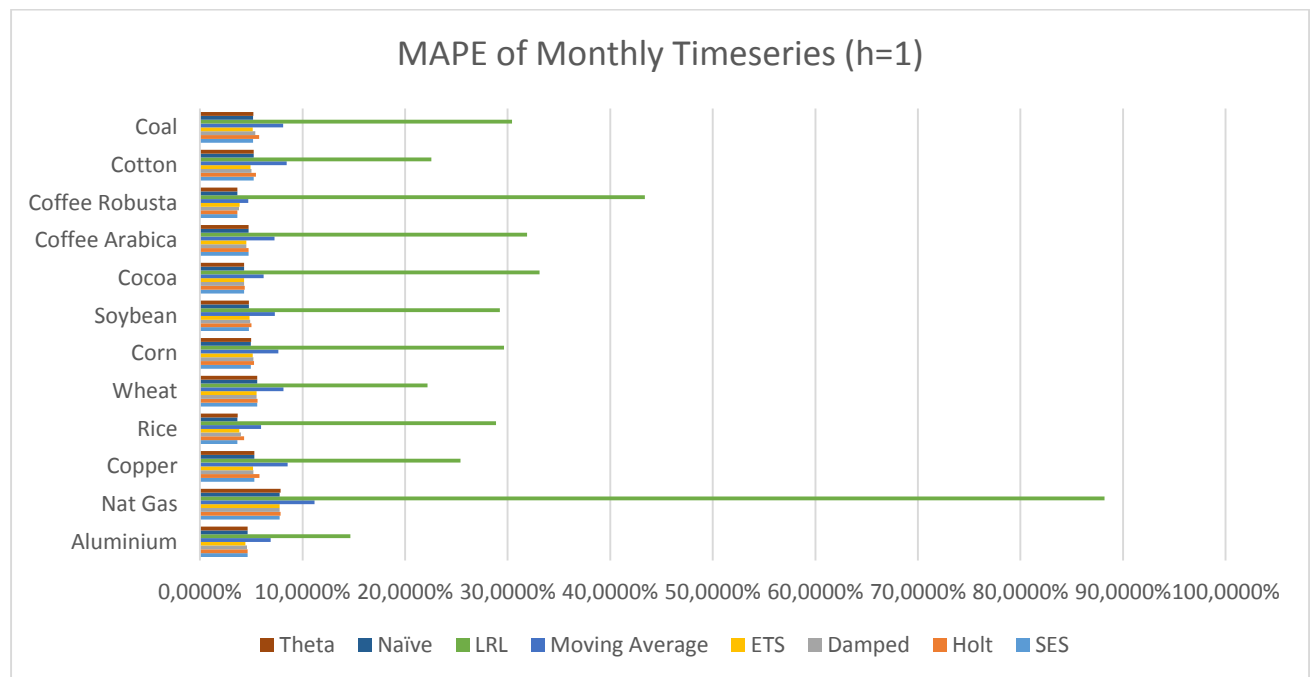
Εδώ μελετάμε την περίπτωση των σφαλμάτων MAPE για την περίπτωση παραγωγής προβλέψεων της επόμενης περιόδου. Φυσικά η σύγκριση είναι πιο σωστό να γίνει στα δεδομένα του ίδιου χρονικού ορίζοντα. Επομένως θα αναλυθεί περαιτέρω σε ημερήσιες και μηνιαίες χρονοσειρές.

##### Ημερήσιες χρονοσειρές



Από το παραπάνω διάγραμμα μπορούμε να βγάλουμε το συμπέρασμα, πως τα ημερήσια δεδομένα που έχουν τα μικρότερα σφάλματα από την εφαρμογή των μεθόδων πρόβλεψης για παραγωγή προβλέψεων της επόμενης περιόδου είναι το βαμβάκι (Cotton) και στη συνέχεια ο χαλκός (Copper). Τα χειρότερα αποτελέσματα συγκριτικά με τα υπόλοιπα αγαθά, τα έχει το φυσικό αέριο (Natural Gas) και ακολουθεί το σιτάρι με το ασήμι.

### Μηνιαίες Χρονοσειρές

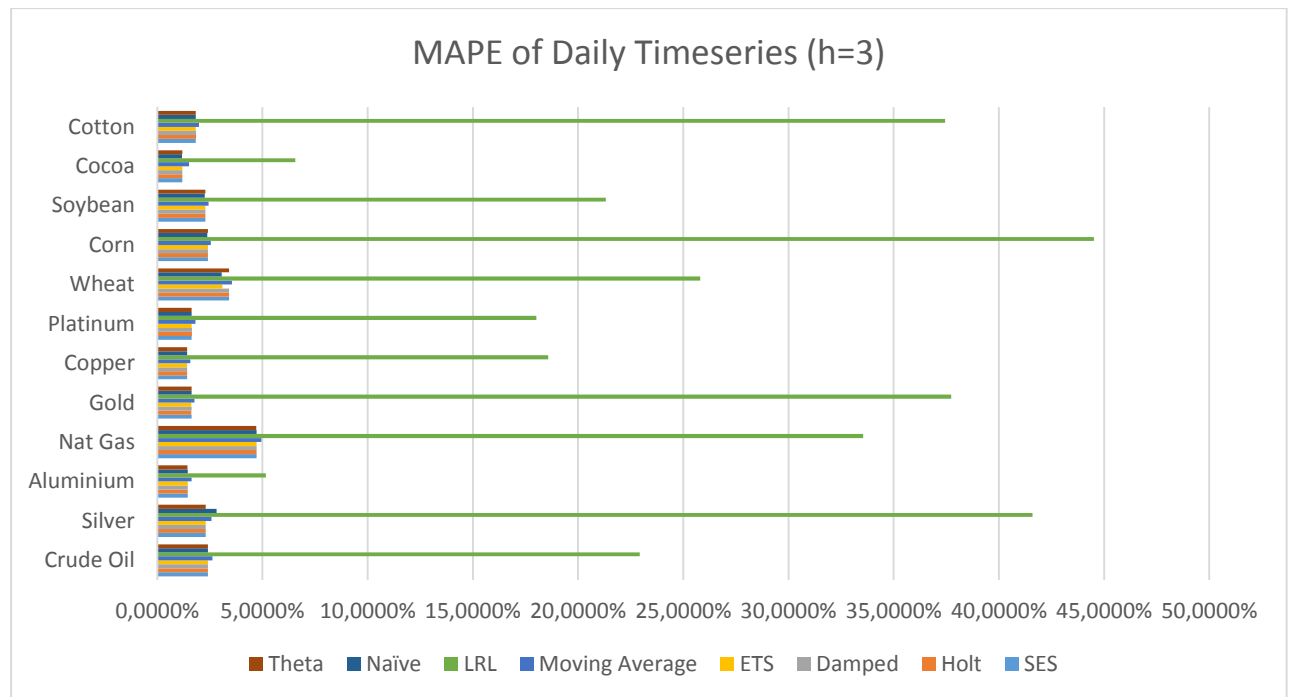


Στην περίπτωση των μηνιαίων χρονοσειρών τα δεδομένα που έχουν τα μικρότερα σφάλματα είναι οι τιμές της ποικιλίας καφέ Robusta και με τα μεγαλύτερα σφάλματα οι τιμές του φυσικού αερίου.

### **1.4.2.2 Υπολογισμός MAPE για παραγωγή προβλέψεων της τρίτης περιόδου (h=3)**

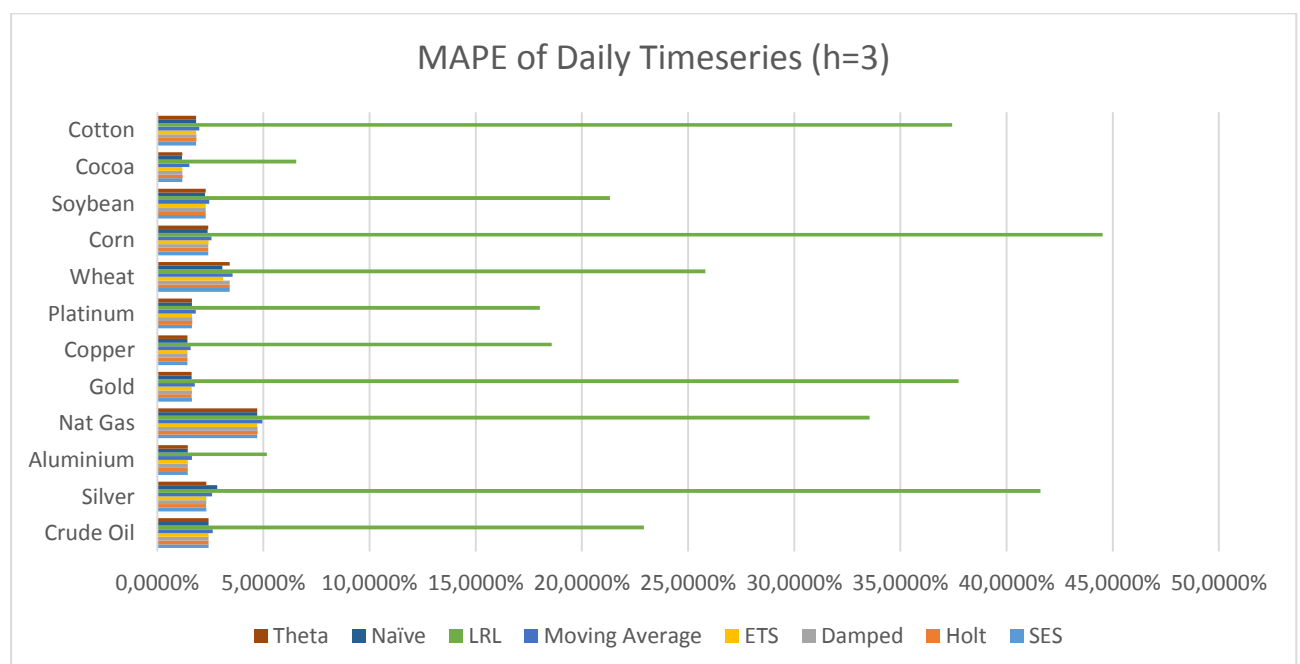
Εδώ μελετάμε την περίπτωση των σφαλμάτων MAPE για την περίπτωση παραγωγής προβλέψεων της τρίτης περιόδου. Φυσικά η σύγκριση είναι πιο σωστό να γίνει στα δεδομένα του ίδιου χρονικού ορίζοντα. Επομένως θα αναλυθεί περαιτέρω σε ημερήσιες και μηνιαίες χρονοσειρές.

### Ημερήσιες χρονοσειρές



Από το παραπάνω διάγραμμα μπορούμε να βγάλουμε το συμπέρασμα, πως τα ημερήσια δεδομένα που έχουν τα μικρότερα σφάλματα από την εφαρμογή των μεθόδων πρόβλεψης για παραγωγή προβλέψεων της επόμενης περιόδου είναι το βαμβάκι (Cotton) και στη συνέχεια ο χαλκός (Copper). Τα χειρότερα αποτελέσματα συγκριτικά με τα υπόλοιπα αγαθά, τα έχει το φυσικό αέριο (Natural Gas) και ακολουθεί το σιτάρι (Wheat) και το ασήμι (Silver).

### Μηνιαίες Χρονοσειρές



Στην περίπτωση των μηνιαίων χρονοσειρών τα δεδομένα που έχουν τα μικρότερα σφάλματα είναι οι τιμές της ποικιλίας καφέ Robusta και ακολουθούν οι τιμές του ρυζιού (Rice) και του κακάο (Cocoa). Οι τιμές με τα μεγαλύτερα σφάλματα είναι οι τιμές του φυσικού αερίου (Natural Gas) και του χαλκού (Copper).

### 1.4.3 Σύγκριση και Επιλογή Μεθόδων

Κατόπιν ανάλυσης και επεξεργασίας των μεθόδων πρόβλεψης προκύπτουν οι παρακάτω πίνακες που αφορούν την επιλογή της κατάλληλης μεθόδου πρόβλεψης για κάθε αγαθό στις περιπτώσεις προβλέψεων για την επόμενη περίοδο και για την τρίτη περίοδο. Κριτήριο επιλογής της μεθόδου είναι η μικρότερη τιμή από όλες τις μεθόδους του δείκτη MAPE. Πιο συγκεκριμένα για τις προβλέψεις της επόμενης περιόδου έχουμε:

<b>Αγαθά</b>	<b>Είδος Χρονοσειράς</b>	<b>Επιλογή Μεθόδου</b>	<b>Τιμή σφάλματος MAPE (h=1)</b>
1) Πετρέλαιο	Ημερήσια	ETS	1,3796%
2) Ασήμι	Ημερήσια	Theta	1,6880%
3) Αλουμίνιο	Ημερήσια	Theta	0,8771%
	Μηνιαία	ETS	4,4282%
4) Φυσικό Αέριο	Ημερήσια	Naïve	2.6329%
	Μηνιαία	Naïve	7,7779%
5) Χρυσός	Ημερήσια	ETS	0,9195%
6) Χαλκός	Ημερήσια	Holt	0,8103%
	Μηνιαία	ETS	5,1642%
7) Λευκόχρυσος	Ημερήσια	Damped	0,8924%
8) Ρύζι	Μηνιαία	Naïve	3,6601%
9) Σιτάρι	Ημερήσια	Naïve	1,6231%
	Μηνιαία	Damped	5,5102%
10) Καλαμπόκι	Ημερήσια	Naïve	1,4289%
	Μηνιαία	Naïve	4,9682%
11) Σόγια	Ημερήσια	Naïve	1,2991%
	Μηνιαία	Naïve	4,7799%
12) Κακάο	Ημερήσια	Naïve	0,3857%
	Μηνιαία	Naïve	4,2891%
13α) Καφές Arabica	Μηνιαία	Damped	4,5141%
13β) Καφές Robusta	Μηνιαία	Holt	3,6356%
14) Βαμβάκι	Ημερήσια	Naïve	1,0323%
	Μηνιαία	ETS	4,9134%
15) Κάρβουνο	Μηνιαία	ETS	5,1486%

Και για την παραγωγή προβλέψεων της τρίτης περιόδου έχουμε εξάγει τον ακόλουθο πίνακα:

<b>Αγαθά</b>	<b>Είδος Χρονοσειράς</b>	<b>Επιλογή μεθόδου</b>	<b>Τιμή σφάλματος MAPE(h=3)</b>
1) Πετρέλαιο	Ημερήσια	Damped	2,4098%
2) Ασήμι	Ημερήσια	Holt	2,3009%
3) Αλουμίνιο	Ημερήσια	Theta	1,4382%
	Μηνιαία	Naïve	9,9325%
4) Φυσικό Αέριο	Ημερήσια	Theta	4,7083%
	Μηνιαία	Damped	15,9203%
5) Χρυσός	Ημερήσια	ETS	1,6152%
6) Χαλκός	Ημερήσια	Naïve	1,4159%
	Μηνιαία	Naïve	12,8624%
7) Λευκόχρυσος	Ημερήσια	Naïve	1,6359%
8) Ρύζι	Μηνιαία	Naïve	8,3315%
9) Σιτάρι	Ημερήσια	Naïve	3,0642%
	Μηνιαία	Naïve	10,9399%
10) Καλαμπόκι	Ημερήσια	Naïve	2,3714%
	Μηνιαία	Naïve	10,9677%
11) Σόγια	Ημερήσια	Naïve	2,2508%
	Μηνιαία	Naïve	10,4691%
12) Κακάο	Ημερήσια	Naïve	1,1678%
	Μηνιαία	ETS	8,8725%
13α) Καφές Arabica	Μηνιαία	Damped	10,2168%
13β) Καφές Robusta	Μηνιαία	Holt	6,7216%
14) Βαμβάκι	Ημερήσια	Naïve	1,8232%
	Μηνιαία	ETS	12,2689%
15) Κάρβουνο	Μηνιαία	ETS	11,7355%

## **1.5 Συμπεράσματα-Προοπτικές**

### **1.5.1 Συμπεράσματα**

Σαν επίλογος της παρούσης διπλωματικής εργασίας παρουσιάζονται τα σημαντικότερα συμπεράσματα της εφαρμογής, αλλά και της διεξαγωγής του πειράματος. Όσον αφορά την εφαρμογή, το συμπέρασμα που βγαίνει είναι η σημαντικότητα της αυτόματης ανανέωσης των δεδομένων, όπως επίσης και των λειτουργιών της. Μία από τις λειτουργίες, που εκμεταλλεύονται κατά κάποιο τρόπο την αυτόματη ανανέωση και την προσθήκη καινούργιων πραγματικών τιμών όταν αυτές είναι διαθέσιμες, είναι η λειτουργία της αξιολόγησης, σύγκρισης και αυτόματης κατάταξης των μεθόδων πρόβλεψης ανά σφάλμα επιλογής του χρήστη.

Με βάση αυτό το σκεπτικό δημιουργήθηκε και το δεύτερο στάδιο της διπλωματικής που αφορά στη περαιτέρω διερεύνηση της αξιολόγησης των μεθόδων, της οποίας τα αποτελέσματα αναφέρθηκαν επιγραμματικά παραπάνω με τα οποία καταλήγουμε στα εξής συμπεράσματα:



- Γενικά η εφαρμογή των μεθόδων πρόβλεψης πάνω στις χρονοσειρές των αγαθών είτε των ημερήσιων δεδομένων, είτε των μηνιαίων είχαν πολύ καλή απόδοση, εκτός των μεθόδων της LRL και του κινητού μέσου όρου που ήταν σταθερά οι μέθοδοι με τα μεγαλύτερα σφάλματα.
- Όπως είναι λογικό τα σφάλματα των προβλέψεων για τις επόμενες τρεις περιόδους ήταν αυξημένα σχετικά με τις προβλέψεις της επόμενης περιόδου. Αυτό είναι συνηθισμένο όταν αυξάνουμε τον ορίζοντα πρόβλεψης μεγαλώνουν και τα σφάλματα.
- Ακόμη ένα συμπέρασμα αφορά την προκατάληψη των μεθόδων πρόβλεψης. Από τα μέσα σφάλματα που υπολογίστηκαν από τις προβλέψεις που παρήχθησαν με εφαρμογή των μεθόδων, προέκυψε πως υπάρχει έντονο το στοιχείο της αισιοδοξίας στις προβλέψεις, κατά ένα μεγάλο ποσοστό και σχεδόν σε όλες τις χρονοσειρές τιμών των αγαθών, ανεξαρτήτου χρονικού ορίζοντα. Η LRL διαφοροποιείται σχετικά με την παραπάνω διαπίστωση.
- Σύμφωνα με τα διαγράμματα της παραγράφου 1.4.2, βγαίνει το συμπέρασμα πως τα αγαθά που ανταποκρίνονται καλύτερα στις προβλέψεις είναι για τις ημερήσιες χρονοσειρές του Κακάο (Cocoa), του βαμβακιού (Cotton) και στη συνέχεια ο χαλκός (Copper). Στον αντίποδα οι χρονοσειρές με τα μεγαλύτερα σφάλματα στις προβλέψεις τους είναι οι χρονοσειρές του Φυσικού Αερίου (Natural Gas) και ακολουθούν το σιτάρι (Wheat) και το ασήμι (Silver). Στην περίπτωση των μηνιαίων δεδομένων τα αγαθά που είχαν τα καλύτερα αποτελέσματα είναι οι τιμές της ποικιλίας καφέ Robusta και ακολουθούν οι τιμές του ρυζιού (Rice) και του κακάο (Cocoa). Αντιθέτως οι τιμές με τα μεγαλύτερα ποσοστιαία σφάλματα (MAPE), είναι οι χρονοσειρές των τιμών του φυσικού αερίου (Natural Gas) και του χαλκού (Copper).
- Σύμφωνα με τους συγκεντρωτικού πίνακες της παραγράφου 1.4.3, φάνηκε πως η Naïve εμφανίζεται πιο συχνά από τις άλλες μεθόδους, ως η πλέον κατάλληλη μέθοδος για την παραγωγή προβλέψεων των αγαθών, είτε αυτό αφορά τις ημερήσιες ή τις μηνιαίες τιμές, είτε ακόμα και τον ορίζοντα πρόβλεψης. Μετά την Naïve, η μέθοδος που είχε τα πιο μικρά σφάλματα MAPE είναι η ETS για τα ημερήσια και για τα μηνιαία δεδομένα ανεξαρτήτου του ορίζοντα πρόβλεψης. Στην περίπτωση παραγωγής προβλέψεων της τρίτης περιόδου, η μέθοδος Naïve έχει και πάλι τα μικρότερα σφάλματα MAPE 13 φορές, συγκριτικά με τις υπόλοιπες μεθόδους, το οποίο μεταφράζεται στο 54,17% των συνολικών χρονοσειρών.
- Η Naïve για την παραγωγή προβλέψεων της επόμενης περιόδου ( $h=1$ ) ανεξαρτήτως του χρονικού ορίζοντα της χρονοσειράς, σε 24 ημερήσιες και μηνιαίες χρονοσειρές είχε το μικρότερο δείκτη σφάλματος MAPE συγκριτικά

με όλες τις μεθόδους, 11 φορές, που πρακτικά σημαίνει στο 45,83% των χρονοσειρών.

- Επίσης υπήρξε η προσπάθεια κατάταξης και βαθμολόγησης των μεθόδων στο σύνολο των αγαθών, με διαχωρισμό του χρονικού ορίζοντα των αγαθών, αλλά και του ορίζοντα πρόβλεψης των αγαθών. Η κατάταξη και η βαθμολόγηση φαίνονται στην παράγραφο 7.5 και έγινε με βάση κάθε είδους σφάλματος. Για την περίπτωση παραγωγής προβλέψεων της επόμενης περιόδου, είναι εμφανές πως οι μέθοδοι που ανταποκρίνονται καλύτερα με βάση την βαθμολόγηση και των πέντε σφαλμάτων είναι, πρώτη από όλες η μέθοδος Naïve και ακολουθεί η μέθοδος ETS. Εδώ η μέθοδος Theta συνολικά δεν παρουσιάζει πολύ καλά αποτελέσματα. Αυτό βέβαια αλλάζει στην περίπτωση παραγωγής προβλέψεων για την τρίτη περίοδο όπου η αποτελεσματικότητα της μεθόδου Theta έχει ανέβει αισθητά και σε σχέση με τα σφάλματα καταλαμβάνει την δεύτερη και την τρίτη θέση, ανάλογα με το είδος του σφάλματος. Αντίστοιχα μία αισθητή μείωση της αποτελεσματικότητας εμφανίζεται στην περίπτωση της μεθόδου ETS, ενώ η μέθοδος SES έχει σταθερά καλά αποτελέσματα. Και στην περίπτωση αυτή η μέθοδος που βαθμολογείται καλύτερα είναι η Naïve, τόσο για τα ημερήσια, όσο και για τα μηνιαία δεδομένα.
- Ακόμη παρατηρώντας κανείς την κατάταξη των μεθόδων για τις μηνιαίες χρονοσειρές των τιμών του ρυζιού, του καλαμποκιού και τις σόγιας υπάρχουν πολλές ομοιότητες σχετικά με την κατάταξη των μεθόδων πρόβλεψης.

### **1.5.2 Προοπτικές**

Φυσικά, η εφαρμογή έχει πολλές δυνατότητες και προοπτικές. Μία από αυτές, σαφώς και είναι η προσθήκη περισσότερων μεθόδων παραγωγής προβλέψεων, ακόμα και συνδυασμός αυτών. Επίσης, σίγουρα το κομμάτι που χρήζει μεγάλης βελτίωσης είναι το αισθητικό και της γραφικής απεικόνισης των δεδομένων με περισσότερες επιλογές. Γενικότερα σε αυτό το κεφάλαιο παρουσιάζονται και άλλες ιδέες που θα μπορούσαν να προστεθούν στις ήδη υπάρχουσες λειτουργίες της εφαρμογής πιο αναλυτικά.

Η πρόβλεψη των τιμών των εμπορευμάτων είναι ένα θέμα που χρήζει περαιτέρω ανάλυσης. Η έρευνα γύρω από τις τιμές των αγαθών αφορά όλο και περισσότερο κόσμο, διότι πλέον ενδιαφέρουν επενδυτικά περισσότερο κόσμο, αναφορικά με τις τιμές του χρυσού, του ασημιού, κλπ. Επίσης αφορά τα οικονομικά σε παγκόσμιο επίπεδο, όσον αφορά τις τιμές των αγαθών που συνδέονται με την ενέργεια και τις μετακινήσεις, όπως είναι το αργό πετρέλαιο, το φυσικό αέριο, το κάρβουνο κλπ. Τα σημεία στα οποία θα μπορούσαν να επικεντρωθούν επόμενες μελέτες ή έρευνες αφορούν:

- Μελέτη απόδοσης κι άλλων μοντέλων ή μεθόδων πρόβλεψης για τα δεδομένα των τιμών των αγαθών.

- Παραγωγή Κριτικών Προβλέψεων και συνδυασμός τους με τα αποτελέσματα κλασσικών μεθόδων πρόβλεψης. Οι κριτικές προβλέψεις μπορούν να βασίζονται σε σχέση με τους παράγοντες διαμόρφωσης των τιμών των αγαθών που αναφέρθηκαν στο 5<sup>ο</sup> κεφάλαιο, όπως επίσης και το οικονομικό περιβάλλον γενικότερα.
- Πρόβλεψη των τιμών των αγαθών για εβδομαδιαία και για ετήσια δεδομένα.
- Χρησιμοποίηση διαφορετικών συντελεστών βαρύτητας στους συνδυασμούς των μεθόδων πρόβλεψης. Οι μέθοδοι με τα περισσότερο ακριβή αποτελέσματα θα μπορούσαν να έχουν μεγαλύτερους συντελεστές βαρύτητας σε σχέση με τις υπόλοιπες.
- Αναζήτηση και εφαρμογή προβλέψεων και σε άλλες χρονοσειρές αγαθών, όπως είναι το παλλάδιο, ο σίδηρος κλπ.
- Χρησιμοποίηση περισσότερων μεθόδων, όπως είναι τα *arima models*, τα νευρωνικά κλπ.
- Αξιολόγηση των μεθόδων με περισσότερους στατιστικούς δείκτες, όπως είναι το μέσο γεωμετρικό σφάλμα.



## **2. Εισαγωγή**

### **2.1 Χρήση Διαδικτυακών Εφαρμογών**

Τα τελευταία χρόνια με τη διάδοση του ίντερνετ σε κάθε μεριά του πλανήτη και τη συνεχή χρήση του, μπορεί κάποιος να πει με σιγουριά πως το ίντερνετ έχει γίνει μέρος της καθημερινότητάς μας. Μεγάλο ποσοστό ανθρώπων χρησιμοποιεί το διαδίκτυο και γενικά τον υπολογιστή σαν ένα εργαλείο για την εργασία του με τις τόσες εφαρμογές που υπάρχουν. Ένα είδος τέτοιων εφαρμογών είναι οι διαδικτυακές εφαρμογές.

Διαδικτυακή εφαρμογή (web application ή web app) ονομάζεται κάθε εφαρμογή η οποία είναι διαθέσιμη στους χρήστες της μέσω του Διαδικτύου (Internet) και ο χρήστης χρησιμοποιεί μόνο τον περιηγητή του για να την χρησιμοποιήσει. Οι εφαρμογές αυτές συνήθως εκτελούνται σε ισχυρές υπολογιστικές μηχανές οι οποίες έχουν τον ρόλο του σταθμού εξυπηρέτησης και παρέχουν τις υπηρεσίες τους σε περισσότερους του ενός χρήστη.

#### **2.1.1 Πλεονεκτήματα Χρήσης Διαδικτυακών Εφαρμογών**

Σήμερα οι διαδικτυακές εφαρμογές γίνονται όλο και περισσότερο δημοφιλείς λόγω των πολλών πλεονεκτημάτων που διαθέτουν. Τα σημαντικότερα από αυτά είναι τα εξής:

- **Άμεση πρόσβαση από οποιαδήποτε συσκευή:** Οι χρήστες των διαδικτυακών εφαρμογών έχουν άμεση προσβασιμότητα στις εφαρμογές που θέλουν να χρησιμοποιήσουν από οποιονδήποτε υπολογιστή ή άλλη συσκευή έχει ίντερνετ χωρίς την εγκατάσταση κάποιου επιπρόσθετου λογισμικού. Η μόνη απαραίτητη εφαρμογή είναι ο περιηγητής διαδικτύου ο οποίος είναι εγκατεστημένος σε όλα τα λειτουργικά συστήματα ακόμα και στις φορητές συσκευές αλλά και στα κινητά τηλέφωνα.
- **Δυνατότητα χρήσης ανεξαρτήτου τοποθεσίας:** Η δυνατότητα αυτή δίνει ευελιξία στους χρήστες ώστε να χρησιμοποιούν τις εφαρμογές οπουδήποτε αυτοί επιθυμούν επιτρέποντας τους ακόμα και να εργάζονται από απομακρυσμένες περιοχές ή και από το σπίτι τους.

- **Συμβατές με όλα τα λειτουργικά συστήματα:** Ένα ακόμα πλεονέκτημα των διαδικτυακών εφαρμογών είναι ότι είναι συμβατές με όλα τα λειτουργικά συστήματα. Καθώς η εφαρμογή εκτελείται μέσω του περιηγητή του διαδικτύου και όχι στον υπολογιστή του χρήστη την κάνει ικανή να εκτελείται σε όλα τα λειτουργικά συστήματα.
- **Δεν καταλαμβάνουν χώρο:** Οι εφαρμογές αυτές δεν καταλαμβάνουν καθόλου ή σχεδόν καθόλου χώρο στον δίσκο του χρήστη αφού το σύνολο της εφαρμογής είναι αποθηκευμένο στον εξυπηρετητή και μόνο κατά την χρήση της εφαρμογής μπορεί να υπάρχει μεταφορά δεδομένων προς την υπολογιστική μονάδα του χρήστη και μόνο στην περίπτωση που ο χρήστης το επιθυμεί.
- **Νέο, βελτιωμένο περιβάλλον:** Ένα ακόμα πλεονέκτημα των διαδικτυακών εφαρμογών είναι ότι πλέον με την εμφάνιση της HTML5 είναι δυνατό ο δημιουργός της εφαρμογής να την εμπλουτίσει έτσι ώστε να είναι πιο φιλική, εύχρηστη και ευχάριστη προς τον χρήστη με εύκολο τρόπο. Παλαιότερα οι εφαρμογές αυτές υστερούσαν στην εμφάνιση ωστόσο πλέον είναι ιδιαίτερα εύκολα να εμπλουτιστούν.

### **2.1.2 Μειονεκτήματα Διαδικτυακών Εφαρμογών**

Όπως είναι λογικό κάθε εφαρμογή έχει εκτός από πλεονεκτήματα και κάποια μειονεκτήματα τα οποία είναι τα εξής:

- **Αδυναμία χρήσης χωρίς σύνδεση στο διαδίκτυο:** Αποτελεί και το κύριο μειονέκτημα των διαδικτυακών εφαρμογών, καθώς αν για κάποιο λόγο η σύνδεση του χρήστη διακοπεί τότε αυτός δεν μπορεί να χρησιμοποιήσει καθόλου την εφαρμογή.
- **Πιθανή μη συμβατότητα κάποιων στοιχείων της εφαρμογής με μια μελλοντική έκδοση του περιηγητή μας:** Τέλος μειονέκτημα είναι και η πιθανή μη συμβατότητα κάποιων στοιχείων του προγράμματος με μια μελλοντική έκδοση του περιηγητή μας. Αυτό συμβαίνει διότι συχνά παρατηρείται το φαινόμενο να εγκαταλείπονται κάποια υποστηριζόμενα στοιχεία από τους περιηγητές αν η δημιουργός εταιρία κρίνει ότι αυτά δεν έχουν μέλλον στις εφαρμογές διαδικτύου. Το αποτέλεσμα αυτών των αποφάσεων είναι η δυσλειτουργία κάποιων παλαιότερων εφαρμογών καθιστώντας αναγκαία την αναβάθμισή τους.

## **2.2 Σύνδεση Προβλέψεων-Διαδικτυακών Εφαρμογών**

Με τη πάροδο του χρόνου και με τις εξελίξεις στον τομέα της τεχνολογίας, όλο και περισσότερες επιχειρήσεις θεωρούν απαραίτητο το πεδίο των προβλέψεων για την εξαγωγή των πολιτικών που θα ακολουθήσουν σχετικά με το μάρκετινγκ, τις

πωλήσεις, το ύψος της παραγωγής κλπ. Για αυτό το λόγο έχει δημιουργηθεί η τάση της μελέτης της επιστήμης των προβλέψεων είτε από τους ερευνητές, ή από τις επιχειρήσεις, είτε ακόμα και από ανθρώπους χωρίς το αντίστοιχο επιστημονικό υπόβαθρο ή ενδιαφέρον.

Για αυτό το λόγο, υπάρχει έντονο ενδιαφέρον για δημιουργία και χρήση διάφορων λογισμικών προβλέψεων. Στη παρούσα διπλωματική έγινε μία προσπάθεια δημιουργίας μίας διαδικτυακής εφαρμογής που σαν κύριο πλεονέκτημα είναι ότι δεν απαιτεί καθόλου κάποιο χώρο αποθήκευσης και που διαθέτει στον χρήστη ένα μεγάλο πλήθος χρονοσειρών με σκοπό να τις επεξεργαστεί.

Η καινοτομία της έγκειται στο γεγονός πως τα δεδομένα ανανεώνονται συνεχώς, κάτι το οποίο δίνει την δυνατότητα στον χρήστη να μελετάει τις καινούργιες διαθέσιμες τιμές με αυτόματο τρόπο. Επιπρόσθετα με βάση αυτό το χαρακτηριστικό υλοποιήθηκε η ιδέα της αυτόματης αξιολόγησης ορισμένων μεθόδων πρόβλεψης και αυτόματης κατάταξης τους. Βεβαίως το συγκεκριμένο χαρακτηριστικό της εφαρμογής, λόγω της υπολογιστικής πολυπλοκότητας ορισμένων μεθόδων και κυρίως του υπολογιστικού χρόνου που χρειάζεται, υλοποιήθηκε για συγκεκριμένες χρονοσειρές.

Στα επόμενα κεφάλαια αναφέρονται οι μέθοδοι που χρησιμοποιήθηκαν, με ποιους δείκτες έγινε η αξιολόγηση των μεθόδων, η υλοποίηση της εφαρμογής, η περιγραφή των δεδομένων που χρησιμοποιήθηκαν, αλλά και η υλοποίηση ενός πειράματος σχετικά με την σύγκριση των μεθόδων πρόβλεψης.





## 3. Τεχνικές Προβλέψεων

### 3.1 Εισαγωγή

Κάθε απόφαση που λαμβάνουμε σχετικά με τις μελλοντικές μας ενέργειες στηρίζεται σε μια πρόβλεψη. Γνωρίζουμε από την καθημερινή μας εμπειρία ότι πολλές από τις προβλέψεις που είμαστε υποχρεωμένοι να κάνουμε, θα αποδειχθούν λανθασμένες. Ωστόσο, αυτό δεν αναιρεί την υπόθεση μας ώστε να βασίζουμε στις προβλέψεις, τις αποφάσεις που πρόκειται να πάρουμε. Είμαστε, κατά κάποιο τρόπο, υποχρεωμένοι να πραγματοποιούμε προβλέψεις οποιουδήποτε είδους, τις οποίες τις χρησιμοποιούμε σαν μέσο για τον καθορισμό των μελλοντικών μας δράσεων.

Σίγουρα, το ενδιαφέρον των προβλέψεων έχει αυξηθεί ραγδαία τα τελευταία τριανταπέντε χρόνια. Αυτό το ενδιαφέρον δημιουργήθηκε κυρίως από την ανάγκη λήψης κάποιας απόφασης αντιμετωπίζοντας όσο γίνεται πιο αποτελεσματικά την αβεβαιότητα το μέλλοντος. Η σημασία των προβλέψεων έχει γίνει έντονη τόσο στο επίπεδο των επιχειρήσεων, όσο και στον ακαδημαϊκό χώρο. Ειδικότερα, ο ακαδημαϊκός κόσμος έχει προσφέρει τα μέγιστα, με την δημιουργία μεγάλου αριθμού μεθόδων πρόβλεψης, οι οποίες βρίσκουν πρακτική εφαρμογή στις επιχειρήσεις, αλλά και στον κόσμο της έρευνας.

Οι προβλέψεις είναι ένα βασικό εργαλείο το οποίο χρησιμοποιείται στο σχεδιασμό και τη λειτουργία μιας επιχείρησης. Η επιτυχία εξαρτάται, σε μεγάλο βαθμό, στις σωστές προβλέψεις. Γνωρίζουμε, ωστόσο, ότι το μέλλον είναι εξαιρετικά αβέβαιο. Όλοι βρισκόμαστε καθημερινά αντιμέτωποι με την αβεβαιότητα, η οποία γίνεται όλο και πιο έντονη και έχει επιβάλλει μία πιο προσεκτική και συστηματική έρευνα του μέλλοντος. Επομένως, η πρόβλεψη αποκτά κεντρικό ρόλο και είναι απαραίτητη. Η μεγαλύτερη πρόκληση στον τομέα της πρόβλεψης είναι να γίνει η διαδικασία των προβλέψεων περισσότερο χρήσιμη, αποδοτική και όσο το δυνατόν πιο ακριβής.

Σύμφωνα με τον Μακρυδάκη και τους συνεργάτες του στο βιβλίο «Χορεύοντας με την Τύχη», η αβεβαιότητα που συναντάμε καθημερινά έχει κατηγοριοποιηθεί σε δύο κατηγορίες, την «αβεβαιότητα του μετρό» και την «αβεβαιότητα της καρύδας». Η «αβεβαιότητα του μετρό» αναφέρεται σε συνεχείς μικρές τυχαίες διακυμάνσεις της ιδιωτικής και επιχειρηματικής καθημερινότητας, όπως ακριβώς συμβαίνει με τον χρόνο που θα χρειαστεί ένας συρμός για να μεταβεί από τον έναν σταθμό στον άλλο, λόγω κάποιου τυχαίου προβλήματος, όπως μία τεχνική βλάβη ή πολυκοσμία. Από την άλλη, η «αβεβαιότητα της καρύδας» αναφέρεται σε ένα αναπάντεχο και σπάνιο γεγονός το οποίο μπορεί να έχει σημαντικές συνέπειες, όπως το να πέσει στο κεφάλι

μία καρύδα καθώς περπατάμε. Συνήθως η αβεβαιότητα αυτή αναφέρεται κυρίως στις φυσικές καταστροφές, όπως σεισμός ή πλημμύρες, που είναι δύσκολο έως απίθανο να γνωρίζουμε το που και το πότε θα συντελεστούν, αλλά και ποιο θα είναι το αποτέλεσμα τους.

Ουσιαστικά, το αντικείμενο των προβλέψεων είναι να ελαχιστοποιηθεί η αβεβαιότητα και όσο είναι δυνατόν, να εντοπιστεί και να αξιολογηθεί ο κίνδυνος. Αυτός που παίρνει τις αποφάσεις βρίσκεται αντιμέτωπος με τις αμφιβολίες σχετικά με το μέλλον και για αυτό το λόγο είναι απαραίτητες οι πληροφορίες τόσο για το παρελθόν, όσο και για το μέλλον, ώστε να πάρει την καταλληλότερη απόφαση.

Για τους λόγους αυτούς, το πρόβλημα της πρόβλεψης πρέπει να προσεγγίζεται, όσο το δυνατόν, επιστημονικά. Ωστόσο, θα πρέπει να λαμβάνεται υπόψη η ασάφεια των αντικειμένων που πρέπει να προβλεφθούν. Δεν είναι μόνο ότι η γνώση μας για το μέλλον είναι εξαιρετικά περιορισμένη, αλλά και η γνώση του παρελθόντος δεν έχει ακόμη ολοκληρωθεί. Υπάρχουν πολύ λίγα αξιόπιστα στατιστικά στοιχεία που περιγράφουν το παρελθόν και επομένως, το περιθώριο σφάλματος σε πολλά από τα διαθέσιμα ιστορικά δεδομένα, είναι μεγάλο και σε ορισμένες περιπτώσεις τόσο μεγάλο ώστε να καθίσταται άκυρη οποιαδήποτε στατιστική ανάλυση.

Διαχρονικά, ο τομέας των προβλέψεων έχει δεχτεί δυσμενείς κριτικές λόγω της αδυναμίας των μεθόδων να προειδοποιήσουν έγκαιρα για επερχόμενες αλλαγές, καθώς και για τα μεγάλα σφάλματα στις προβλέψεις. Την ίδια στιγμή, όμως, λανθασμένες ενέργειες που προκαλούνται από ασταθή περιβάλλοντα, μη αναμενόμενες εξελίξεις, ασυνέχειες και άλλα αυξάνουν τις ανάγκες για πραγματοποίηση προβλέψεων. Όταν δεν υπάρχει αβεβαιότητα και τα πάντα κυλούν ομαλά, δεν υπάρχει ουσιαστική ανάγκη για προβλέψεις. Αντιθέτως, σε περιόδους που υπάρχουν συνεχείς και απότομες μεταβολές και η ικανότητα πρόβλεψης είναι αισθητά μειωμένη, η ζήτηση για πραγματοποίηση προβλέψεων αυξάνεται κατακόρυφα. Βεβαίως, η κύρια αιτία των δυσμενών κριτικών που αντιμετωπίζει η επιστήμη των προβλέψεων είναι οι λανθασμένες προσδοκίες που έχουν οι χρήστες της. Η ύπαρξη σφαλμάτων στις προβλέψεις είναι αναπόφευκτη. Αυτό που έχει σημασία είναι να εκτιμώνται όσο πιο ρεαλιστικά γίνεται τα πλεονεκτήματα και οι περιορισμοί των διάφορων μεθόδων πρόβλεψης και να λαμβάνονται υπόψη όταν χρησιμοποιούνται οι προβλέψεις στη λήψη αποφάσεων.

Όλα τα παραπάνω συντελούν στην αναγκαιότητα ύπαρξης εφαρμογών προβλέψεων οι οποίες δεν έχουν σαν προαπαιτούμενο την ευρεία γνώση των τεχνικών προβλέψεων και μπορούν να χρησιμοποιηθούν από κάθε χρήστη.

## 3.2 Κατηγορίες Μεθόδων Πρόβλεψης

Υπάρχουν δύο μεγάλες κατηγορίες μεθόδων πρόβλεψης στις οποίες εντάσσονται οι τεχνικές προβλέψεων που έχουν αναπτυχθεί μέχρι σήμερα. Αυτές είναι οι εξής:

- Ποσοτικές προβλέψεις (quantitative forecasting)
- Ποιοτικές προβλέψεις (qualitative ή judgmental forecasting)

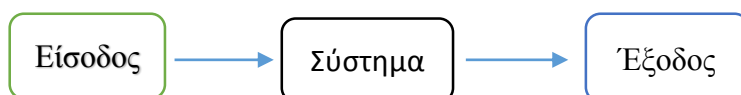
Στα πλαίσια της διπλωματικής θα ασχοληθούμε με τις ποσοτικές μεθόδους πρόβλεψης και ειδικότερα με τις μεθόδους των χρονοσειρών.

### 3.2.1 Ποσοτική πρόβλεψη (quantitative forecasting)

Οι μέθοδοι αυτές βασίζονται στη μαθηματική μοντελοποίηση και άρα είναι αντικειμενικές και επαναλήψιμες (δηλαδή παράγουν το ίδιο αποτέλεσμα κάθε φορά που εισάγουμε τα ίδια δεδομένα). Οι ποσοτικές μέθοδοι απαιτούν μία σειρά από αριθμητικά δεδομένα που όμως δεν είναι πάντα διαθέσιμα ή αξιόπιστα. Οι ποσοτικές μέθοδοι μπορούν να διακριθούν σε *μοντέλα χρονοσειρών (time series models)* και σε αυτές που βασίζονται σε *αιτιοκρατικά ή επεξηγηματικά μοντέλα (causal relationship ή explanatory models)*.

#### 3.2.1.1 Μοντέλο Χρονοσειρών

Πρόκειται για το πιο διαδεδομένο είδος ποσοτικού μοντέλου πρόβλεψης. Το μοντέλο χρονοσειρών προϋποθέτει ότι η απαραίτητη πληροφορία για την πρόβλεψη περιέχεται στη χρονοσειρά των στοιχείων. Χρονοσειρά είναι μία σειρά παρατηρήσεων που λαμβάνονται σε κανονικά διαστήματα μέσα σε ένα καθορισμένο χρονικό διάστημα. Για παράδειγμα, αν καταγράφουμε τις μηνιαίες τιμές του χρυσού στο χρηματιστήριο μετάλλων του Λονδίνου για χρονικό διάστημα 5 ετών, τότε έχουμε στη διάθεσή μας μια χρονοσειρά των μηνιαίων τιμών του χρυσού. Η ανάλυση χρονοσειράς κάνει την υπόθεση ότι μπορεί να γίνει πρόβλεψη με βάση τα μοτίβα (patterns) των διαθέσιμων δεδομένων. Έτσι, η ανάλυση αυτή αναζητάει τάσεις, κυκλικότητα, περιοδικότητα κτλ. στα δεδομένα προκειμένου να δημιουργήσει ένα μοντέλο πρόβλεψης. . Στα μοντέλα χρονοσειρών περιλαμβάνονται οι μέθοδοι αποσύνθεσης, οι μέθοδοι εξομάλυνσης καθώς και οι αυτό-παλινδρομικές μέθοδοι κινητού μέσου όρου.



➤ **Μέθοδοι αποσύνθεσης**

Για την μέθοδο της αποσύνθεσης εντοπίζουμε τις συνιστώσες των διαθέσιμων δεδομένων (δηλ., της χρονοσειράς) προκειμένου να καταλάβουμε τον τρόπο με τον οποίο συμπεριφέρονται. Ακριβώς σε αυτή τη συμπεριφορά βασίζεται η πρόβλεψη. Οι συνιστώσες είναι η τάση (T), ο κύκλος (C), η εποχικότητα (S) και η τυχαιότητα (R). Αν εξετάσουμε μια οποιαδήποτε χρονοσειρά, διαπιστώνουμε ότι αποτελεί συνδυασμό ενός ή περισσότερων από τα παραπάνω στοιχεία. Σκοπός των μεθόδων αποσύνθεσης είναι η απομόνωση των συνιστωσών αυτών με τη μέγιστη δυνατή ακρίβεια.

➤ **Μέθοδοι εξομάλυνσης**

Οι μέθοδοι εξομάλυνσης μπορούν να εφαρμοστούν πολύ εύκολα και είναι κατάλληλες για προβλέψεις με βραχυπρόθεσμο ορίζοντα. Σε αυτές τις μεθόδους γίνεται χρήση των ιστορικών δεδομένων για τον προσδιορισμό μίας εξομαλυμένης τιμής για τις χρονοσειρές. Στη συνέχεια, αυτή η εξομαλυμένη τιμή προεκτείνεται ώστε να αποτελέσει πρόβλεψη για τη μελλοντική τιμή της χρονοσειράς. Η βασική ιδέα των μεθόδων εξομάλυνσης είναι ότι υπάρχει ένα λανθάνον πρότυπο συμπεριφοράς το οποίο ακολουθούν οι τιμές των μεταβλητών πρόβλεψης. Έτσι, ελαχιστοποιείται η τυχαιότητα που υπάρχει στην χρονοσειρά με την πρόβλεψη πλέον να βασίζεται σε ένα εξομαλυμένο πρότυπο συμπεριφοράς. Υπάρχουν δύο κατηγορίες μεθόδων εξομάλυνσης: οι κινητού μέσου όρου και οι εκθετικής εξομάλυνσης.

➤ **Αυτοπαλινδρομικές μέθοδοι κινητού μέσου όρου (ARMA)**

Οι συγκεκριμένες μέθοδοι είναι στοχαστικά μαθηματικά μοντέλα με τα οποία περιγράφεται η χρονική εξέλιξη ενός φυσικού μεγέθους. Τα μοντέλα περιέχουν το τυχαίο παράγοντα, τα ιστορικά δεδομένα και άλλους στοχαστικούς παράγοντες. Το μοντέλο που προκύπτει τελικά είναι ένας γραμμικός συνδυασμός των παραπάνω ποσοτήτων. Τα αυτοπαλινδρομικά μοντέλα βασίζονται στην παραδοχή της αλληλεξάρτησης μεταξύ των τιμών που λαμβάνει η χρονοσειρά τις διάφορες χρονικές στιγμές.

### **3.2.1.2 Αιτιοκρατικό ή επεξηγηματικό μοντέλο**

Στα αιτιοκρατικά μοντέλα χρησιμοποιείται μία αρκετά διαφορετική προσέγγιση για την δημιουργία πρόβλεψης. Θεωρούν ότι η μεταβλητή για την οποία θέλουμε να κάνουμε πρόβλεψη είναι εξαρτημένη με κάποιο τρόπο από μία ή περισσότερες παραμέτρους. Η δυσκολία έγκειται στην εύρεση της μαθηματικής σχέσης με την οποία επηρεάζεται η ζητούμενη μεταβλητή από τις παραμέτρους αυτές. Για παράδειγμα, αν θεωρήσουμε ότι η ζήτηση για σίδηρο οπλισμού εξαρτάται: α) από τη χρηματική αξία των συβάσεων για δημόσια έργα που υπογράφονται μεταξύ του ΥΠΕΧΩΔΕ και των κατασκευαστικών εταιρειών και β) από τον αριθμό των οικοδομικών αδειών που εκδίδονται από τις πολεοδομίες της χώρας, τότε αν βρούμε τη μαθηματική τους σχέση μπορούμε να κατασκευάσουμε ένα μαθηματικό μοντέλο πρόβλεψης για την ζήτηση σε σίδηρο οπλισμού. Είναι προφανές ότι τα αιτιοκρατικά μοντέλα μπορεί να είναι πολύ περίπλοκα, ειδικά στην περίπτωση που λαμβάνονται

υπόψιν πολλές παράμετροι. Στα αιτιοκρατικά μοντέλα ανήκουν οι μέθοδοι παλινδρόμησης και οι οικονομετρικές μέθοδοι.



➤ **Μέθοδοι παλινδρόμησης**

Η παλινδρόμηση περιλαμβάνει διαδικασίες και τεχνικές μοντελοποίησης και ανάλυσης διαφόρων μεταβλητών, όπου το ζητούμενο είναι εύρεση συσχετίσεων μεταξύ μίας εξαρτημένης και μίας ή και περισσότερων ανεξάρτητων μεταβλητών. Ειδικότερα, η ανάλυση της παλινδρόμησης μας βοηθά να καταλάβουμε με ποιο τρόπο αλλάζει η τιμή της εξαρτημένης μεταβλητής, όταν αλλάζουμε την τιμή μίας ανεξάρτητης μεταβλητής, ενώ οι υπόλοιπες διατηρούνται σταθερές. Στην περίπτωση ύπαρξης μίας ανεξάρτητης μεταβλητής η μέθοδος ονομάζεται “Απλή Παλινδρόμηση” ενώ στην περίπτωση περισσότερων, “Πολλαπλή Παλινδρόμηση”.

➤ **Οικονομετρικές μέθοδοι**

Εάν οι ανεξάρτητες μεταβλητές συσχετίζονται μεταξύ τους τότε προκύπτει ένα σύστημα ταυτόχρονων εξισώσεων. Αυτό το σύστημα αποτελεί ένα μοντέλο, το οποίο ονομάζεται οικονομετρικό και συναντάται συχνά σε περιπτώσεις οικονομικών ή επιχειρησιακών σχέσεων.

### **3.2.2 Ποιοτική πρόβλεψη (*qualitative ή judgmental forecasting*)**

Στην κατηγορία αυτή περιλαμβάνονται μέθοδοι στις οποίες η πρόβλεψη γίνεται από έναν ή περισσότερους ειδικούς με βάση την γνώση, την εμπειρία και το ένστικτό τους. Αυτού του είδους η πρόβλεψη είναι υποκειμενική και συχνά περιλαμβάνει συγκεκριμένους τύπους μεροληψιών και προκαταλήψεων που έχουν αρνητική απόδοση στην ακρίβεια των προβλέψεων.

#### **Πάνελ**

Η μέθοδος αυτή βασίζεται στο γεγονός ότι μια πολυπληθής ομάδα ανθρώπων από διαφορετικές θέσεις μπορεί να κάνει μία πιο αξιόπιστη πρόβλεψη απ’ ότι ένας μεμονωμένος ή λίγοι άνθρωποι. Έτσι, διοργανώνονται ανοιχτές συναντήσεις με ελεύθερη ανταλλαγή απόψεων μεταξύ ανθρώπων από όλο το φάσμα των θέσεων ενός οργανισμού. Ένα μειονέκτημα της μεθόδου αυτής είναι ότι η άποψη των υφιστάμενων «σκεδάζεται» ή υποβαθμίζεται από την άποψη των ανώτερων στην ιεραρχία. Αυτό το μειονέκτημα προσπαθεί να διορθώσει η μέθοδος Delphi.

## **Μέθοδος Delphi**

Η μέθοδος αυτή είναι μια τεχνική πρόβλεψης, στόχος της οποίας είναι η προσέγγιση μιας συμφωνίας μεταξύ μιας ομάδας ειδικών, διατηρώντας την ανωνυμία τους. Η ιδέα πίσω από αυτήν είναι ότι ενώ οι ειδικοί δεν θα συμφωνήσουν σε όλα τα ζητήματα, εντούτοις σε ότι συμφωνήσουν αυτά κατά πάσα πιθανότητα που θα συμβούν. Σύμφωνα με τη μέθοδο αυτή:

1. γίνεται επιλογή των συμμετεχόντων, συνήθως ειδικών από διαφορετικές θέσεις ή επιστημονικά υπόβαθρα,
2. μέσω ενός ερωτηματολογίου συλλέγονται οι απόψεις όλων (χωρίς ο ένας να δει ή να γνωρίζει τους υπόλοιπους συμμετέχοντες),
3. οι απαντήσεις όλων ταξινομούνται και επανατροφοδοτούνται (feedback) στους συμμετέχοντες μαζί με ένα καινούργιο ερωτηματολόγιο,
4. το βήμα 3 επαναλαμβάνεται όσες φορές κρίνεται απαραίτητο προκειμένου να επιτευχθεί μια συμφωνία μεταξύ των συμμετεχόντων. Συνήθως 3 ή 4 «γύροι» είναι αρκετοί.

Λόγω της ανωνυμίας του καθενός και της ίδιας βαρύτητας όλων των απόψεων, με τη μέθοδο Delphi αποφεύγεται το μειονέκτημα της πρώτης μεθόδου (πάνελ). Από την άλλη, η μέθοδος είναι σχετικά χρονοβόρα.

## **Έρευνα αγοράς**

Αποτελεί μια προσέγγιση που χρησιμοποιεί ερωτηματολόγια και συνεντεύξεις για τον καθορισμό των αναγκών, των προτιμήσεων, των επιλογών κτλ., μιας ομάδας στόχου (π.χ. των καταναλωτών). Η μέθοδος χρησιμοποιείται ευρέως για την βελτίωση και την δημιουργία καινούργιων προϊόντων. Σημαντικό στοιχείο για την επιτυχία της μεθόδου είναι ο σχεδιασμός των ερωτηματολογίων (ή των συνεντεύξεων).

## **3.3 Χρονοσειρές**

### **3.3.1 Έννοια της Χρονοσειράς**

Οι χρονοσειρές αποτελούν ένα σύνολο διαδοχικών παρατηρήσεων της τιμής κάποιου φυσικού ή άλλου μεγέθους ανηγμένες στο χρόνο. Οι χρονοσειρές μπορούν να κατηγοριοποιηθούν με βάση τον τρόπο προσδιορισμού των μελλοντικών τιμών τους, σε ντετερμινιστικές και στοχαστικές χρονοσειρές αντίστοιχα. Όταν οι διαδοχικές παρατηρήσεις δεν είναι ανεξάρτητες μεταξύ τους και οι μελλοντικές τιμές μπορούν να προσδιοριστούν ακριβώς από τις προηγούμενες, τότε αυτές οι χρονοσειρές είναι ντετερμινιστικές. Ωστόσο, κάτι τέτοιο δεν συμβαίνει με τις πραγματικές χρονοσειρές, καθώς το μέλλον καθορίζεται μερικώς από το παρελθόν. Αυτό συμβαίνει διότι στον

πραγματικό κόσμο οι χρονοσειρές, δεν περιγράφουν την διαχρονική εξέλιξη κάποιου φυσικού μεγέθους, για το οποίο υπάρχει πλήρης γνώση όλων των παραγόντων που το επηρεάζουν, αλλά αντιθέτως επηρεάζονται από τον λεγόμενο «τυχαίο παράγοντα», ο οποίος αντιπροσωπεύει μία στατιστική μεταβλητή.

Ωστόσο, ο διαχωρισμός των χρονοσειρών σύμφωνα με την παραπάνω κατηγοριοποίηση δεν είναι τόσο εύκολος και προφανής. Παρόλα αυτά, θα πρέπει να γίνει η κατηγοριοποίηση ώστε να γίνεται πιο εύκολα η αναγνώριση και η κατανόηση των παραμέτρων που παίζουν ρόλο στη χρονική εξέλιξη της χρονοσειράς ανεξαρτήτως των δεδομένων της.

### **3.3.2 Ποιοτικά Χαρακτηριστικά της Χρονοσειράς**

Οι χρονοσειρές δομούνται από κάποια βασικά χαρακτηριστικά, όπου με τις κατάλληλες τεχνικές μπορούν να αναλυθούν σε αυτά. Οι παραδοσιακές μέθοδοι ανάλυσης χρονοσειρών ασχολούνται με την αποσύνθεση της χρονοσειράς σε τέσσερα βασικά της στοιχεία: την τάση, την κυκλικότητα, την εποχιακότητα και τις μη κανονικές διακυμάνσεις. Αυτή η προσέγγιση είναι χρήσιμη όχι μόνο για την εφαρμογή της κλασικής μεθόδου αποσύνθεσης, αλλά και για την περαιτέρω ανάλυση των ποιοτικών χαρακτηριστικών της χρονοσειράς, ανεξαρτήτως της μεθόδου πρόβλεψης που θα χρησιμοποιηθεί.

Αρχικά, η **τάση** ορίζεται σαν μία «μακροπρόθεσμη» μεταβολή του μέσου επιπέδου τιμών της χρονοσειράς. Βέβαια, ενώ αυτός ο όρος είναι κοινά αποδεκτός, υπάρχει ένα πρόβλημα σχετικά με την αποσαφήνιση του ποια μεταβολή θεωρείται μακροπρόθεσμη. Για παράδειγμα βλέπουμε ότι οι περιβαλλοντολογικές και κλιματικές μεταβλητές εμφανίζουν μία κυκλική διακύμανση 50 ετών. Επομένως, αν κάποιος διαθέτει στοιχεία για τα πρώτα 20 έτη θα θεωρούσε την παρατηρούμενη αύξηση του μέσου επιπέδου σαν τάση, ενώ αν εξέταζε δεδομένα μερικών εκατοντάδων ετών θα συμπέραινε πως πρόκειται, για κυκλική διακύμανση. Άρα, για να μπορέσει κανείς να διαπιστώσει την τάση θα πρέπει να έχει στην διάθεσή του έναν ικανό αριθμό παρατηρήσεων και ταυτόχρονα να εκτιμήσει το κατάλληλο μήκος της περιόδου μέσα στην οποία θα αναζητήσει την ύπαρξη της.

Η **κυκλικότητα** είναι το δεύτερο συστατικό μίας χρονοσειράς και αντιπροσωπεύει μια “κυματοειδή” μεταβολή που οφείλεται σε ειδικές εξωγενείς συνθήκες και εμφανίζεται κατά περιόδους, οι οποίες δεν είναι σταθερές και είναι συνήθως είναι μεγαλύτερες του ενός έτους. Κυκλικότητα εμφανίζουν οι χρονοσειρές, κατά κύριο λόγο των περισσότερων οικονομικών μεγεθών όπως για παράδειγμα του Ακαθάριστου Εγχώριου Προϊόντος, των τιμών των μετοχών και των τιμών του πετρελαίου. Αυτές οι κυκλικές μεταβολές οφείλονται στις γενικότερες οικονομικές συνθήκες που επικρατούν και που χαρακτηρίζονται από περιόδους με διαδοχικές ανόδους και υφέσεις και είναι γνωστές με την ονομασία επιχειρηματικός κύκλος.

Η **εποχιακότητα** είναι μία περιοδική διακύμανση που έχει σταθερό μήκος και μικρότερο του έτους μήκος. Η εποχιακότητα μπορεί εύκολα να αναγνωριστεί παρατηρώντας την γραφική της αναπαράσταση. Επειδή οι αλλαγές που προκαλεί στα δεδομένα της χρονοσειράς είναι σταθερές με την πάροδο του χρόνου, είναι εύκολο να αντιμετωπίσουμε την επίδρασή της. Αυτό συμβαίνει με την εύρεση των δεικτών εποχιακότητας για τα αντίστοιχα χρονικά διαστήματα, όπου με την διαίρεση αυτών των δεικτών με τα πραγματικά δεδομένα παίρνουμε την λεγόμενη αποεποχικοποιημένη χρονοσειρά, η οποία είναι απαλλαγμένη από την επίδραση της εποχιακότητας. Ένα πολύ χαρακτηριστικό παράδειγμα εποχιακής χρονοσειράς είναι οι πωλήσεις παγωτών που εμφανίζουν μία έντονη άνοδο τους καλοκαιρινούς μήνες και οι πωλήσεις καυσίμων τους χειμερινούς.

Ως **ασυνέχειες** ορίζονται οι απομονωμένες παρατηρήσεις που εμφανίζονται στη γραφική αναπαράσταση κάποιας χρονοσειράς ως απότομες αλλαγές στο πρότυπο συμπεριφοράς της. Τέτοιες αλλαγές έχουν είτε παροδικό είτε μόνιμο χαρακτήρα. Αυτές που έχουν παροδικό χαρακτήρα και που η επίδρασή τους στη χρονοσειρά έχει μικρή χρονική διάρκεια ονομάζονται *outliers* ή *special events*. Πρέπει να τονιστεί εδώ πως η αναγνώρισή τους είναι δύσκολη και απαιτεί και θεωρητική γνώση, αλλά και κριτική ικανότητα. Ένα *outlier* αποτελεί μία ασυνήθιστη παρατήρηση που οφείλεται σε κάποιο εξαιρετικό ή απρόβλεπτο γεγονός. Για παράδειγμα, μία απεργία μπορεί να προκαλέσει δραματική πτώση στη παραγωγική διαδικασία, το οποίο έχει σαν αποτέλεσμα μία μικρή πτώση στο γράφημα παραγόμενων προϊόντων της εταιρείας.

Στον αντίποδα, οι αλλαγές με μόνιμο χαρακτήρα ονομάζονται *level-shifts* αφού εμφανίζονται ως απότομες αλλαγές στο μέσο επίπεδο των τιμών της χρονοσειράς. Ένα παράδειγμα τέτοιας μόνιμης αλλαγής είναι η πτώση των πωλήσεων μίας εταιρείας εξαιτίας της εισόδου στην αγορά μίας ανταγωνίστριας εταιρείας. Με αυτές τις συνθήκες θα υπάρξει μία απότομη μείωση του μέσου επιπέδου τιμών, το οποίο θα σταθεροποιηθεί σε ένα πιο χαμηλό σημείο.

**Μη κανονικές διακυμάνσεις** θεωρούνται εκείνες που απομένουν όταν έχουν απομονωθεί όλα τα υπόλοιπα χαρακτηριστικά της χρονοσειράς. Οι διακυμάνσεις αυτές αντιπροσωπεύουν μία τυχαία μεταβλητή που εκφράζει τον τυχαίο παράγοντα μιας στοχαστικής διαδικασίας. Επίσης μπορεί να αντιπροσωπεύουν κάποια ασυνέχεια που συνδέεται με κάποιο εξαιρετικό γεγονός.

Τέλος, θα πρέπει να τονιστεί πως η οπτικοποίηση των δεδομένων είναι πολύ σημαντική για την ευκολότερη εύρεση των ποιοτικών χαρακτηριστικών των χρονοσειρών, όπως φαίνεται στα παρακάτω παραδείγματα.

### **3.3.3 Αναπαράσταση Χρονοσειρών**

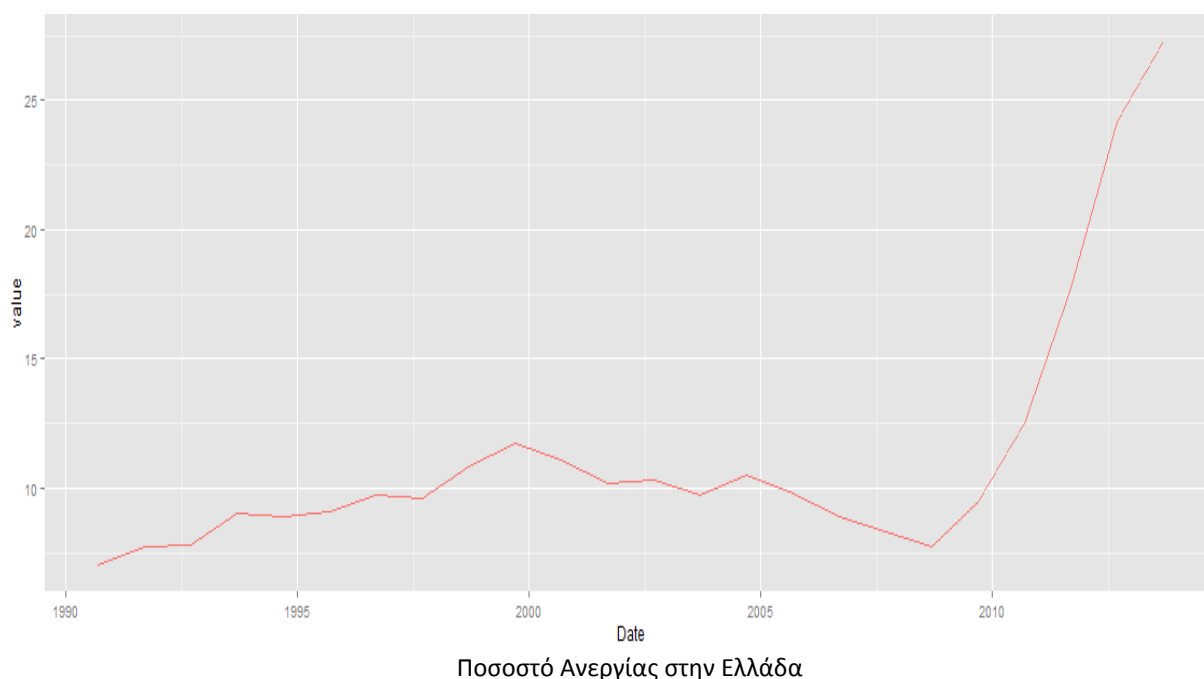
Η χρονοσειρά δεν είναι τίποτα παραπάνω από μία σειρά παρελθουσών τιμών για την περιγραφή ενός μεγέθους. Η γραφική αναπαράσταση των δεδομένων αποτελεί ένα



σημαντικό εργαλείο που βοηθάει στην ανάλυση της χρονοσειράς, αλλά και τη διαδικασία της πρόβλεψης. Από την αναπαράσταση των δεδομένων γίνονται πιο εμφανή τα βασικά στοιχεία της χρονοσειράς, κάτι το οποίο βοηθάει τον αναλυτή στην επιλογή της κατάλληλης μεθοδολογίας και εργαλείων με σκοπό να έχει τα βέλτιστα αποτελέσματα με το μικρότερο σφάλμα. Επίσης, με το γράφημα μπορεί να γίνει εύκολα ο εντοπισμός ακραίων εσφαλμένων τιμών, και έτσι ο αναλυτής μπορεί κάνοντας τις κατάλληλες κινήσεις να διορθώσει τις εσφαλμένες τιμές.

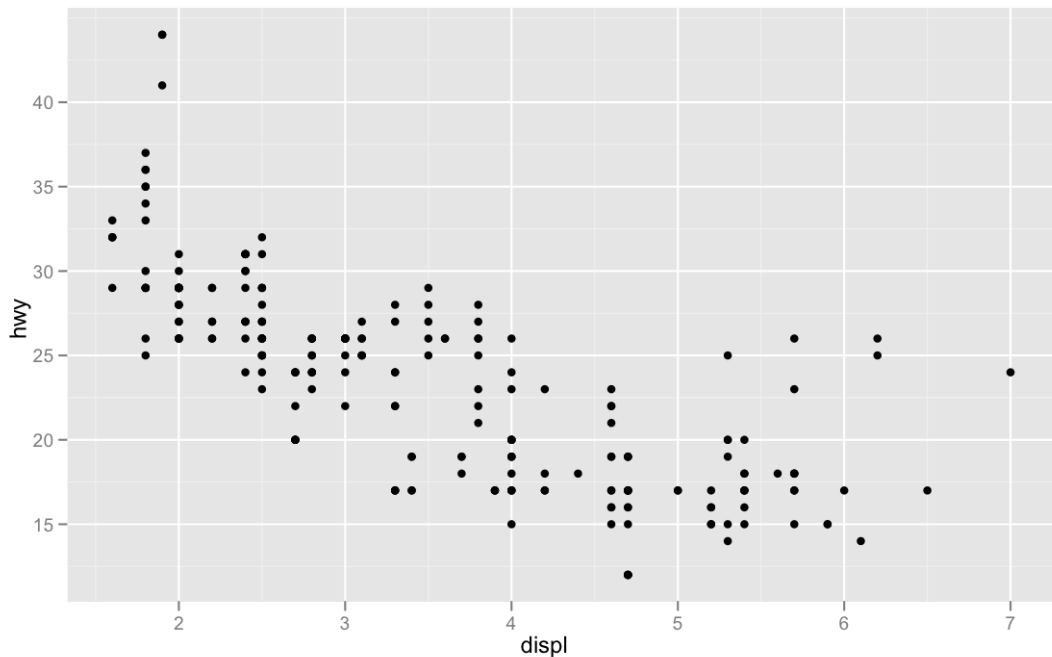
Η αναπαράσταση ουσιαστικά έγκειται σε δισδιάστατη γραφική απεικόνιση των πραγματικών τιμών των διαθέσιμων ιστορικών δεδομένων ως προς το χρόνο. Οι κύριοι τύποι γραφημάτων που χρησιμοποιούνται είναι:

- 1. Διαγράμματα Χρόνου (time plots):** Είναι το πιο χρησιμοποιούμενο διάγραμμα και αναπαριστά τα διαθέσιμα ιστορικά δεδομένα στην πάροδο του χρόνου. Με την χρησιμοποίηση αυτών των διαγραμμάτων είναι πολύ πιο εύκολο να αντιληφθούμε τα βασικά χαρακτηριστικά των χρονοσειρών. Λόγω της ευκολίας της κατανόησης του συγκεκριμένου είδους γραφήματος από την πλειοψηφία του κόσμου, οδήγησε στην χρησιμοποίηση αυτού του τύπου γραφήματος για την απεικόνιση των δεδομένων στην εφαρμογή που κατασκευάστηκε.



- 2. Εποχιακά Διαγράμματα (seasonal plots):** Αυτά τα διαγράμματα χρησιμοποιούνται για χρονοσειρές που παρουσιάζουν έντονη εποχιακότητα.
- 3. Διαγράμματα Διασποράς (scatter plots):** Το διάγραμμα διασποράς είναι το κατάλληλο γράφημα που δημιουργούμε ως πρώτο βήμα για να μελετήσουμε τη σχέση που υπάρχει μεταξύ δύο συνεχών αριθμητικών μεταβλητών, ιδιαίτερα αν αποσκοπούμε στη δημιουργία μοντέλου πρόγνωσης των τιμών της μίας μεταβλητής από την άλλη. Με το διάγραμμα διασποράς και μια

έμπειρη στατιστική ματιά ανιχνεύεται η σχέση που ενδεχομένως να υπάρχει μεταξύ των δύο μεταβλητών.



### 3.4 Βασικά Βήματα Προβλέψεων

Για την επίτευξη της μεγαλύτερης δυνατής ακρίβειας και αξιοπιστίας κατά την διάρκεια της πρόβλεψης, είναι απαραίτητο ο σχεδιασμός και η εκτέλεση να γίνεται με ένα πολύ προσεκτικό και καλά σχεδιασμένο τρόπο.

#### **Βήμα 1<sup>ο</sup>: Ορισμός του προβλήματος**

Ο ορισμός του προβλήματος είναι αρκετές φορές το πιο δύσκολο μέρος στην διαδικασία πρόβλεψης. Κατά τη διάρκεια αυτού του βήματος θα πρέπει να γίνει απολύτως σαφές τι θέλουμε να προβλέψουμε, αλλά και να καθορίσουμε όλες τις παραμέτρους του προβλήματος που μελετάται. Επίσης θα πρέπει να γίνει αποτίμηση, όλων των συνιστωσών που πιθανώς επηρεάζουν το τελικό αποτέλεσμα.

#### **Βήμα 2<sup>ο</sup>: Συλλογή των πληροφοριών**

Γενικά υπάρχουν δύο τύποι δεδομένων, που μπορεί να χρησιμοποιηθούν σε προβλήματα προβλέψεων. Από τη μία είναι τα στατιστικά ιστορικά δεδομένα και από την άλλη η συσσωρευμένη εμπειρία και κριτική ικανότητα διάφορων ανθρώπων. Και

οι δύο τύποι πληροφοριών θα πρέπει να λαμβάνονται υπόψιν στις προβλέψεις. Η συλλογή των ιστορικών δεδομένων είναι πολύ σημαντική, διότι αποτελούν την βάση με τα οποία θα εφαρμόσουμε τις μεθόδους πρόβλεψης και θα εξάγουμε τις προβλέψεις μας για αυτά.

### **Βήμα 3<sup>ο</sup>: Προετοιμασία χρονοσειρών**

Με αυτό το βήμα μπορούμε να αποκτήσουμε μία ολοκληρωμένη αίσθηση των διαθέσιμων δεδομένων, και έτσι να αναγνωρίσουμε πιθανά λανθάνοντα πρότυπα, όπως σημαντική τάση ή εποχιακότητα και ασυνέχειες. Η ανάλυση αυτή βοηθά και στην αναγνώριση πιθανής ανάγκης για χρήση κριτικής αξιολόγησης από μέρους ειδικών. Η προσαρμογή των δεδομένων, η απομόνωση των συνιστωσών της χρονοσειράς και η αναγνώριση και απομόνωση ειδικών γεγονότων και ενεργειών μας οδηγεί σε μία εξομλυμένη σειρά δεδομένων, έτοιμη για την εφαρμογή μοντέλων πρόβλεψης.

### **Βήμα 4<sup>ο</sup>: Επιλογή μεθόδων πρόβλεψης**

Η κατασκευή ενός μοντέλου πρόβλεψης περιλαμβάνει την προσαρμογή των δεδομένων που έχουν συλλεχθεί σε ένα υπάρχον μοντέλο, με στόχο την ελαχιστοποίηση των σφαλμάτων. Συνεπώς απαιτείται η κατάλληλη επιλογή μοντέλου που πρέπει ιδανικά να έχει τη κατάλληλη ισορροπία μεταξύ ενός πολύπλοκου μοντέλου και ενός πιο απλοϊκού. Και αυτό διότι το πιο πολύπλοκο μοντέλο μπορεί να έχει μεγαλύτερη ακρίβεια, αλλά το πιο απλό μοντέλο είναι πιο εύκολο να γίνει κατανοητό. Άρα επιτυγχάνεται η σωστή επιλογή των κατάλληλων μεθόδων πρόβλεψης, αλλά και επιλογή των παραμέτρων τους, σύμφωνα με τα ιστορικά δεδομένα τους.

### **Βήμα 5<sup>ο</sup>: Χρήση και αξιολόγηση των μοντέλων πρόβλεψης**

Αφού επιλέξουμε το κατάλληλο μοντέλο πρόβλεψης και κάνουμε σωστή επιλογή των παραμέτρων του, στο τελευταίο στάδιο το επιλεγμένο μοντέλο χρησιμοποιείται ώστε να παραχθούν οι ζητούμενες προβλέψεις. Η αξιολόγηση του μοντέλου μπορεί να επιτευχθεί με το πέρασμα του χρόνου, καθώς νέα δεδομένα γίνονται διαθέσιμα. Και αυτό διότι έτσι μπορεί να γίνει σύγκριση ανάμεσα στις παραχθείσες προβλέψεις και στα καινούργια διαθέσιμα δεδομένα. Αυτή η αξιολόγηση είναι δυνατή χρησιμοποιώντας εξειδικευμένους στατιστικούς δείκτες. Τέλος, η παρακολούθηση του προτύπου χρονοσειράς και των σφαλμάτων της πρόβλεψης συμβάλλει στην έγκαιρη αντιμετώπιση διαρθρωτικών αλλαγών και στην εξάλειψη της προκατάληψης στις προβλέψεις.

### **3.5 Πεδία εφαρμογής των Προβλέψεων**

Όπως έχει ήδη αναφερθεί οι προβλέψεις είναι πάρα πολύ σημαντικές και βρίσκουν εφαρμογή σε πληθώρα επιστημονικών και ερευνητικών πεδίων, αλλά και σε καθημερινές πτυχές της ζωής μας. Παρακάτω παρατίθενται μερικά πεδία εφαρμογής των τεχνικών προβλέψεων.

#### **Οικονομία**

Οι κυβερνήσεις κυρίως, αλλά και άλλοι οργανισμοί έχουν ανάγκη να προβλέψουν βασικούς οικονομικούς δείκτες, όπως το Ακαθάριστο Εγχώριο Προϊόν, την ανεργία, τον πληθωρισμό κλπ. Επομένως οι προβλέψεις είναι απαραίτητες για τον καθορισμό των δημοσιονομικών και νομισματικών πολιτικών τους.

#### **Χρηματοοικονομικά**

Στον κόσμο των επιχειρήσεων, αλλά κυρίως των επενδύσεων είναι απαραίτητες οι προβλέψεις για τις τιμές των μετοχών, των επιτοκίων δανεισμού, της ισοτιμίας και των τιμών των αγαθών. Συνεπώς οι τεχνικές προβλέψεων έχουν εφαρμοσθεί αι μελετηθεί ευρέως στην πρόβλεψη γενικών τιμών χρηματιστηρίων αλλά και στην πρόβλεψη μεταβλητότητας.

#### **Παραγωγή**

Οι επιχειρήσεις έχουν ανάγκη να προβλέπουν, εκτός από τα επίπεδα τιμών που θα διαμορφωθούν μελλοντικά, αλλά και τις ποσότητες των πρώτων υλών που θα χρειαστούν για την παραγωγική διαδικασία. Αυτό για να γίνει θα πρέπει να μελετηθούν πρώτα από όλα οι προβλέψεις που αφορούν την ζήτηση.

#### **Marketing**

Και στον τομέα του marketing οι προβλέψεις παίζουν σπουδαίο ρόλο. Αποφάσεις για το κόστος της διαφήμισης εξαρτάται άμεσα από τις προβλέψεις για την ζήτηση στην αγορά.

#### **Περιβάλλον και Κλίμα**

Για θέματα όπως η ενεργειακή ζήτηση, η διαχείριση υδάτινων πόρων, η μετεωρολογία και η ρύπανση αντιμετωπίζονται συχνά με χρήση τεχνικών προβλέψεων.

#### **Κοινωνία**

Δημογραφικές προβλέψεις διεξάγονται συχνά με αντικείμενο τον πληθυσμό και την πληθυσμιακή σύνθεση διαφόρων περιοχών του κόσμου. Επίσης προβλέψεις γίνονται και τομείς όπως η εγκληματικότητα και η διάδοση επιδημιών.

## **Τουρισμός**

Ο τουρισμός αποτελεί ως γνωστόν την βασικότερη υπηρεσία που εξάγει η χώρα μας. Για αυτό το λόγο χρησιμοποιούνται μοντέλα προβλέψεων για την εκτίμηση των συνολικών αφίξεων τουριστών, τις διανυκτερεύσεις ανά περιοχή και το ταξιδιωτικό συνάλλαγμα.

## **Μεταφορές και μετακινήσεις**

Η πρόβλεψη του κυκλοφοριακού φόρτου και η πρόβλεψη καταληψιμότητας ανά μονάδα χρόνου σε διαφορετικά γεωγραφικά σημεία και προς συγκεκριμένη κατεύθυνση είναι μερικά προβλήματα που μπορούν να αντιμετωπιστούν με εξειδικευμένα μοντέλα και τεχνικές προβλέψεων.

## **3.6 Μέθοδοι Πρόβλεψης Συνεχούς Ζήτησης**

Στην παρούσα διπλωματική, χρησιμοποιήσαμε 8 μεθόδους πρόβλεψης χρονοσειρών, στις οποίες έγινε αργότερα σύγκριση για το ποια μέθοδος δίνει τα καλύτερα αποτελέσματα με βάση τα δεδομένα και τους στατιστικούς δείκτες. Επίσης εξετάστηκε και η σειρά κατάταξης της κάθε μεθόδου βάσει των δεικτών που εξετάστηκαν. Επομένως στην εφαρμογή ενσωματώθηκαν οι παρακάτω μέθοδοι πρόβλεψης:

1. Απλοϊκή μέθοδος - Naive
2. Κινητός μέσος όρος – KMO
3. Απλή γραμμική παλινδρόμηση – LRL
4. Απλή εκθετική εξομάλυνση – SES
5. Εκθετική εξομάλυνση γραμμικής τάσης – Holt
6. Εκθετική εξομάλυνση μη γραμμικής τάσης – Damped
7. Error-Trend-Seasonal ή Exponential Smoothing method – ETS
8. Κλασική μέθοδος Theta

Στις παρακάτω παραγράφους αναφέρονται αναλυτικά τα παραπάνω μοντέλα, αλλά και μερικά ακόμα που δεν έχουν ενσωματωθεί στην εφαρμογή.

### **3.6.1 Απλοϊκή μέθοδος – Naive**

Στην ουσία η Naive είναι η πιο απλή μέθοδος πρόβλεψης. Η συγκεκριμένη στατιστική μέθοδος δίνει ως πρόβλεψη για την επόμενη χρονική περίοδο, την ίδια τιμή με την πραγματική τιμή της τωρινής περιόδου. Η αποδοτικότητά της είναι καλή για αποεποχικοποιημένες χρονοσειρές και προβλέψεις μίας περιόδου, διότι συνήθως η

επόμενη τιμή της χρονοσειράς δεν διαφέρει σημαντικά από την προηγούμενη παρατήρηση που είναι διαθέσιμη. Συνήθως όμως δεν δίνει ακριβείς προβλέψεις και για αυτό το λόγο χρησιμοποιείται περισσότερο ως σημείο αναφοράς (benchmark) για άλλες πιο πολύπλοκες μεθόδους. Ο τύπος που χρησιμοποιείται για την περιγραφή της μεθόδου είναι:

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

### **3.6.2 Μέθοδοι Μέσου Όρου**

Με τις μεθόδους των μέσων όρων γίνεται χρήση των ιστορικών δεδομένων, ώστε να προσδιοριστεί μία εξομαλυμένη τιμή για τις χρονοσειρές κάτι το οποίο ομαλοποιεί την χρονοσειρά και έτσι μπορούν να χρησιμοποιηθούν και ως μέθοδοι πρόβλεψης. Αυτό γίνεται πολύ εύκολα, προεκτείνοντας αυτή τη τιμή για την χρονική περίοδο που θέλουμε να κάνουμε την πρόβλεψη, και ως συνέπεια αυτή η τιμή αποτελεί την μελλοντική τιμή της χρονοσειράς. Μερικοί από τους μέσους όρους που χρησιμοποιούμε για την παραγωγή προβλέψεων είναι ο απλός μέσος όρος και ο κινητός μέσος όρος.

#### **3.6.2.1 Απλός μέσος όρος**

Η μέθοδος του απλού μέσου όρου βασίζεται στη λογική της εύρεσης του μέσου όρου όλων των διαθέσιμων ιστορικών δεδομένων και η χρησιμοποίηση αυτής της τιμής ως πρόβλεψη. Γενικά η χρήση αυτής της μεθόδου προκρίνεται για περιπτώσεις που τα δεδομένα δεν παρουσιάζουν τάση ή έντονη εποχιακότητα, αλλά εμφανίζουν μία σταθερότητα με την πάροδο του χρόνου. Επομένως ο τύπος που μας δίνει αυτή την πρόβλεψη είναι εξής:

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

#### **3.6.2.2 Κινητός μέσος όρος**

Για να είναι εφικτός ο υπολογισμός του κινητού μέσου όρου των δεδομένων μίας χρονοσειράς θα πρέπει να καθοριστεί αρχικά ο αριθμός των δεδομένων που θα συμπεριληφθεί στον υπολογισμό αυτό. Ο όρος κινητός μέσος όρος περιγράφει την διαδικασία κατά την οποία καθώς μία καινούργια παρατήρηση γίνεται διαθέσιμη, μπορεί να υπολογιστεί ένας καινούργιος μέσος όρος, στον οποίο παραλείπεται η πιο παλιά παρατήρηση προκειμένου να συμπεριληφθεί η πιο πρόσφατη. Αυτός ο νέος

κινητός μέσος όρος θα είναι η πρόβλεψη για την επόμενη περίοδο. Θα πρέπει να τονιστεί πως ο αριθμός των παρατηρήσεων που χρησιμοποιούνται για την εξαγωγή του μέσου όρου πρέπει οπωσδήποτε να παραμείνει σταθερός καθ' όλη την διαδικασία πρόβλεψης και εμπεριέχει πάντα τις πιο πρόσφατες παρατηρήσεις. Αν η χρονοσειρά αποτελείται από  $n$  δεδομένα και το μήκος των παρατηρήσεων που θα χρησιμοποιηθούν για τον υπολογισμό κάθε μέσου όρου είναι  $k$ , τότε ονομάζεται κινητός μέσος όρος μήκους  $k$  ή ΚΜΟ( $k$ ) ή ΜΑ( $k$ ). Η σχέση που χρησιμοποιείται για την εφαρμογή της μεθόδου του κινητού μέσου όρου είναι:

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

Είναι προφανές πως αν το  $k$  τεθεί ίσο με τη μονάδα, τότε ουσιαστικά η τιμή του ΚΜΟ(1) ταυτίζεται με την απλοϊκή μέθοδο (Naive). Είναι επίσης ξεκάθαρο πως ο κινητός μέσος όρος δίνει ιδιαίτερη σημασία στις τελευταίες ίσου μήκους παρατηρήσεις, αλλά δυστυχώς κάτι τέτοιο απαιτεί περισσότερο χώρο αποθήκευσης, διότι πρέπει να αποθηκεύονται όλες οι παρατηρήσεις από τις οποίες γίνεται η εξαγωγή του κινητού μέσου όρου κάθε φορά και όχι μόνο η τιμή του μέσου όρου.

Ένα από τα μειονεκτήματα των μεθόδων μέσων όρων είναι ότι δεν γίνεται να διαχειριστεί επιτυχώς ορισμένα κύρια χαρακτηριστικά των χρονοσειρών, όπως είναι η τάση και η εποχιακότητα. Για αυτό το λόγο το μοντέλο μέσων όρων χρησιμοποιείται κυρίως για την εξομάλυνση των συνιστωσών της εποχιακότητας και της τάσης των χρονοσειρών, ώστε να προκύψει μία εκτίμηση της γραμμής τάσης-κύκλου. Άρα τα μοντέλα αυτά θα μπορούσαμε να πούμε ότι είναι κατάλληλα για εργαλεία αποσύνθεσης παρά για μεθόδους πρόβλεψης.

### **3.6.3 Μέθοδοι Εκθετικής Εξομάλυνσης**

Οι μέθοδοι πρόβλεψης της εκθετικής εξομάλυνσης είναι στην ουσία μία προέκταση του προτύπου των ιστορικών δεδομένων, όπως τάση και εποχιακότητα. Οι προβλέψεις παράγονται αφού γίνει η εξομάλυνση των παρατηρήσεων, προκειμένου να απομονωθούν τα πραγματικά πρότυπα από τις τυχαίες διακυμάνσεις. Ουσιαστικά, οι μέθοδοι εκθετικής εξομάλυνσης είναι μία προέκταση των κινητών μέσων όρων και ως εκ τούτου, έχουν ως βασική αρχή πως τα πιο πρόσφατα ιστορικά δεδομένα εμπεριέχουν και την πιο σημαντική πληροφορία. Άρα τα πιο πρόσφατα δεδομένα αποτελούν τον οδηγό για την εξαγωγή των προβλέψεων, ενώ στα πιο παλιά μειώνεται εκθετικά ο συντελεστής βαρύτητας. Για αυτό εξάλλου καλούνται και μέθοδοι εκθετικής εξομάλυνσης.

Οι μέθοδοι εκθετικής εξομάλυνσης παρουσιάστηκαν για πρώτη φορά στις αρχές του 1940 και γνώρισαν ευρεία εφαρμογή λόγω της επανάστασης της πληροφορικής το

1960. Οι μέθοδοι αυτοί είναι αρκετά δημοφιλείς στο πεδίο των προβλέψεων λόγω της απλότητας των μοντέλων που υιοθετούν, των περιορισμένων απαιτήσεών τους για αποθήκευση δεδομένων και το μειωμένο υπολογιστικό φόρτο που απαιτούν. Από ότι φαίνεται παρά την απλότητά τους, σύμφωνα με τα αποτελέσματα εμπειρικών μελετών, παρουσιάζουν πολύ καλά αποτελέσματα ακρίβειας σε σχέση με πιο πολύπλοκες μεθόδους. Αυτό συμβαίνει διότι οι μέθοδοι εκθετικής εξομάλυνσης δεν επηρεάζονται από τις ιδιομορφίες των προτύπων των δεδομένων, ούτε από εμφανιζόμενες ακραίες τιμές, οι οποίες είναι συχνό φαινόμενο στα επιχειρησιακά δεδομένα.

Τα μοντέλα εκθετικής εξομάλυνσης κατηγοριοποιούνται σύμφωνα με τη γενική μορφή της γραφικής αναπαράστασης των δεδομένων με οριζόντιο άξονα τον χρόνο. Σύμφωνα με αυτό τον τρόπο της κατηγοριοποίησης προκύπτουν τέσσερα μοντέλα τάσης, τα μοντέλα σταθερού επιπέδου, γραμμικής τάσης, εκθετικής τάσης και φθίνουσας τάσης. Πριν προβούμε σε περαιτέρω ανάλυση του κάθε μοντέλου, κρίνεται σκόπιμο να αναφερθούν ορισμένα χαρακτηριστικά και για τα τέσσερα μοντέλα, ώστε να υπάρχει μία γενική εικόνα για την χρήση τους.

Αρχικά, το μοντέλο σταθερού επιπέδου υποθέτει την απουσία τάσης από τα δεδομένα. Συνήθως αυτά τα μοντέλα χρησιμοποιούνται ευρέως για προβλέψεις ενός βήματος ή όταν οι χρονοσειρές χαρακτηρίζονται από αυξημένο θόρυβο ή τυχαιότητα. Εν συνεχεία, το μοντέλο γραμμικής τάσης είναι πιο διαδεδομένο, διότι η πρόβλεψη γίνεται προκύπτει απλά από την προέκταση μίας ευθείας γραμμής για οποιαδήποτε χρονική στιγμή στο μέλλον. Ένα χαρακτηριστικό παράδειγμα που ενδείκνυται η χρήση του μοντέλου της εκθετικής τάσης είναι όταν βρισκόμαστε στην αρχή του κύκλου ζωής ενός προϊόντος, καθώς το ποσοστό αύξησης των πωλήσεων δύναται να μεγαλώνει συνεχώς.

Πολλές μελέτες έχουν δείξει πως τα μοντέλα γραμμικής και εκθετικής τάσης δίνουν γενικά υπεραισιόδοξα αποτελέσματα και οδηγούν σε ιδιαίτερα υψηλές τιμές προβλέψεων και συνεπώς σε αυξημένες τιμές του δείκτη Mean Error. Έτσι προέκυψε η ανάγκη εφαρμογής ενός άλλου μοντέλου, όπως το μοντέλο φθίνουσας τάσης, το οποίο στις περιπτώσεις μεσοπρόθεσμων και μακροπρόθεσμων προβλέψεων αποτελεί την καλύτερη επιλογή, διότι μειώνεται σταδιακά το μέγεθος κατά το οποίο αυξάνονται οι τιμές της χρονοσειράς κάθε χρονική περίοδο. Παρακάτω αναλύονται και παρουσιάζονται τα μοντέλα για την καλύτερη κατανόησή τους.

### **3.6.3.1 Μοντέλο σταθερού επιπέδου (Simple Exponential Smoothing)**

Το μοντέλο σταθερού επιπέδου ή αλλιώς απλή εκθετική εξομάλυνση, περιγράφεται από τις παρακάτω εξισώσεις:

$$e_t = Y_t - F_t$$

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



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

Όπου  $e_t$  είναι το σφάλμα που προκύπτει από τη διαφορά της πραγματικής με την τιμή της πρόβλεψης για την ίδια χρονική περίοδο  $t$ . Το  $S_t$  δηλώνει το επίπεδο, η  $F_t$  είναι η πρόβλεψη που πραγματοποιείται και  $t$  είναι χρονική περίοδος. Η παράμετρος  $a$  αποτελεί τον συντελεστή εξομάλυνσης της μεθόδου και λαμβάνει τιμές στο διάστημα  $[0, 1]$ .

Από τις παραπάνω εξισώσεις βλέπουμε πως σε κάθε πρόβλεψη υπολογίζουμε το σφάλμα, διότι με αυτό κρατάμε την τιμή της πρόβλεψης αρκετά κοντά στο επίπεδο της πραγματικής χρονοσειράς. Αυτό βέβαια προϋποθέτει τον υπολογισμό της τιμής του επιπέδου από την πραγματική χρονοσειρά για την προηγούμενη χρονική στιγμή.

### **Αρχικό Επίπεδο**

Βέβαια από τα παραπάνω προκύπτει το εξής ερώτημα σχετικά με την λειτουργία του μοντέλου. Αυτό σχετίζεται με το τι θα γίνει με την πρόβλεψη για την πρώτη χρονική περίοδο για την οποία δεν υπάρχουν διαθέσιμα ιστορικά δεδομένα. Η απάντηση που δίνεται είναι ότι η πρώτη πρόβλεψη θα πάρει τη τιμή του αρχικού επιπέδου. Επομένως θα πρέπει να δοθεί ιδιαίτερη προσοχή στην σωστή επιλογή του αρχικού επιπέδου του μοντέλου πρόβλεψης για την εξαγωγή των προβλέψεων με όσο το δυνατόν μεγαλύτερη ακρίβεια.

Μερικές συνηθισμένες μεθοδολογίες για την πρώτη τιμή του επιπέδου ( $S_0$ ) της απλής εκθετικής εξομάλυνσης είναι:

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

Η επιλογή του αρχικού επιπέδου για την εφαρμογή του μοντέλου, αφήνεται στον ερευνητή και εξαρτάται από τα ποιοτικά χαρακτηριστικά της εκάστοτε χρονοσειράς. Είναι λογικό πως αν το αρχικό επίπεδο δεν ανταποκρίνεται στα δεδομένα, τότε θα έχει αρνητική επίδραση στις τιμές των δεδομένων. Επίσης, το αρχικό επίπεδο επηρεάζει ακολούθως την επιλογή του συντελεστή εξομάλυνσης.

### **Συντελεστής Εξομάλυνσης**

Ο βέλτιστος συντελεστής εξομάλυνσης καθορίζεται από δύο παράγοντες. Ο πρώτος παράγοντας είναι το ποσοστό θορύβου στην χρονοσειρά. Αν υπάρχει, δηλαδή, μεγάλο ποσοστό θορύβου, τότε θα πρέπει να επιλεγεί μικρής τιμής για τον συντελεστή εξομάλυνσης για να αποφύγουμε την υπερβολική αντίδραση στο θόρυβο. Ο δεύτερος παράγοντας είναι η σταθερότητα του μέσου όρου της χρονοσειράς, Όσο πιο πολύ μεταβάλλεται ο μέσος όρος, τόσο μεγαλύτερη πρέπει να είναι η τιμή του

συντελεστή εξομάλυνσης, ώστε οι προβλέψεις να παρακολουθούν τις μεταβολές που παρουσιάζουν τα δεδομένα.

Οι τιμές που λαμβάνει ο συντελεστής εξομάλυνσης βρίσκονται στο διάστημα  $[0, 1]$  και γενικά το κριτήριο προσδιορισμού είναι η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE). Βέβαια η τιμή της παραμέτρου μπορεί να αλλάξει δεδομένου πως μπορεί να επιλεχθούν άλλα κριτήρια για την επιλογή του συντελεστή, όπως είναι η ελαχιστοποίηση κάποιου άλλου σφάλματος. Στην εφαρμογή χρησιμοποιήθηκε για την εύρεση τόσο του αρχικού επιπέδου, όσο και για την εύρεση του συντελεστή εξομάλυνσης, το κριτήριο της ελαχιστοποίησης του αθροίσματος του μέσου τετραγωνικού σφάλματος (SSE), του οποίου ο τύπος είναι:

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

Η εύρεση της σταθεράς εξομάλυνσης βρίσκεται με αλγοριθμικό τρόπο μέσω της γραμμικής αναζήτησης η οποία ελαχιστοποιεί το σφάλμα, είτε με κάποια άλλη βελτιστοποίηση. Το μειονέκτημα της γραμμικής αναζήτησης είναι ο χρόνος υπολογισμού ειδικά αν υπάρχουν διαθέσιμα πολλά ιστορικά δεδομένα.

Είναι σημαντικό να εξετάσουμε το τι ακριβώς συμβαίνει με τις ακραίες τιμές του συντελεστή εξομάλυνσης. Όταν η τιμή του συντελεστή είναι κοντά στη μονάδα, τότε το σφάλμα της τελευταίας περιόδου διαδραματίζει σημαντικό ρόλο στον υπολογισμό της επόμενης πρόβλεψης, ενώ στην αντίθετη περίπτωση η νέα πρόβλεψη θα μεταβληθεί σε πολύ μικρό βαθμό. Στην ακραία περίπτωση όπου  $a = 1$ , το μοντέλο ταυτίζεται με την απλοϊκή μέθοδο (Naive), ενώ όταν  $a = 0$  τότε κάθε πρόβλεψη είναι ίση με το αρχικό επίπεδο.

### **3.6.3.2 Μοντέλο γραμμικής τάσης (Holt Exponential Smoothing)**

Το μοντέλο εξομάλυνσης για γραμμική τάση είναι στην ουσία μία επέκταση του μοντέλου της απλής εκθετικής εξομάλυνσης η οποία μπορεί επιπρόσθετα να διαχειριστεί την συνιστώσα της τάσης που συχνά συναντάμε στα επιχειρησιακά δεδομένα. Το μοντέλο αυτό προτάθηκε το 1957 από τον Holt και για το λόγο αυτό η μέθοδος διατηρεί μέχρι σήμερα το όνομά του. Η μέθοδος περιγράφεται μαθηματικά από τις εξής εξισώσεις:

$$e_t = Y_t - F_t$$

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

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

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

Όπου  $e_t$  είναι το σφάλμα της πρόβλεψης το οποίο είναι η απόκλιση της πραγματικής τιμής από την πρόβλεψη για την ίδια χρονική περίοδο  $t$ . Το  $S_t$ , είναι το επίπεδο της χρονοσειράς και  $T_t$  η τάση. Η  $F_{t+m}$  είναι η πρόβλεψη που πραγματοποιείται στο τέλος της περιόδου  $t$  και αναφέρεται σε  $m$  περιόδους μπροστά,  $t$  η χρονική περίοδος και  $m$  ο χρονικός ορίζοντας της πρόβλεψης. Η παράμετρος  $\alpha$  αποτελεί τον συντελεστή εξομάλυνσης για το επίπεδο, ενώ η παράμετρος  $\beta$  είναι ο συντελεστής εξομάλυνσης της τάσης, όπου και οι δύο συντελεστές παίρνουν τιμές στο διάστημα  $[0, 1]$ .

### **Αρχικό Επίπεδο και Αρχική Τάση**

Όπως είναι εμφανές από την αναδρομικότητα του μοντέλου, είναι αναγκαίο να οριστεί ένα αρχικό επίπεδο ( $S_0$ ) και μία αρχική τάση ( $T_0$ ), ώστε να μπορέσει να εφαρμοστεί το μοντέλο πρόβλεψης αρχικά και κατόπιν, των ζητούμενων σημειακών προβλέψεων για δεδομένο χρονικό ορίζοντα. Το αρχικό επίπεδο ( $S_0$ ) υπολογίζεται όπως και στην απλή εκθετική εξομάλυνση. Σαν αρχική τάση συνήθως χρησιμοποιείται:

- Η διαφορά της δεύτερης και πρώτης παρατήρησης ( $Y_2 - Y_1$ )
- Η διαφορά της νιοστής και πρώτης παρατήρησης διαιρεμένης με  $n - 1$  (π.χ.  $(Y_4 - Y_1)/3$ )
- Η σταθερά της κλίσης από το μοντέλο της απλής γραμμικής παλινδρόμησης

Και εδώ θα πρέπει να τονιστεί πως απαιτείται προσεκτική επιλογή τόσο του αρχικού επιπέδου όσο και της αρχικής τάσης, αφού όπως είναι λογικό οι τιμές αυτές επηρεάζουν το μοντέλο της πρόβλεψης και κατ' επέκταση τις σημειακές προβλέψεις. Η τελική επιλογή γίνεται πάντα σύμφωνα με τα ποιοτικά χαρακτηριστικά των δεδομένων της χρονοσειράς που θέλουμε να μελετήσουμε.

### **Συντελεστές Εξομάλυνσης**

Όσον αφορά την εύρεση των τιμών των συντελεστών εξομάλυνσης, όπως στο μοντέλο της απλής εκθετικής εξομάλυνσης, οι τιμές που παίρνουν είναι στο διάστημα  $[0, 1]$  και ποικίλουν ανάλογα με το κριτήριο επιλογής που θα χρησιμοποιηθεί. Η πιο δημοφιλής διαδικασία που χρησιμοποιείται είναι η γραμμική αναζήτηση των συντελεστών που ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα ( $MSE$ ). Συνήθως η βέλτιστη τιμή του συντελεστή  $\alpha$  για το επίπεδο είναι μεγαλύτερη από τη τιμή του συντελεστή  $\beta$  για την τάση. Ο λόγος είναι ότι η τιμή της τάσης είναι συνήθως πολύ μικρότερη από την τιμή του επιπέδου για κάθε περίοδο., καθώς αν προστεθεί το ίδιο ποσοστό σφάλματος και στη τάση και στο επίπεδο τότε οι προβλέψεις που προκύπτουν μπορεί να είναι ασταθείς. Για την ανάπτυξη της εφαρμογής σαν κριτήριο υπολογισμού των συντελεστών έχει χρησιμοποιηθεί το κριτήριο ελαχιστοποίησης του αθροίσματος των τετραγωνικών σφαλμάτων ( $SSE$ )

### 3.6.3.3 Μοντέλο μη γραμμικής τάσης (*Damped Trend Method*)

Το μοντέλο γραμμικής τάσης που περιεγράφηκε παραπάνω, έχει μία τάση για θετική προκατάληψη ιδιαίτερα στις περιπτώσεις μεσοπρόθεσμων και μακροπρόθεσμων προβλέψεων. Επίσης το μοντέλο μπορεί να μεταβληθεί κατάλληλα ώστε να προσαρμόζεται και σε μη γραμμικές τάσεις. Αυτό μπορεί να επιτευχθεί με την προσθήκη μίας παραμέτρου η οποία ελέγχει το ρυθμό αύξησης των τιμών των προβλέψεων. Αυτή η παράμετρος ονομάζεται παράμετρος διόρθωσης τάσης (*trend – modification parameter*), συμβολίζεται με  $\varphi$  και περιεγράφηκε από τους Gardner και McKenzie το 1985. Οι εξισώσεις που περιγράφουν μαθηματικά το μοντέλο μη γραμμικής τάσης είναι:

$$e_t = Y_t - F_t$$
$$S_t = S_{t-1} + \varphi \cdot T_{t-1} + a \cdot e_t$$
$$T_t = \varphi \cdot T_{t-1} + \alpha \cdot \beta \cdot e_t$$
$$F_{t+m} = S_t + \sum_{i=1}^m \varphi^i \cdot T_t$$

Όπως βλέπουμε οι παραπάνω εξισώσεις είναι παρόμοιες με αυτές του γραμμικού μοντέλου εκτός από την τελευταία, όπου αντί να υπολογίζεται μία γραμμική αύξηση της τάσης μέσω του συντελεστή  $m$ , γίνεται ένας μη γραμμικός υπολογισμός αυτής, γεγονός που οφείλεται στην παράμετρο  $\varphi$ . Η συγκεκριμένη παράμετρος δεν έχει κάποιο άνω όριο και μπορεί να λάβει οποιαδήποτε τιμή άνω του μηδενός. Βέβαια είναι πολύ σημαντική η επιβολή άνω και κάτω ορίων ανάλογα με την εκάστοτε περίπτωση. Το μοντέλο της μη γραμμικής τάσης, ανάλογα με την τιμή που παίρνει η παράμετρος εξομάλυνσης, μπορεί να πάρει τις μορφές:

- Αν  $\varphi = 0$ , τότε προκύπτει το μοντέλο της απλής εκθετικής εξομάλυνσης, αφού η τάση δεν έχει πλέον καμία επίδραση στον καθορισμό των στατιστικών σημειακών προβλέψεων.
- Αν  $0 < \varphi < 1$ , τότε προκύπτει το μοντέλο της φθίνουσας τάσης, το οποίο αναλύουμε σε αυτή την παράγραφο και είναι κατάλληλο για την παραγωγή μεσοπρόθεσμων και μακροπρόθεσμων προβλέψεων.
- Αν  $\varphi = 1$ , τότε προκύπτει το μοντέλο της γραμμικής τάσης, αφού στην εξίσωση υπολογισμού της πρόβλεψης, τη θέση του αθροίσματος παίρνει το γινόμενο της μεταβλητής χρονικού ορίζοντα  $m$  και της προηγούμενης τάσης  $T_t$ .
- Αν  $\varphi > 1$ , τότε προκύπτει το μοντέλο της εκθετικής τάσης, το οποίο είναι κατάλληλο στις περιπτώσεις όπου το ζητούμενο είναι η πρόβλεψη ζήτησης στην αρχή του κύκλου ζωής ενός προϊόντος ή μιας υπηρεσίας. Αυτό συμβαίνει, διότι το συγκεκριμένο μοντέλο χαρακτηρίζεται από μεγάλη θετική προκατάληψη.

Γενικά όσον αφορά την εύρεση των συντελεστών εξομάλυνσης και της παραμέτρου διόρθωσης τάσης, το συνηθέστερο κριτήριο που χρησιμοποιείται είναι η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος. Είναι επόμενο πως όσο αυξάνεται το πλήθος των παραμέτρων, τόσο αυξάνεται και η υπολογιστική πολυπλοκότητα του προβλήματος και συνεπώς και του χρόνου υπολογισμού των παραμέτρων. Αναφορικά με την επιλογή του αρχικού επιπέδου και της αρχικής τάσης ισχύουν ακριβώς τα ίδια που περιγράφηκαν και στην περίπτωση του μοντέλου γραμμικής τάσης.

Όμως σε περίπτωση λανθασμένης βελτιστοποίησης του συντελεστή  $\varphi$ , υπάρχει το σενάριο της παραγωγής προκατειλημμένων προβλέψεων. Για αυτό το λόγο, το εύρος τιμών του συντελεστή  $\varphi$  το περιορίζουμε στο διάστημα  $[0, 1]$ . Μέσα από εμπειρικές έρευνες έχει αποδειχθεί πως η χρήση του μοντέλου μη γραμμικής τάσης δίνει μεγαλύτερη ακρίβεια προβλέψεων, από το μοντέλο γραμμικής τάσης. Σε γενικές γραμμές, το μοντέλο μη γραμμικής τάσης δίνει καλύτερα αποτελέσματα σε περιπτώσεις όπου είναι αδύνατη η επιλογή κάποιου μοντέλου για περιγραφή της χρονοσειράς ή σε περιπτώσεις όπου απαιτείται ένα πλήρως αυτοματοποιημένο σύστημα για την παραγωγή προβλέψεων μεγάλου πλήθους χρονοσειρών. Επίσης, όσο μεγαλύτερος είναι ο ορίζοντας πρόβλεψης τόσο πλεονεκτεί το μοντέλο μη γραμμικής τάσης.

#### **3.6.3.4 Error-Trend-Seasonal ή Exponential Smoothing method – ETS**

Είναι ένα state space model, το οποίο δημιουργήθηκε από τον Hyndman και τους συνεργάτες του (2008). Το State Space Model (SSM) αναφέρεται σε μια κατηγορία πιθανολογικών γραφικών μοντέλων, που περιγράφει την πιθανολογική εξάρτηση μεταξύ της μεταβλητής λανθάνουσας κατάστασης και την παρατηρούμενη μέτρηση. Η μέτρηση μπορεί να είναι είτε συνεχής ή διακριτή. Ένα στατιστικό μοντέλο είναι μια στοχαστική (ή τυχαία) διαδικασία δημιουργίας δεδομένων, όπου μπορεί να παραχθεί μία ολόκληρη κατανομή προβλέψεων. Κάθε μοντέλο αποτελείται από μια εξίσωση που περιγράφει τα δεδομένα και μερικές εξισώσεις μετάβασης που δείχνουν πως τα βασικά στοιχεία της χρονοσειράς (επίπεδο, τάση, εποχιακότητα) αλλάζουν με την πάροδο του χρόνου. Ως εκ τούτου, αυτά αναφέρονται ως «state space models».

Για κάθε μέθοδο υπάρχουν δύο μοντέλα, ένα με προσθετικά σφάλματα (Additive errors) και ένα με πολλαπλασιαστικά σφάλματα (Multiplicative errors). Οι σημειακές προβλέψεις που παράγονται από τα μοντέλα είναι ίδια αν είναι ίδιες οι τιμές των παραμέτρων εξομάλυνσης. Θα μπορούν, ωστόσο, να δημιουργήσουν διαφορετικά χρονικά διαστήματα πρόβλεψης. Ουσιαστικά η μέθοδος ETS διαθέτει τρία βασικά στοιχεία για την εξαγωγή προβλέψεων, το είδος των σφαλμάτων, το είδος της τάσης και το είδος της εποχιακότητας και επιλέγονται τα είδη αυτά με αυτόν τον τρόπο *ETS (Error, Trend, Seasonal)*. Το στοιχείο των σφαλμάτων και το στοιχείο της εποχιακότητας μπορεί να είναι είτε προσθετικά (Additive-A), είτε πολλαπλασιαστικά (Multiplicative-M). Επίσης το στοιχείο της εποχιακότητας μπορεί να μην λαμβάνεται

υπόψη (None-N). Αν ληφθεί υπόψη και το στοιχείο της τάσης, τότε μπορεί να είναι ένα από τα παρακάτω είδη:

- Δεν λαμβάνεται υπόψη το στοιχείο της τάσης-None ( $N$ )
- Προσθετική-Additive ( $A$ )
- Προσθετική φθίνουσα τάση-Additive damped ( $A_d$ )
- Πολλαπλασιαστική τάση-Multiplicative ( $M$ )
- Πολλαπλασιαστική φθίνουσα τάση-Multiplicative Damped ( $M_d$ )

Επομένως οι περιπτώσεις ανάλογα με αυτές τις συνιστώσες χωρίζονται αρχικά με βάση το είδος των σφαλμάτων, δηλαδή σε προσθετικά ή πολλαπλασιαστικά. Κατόπιν ο διαχωρισμός γίνεται με βάση τις περιπτώσεις των άλλων δύο συνιστωσών, της τάσης και της εποχιακότητας, όπως φαίνεται στον παρακάτω πίνακα:

<b>Seasonal</b>			
<b>Trend</b>	<b><math>N</math> (None)</b>	<b><math>A</math> (Additive)</b>	<b><math>M</math> (Multiplicative)</b>
$N$ (None)	(N, N)	(N, A)	(N, M)
$A$ (Additive)	(A, N)	(A, A)	(A, M)
$A_d$ (Additive Damped)	( $A_d$ , N)	( $A_d$ , A)	( $A_d$ , M)
$M$ (Multiplicative)	(M, N)	(M, A)	(M, M)
$M_d$ (Multiplicative Damped)	( $M_d$ , N)	( $M_d$ , A)	( $M_d$ , M)

Άρα υπάρχουν διαθέσιμα 30 state models, 15 με προσθετικά σφάλματα και 15 με πολλαπλασιαστικά.

### **Απλή Εκθετική Εξομάλυνση με Προσθετικά Σφάλματα-ETS(A, N, N)**

Όπως έχουμε αναφέρει και παραπάνω για την απλή γραμμική παλινδρόμηση ισχύουν οι τύποι:

$$e_t = Y_t - F_t$$

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

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

Όπου επιλύοντας τις παραπάνω εξισώσεις η εξίσωση της πραγματικής τιμής γράφεται ως εξής:

$$Y_t = S_{t-1} + e_t$$

Για να γίνει αυτό σε ένα state model, το μόνο που χρειάζεται είναι να καθοριστεί η κατανομή πιθανοτήτων για  $e_t$ . Στο μοντέλο με προσθετικά σφάλματα, υποθέτουμε ότι τα σφάλματα πρόβλεψης ενός βήματος (one-step forecast errors)  $e_t$  διανέμονται κανονικά με μέση τιμή 0 και διακύμανση  $\sigma^2$ . Μία συντομογραφία για αυτό είναι το  $e_t = \varepsilon_t \sim NID(0, \sigma^2)$ , όπου NID σημαίνει πως είναι κατανεμημένα με κανονικό και

ανεξάρτητο τρόπο (Normally and Independently Distributed). Επομένως τώρα οι εξισώσεις γράφονται:

$$Y_t = S_{t-1} + \varepsilon_t$$

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

Η πρώτη εξίσωση αναφέρεται ως εξίσωση μέτρησης και η δεύτερη ως εξίσωση κατάστασης ή μετάβασης. Αυτές οι δύο εξισώσεις, μαζί με την στατιστική κατανομή των σφαλμάτων, σχηματίζουν ένα πλήρως καθορισμένο στατιστικό μοντέλο. Συγκεκριμένα, αποτελούν ένα state model, οι οποίες είναι υποκείμενες στην απλή εκθετική εξομάλυνση.

Η εξίσωση μέτρησης δείχνει τη σχέση μεταξύ των παρατηρήσεων και των μη παρατηρούμενων καταστάσεων. Σε αυτή την περίπτωση η παρατήρηση  $Y_t$  είναι μια γραμμική συνάρτηση του επιπέδου  $S_t$  και το τυχαίο σφάλμα  $\varepsilon_t$ , το απρόβλεπτο μέρος του  $Y_t$ . Για τα άλλα μοντέλα, αυτή η σχέση μπορεί να είναι μη γραμμική.

Η εξίσωση της μετάβασης δείχνει την εξέλιξη της κατάστασης μέσα στο χρόνο. Η επίδραση της παραμέτρου εξομάλυνσης  $a$  είναι η ίδια όπως και για τις μεθόδους που συζητήθηκαν προηγουμένως. Για παράδειγμα η παράμετρος  $a$  ρυθμίζει τον βαθμό της αλλαγής στα διαδοχικά επίπεδα. Όσο υψηλότερη είναι η τιμή του  $a$ , τόσο πιο γρήγορες θα είναι οι μεταβολές του επιπέδου.

### **Απλή Εκθετική Εξομάλυνση με Πολλαπλασιαστικά Σφάλματα-ETS (M, N, N)**

Με τον ίδιο τρόπο με πριν μπορούμε να εξάγουμε το μοντέλο με χρήση πολλαπλασιαστικών σφαλμάτων. Αυτό γίνεται γράφοντας τα τυχαία σφάλματα σαν σχετικά σφάλματα:

$$\varepsilon_t = \frac{Y_t - F_t}{F_t}$$

Όπου  $\varepsilon_t \sim NID(0, \sigma^2)$ , και αντικαθιστώντας  $F_t = S_{t-1}$  μας δίνει ότι:

$$Y_t = S_{t-1} + S_{t-1} \cdot \varepsilon_t \text{ και } e_t = Y_t - F_t = S_{t-1} \cdot \varepsilon_t$$

Τότε μπορούμε να γράψουμε την πολλαπλασιαστική μορφή του μοντέλου ως εξής:

$$Y_t = S_{t-1} \cdot (1 + \varepsilon_t)$$

$$S_t = S_{t-1} \cdot (1 + a \cdot \varepsilon_t)$$

### **Προσδιορισμός παραμέτρων**

Μια εναλλακτική λύση για τον υπολογισμό των παραμέτρων, εκτός από την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων, είναι να μεγιστοποιήσουμε την «πιθανότητα». Με τον όρο αυτό εννοούμε συγκεκριμένα την πιθανότητα των στοιχείων που προκύπτουν από το καθορισμένο μοντέλο. Έτσι, μία μεγάλη πιθανότητα συνδέεται με ένα καλό μοντέλο. Για ένα μοντέλο προσθετικών

σφαλμάτων, μεγιστοποιώντας την πιθανότητα παίρνουμε τα ίδια αποτελέσματα με την ελαχιστοποίηση του αθροίσματος των τετραγώνων των σφαλμάτων. Ωστόσο, θα διαφέρει αν χρησιμοποιηθεί κάποιο μοντέλο πολλαπλασιαστικών σφαλμάτων. Άρα οι παράμετροι εξομάλυνσης  $\alpha$ ,  $\beta$ ,  $\gamma$  και  $\phi$ , και οι αρχικές καταστάσεις  $S_t$  και  $T_t$  υπολογίζονται με βάση τη μεγιστοποίηση της πιθανότητας. Οι παράμετροι εξομάλυνσης παίρνουν τιμές στο διάστημα  $[0,1]$ .

### **Επιλογή μοντέλου**

Η επιλογή του κατάλληλου μοντέλου γίνεται επιλέγοντας μέσα από τρία κριτήρια. Το πρώτο κριτήριο είναι το κριτήριο πληροφορίας του Akaike (Akaike's Information Criteria) και δίνεται από τον τύπο:

$$AIC = -2 \log(L) + 2 \cdot k$$

Όπου το  $L$  είναι η πιθανότητα του μοντέλου και  $k$  ο συνολικός αριθμός των παραμέτρων και των αρχικών καταστάσεων που θα πρέπει να εκτιμηθούν. Το δεύτερο είναι η διορθωμένη μορφή του προηγούμενου κριτηρίου και υπολογίζεται ως εξής:

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

και τρίτο είναι το κριτήριο πληροφορίας του Bayes (Bayes Information Criterion):

$$BIC = AIC + k[\log(T) - 2]$$

Θα πρέπει να τονιστεί πως στην εφαρμογή που υλοποιήθηκε, η επιλογή του κατάλληλου μοντέλου γίνεται αυτόματα χρησιμοποιώντας το δεύτερο κριτήριο.

### **3.6.4 Απλή Γραμμική Παλινδρόμηση**

Η παλινδρόμηση εμπεριέχει διαδικασίες και τεχνικές μοντελοποίησης και ανάλυσης διαφόρων μεταβλητών, όπου το ζητούμενο είναι η εύρεση συσχετίσεων μεταξύ μίας εξαρτημένης και μίας ή και περισσότερων ανεξάρτητων μεταβλητών. Ουσιαστικά με τη μέθοδο αυτή μελετάμε τον τρόπο συμπεριφοράς της εξαρτημένης μεταβλητής, όταν μεταβληθεί η τιμή μίας ανεξάρτητης μεταβλητής όταν οι υπόλοιπες ανεξάρτητες μεταβλητές παραμένουν σταθερές.

Επιπρόσθετα, με την χρήση της ανάλυσης παλινδρόμησης μπορούμε να κάνουμε μία εκτίμηση της τιμής της εξαρτημένης μεταβλητής, δεδομένου των τιμών των ανεξάρτητων μεταβλητών. Πρόκειται ουσιαστικά για την μέση τιμή της εξαρτημένης μεταβλητής, διατηρώντας σταθερές τις ανεξάρτητες μεταβλητές. Η πρόβλεψη της τιμής της εξαρτημένης μεταβλητής είναι μία μαθηματική συνάρτηση των ανεξάρτητων μεταβλητών η οποία ονομάζεται εξίσωση της παλινδρόμησης.



Παρόλο που η παλινδρόμηση χρησιμοποιείται σε θέματα πρόβλεψης, ο κυριότερος λόγος χρησιμοποίησης είναι η ανάλυση και η κατανόηση των σχέσεων που μπορεί να έχουν οι ανεξάρτητες και οι εξαρτημένες μεταβλητές. Πιο συγκεκριμένα μέσα από τη παλινδρόμηση μπορούμε να συμπεράνουμε ποιες από τις εξεταζόμενες μεταβλητές συσχετίζονται, αλλά και να βρούμε τη μορφή αυτής της συσχέτισης. Προκειμένου να επιτευχθεί η ανάλυση της παλινδρόμησης, έχουν αναπτυχθεί ποικίλες τεχνικές. Μία από τις πιο διαδεδομένες είναι και η απλή γραμμική παλινδρόμηση.

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

$$\hat{Y}_i = a + b \cdot X_i$$

Η οποία χρησιμοποιείται στην πράξη. Όπου το  $a$  είναι η τεταγμένη του σημείου τομής της ευθείας με τον άξονα των εξαρτημένων μεταβλητών (για  $b = 0$ ) και  $b$  είναι η κλίση της ευθείας, δηλαδή ο ρυθμός μεταβολής του  $\hat{Y}$  ανά μοναδιαία αύξηση του  $X$ .

Στόχος της απλής γραμμικής παλινδρόμησης είναι η εκτίμηση των παραμέτρων  $a$  και  $b$  έτσι ώστε η παραπάνω ευθεία να αποτελεί τη βέλτιστη, δηλαδή να εφαρμόζεται όσο γίνεται καλύτερα στα δεδομένα. Το σφάλμα θεωρείται η κατακόρυφη απόκλιση της παρατήρησης από την ευθεία και δίδεται ως εξής:

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

Όπου η τιμή  $\hat{Y}_i$  αντιπροσωπεύει την εκτιμώμενη τιμή από την ευθεία της παλινδρόμησης και η τιμή  $Y_i$  αντιστοιχεί στην πραγματική παρατήρηση. Σαν βέλτιστη ευθεία παλινδρόμησης, επιλέγεται αυτή για την οποία το άθροισμα των τετραγώνων των σφαλμάτων γίνεται ελάχιστο. Κατά συνέπεια με τη χρήση της μεθόδου αυτής που είναι γνωστή και σαν μέθοδος των ελαχίστων τετραγώνων οι τιμές των συντελεστών  $a$  και  $b$  υπολογίζονται από τις ακόλουθες σχέσεις:

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

Ή ισοδύναμα:

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

$$a = \bar{Y} - b \cdot \bar{X}$$

Όπου  $\bar{X}$  και  $\bar{Y}$  οι μέσες τιμές των διανυσμάτων  $X$  και  $Y$  και  $n$  ο αριθμός των δεδομένων με βάση των οποίων υπολογίζεται η ευθεία παλινδρόμησης.

### 3.6.5 Μοντέλο *Theta*

Η μέθοδος *Theta* υλοποιήθηκε από δύο μέλη της Μονάδας Προβλέψεων και Στρατηγικής το 2000 (Assimakorou και Nikolopoulos, 2000; Νικολόπουλος, 2002). Η μέθοδος αυτή είναι μία μονοδιάστατη μέθοδος πρόβλεψης, η οποία βασίζεται στη μεταβολή των τοπικών καμπυλοτήτων μίας χρονοσειράς μέσα από την παράμετρο  $\theta$  (*theta*) και εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των δεδομένων. Η καινούργια χρονοσειρά που δημιουργείται έχει την ίδια μέση τιμή και κλίση της αρχικής χρονοσειράς, αλλά όχι και τις τοπικές καμπυλότητες και τη διακύμανση. Οι παραγόμενες χρονοσειρές αυτής της μεθόδου ονομάζονται γραμμές *Theta*. Ανάλογα με την τιμή της παραμέτρου  $\theta$  (μικρότερη ή μεγαλύτερη της μονάδας), μπορεί να γίνει καλύτερη προσέγγιση της μακροπρόθεσμης συμπεριφοράς των δεδομένων ή ανάδειξη και τονισμός των βραχυπρόθεσμων χαρακτηριστικών αντίστοιχα.

Η προτεινόμενη μέθοδος αποσυνθέτει (διαχωρίζει) την αρχική χρονοσειρά σε δύο ή περισσότερες γραμμές *Theta*. Η κάθε γραμμή *Theta* προεκτείνεται στο μέλλον ξεχωριστά, με την ίδια ή και με διαφορετικές μεθόδους πρόβλεψης και οι παραγόμενες προβλέψεις συνδυάζονται για να προκύψει η τελική πρόβλεψη. Στον διεθνή διαγωνισμό προβλέψεων M3 χρησιμοποιήθηκε, για την παραγωγή προβλέψεων, ο απλός συνδυασμός δύο γραμμών *Theta*, για  $\theta = 0$  (ευθεία γραμμή) και  $\theta = 2$  (διπλασιασμός των τοπικών καμπυλοτήτων) όπου παρήγαγε πολύ καλά αποτελέσματα, με μικρά σφάλματα προβλεπτικής ακρίβειας, ιδιαίτερα για τις μηνιαίες χρονοσειρές και τα μικροοικονομικά δεδομένα.

Γενικά, έχουν γίνει πολλές προσπάθειες να παραχθούν προβλέψεις βασισμένες απευθείας στην κλασική αποσύνθεση. Όπως έχει ήδη αναφερθεί τα χαρακτηριστικά που απομονώνονται είναι οι συνιστώσες τάσης-κύκλου, εποχιακότητας και τυχαιότητας, όπου η κάθε συνιστώσα προεκτείνεται ξεχωριστά στο μέλλον και οι προβλέψεις που παράγονται συνδυάζονται με ίσα βάρη από όπου προκύπτουν οι τελικές προβλέψεις τις χρονοσειράς. Μία τεχνική, η οποία βασίζεται στην παραπάνω διαδικασία, εφαρμόζει την μέθοδο πρόβλεψης στα αποεποχικοποιημένα δεδομένα, ενώ οι τελικές προβλέψεις προκύπτουν με την επαναχικοποίηση των παραγόμενων προβλέψεων.

Η μέθοδος *Theta* εισήγαγε μία διαφορετική προσέγγιση της αποσύνθεσης, όπου ο διαχωρισμός των αποεποχικοποιημένων δεδομένων γίνεται σε συνιστώσες (γραμμές *Theta*) μακροπρόθεσμης και βραχυπρόθεσμης τάσης. Η πρόκληση αυτής της μεθόδου ήταν να αυξήσει τον βαθμό αξιοποίησης της χρήσιμης πληροφορίας που βρίσκεται κρυμμένη μέσα στα δεδομένα, πριν την επέκταση των δεδομένων στο μέλλον. Ουσιαστικά, η μέθοδος *Theta* λειτουργεί σαν ένας μεγεθυντικός φακός μέσα από τον οποίο οι διακυμάνσεις της χρονοσειράς μεγεθύνονται ή μικραίνουν. Ο γραμμικός συνδυασμός των προβλέψεων των συνιστωσών γίνεται πιο αποδοτικός μέσα από αυτή την διαδικασία.

Τα βήματα που ουσιαστικά περιγράφουν τη μεθοδολογία της κλασικής μεθόδου *Theta* και που χρησιμοποιήθηκαν στον διεθνή διαγωνισμό προβλέψεων M3, είναι τα παρακάτω:

**Βήμα 0: Έλεγχος εποχιακότητας.** Ελέγχεται η κάθε χρονοσειρά για στατιστικά σημαντική εποχιακή συμπεριφορά.

**Βήμα 1: Αποεποχικοποίηση.** Εφόσον η χρονοσειρά έχει σημαντική εποχιακή συμπεριφορά, αποεποχικοποιείται μέσω της κλασικής μεθόδου πολλαπλασιαστικής αποσύνθεσης.

**Βήμα 2: Αποσύνθεση.** Η κάθε χρονοσειρά αποσυντίθεται σε δύο γραμμές *Theta*, την ευθεία γραμμικής παλινδρόμησης, όπου  $\theta = 0$  και τη γραμμή *Theta* με παράμετρο  $\theta = 2$ .

**Βήμα 3: Πρόβλεψη.** Η γραμμή *Theta* με παράμετρο  $\theta = 0$ , που αναπαριστά την ευθεία γραμμικής παλινδρόμησης, προεκτείνεται με το συνηθισμένο τρόπο, ενώ η δεύτερη γραμμή προεκτείνεται μέσω της απλής εκθετικής εξομάλυνσης.

**Βήμα 4: Συνδυασμός.** Οι παραγόμενες προβλέψεις των δύο γραμμών *Theta* συνδυάζονται με ίσα βάρη.

**Βήμα 5: Εποχικοποίηση.** Οι τελικές προβλέψεις εποχικοποιούνται, χρησιμοποιώντας τους δείκτες εποχιακότητας που υπολογίστηκαν στο Βήμα 1.

Το δυνατό σημείο του μοντέλου έγκειται στην αποσύνθεση των αρχικών δεδομένων. Ακολούθως περιγράφεται ο υπολογισμός των δύο γραμμών *Theta*, όπως περιγράφεται από την κλασική μέθοδο *Theta*, όπου η αρχική χρονοσειρά των δεδομένων διαχωρίζεται σε δύο γραμμές *Theta* με  $\theta = 0$  και  $\theta = 2$ . Δεδομένου πως η *Theta Line(0)* ισοδυναμεί με τη ευθεία ελαχίστων τετραγώνων (*LRL*) που περιεγράφηκε στην προηγούμενη παράγραφο, απομένει ο υπολογισμός της *Theta Line(2)*. Έτσι έχουμε:

$$Y_t = \frac{1}{2} \cdot (Y_t^{\theta=1+a} + Y_t^{\theta=1-a}) \stackrel{a=1}{\Leftrightarrow}$$

$$Y_t = \frac{1}{2} \cdot (Y_t^{\theta=0} + Y_t^{\theta=2}) \stackrel{Y_t^{\theta=0} \equiv LRL_t}{\Leftrightarrow}$$

$$Y_t = \frac{1}{2} \cdot (LRL_t + Y_t^{\theta=2}) \Leftrightarrow$$

$$2 \cdot Y_t = LRL_t + Y_t^{\theta=2} \Leftrightarrow$$

$$Y_t^{\theta=2} = 2 \cdot Y_t - LRL_t$$

Η τελευταία εξίσωση οδηγεί σε έναν εναλλακτικό τρόπο υπολογισμού της *Theta Line(2)* αφού η *LRL* μπορεί εύκολα να παραχθεί σύμφωνα με τη θεωρία της παλινδρόμησης. Ο Νικολόπουλος και οι συνεργάτες του παρουσίασαν το 2008 μία απλοποιημένη διαδικασία υπολογισμού των διαφόρων γραμμών *Theta*. Πιο

συγκεκριμένα, ο υπολογισμός κάθε γραμμής *Theta* έγκειται στο άθροισμα της ευθείας γραμμικής παλινδρόμησης με  $\theta$ -φορές το σφάλμα της αρχικής χρονοσειράς από τη γραμμή *Theta Line*(0):

$$Theta\ Line(\theta)_t = Y_t^\theta = LRL_t + \theta \cdot e_t$$

Όπου:

$$e_t = Y_t - LRL_t$$

Και  $Y_t^\theta$  είναι η υπολογισμένη γραμμή *Theta*, η *LRL* δηλώνει την ευθεία γραμμικής παλινδρόμησης και  $Y$  δηλώνει την αρχική χρονοσειρά. Η απόδειξη της απλοποιημένης σχέσης υπολογισμού των γραμμών *Theta* έχει ως εξής:

$$Y_t^\theta = LRL_t + \theta \cdot e_t \Leftrightarrow$$

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

$$Y_t^\theta = \theta \cdot Y_t + LRL_t - \theta \cdot LRL_t \Leftrightarrow$$

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

Η τελευταία εξίσωση είναι ισοδύναμη με την εξίσωση υπολογισμού γραμμών *Theta* που παρουσιάστηκε από τους Hyndman και Billah:

$$Y_t^\theta = \theta \cdot Y_t + a_\theta + b_\theta \cdot (t - 1)$$

### 3.7 Δείκτες Αξιολόγησης Προβλέψεων

Πέρα από την γραφική απεικόνιση των δεδομένων, είναι χρήσιμο να παρέχονται και κάποια αριθμητικά δεδομένα που με βάση αυτά μπορούμε να εξάγουμε συμπεράσματα και να αξιολογήσουμε με βάση τα διαθέσιμα ιστορικά δεδομένα. Επίσης με τον υπολογισμό των βασικών στατιστικών δεικτών μπορούμε να επιλέξουμε τις σωστές μεθοδολογίες και μεθόδους πρόβλεψης. Παρακάτω αναλύονται οι δείκτες που χρησιμοποιήθηκαν για την πραγματοποίηση του πειράματος, αλλά και μερικοί ακόμα για λόγους βιβλιογραφικής πληρότητας.

#### 3.7.1 Βασική Στατιστική Ανάλυση

Η κατηγορία αυτή περιλαμβάνει τους βασικούς στατιστικούς δείκτες οι οποίοι μπορούν να υπολογισθούν άμεσα σε δεδομένη χρονοσειρά  $Y$  μεγέθους  $n$  παρατηρήσεων.

### Μέση τιμή (Average)

Αναφέρεται στον υπολογισμό της μέσης τιμής των παρατηρήσεων. Αυτό βοηθάει στον εντοπισμό του επιπέδου γύρω από το οποίο κινούνται οι πραγματικές τιμές της χρονοσειράς. Η μέση τιμή υπολογίζεται από τον ακόλουθο τύπο:

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

### Μέγιστη και ελάχιστη τιμή (Maximum και Minimum)

Αναφέρονται στις ακραίες τιμές μίας χρονοσειράς και αποτελούν μια εκτίμηση της διακύμανσης των δεδομένων της, αλλά και της τυχαιότητας που μπορεί να περιέχουν.

### Τυπική απόκλιση (Standard Deviation)

Η Τυπική Απόκλιση προσπαθεί να δείξει ποια είναι η μέση απόσταση που απέχει ο κάθε αριθμός από το μέσο όρο του αριθμητικού συνόλου που ανήκει. Η τυπική απόκλιση δίνεται από τον ακόλουθο τύπο:

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

$$\sigma_{\text{πληθυσμού}} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}$$

### Διακύμανση (Variance)

Ορίζεται ως το τετράγωνο της τυπικής απόκλισης και είναι ένας ακόμα δείκτης ο οποίος χρησιμεύει για τον εντοπισμό της διασποράς των παρατηρήσεων γύρω από τη μέση τιμή.

### Συνδιακύμανση

Αν υποθέσουμε πως έχουμε δύο διακριτές τυχαίες μεταβλητές X και Y, ορίζουμε ένα μέτρο σχέσης μεταξύ των δύο περιοχών δεδομένων ώστε να καθορισθεί κατά πόσο οι δύο μεταβλητές μεταβάλλονται ανάλογα (θετική συνδιακύμανση), αντιστρόφως ανάλογα (αρνητική συνδιακύμανση) ή κατά πόσο οι τιμές είναι ασυσχέτιστες (μηδενική συνδιακύμανση). Η συνδιακύμανση υπολογίζεται από τον ακόλουθο τύπο:

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

### Συντελεστής γραμμικής συσχέτισης (Linear Correlation Coefficient)

Ονομάζεται και συντελεστής Pearson και εκφράζει την συγκέντρωση των σημείων ενός διαγράμματος διασποράς γύρω από την γραμμή παλινδρόμησης, δίνοντας ένα μέτρο της γραμμικής συσχέτισης μεταξύ των δύο μεταβλητών. Ο τύπος του συντελεστή γραμμικής συσχέτισης είναι ο ακόλουθος:

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

Ο συντελεστής Pearson παίρνει τιμές στο διάστημα [-1, 1] και αναλόγως με την τιμή έχουμε:

- Αν  $r = \pm 1$ , υπάρχει τέλεια γραμμική συσχέτιση.
- Αν  $-0,3 < r < 0,3$ , δεν υπάρχει γραμμική συσχέτιση.
- Αν  $-0,5 < r \leq -0,3$  ή  $0,3 \leq r < 0,5$ , υπάρχει ασθενής γραμμική συσχέτιση.
- Αν  $-0,7 < r \leq -0,5$  ή  $0,5 \leq r < 0,7$ , υπάρχει μέση γραμμική συσχέτιση.
- Αν  $-0,8 < r \leq -0,7$  ή  $0,7 \leq r < 0,8$ , υπάρχει ισχυρή γραμμική συσχέτιση.
- Αν  $-1 < r \leq -0,8$  ή  $0,8 \leq r < 1$ , Υπάρχει πολύ ισχυρή γραμμική συσχέτιση.

### Συντελεστής αυτοσυσχέτισης (Autocorrelation Coefficient)

Ο συντελεστής αυτός εκφράζει τη συσχέτιση μεταξύ παρατηρήσεων της ίδιας μεταβλητής με χρονική υστέρηση  $k$  περιόδους και παίρνει τιμές στο διάστημα [0, 1]. Αν η τιμή του συντελεστή τείνει στο μηδέν, τότε υπάρχει μηδενική συσχέτιση των παρατηρήσεων χρονικής υστέρησης  $k$ , ενώ αν είναι κοντά στην μονάδα, έχει μεγάλη συσχέτιση. Ο τύπος υπολογισμού του συντελεστή είναι ο εξής:

$$ACF_k = \frac{\sum_{i=1+k}^n [(Y_i - \bar{Y}) \cdot (Y_{i-k} - \bar{Y})]}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

### Συντελεστής μεταβλητότητας (Coefficient of Variation)

Πρόκειται για ένα κανονικοποιημένο μέτρο της διασποράς των παρατηρήσεων ενός δείγματος ή ενός πληθυσμού. Το πλεονέκτημά του, σε σχέση με την τυπική απόκλιση είναι ότι είναι απαλλαγμένο από την επίδραση του επιπέδου των παρατηρήσεων. Ο υπολογισμός μπορεί να γίνει χρησιμοποιώντας τον ακόλουθο τύπο:

$$c_v = \frac{\sigma}{\bar{Y}} \cdot 100 (\%)$$

### Μέση τιμή διαστήματος μεταξύ ζητήσεων (Intermittent Demand Interval)

Ο δείκτης αυτός εφαρμόζεται σε χρονοσειρές διακοπτόμενης ζήτησης και εκφράζει τη μέση τιμή των αποστάσεων διαδοχικών περιόδων με μη μηδενική τιμή. Αν η τιμή του δείκτη προκύψει ίση με τη μονάδα, τότε αναφερόμαστε σε χρονοσειρά συνεχούς ζήτησης, ενώ όταν πρόκειται για χρονοσειρά διακοπτόμενης ζήτησης παίρνει τιμές μεγαλύτερης της μονάδας. Μεγάλη τιμή του δείκτη συνεπάγεται και μεγάλα, κατά μέσο όρο, μεσοδιαστήματα μεταξύ των μη μηδενικών παρατηρήσεων.

### 3.7.2 Στατιστική ανάλυση ακρίβειας προβλέψεων

Σε αυτή την κατηγορία, σε αντίθεση με τη βασική στατιστική ανάλυση όπου απαιτείται μονάχα η πραγματική χρονοσειρά των παρατηρήσεων, ανήκουν οι στατιστικοί δείκτες που, για να υπολογιστούν, απαιτείται και μία σειρά πρόβλεψης, που προκύπτει από εφαρμογή κάποιας μεθόδου πρόβλεψης επί της πραγματικής χρονοσειράς. Η συγκεκριμένη κατηγορία έχει μεγάλη σημασία στον κλάδο των προβλέψεων καθώς αποτελεί ένα τρόπο αξιολόγησης των μεθόδων.

Οι στατιστικοί δείκτες της κατηγορίας αυτής μπορούν να υπολογιστούν, αφού πρώτα ορίσουμε το σφάλμα, δηλαδή τη διαφορά μεταξύ πραγματικής τιμής και πρόβλεψης για μία περίοδο:

$$e_i = Y_i - F_i$$

Είναι προφανές πως η τιμή του σφάλματος μπορεί να υπολογισθεί μόνο για τις τιμές όπου έχουμε υπολογίσει μοντέλο πρόβλεψης (in-sample error). Αφού όμως παρέλθει το απαραίτητο χρονικό διάστημα και αποκτήσουμε τα δεδομένα που χρειαζόμαστε, μπορούμε να υπολογίσουμε τόσο το σφάλμα του μοντέλου πρόβλεψης, όσο και το πραγματικό σφάλμα (out-of-sample error). Οι παρακάτω στατιστικοί δείκτες είναι ορισμένοι έτσι ώστε να αναφέρονται στο σφάλμα του μοντέλου πρόβλεψης για συγκεκριμένες περιόδους, αλλά με κατάλληλες αλλαγές στις περιόδους, μπορεί να εκφράσουν την ακρίβεια πρόβλεψης και πέρα των συγκεκριμένων περιόδων.

#### Μέσο σφάλμα (Mean Error)

Στην ουσία αποτελεί ένα απλό προσημασμένο μέσο όρο των σφαλμάτων και χρησιμοποιείται ως ένα μέτρο εκτίμησης της συστηματικότητας του σφάλματος. Πιο πρακτικά αν ο συντελεστής είναι κοντά στο μηδέν, τόσο τα σφάλματα είναι τυχαία και όχι συστηματικά. Αν παίρνει αρνητικές τιμές σημαίνει αισιοδοξία στις προβλέψεις, ενώ αν παίρνει θετικές τιμές δηλώνει απαισιοδοξία. Ο δείκτης μπορεί να υπολογιστεί με τον παρακάτω τύπο:

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

#### Μέσο απόλυτο σφάλμα (Mean Absolute Error)

Αποτελεί ένα μέτρο της ακρίβειας της πρόβλεψης έναντι των πραγματικών τιμών. Χρησιμοποιείται ως ένα μέτρο της αστοχίας της πρόβλεψης, χωρίς να δίνεται έμφαση στην κατεύθυνση της πρόβλεψης. Όσο πιο μεγάλη η τιμή του δείκτη, τόσο πιο μικρή η ακρίβεια της μεθόδου πρόβλεψης που εφαρμόστηκε. Υπολογίζεται ως εξής:

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

### Μέσο τετραγωνικό σφάλμα (Mean Squared Error)

Είναι ένα μέτρο ακρίβειας της πρόβλεψης, το οποίο δίνει πολύ μεγαλύτερο βάρος στα μεγάλα σφάλματα και μικρότερο βάρος στα μικρά. Χρησιμοποιείται συνήθως για τον υπολογισμό των βέλτιστων παραμέτρων εξομάλυνσης. Δίνεται από τον ακόλουθο τύπο:

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

### Ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error)

Προκύπτει από τον υπολογισμό της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος. Ο τύπος είναι ο ακόλουθος:

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

### Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error)

Αρκετές φορές είναι πιο χρήσιμος ο υπολογισμός των σφαλμάτων πρόβλεψης σε καθαρά ποσοστιαία μορφή. Αυτό γίνεται διότι διευκολύνει τη σύγκριση μίας μεθόδου πρόβλεψης, όταν εφαρμόζεται σε διαφορετικές χρονοσειρές. Το μέσο απόλυτο σφάλμα είναι εκφρασμένο επί τοις εκατό και δίνεται από τον ακόλουθο τύπο:

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

### Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (Symmetric Mean Absolute Percentage Error)

Πρόκειται για μία παραλλαγή του MAPE και υπολογίζεται με τον ακόλουθο τύπο:

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

### Μέσο απόλυτο κανονικοποιημένο σφάλμα (Mean Absolute Scaled Error)

Ο συντελεστής αυτός δημιουργήθηκε με σκοπό να αντιμετωπιστούν όσο το δυνατόν οι περιπτώσεις απροσδιοριστίας των δεικτών MAPE και sMAPE, αλλά και να δοθεί η ίδια βαρύτητα και για τα μικρά και για τα μεγάλα σφάλματα. Η παρακάτω εξίσωση μας δίνει την τιμή του σφάλματος:



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

### 3.7.3 Ρυθμός Ανάπτυξης

Ο δείκτης αυτός είναι ένα μέτρο το οποίο εκφράζει την αυξητική ή την φθίνουσα πορεία μίας σειράς δεδομένων μέσα σε ένα συγκεκριμένο χρονικό διάστημα. Η μαθηματική έκφραση του ρυθμού ανάπτυξης έχει ως εξής:

$$Growth Rate = \frac{\frac{1}{ppy} \cdot \sum_{i=n-ppy+1}^n Y_i - \frac{1}{n-ppy} \cdot \sum_{i=1}^{n-ppy} Y_i}{\frac{1}{n-ppy} \cdot \sum_{i=1}^{n-ppy} Y_i} \cdot 100 (\%)$$

Όπου  $Y$  το διάνυσμα των  $n$  παρατηρήσεων και  $ppy$  το πλήθος των περιόδων στο μήκος ενός έτους.



# 4. Περιγραφή της Εφαρμογής

## 4.1 Τεχνολογία Λογισμικού

Η τεχνολογία λογισμικού είναι η βελτίωση της διαδικασίας μεταξύ της ανθρωπότητας και της τεχνολογίας των υλικών (hardware technology). Η τεχνολογία λογισμικού απαρτίζεται από την μηχανική λογισμικού (software engineering), τις επιστημονικές έρευνες λογισμικού και κάθε είδους εξελίξεις, όπως είναι στα οικονομικά, κοινωνικά, ηθικά, βιολογικά, ψυχολογικά και σχεδόν σε όλες οι προσπάθειες των ανθρώπων στη δημιουργία καινοτομιών, οι οποίες συμβάλλουν στην βελτίωση της ανθρώπινης ζωής.

Ένας από τους στόχους της τεχνολογίας λογισμικού είναι να αυξήσει τις αρετές του λογισμικού μέσω ερευνών, την δημιουργία εντολών και προτύπων, και κυρίως να βελτιώνει συνεχώς την αλληλεπίδραση ανάμεσα στους ανθρώπους και στους υπολογιστές. Επίσης άλλος ένας στόχος της τεχνολογίας λογισμικού είναι η εισαγωγή άρτιων και βελτιωμένων τεχνολογιών ασφαλείας.

Βέβαια μεγάλο κομμάτι της τεχνολογίας λογισμικού παίζουν οι γλώσσες προγραμματισμού και για αυτό το λόγο γίνεται συνεχής προσπάθεια βελτίωσής τους. Μια γλώσσα προγραμματισμού είναι μια επίσημη γλώσσα, η οποία έχει κατασκευαστεί και σχεδιαστεί για να επικοινωνεί ο άνθρωπος, μέσω συγκεκριμένων και καθορισμένων εντολών, με μία μηχανή, και πιο συγκεκριμένα έναν υπολογιστή. Οι γλώσσες προγραμματισμού μπορούν να χρησιμοποιηθούν για τη δημιουργία κατάλληλων προγραμμάτων είτε για τον έλεγχο της συμπεριφοράς μιας μηχανής, είτε για την έκφραση αλγορίθμων.

Οι πρώτες γλώσσες προγραμματισμού προηγήθηκαν της εφεύρεσης του ψηφιακού υπολογιστή και χρησιμοποιήθηκαν για να κατευθύνουν τη συμπεριφορά των μηχανών όπως είναι τα μηχανήματα Jacquard και τα αυτόματα πιάνο. Χιλιάδες διαφορετικές γλώσσες προγραμματισμού έχουν δημιουργηθεί, κυρίως στον τομέα της πληροφορικής, και πολλές άλλες ακόμη δημιουργούνται κάθε χρόνο. Πολλές γλώσσες προγραμματισμού απαιτούν τον καθορισμό των υπολογιστικών διαδικασιών σε μια επιτακτική μορφή (δηλαδή, ως μια ακολουθία από λειτουργίες προς εκτέλεση), ενώ άλλες γλώσσες χρησιμοποιούν διαφορετικές μορφές περιγραφής του προγράμματος, όπως τη δηλωτική μορφή (δηλαδή έχει καθοριστεί το επιθυμητό αποτέλεσμα, και όχι ο τρόπος για να επιτευχθεί αυτό).

## **4.2 Λογισμικό της Εφαρμογής**

### **4.2.1 Γλώσσα Προγραμματισμού της Εφαρμογής**

Η εφαρμογή υλοποιήθηκε με την γλώσσα προγραμματισμού *R*. Η *R* είναι μία γλώσσα προγραμματισμού και ένα περιβάλλον λογισμικού για εφαρμογές στατιστικής και γραφικών. Η γλώσσα αυτή συμπεριφέρεται σαν μία εναλλακτική μορφή των παραδοσιακών στατιστικών πακέτων, όπως είναι το SPSS, το SAS και το Stata και λειτουργεί στις πλατφόρμες των Windows, Macintosh, UNIX, και Linux. Αυτό το λογισμικό επιτρέπει στον χρήστη ελεύθερα να συνεισφέρει, να μελετήσει, να αλλάξει και να βελτιώσει το λογισμικό σύμφωνα με τη Γενική Άδεια Δημόσιας Χρήσης GNU (GNU General Public License) του Ιδρύματος Ελεύθερου Λογισμικού (Free Software Foundation). Πρόκειται για μία δωρεάν εφαρμογή της γλώσσας προγραμματισμού *S* που αρχικά δημιουργήθηκε και διανεμήθηκε από τα εργαστήρια της Bell (Bell Laboratories), από τον John Chambers και τους συναδέλφους του. Υπάρχουν ορισμένες σημαντικές διαφορές, αλλά μεγάλο μέρος του κώδικα, που είναι γραμμένο σε γλώσσα *S*, μπορεί να “τρέξει” αυτούσιο στη γλώσσα *R*.

Η *R* παρέχει μια ευρεία ποικιλία στατιστικών (γραμμικών και μη γραμμικών μοντέλων, κλασικούς στατιστικούς ελέγχους, ανάλυση χρονοσειρών, ταξινόμηση, ομαδοποίηση) και γραφικών τεχνικών, και είναι εξαιρετικά επεκτάσιμη. Η γλώσσα *S* είναι συχνά το όχημα της επιλογής για την έρευνα της στατιστικής μεθοδολογίας, και η *R* αποτελεί μια εναλλακτική λύση ανοιχτού κώδικα για τη συμμετοχή στην εν λόγω δραστηριότητα.

Ένα από τα δυνατά σημεία της *R* είναι η ευκολία με την οποία μπορούν να παραχθούν καλά σχεδιασμένα διαγράμματα, συμπεριλαμβανομένων των μαθηματικών συμβόλων και τύπων, όπου χρειάζεται. Μεγάλη προσοχή έχει ληφθεί κατά τη διάρκεια των προεπιλογών για τις δευτερεύουσες επιλογές κατά το σχεδιασμό των γραφικών, αλλά ο χρήστης διατηρεί τον πλήρη έλεγχο.

#### **4.2.1.1 Το περιβάλλον της *R***

Η *R* είναι ένα ενοποιημένο πακέτο λογισμικού για χειρισμό δεδομένων, υπολογισμούς και γραφική απεικόνιση. Η γλώσσα αυτή περιλαμβάνει:

- Ένα αποτελεσματικό χειρισμό δεδομένων και διευκολύνσεις στην αποθήκευσή τους.
- Μία σειρά από λειτουργίες για τον υπολογισμό πινάκων, και ιδίως μητρών.
- Μία μεγάλη, συνεκτική και ολοκληρωμένη συλλογή των απαραίτητων εργαλείων για την ανάλυση των δεδομένων.

- Δυνατότητες γραφικής απεικόνισης για ανάλυση των δεδομένων και η απεικόνισή τους γίνεται είτε στην οθόνη ή σε έντυπη μορφή.
- Μία καλά αναπτυγμένη, απλή και αποτελεσματική γλώσσα προγραμματισμού, η οποία περιλαμβάνει εντολές συνθήκης, βρόχους (loops), αναδρομικές συναρτήσεις καθοριζόμενες από το χρήστη και εντολές εισόδου και εξόδου.

Ο όρος περιβάλλον έχει σκοπό να το χαρακτηρίζει ως ένα πλήρως σχεδιασμένο και συνεκτικό σύστημα, παρά μία σταδιακή προσαύξηση από πολύ συγκεκριμένα και άκαμπτα εργαλεία, όπως συμβαίνει συχνά με άλλα λογισμικά ανάλυσης δεδομένων.

Η R, όπως και η S, έχει σχεδιαστεί γύρω από μία πραγματική γλώσσα του υπολογιστή, και επιτρέπει στους χρήστες να προσθέσουν επιπλέον λειτουργικότητα καθορίζοντας νέες λειτουργίες. Μεγάλο μέρος του συστήματος είναι γραμμένο στην R, η οποία είναι διάλεκτος της S, κάτι το οποίο βοηθάει τους χρήστες να ακολουθούν τις αλγοριθμικές επιλογές που γίνονται. Για εργασίες οι οποίες απαιτούν έντονους υπολογισμούς, κώδικες γραμμένοι σε C, C++ και Fortran μπορούν να συνδεθούν με την R και να κληθούν κατά τη διάρκεια εκτέλεσης του προγράμματος. Οι προχωρημένοι χρήστες μπορούν να γράψουν κώδικα στη γλώσσα C και μπορούν να χειριστούν απευθείας αντικείμενα της R.

Πολλοί χρήστες θεωρούν ότι η R είναι ένα στατιστικό σύστημα. Όμως, είναι προτιμότερο να το θεωρούμε σαν ένα περιβάλλον εντός του οποίου εφαρμόζονται στατιστικές τεχνικές. Επίσης, η R μπορεί να επεκταθεί εύκολα μέσω πακέτων (packages). Υπάρχουν περίπου οχτώ πακέτα που παρέχονται με τη διανομή της R και πολλά άλλα είναι διαθέσιμα μέσα από την οικογένεια CRAN των δικτυακών τόπων που καλύπτουν ένα ευρύτατο φάσμα σύγχρονων στατιστικών.

Η γλώσσα R, έχει τη δική της μορφή της σχετικής τεκμηρίωσης (documentation format), η οποία χρησιμοποιείται για να παρέχει πλήρη τεκμηρίωση, τόσο on-line σε διάφορες μορφές και σε έντυπη μορφή.

#### **4.2.1.2 R-shiny**

Το shiny είναι ένα πακέτο της R ανοιχτού κώδικα που παρέχει ένα κομψό και ισχυρό πλαίσιο διαδικτύου (web framework) για τη δημιουργία διαδικτυακών εφαρμογών χρησιμοποιώντας τη γλώσσα προγραμματισμού R. Το πακέτο shiny καθιστά εξαιρετικά απλό για τους χρήστες της R να μετατρέψουν τις αναλύσεις τους σε διαδραστικές εφαρμογές διαδικτύου (interactive web applications) που ο καθένας μπορεί να χρησιμοποιήσει. Επιτρέπει στους χρήστες της εφαρμογής να επιλέξουν την είσοδο των παραμέτρων με τη χρήση φιλικών προς το χρήστη λειτουργιών, όπως sliders, drop-downs και πεδία κειμένου. Ενσωματώνει πολύ εύκολα οποιοδήποτε αριθμό αποτελεσμάτων, όπως διαγράμματα, πίνακες και οποιαδήποτε σύνοψη θέλουμε να κάνουμε.

Οι εφαρμογές R-shiny έχουν δύο μέρη, το ui.R και το server.R, δηλαδή την διεπαφή του χρήστη με το πρόγραμμα και τη δέσμη ενεργειών που απαιτούνται για την λειτουργία της εφαρμογής, αντίστοιχα. Επίσης μπορεί να περιλαμβάνουν και άλλα δεδομένα ή άλλα αρχεία που μπορεί να απαιτούνται από την εφαρμογή, ανάλογα με την πολυπλοκότητά της. Ουσιαστικά με το ui.R ελέγχεται η διάταξη και η εμφάνιση της εφαρμογής μετατρέποντας των κώδικα από την γλώσσα R σε HTML. Όπως είναι φυσικό δεν χρειάζεται η γνώση των γλωσσών προγραμματισμού HTML, JavaScript ή CSS που συνήθως χρησιμοποιούνται στις διαδικτυακές εφαρμογές. Ωστόσο, μία στοιχειώδης γνώση της γλώσσας HTML μπορεί να κάνει τις εφαρμογές πιο ενδιαφέρουσες και σίγουρα οπτικά καλύτερες.

#### **4.2.2 Τεχνολογία που χρησιμοποιήθηκε για την άντληση των δεδομένων**

Τα δεδομένα που χρησιμοποιούνται και απεικονίζονται στην εφαρμογή αντλούνται από διάφορες ιστοσελίδες οι οποίες διαθέτουν στις βάσεις δεδομένων τους αυτές τις χρονοσειρές. Η άντληση των δεδομένων αυτών γίνεται μέσω διάφορων πακέτων, τα οποία θα εξηγηθούν εκτενώς παρακάτω, τα οποία χρησιμοποιούν την Διασύνδεση Προγράμματος Εφαρμογής, γνωστή και ως API (Application Program Interface).

Το API είναι ένα σύνολο από ρουτίνες, πρωτόκολλα και εργαλεία για την δημιουργία εφαρμογών λογισμικού. Το API καθορίζει πως οι συνιστώσες λογισμικού (software components) θα πρέπει να χρησιμοποιούνται κατά τον προγραμματισμό των συνιστωσών της διασύνδεσης γραφικών (GUI components-Graphic User Interface). Ένα καλό API κάνει ευκολότερη την ανάπτυξη ενός προγράμματος με την παροχή όλων των δομικών στοιχείων και ο προγραμματιστής βάζει τα στοιχεία μαζί.

Εκτός από την πρόσβαση σε βάσεις δεδομένων ή υλικού υπολογιστών, όπως σκληρούς δίσκους ή κάρτες γραφικών, ένα API μπορεί να διευκολύνει το έργο για τον προγραμματισμό των συνιστωσών GUI. Για παράδειγμα, ένα API μπορεί να διευκολύνει την ενσωμάτωση των νέων χαρακτηριστικών σε ήδη υπάρχουσες εφαρμογές (το λεγόμενο «plug-in API»). Ένα API μπορεί επίσης να βοηθήσει με άλλο τρόπο τις διακριτές εφαρμογές με την ανταλλαγή δεδομένων, κάτι το οποίο βοηθάει ώστε να ολοκληρώσει και να ενισχύσει τις λειτουργίες των εφαρμογών.

Στον προγραμματισμό υπολογιστών, ένα API καθορίζει ένα στοιχείο λογισμικού με βάση τις λειτουργίες του, τις εισόδους και εξόδους του. Ο κύριος σκοπός του είναι να ορίζει ένα σύνολο από λειτουργίες που είναι ανεξάρτητες από την κύρια εφαρμογή τους, επιτρέποντας τόσο τον καθορισμό όσο και την επιλογή να ποικίλλουν.

## **Web API**

Στην εφαρμογή χρησιμοποιούμε επίσης και το web-API. Τα web-APIs είναι οι καθορισμένες διεπαφές μέσω των οποίων συμβαίνουν οι αλληλεπιδράσεις μεταξύ μιας επιχείρησης και των εφαρμογών που χρησιμοποιούν τα στοιχεία που διαθέτουν. Μια προσέγγιση του API είναι μια αρχιτεκτονική προσέγγιση που περιστρέφεται γύρω από την παροχή προγραμματιστικών διεπαφών σε ένα σύνολο υπηρεσιών σε διαφορετικές εφαρμογές που εξυπηρετούν διαφορετικούς τύπους καταναλωτών.

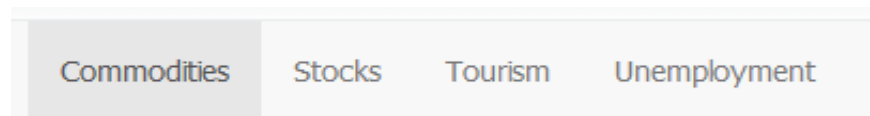
Η πρακτική της δημοσίευσης APIs επέτρεψε στις διαδικτυακές κοινότητες να δημιουργήσουν μία ανοιχτή αρχιτεκτονική για την ανταλλαγή περιεχομένου και δεδομένων μεταξύ των κοινοτήτων και των εφαρμογών. Με αυτόν τον τρόπο, το περιεχόμενο που δημιουργείται σε ένα μέρος μπορεί να αναρτηθεί και να ενημερώνεται ταυτόχρονα σε πολλαπλές θέσεις στο διαδίκτυο και όπου χρησιμοποιείται. Για παράδειγμα:

- Είναι εύκολο το μοίρασμα εικόνων από ιστότοπους, όπως το Flickr και το Photobucket στους ιστότοπους κοινωνικής δικτύωσης, όπως το Facebook και το MySpace.
- Μπορεί να ενσωματωθεί ένα περιεχόμενο, π.χ. μία παρουσίαση από το SlideShare μπορεί να ενσωματωθεί σε ένα προφίλ LinkedIn.
- Το περιεχόμενο μπορεί να αναρτηθεί δυναμικά. Όπως είναι η κοινή χρήση των σχόλιων που μόλις έγιναν στο Twitter με ένα λογαριασμό στο Facebook, για παράδειγμα, είναι ενεργοποιημένη από τα API τους.
- Μπορούν να ενσωματωθούν βίντεο σε οποιουδήποτε ιστότοπους, ενώ φιλοξενούνται από άλλον.
- Οι πληροφορίες των χρηστών μπορούν να μοιραστούν από τις διαδικτυακές κοινότητες προς τρίτες εφαρμογές, προσφέροντας νέες λειτουργίες στην δικτυακή κοινότητα που μοιράζεται τα δεδομένα του χρήστη μέσω του ανοιχτού API. Ένα από τα καλύτερα παραδείγματα αυτού είναι η πλατφόρμα Facebook Application. Ένα άλλο παράδειγμα είναι η Ανοικτή Κοινωνική πλατφόρμα (Open Social Platform).
- Εάν το περιεχόμενο αφορά τον φυσικό κόσμο (π.χ., θερμοκρασία σε γεωχωρική θέση στη γη) τότε ένα API μπορεί να θεωρηθεί ως «Διασύνδεση Προγραμματισμού Περιβάλλοντος», γνωστό και ως EPI (Environmental Programming Interface). Τα EPIs χαρακτηρίζονται από την ικανότητά τους να παρέχουν ένα μέσο για τον προσδιορισμό της αλληλουχίας γεγονότων τα οποία είναι καθολικά επαρκή ώστε τα δεδομένα να αξιοποιηθούν για τη λήψη αποφάσεων.

## 4.3 Υλοποίηση της Εφαρμογής

### 4.3.1 Είσοδοι – Έξοδοι Προγράμματος

Όπως φαίνεται στην παρακάτω εικόνα υπάρχουν διαθέσιμες τέσσερις καρτέλες (Tabs) που είναι διαθέσιμες προς επιλογή από τον χρήστη. Αυτές είναι τα αγαθά (commodities), οι μετοχές (stocks), ο τουρισμός (tourism) και η ανεργία (unemployment).



Όπως σε κάθε υλοποιημένη εφαρμογή, έτσι και εδώ το σύστημα δέχεται κάποιες εισόδους από τον χρήστη, γίνεται η επεξεργασία μέσω του προγράμματος και βγάζει κάποιες εξόδους.

#### 4.3.1.1 Κοινές λειτουργίες όλων των dataset

Η είσοδος που δέχεται το πρόγραμμα εξαρτάται από το τι ακριβώς θέλει να μελετήσει ο χρήστης. Όσον αφορά τα αγαθά ο χρήστης μπορεί να επιλέξει μέσα από μία λίστα δεκαπέντε βασικών αγαθών, αυτό το οποίο τον ενδιαφέρει. Πηγαίνοντας στις μετοχές ο χρήστης βάζει σαν είσοδο στο ανάλογο πλαίσιο το σύμβολο της μετοχής που θέλει να μελετήσει. Τέλος, αν ο χρήστης θέλει να μελετήσει τις αφίξεις των τουριστών ή τα ποσοστά ανεργίας κάποιας χώρας, θα πρέπει να θέσει σαν είσοδο το σύμβολο της χώρας που θέλει να επεξεργαστεί. Για παράδειγμα για την Ελλάδα είναι το σύμβολο *GR* (Greece), για την Ιταλία το σύμβολο *IT* (Italy), για την Κίνα το σύμβολο *CN* (China), για την Γαλλία το σύμβολο *FR* (France) και ούτω καθεξής.

Επιπρόσθετα, δίνεται η δυνατότητα στον χρήστη να επιλέξει το χρονικό διάστημα της χρονοσειράς που θέλει να μελετήσει επιλέγοντας το ανάλογο πλαίσιο. Από εκεί και πέρα κάνοντας τα συγκεκριμένα βήματα ο χρήστης, η εφαρμογή αναπαριστά τα συγκεκριμένα δεδομένα, είτε σε μορφή πίνακα, είτε σε διαγραμματική μορφή, όπου στον κατακόρυφο άξονα είναι η τιμή των δεδομένων και στον οριζόντιο άξονα η αντίστοιχη ημερομηνία.

Επίσης, μία λειτουργία του προγράμματος είναι η εφαρμογή των 8 μεθόδων πρόβλεψης, που αναλύθηκαν στο κεφάλαιο 3, πάνω στα δεδομένα που θα επιλέξει ο χρήστης. Η έξοδος από αυτή την λειτουργία γίνεται με δύο τρόπους, διαγραμματικά και σε πίνακα:

- Στην πρώτη περίπτωση δίνεται η δυνατότητα στον χρήστη να επιλέξει μέσα από εννέα κουτάκια επιλογών την αναπαράσταση των δεδομένων. Το ένα κουτάκι αναφέρεται στις πραγματικές παρατηρήσεις της χρονοσειράς που



έχει επιλέξει ο χρήστης. Τα υπόλοιπα οχτώ κουτάκια αναφέρονται στις μεθόδους πρόβλεψης και επιλέγοντάς τα, είτε ένα-ένα, είτε σαν ομάδες, είτε ακόμα και όλα μαζί, αυτό που λαμβάνεται σαν έξοδος, μετά από κατάλληλη επεξεργασία του προγράμματος είναι η διαγραμματική αναπαράσταση της χρονοσειράς κάθε μεθόδου, που έχει σαν δεδομένα τις προσαρμοσμένες τιμές προβλέψεων στις πραγματικές παρατηρήσεις.

- Σε αυτή την περίπτωση, με την επιλογή της κατάλληλης καρτέλας, το πρόγραμμα επεξεργάζεται την χρονοσειρά που έχει επιλέξει από προηγουμένως ο χρήστης και κατόπιν βγάζει σαν έξοδο τις σημειακές προβλέψεις των οχτώ μεθόδων πρόβλεψης συγκεντρωμένες σε ένα πίνακα με επτά δεκαδικά ψηφία.

#### **4.3.1.2 Αξιολόγηση μεθόδων πρόβλεψης των τιμών των αγαθών**

Θα πρέπει να τονιστεί ξεχωριστά, πως για την καρτέλα των αγαθών υπάρχει διαθέσιμη ακόμη μία λειτουργία στην εφαρμογή. Αυτή αφορά την απευθείας αξιολόγηση των οχτώ μεθόδων πρόβλεψης με βάση τρία σφάλματα:

- Το μέσο σφάλμα (*Mean Error – ME*)
- Το μέσο απόλυτο σφάλμα (*Mean Absolute Error – MAE*)
- Το συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (*symmetric Mean Absolute Percentage Error – sMAPE*)

Η επιλογή φυσικά του σφάλματος γίνεται από τον χρήστη και κατόπιν η εφαρμογή μετά από κατάλληλη επεξεργασία εμφανίζει σε ένα πίνακα με σειρά κατάταξης από την καλύτερη μέθοδο πρόβλεψης προς την χειρότερη ανάλογα με τα σφάλματα και δίπλα αναγράφεται η τιμή του σφάλματος που έχει επιλεγεί.

Η αξιολόγηση των μεθόδων γίνεται με τον εξής τρόπο. Αρχικά παίρνουμε το 80% των διαθέσιμων δεδομένων της επιλεχθείσας χρονοσειράς και εφαρμόζουμε τις μεθόδους πρόβλεψης από τις οποίες εξάγουμε την σημειακή πρόβλεψη για την επόμενη χρονική περίοδο. Στη συνέχεια φέρνουμε την πραγματική τιμή της χρονοσειράς για την περίοδο όπου έχουμε εξάγει την πρόβλεψη και υπολογίζουμε την απόκλιση, ανάμεσα στην πραγματική τιμή και την πρόβλεψη. Κατόπιν, έχοντας την συγκεκριμένη πραγματική παρατήρηση εφαρμόζουμε εκ νέου τις μεθόδους πρόβλεψης και εξάγουμε την πρόβλεψη για την επόμενη περίοδο. Ύστερα, φέρνουμε ξανά την πραγματική τιμή της περιόδου που εξάγαμε την πρόβλεψη και υπολογίζουμε πάλι την διαφορά τους. Αυτή την διαδικασία την ακολουθούμε και για το υπόλοιπο 20% των διαθέσιμων παρατηρήσεων όπου κρατάμε τα σφάλματα ανάμεσα στις προβλέψεις και τα δεδομένα, όπου με βάση αυτά τα σφάλματα και ανάλογα το είδος τους κάνουμε την αξιολόγηση των μεθόδων.

Η τεχνική αυτή, η οποία ονομάζεται τεχνική κυλιόμενης πρόβλεψης (*rolling forecasting*), ακολουθήθηκε, διότι προσφέρει την δυνατότητα της αξιολόγησης των μεθόδων σε πραγματικές συνθήκες κατά κάποιο τρόπο, κάτι το οποίο δεν συμβαίνει χρησιμοποιώντας τις προσαρμοσμένες τιμές των μεθόδων πρόβλεψης της εκάστοτε

πραγματικής χρονοσειράς. Τέλος, θα πρέπει να τονιστεί πως κάθε φορά που ανανεώνονται τα δεδομένα και γίνονται συνεχώς καινούργια διαθέσιμα, ανανεώνεται ταυτόχρονα η αξιολόγηση των μεθόδων πρόβλεψης. Εκεί εξάλλου έγκειται η καινοτομία της εφαρμογής, δηλαδή στην αυτόματη ανανέωση των δεδομένων με το που γίνουν διαθέσιμα.

### **4.3.2 Γενική Περιγραφή των Dataset**

Όπως αναφέρθηκε στην προηγούμενη παράγραφο ο χρήστης έχει την δυνατότητα να επιλέξει μέσα από τέσσερις κατηγορίες δεδομένων, τα αγαθά, τις μετοχές, την τουρισμό και την ανεργία. Κρίνεται σκόπιμο να αναφερθούν επιγραμματικά τα χαρακτηριστικά των εκάστοτε χρονοσειρών που δύναται να επιλεγούν από τον χρήστη.

#### **4.3.2.1 Αγαθά (Commodities)**

Σε αυτή την κατηγορία υπάρχουν διαθέσιμα δεκαπέντε βασικά αγαθά εκ των οποίων οι παρατηρούμενες τιμές των αγαθών μπορεί να είναι είτε σε ημερήσια βάση, είτε σε μηνιαία ή και τα δύο. Παρακάτω δίνεται μία γενική περιγραφή κάθε διαθέσιμου αγαθού.

##### **4.3.2.1.1 Αργό Πετρέλαιο (Crude Oil)**

Η χρονοσειρά του αργού πετρελαίου είναι ημερήσια. Διαθέτει ιστορικά δεδομένα από το 1986 και συγκεκριμένα αναφέρεται στην τιμή αξίας του αργού πετρελαίου WTI και μετρείται σε δολάρια ανά βαρέλι.

##### **4.3.2.1.2 Ασήμι (Silver)**

Δεύτερο στη λίστα είναι το ασήμι, όπου και αυτό έχει ημερήσια χρονοσειρά και για το οποίο υπάρχουν διαθέσιμα ιστορικά δεδομένα από το 1968 μέχρι και σήμερα. Στην αναπαράσταση των δεδομένων με τη μορφή πίνακα έχουμε στην πρώτη στήλη την ημερομηνία, στη δεύτερη τη τιμή του ασημιού σε δολάρια ανά ουγκιά, στην τρίτη σε βρετανικές λίρες ανά ουγκιά και στην τέταρτη ευρώ ανά ουγκιά. Στη διαγραμματική αναπαράσταση έγινε η επιλογή της τιμής σε δολάρια ανά ουγκιά.

##### **4.3.2.1.3 Αλουμίνιο (Aluminum)**

Το αλουμίνιο διαθέτει και ημερήσιες, αλλά και μηνιαίες χρονοσειρές. Και στις δύο περιπτώσεις οι τιμές που αναγράφονται είναι σε δολάρια ανά τόνο. Η ημερήσια χρονοσειρά ξεκινάει από 03/01/2012 και τελειώνει 11/06/2015, ενώ η μηνιαία από τις αρχές του 1980 μέχρι και σήμερα.

#### **4.3.2.1.4 Φυσικό Αέριο (Natural Gas)**

Τέταρτο στη λίστα είναι το φυσικό αέριο (Natural Gas) όπου και εδώ είναι διαθέσιμες ημερήσια και μηνιαία χρονοσειρά. Όσον αφορά την ημερήσια χρονοσειρά η τιμή είναι υπολογισμένη σε δολάρια ανά BTU και συγκεκριμένα αναφέρεται στην τιμή φυσικού αερίου που περνάει από τον αγωγό του Henry Hub, η οποία χρησιμοποιείται σαν τιμή αναφοράς για την διαμόρφωση γενικά των τιμών του φυσικού αερίου. Το ίδιο ισχύει και στην περίπτωση της μηνιαίας χρονοσειράς της τιμής του φυσικού αερίου.

#### **4.3.2.1.5 Χρυσός (Gold)**

Επόμενο αγαθό στη λίστα είναι ο χρυσός (Gold), όπου είναι διαθέσιμη μία ημερήσια χρονοσειρά. Η χρονοσειρά αυτή ξεκινάει από το 1968 μέχρι και σήμερα. Στην πρώτη στήλη αναγράφεται η ημερομηνία και στην δεύτερη στήλη η τιμή του χρυσού για την ημερομηνία αυτή, η οποία είναι υπολογισμένη σε δολάρια ανά ουγκιά.

#### **4.3.2.1.6 Χαλκός (Copper)**

Για το συγκεκριμένο αγαθό υπάρχουν διαθέσιμες στην εφαρμογή μία ημερήσια και μία μηνιαία χρονοσειρά. Η πρώτη ξεκινάει από το 2007 και έχει υπολογισμένη την τιμή του χαλκού σε δολάρια ανά λίμπρα. Η δεύτερη χρονοσειρά διαθέτει ιστορικά δεδομένα από το 1980, τα οποία είναι υπολογισμένα σε δολάρια ανά τόνο.

#### **4.3.2.1.7 Λευκόχρυσος ή Πλατίνα (Platinum)**

Έβδομο στη λίστα είναι η τιμή του λευκόχρυσου ή πλατίνα (Platinum). Η χρονοσειρά που είναι διαθέσιμη στην εφαρμογή αναγράφει ημερήσια δεδομένα. Όπως φαίνεται σε κάθε στήλη αναγράφεται, εκτός από την ημερομηνία στην πρώτη στήλη, η τιμή του λευκόχρυσου σε κάποιες αγορές και σε συγκεκριμένες ώρες της ημέρας. Τα διαθέσιμα ιστορικά δεδομένα ξεκινάνε από το 1992.

#### **4.3.2.1.8 Ρύζι (Rice)**

Επόμενο αγαθό, που είναι διαθέσιμο στην εφαρμογή, είναι το ρύζι. Η τιμή των δεδομένων που αναγράφεται είναι υπολογισμένη σε δολάρια ανά τόνο. Η χρονοσειρά είναι μηνιαία με καταγεγραμμένες τις τιμές από το 1980 μέχρι και σήμερα.

#### **4.3.2.1.9 Σιτάρι (Wheat)**

Ένα αγαθό που βρίσκεται στη λίστα και είναι διαθέσιμο προς χρήση από την εφαρμογή είναι το σιτάρι (Wheat). Για το σιτάρι υπάρχουν διαθέσιμες δύο χρονοσειρές, μία ημερήσια και μία μηνιαία. Η ημερήσια χρονοσειρά αναγράφει την τιμή του σιταριού υπολογισμένη σε δολάρια ανά μπούσελ και ξεκινάει από το 2007. Αντιθέτως, η τιμή του σιταριού στη μηνιαία χρονοσειρά είναι σε δολάρια ανά τόνο και υπάρχουν καταγεγραμμένα τα δεδομένα από το 1980 μέχρι και σήμερα.

#### **4.3.2.1.10 Καλαμπόκι (Corn)**

Δέκατο αγαθό που είναι διαθέσιμο είναι το καλαμπόκι, το οποίο διαθέτει και μηνιαία, αλλά και ημερήσια χρονοσειρά. Όσον αφορά τη μηνιαία η τιμή που αναγράφεται είναι υπολογισμένη σε δολάρια ανά τόνο και ξεκινάει από το 1980, ενώ η ημερήσια χρονοσειρά ξεκινάει από το 2000 μέχρι και σήμερα.

#### **4.3.2.1.11 Σόγια (Soybean)**

Ακόμα ένα αγαθό, που υπάρχει στη λίστα και είναι διαθέσιμο για χρησιμοποίηση στην εφαρμογή, είναι η σόγια (Soybean). Και σε αυτή την περίπτωση διακρίνουμε δύο χρονοσειρές, μία με ημερήσια δεδομένα και μία με μηνιαία δεδομένα. Η μηνιαία χρονοσειρά ξεκινάει και σε αυτή την περίπτωση από το 1980 και τα δεδομένα που είναι διαθέσιμα αντιπροσωπεύουν την τιμή πώλησης της σόγιας σε δολάρια ανά τόνο. Τα δεδομένα της ημερήσιας χρονοσειράς ξεκινάνε από τα τέλη του 2000 και εκτός από την τιμή της σόγιας, περιλαμβάνει επίσης και τη βάση τιμών για τα συμβόλαια μελλοντικής εκπλήρωσης της σόγιας (future contracts).

#### **4.3.2.1.12 Κακάο (Cocoa)**

Επόμενο στη λίστα είναι η τιμή πώλησης του κακάο (Cocoa) και σε ημερήσια, αλλά και σε μηνιαία βάση. Και οι δύο χρονοσειρές διαθέτουν τα ιστορικά δεδομένα της τιμής πώλησης του κακάο σε δολάρια ανά τόνο. Η ημερήσια χρονοσειρά ξεκινάει από τα μέσα του 2007, ενώ η μηνιαία από το 1980, όπως και στις προηγούμενες μηνιαίες χρονοσειρές. Οι αναγραφόμενες τιμές είναι για την τιμή πώλησης του κακάο της Ακτής Ελεφαντοστού (Ivory Coast).

#### **4.3.2.1.13 Καφές (Coffee)**

Δέκατο τρίτο αγαθό, το οποίο είναι διαθέσιμο στη λίστα, είναι ο καφές (Coffee) και συγκεκριμένα δύο είδη καφέ, ο καφές Arabica και ο καφές Robusta. Ο καφές Arabica είναι ο καφές που παράγεται στην Λατινική Αμερική και θεωρείται ανώτερος ποιοτικά σε σχέση με τη ποικιλία Robusta, η οποία παράγεται στην Αφρική. Λόγω της ποιότητας του καφέ, η ποικιλία Arabica είναι πιο ακριβή συγκριτικά με την ποικιλία Robusta, όπως είναι λογικό. Τα ιστορικά δεδομένα και των δύο χρονοσειρών ξεκινούν από το 1980 και οι τιμές που αναγράφονται είναι σε δολάρια ανά τόνο.

#### **4.3.2.1.14 Βαμβάκι (Cotton)**

Επόμενο στη σειρά έρχεται το βαμβάκι (Cotton), για το οποίο υπάρχουν και εδώ δύο διαθέσιμες χρονοσειρές, μία ημερήσια και μία μηνιαία. Όσον αφορά τα ημερήσια δεδομένα οι τιμές είναι υπολογισμένες σε δολάρια ανά λίμπρα και τα διαθέσιμες ιστορικές παρατηρήσεις είναι από τα μέσα του 2007. Επίσης, οι τιμές των μηνιαίων δεδομένων είναι σε cent δολαρίων ανά λίβρα και οι παρατηρήσεις που είναι καταγεγραμμένες είναι από το 1980.

#### 4.3.2.1.15 Κάρβουνο (Coal)

Τελευταίο αγαθό, που είναι διαθέσιμο, είναι το κάρβουνο (Coal). Το κάρβουνο διαθέτει μία χρονοσειρά, η οποία είναι μηνιαία με τα διαθέσιμα ιστορικά δεδομένα σχετικά με την τιμή του να είναι από τις αρχές του 1980. Οι παρατηρήσεις που έχουν καταγραφεί υπολογίζουν την τιμή του κάρβουνου σε δολάρια ανά τόνο.

Στον παρακάτω πίνακα αναφέρονται συνοπτικά τα χαρακτηριστικά των χρονοσειρών των αγαθών.

<b>Αγαθά</b>	<b>Είδος Χρονοσειράς</b>	<b>Από</b>	<b>Έως</b>
1) Πετρέλαιο	Ημερήσια	02/01/1986	Σήμερα
2) Ασήμι	Ημερήσια	02/01/1968	Σήμερα
3) Αλουμίνιο	Ημερήσια	03/01/2012	11/06/2015
	Μηνιαία	31/01/1980	Σήμερα
4) Φυσικό Αέριο	Ημερήσια	01/05/2007	Σήμερα
	Μηνιαία	03/01/1991	Σήμερα
5) Χρυσός	Ημερήσια	01/05/1968	Σήμερα
6) Χαλκός	Ημερήσια	01/05/2007	Σήμερα
	Μηνιαία	31/01/1980	Σήμερα
7) Λευκόχρυσος	Ημερήσια	01/07/1992	Σήμερα
8) Ρύζι	Μηνιαία	31/01/1980	Σήμερα
9) Σιτάρι	Ημερήσια	01/05/2007	Σήμερα
	Μηνιαία	31/01/1980	Σήμερα
10) Καλαμπόκι	Ημερήσια	01/12/2000	17/08/2015
	Μηνιαία	31/08/1980	Σήμερα
11) Σόγια	Ημερήσια	01/12/2000	17/08/2015
	Μηνιαία	31/01/1980	Σήμερα
12) Κακάο	Ημερήσια	01/05/2007	Σήμερα
	Μηνιαία	31/01/1980	Σήμερα
13α) Καφές Arabica	Μηνιαία	31/01/1980	Σήμερα
13β) Καφές Robusta	Μηνιαία	31/01/1980	Σήμερα
14) Βαμβάκι	Ημερήσια	01/05/2007	Σήμερα
	Μηνιαία	31/01/1980	Σήμερα
15) Κάρβουνο	Μηνιαία	31/01/1980	Σήμερα

#### 4.3.2.2 Μετοχές (Stocks)

Οι μετοχές είναι ημερήσιες χρονοσειρές εκ των οποίων ο χρονικός ορίζοντας των παρατηρήσεων είναι αδύνατον να προσδιοριστεί για κάθε μία από τις μετοχές που είναι διαθέσιμες μέσα από την εφαρμογή, διότι ο αριθμός αυτών των μετοχών είναι χιλιάδες. Όπως βλέπουμε και στην παρακάτω εικόνα η αναπαράσταση των δεδομένων σε μορφή πίνακα γίνεται ως εξής:

	YHOO.Open	YHOO.High	YHOO.Low	YHOO.Close	YHOO.Volume	YHOO.Adjusted
2012-12-31	19.43	19.97	19.40	19.90	20645100.00	19.90
2013-01-02	20.20	20.32	20.01	20.08	20463100.00	20.08
2013-01-03	20.05	20.10	19.72	19.78	19504400.00	19.78
2013-01-04	19.76	19.95	19.72	19.86	12489600.00	19.86

- Στην πρώτη στήλη αναγράφεται η ημερομηνία.
- Στην επόμενη η τιμή ανοίγματος της μετοχής.
- Στην στήλη με την ονομασία High αναφέρεται στην υψηλότερη τιμή που έλαβε η μετοχή κατά την διάρκεια της συνεδρίασης του χρηματιστηρίου.
- Στην επόμενη με την ονομασία Low είναι αντίστοιχα η χαμηλότερη τιμή κατά τη διάρκεια της συνεδρίασης.
- Στην στήλη με την ονομασία Close είναι η τιμή κλεισίματος της μετοχής.
- Η προτελευταία στήλη έχει την ονομασία Volume, που αναφέρεται στον όγκο των συναλλαγών. Με τον όρο αυτό εννοούμε τον αριθμό των μετοχών ή των συμβολαίων που συναλλάσσονται για εγγύηση ή σε μια ολόκληρη αγορά κατά τη διάρκεια μιας συγκεκριμένης χρονικής περιόδου. Είναι απλά το ποσό των μετοχών που αλλάζουν χέρια από τους πωλητές στους αγοραστές και αποτελεί μία μέτρηση της δραστηριότητας. Εάν ένας αγοραστής από ένα απόθεμα αγοράζει 100 μετοχές από έναν πωλητή, τότε ο όγκος για την περίοδο αυτή αυξάνεται κατά 100 μετοχές.

Ο όγκος των συναλλαγών αποτελεί σημαντικό δείκτη για την τεχνική ανάλυση, δεδομένου ότι χρησιμοποιείται για να μετρήσει την αξία της κίνησης της αγοράς. Αν οι αγορές έχουν δημιουργήσει ισχυρή κίνηση των τιμών, είτε προς τα πάνω ή προς τα κάτω, η «δύναμη» της κίνησης αυτής εξαρτάται από τον όγκο για την εν λόγω περίοδο. Όσο υψηλότερος είναι ο όγκος κατά τη διάρκεια αυτής της κίνησης της τιμής, τόσο πιο σημαντική είναι η κίνηση.

- Στην τελευταία στήλη αναγράφεται η προσαρμοσμένη τιμή κλεισίματος. Γενικότερα ορίζεται ως η τιμή κλεισίματος μιας μετοχής σε οποιαδήποτε δεδομένη ημέρα διαπραγμάτευσης που έχει τροποποιηθεί ώστε να περιλαμβάνει τυχόν κατανομές και εταιρικές πράξεις που έλαβαν χώρα σε οποιαδήποτε στιγμή πριν από το άνοιγμα της επόμενης ημέρας. Συχνά η προσαρμοσμένη τιμή κλεισίματος χρησιμοποιείται κατά την εξέταση των ιστορικών αποδόσεων ή εκτελείται μία λεπτομερή ανάλυση των ιστορικών αποδόσεων.

Η προσαρμοσμένη τιμή κλεισίματος είναι ένα χρήσιμο εργαλείο κατά την εξέταση των ιστορικών αποδόσεων, διότι δίνει στους αναλυτές μια ακριβή αναπαράσταση της αξίας των ιδίων κεφαλαίων της επιχείρησης πέρα από την απλή τιμή της αγοράς. Ευθύνεται για όλες τις εταιρικές πράξεις, όπως η διάσπαση μετοχών, μερισμάτων, κατανομών και προσφορών δικαιωμάτων.

Στη διαγραμματική αναπαράσταση των δεδομένων, λαμβάνεται υπόψιν η τιμή κλεισίματος της μετοχής.

### **4.3.2.3 Τουρισμός (Tourism)**

Οι τιμές που αναφέρονται σε αυτή την κατηγορία, αφορούν τις αφίξεις των τουριστών σε μία χώρα κατά την διάρκεια ενός έτους. Άρα είναι ετήσιες οι χρονοσειρές και ξεκινάνε από το 1995 και φτάνουν μέχρι το 2013.

### **4.3.2.4 Ανεργία (Unemployment)**

Σε αυτή την κατηγορία οι τιμές, που αναγράφονται, αναφέρονται στο ποσοστό της ανεργίας σε κάθε χώρα στη διάρκεια του έτους. Και σε αυτή την περίπτωση οι χρονοσειρές είναι ετήσιες ξεκινώντας από το 1980 και φτάνουν μέχρι το 2013.

## **4.3.3 Βασικά Πακέτα της Εφαρμογής**

Στην παράγραφο αυτή κρίθηκε σκόπιμο να αναφερθούν ορισμένα από τα βασικά πακέτα (packages) που βοήθησαν και χρησιμοποιήθηκαν για την υλοποίηση της εφαρμογής.

### **4.3.3.1 Πακέτα άντλησης δεδομένων**

#### **Αγαθά (Commodities)**

Η άντληση των ιστορικών δεδομένων των τιμών των αγαθών γίνεται από το API του ιστότοπου Quandl, μέσω του πακέτου της R το οποίο ονομάζεται Quandl. Το Quandl είναι μια μηχανή αναζήτησης για αριθμητικά δεδομένα. Η ιστοσελίδα προσφέρει πρόσβαση σε εκατομμύρια χρηματοοικονομικές, οικονομικές και κοινωνικές συλλογές δεδομένων. Το Quandl κατηγοριοποιεί τα δεδομένα από πολλαπλές πηγές, κάτι το οποίο επιτρέπει στους χρήστες να τα βρει εύκολα και να τα κατεβάσει σε διάφορες μορφές. Όλα τα δεδομένα του Quandl είναι προσβάσιμα μέσω ενός API. Η πρόσβαση στο API είναι δυνατή μέσω πακέτων για πολλές γλώσσες προγραμματισμού, συμπεριλαμβανομένων της R, της Python, του Matlab και Stata. Μακροπρόθεσμα το Quandl θέλει να κάνει όλα τα αριθμητικά δεδομένα του διαδικτύου διαθέσιμα στο ευρύ κοινό και να είναι εύκολο στο να τα βρει κανείς και στη χρήση τους.

Το Quandl έχει κατηγοριοποιήσει εκατομμύρια των αριθμητικών συνόλων δεδομένων από όλο το διαδίκτυο. Όταν κάνετε κλικ σε ένα συγκεκριμένο σύνολο δεδομένων που αναφέρεται στο ευρετήριο, το Quandl εξάγει την πιο πρόσφατη έκδοση των εν λόγω δεδομένων από την πηγή και τα κάνει διαθέσιμα σε οποιαδήποτε μορφή.

### **Μετοχές (Stocks)**

Οι τιμές των μετοχών αντλούνται μέσα από το διαθέσιμο API του Yahoo Finance χρησιμοποιώντας το πακέτο της R το `quantmod`. Το Yahoo Finance είναι ένα μέσο ενημέρωσης που αποτελεί μέρος του δικτύου της Yahoo. Παρέχει οικονομικές ειδήσεις, τα δεδομένα και τα σχόλια, συμπεριλαμβανομένων τιμές μετοχών, δελτία τύπου, οικονομικές εκθέσεις, και τον αρχικό προγραμματισμό. Προσφέρει επίσης ορισμένα διαδικτυακά εργαλεία για την διαχείριση των προσωπικών οικονομικών.

### **Τουρισμός και Ανεργία**

Η άντληση των δεδομένων γίνεται από το API της Παγκόσμιας Τράπεζας (World Bank) μέσω του πακέτου της R το `WDI` (World Development Indicators). Η Παγκόσμια Τράπεζα είναι ένα διεθνές χρηματοπιστωτικό ίδρυμα το οποίο παρέχει οικονομική και τεχνική βοήθεια προς τις αναπτυσσόμενες χώρες για αναπτυξιακά έργα, όπως είναι η δημιουργία δρόμων, χτίσιμο νοσοκομείων και σχολείων κλπ. Η Παγκόσμια Τράπεζα περιλαμβάνει δύο ιδρύματα, τη Διεθνή Τράπεζα Ανασυγκρότησης και Ανάπτυξης (International Bank for Reconstruction and Development-IBRD) και τη Διεθνή Ένωση Ανάπτυξης (International Development Association-IDA).

Η Παγκόσμια Τράπεζα δεν θα πρέπει να συγχέεται με τον Όμιλο της Παγκόσμιας Τράπεζας των Ηνωμένων Εθνών (World Bank Group), ο οποίος αποτελεί μέλος του Οικονομικού και Κοινωνικού Συμβουλίου των Ηνωμένων Εθνών (United Nations Development Group), και μια οικογένεια πέντε διεθνών οργανισμών που κάνουν μόχλευση δανείων προς τις φτωχές χώρες:

- Διεθνή Τράπεζα Ανασυγκρότησης και Ανάπτυξης (International Bank for Reconstruction and Development-IBRD)
- Διεθνής Ένωση Ανάπτυξης (International Development Association-IDA)
- Διεθνής Οργανισμός Χρηματοδότησης (International Finance Corporation-IFC).
- Πολυμερής Οργανισμός Εγγύησης Επενδύσεων (Multilateral Investment Guarantee Agency-MIGA)

Διεθνές Κέντρο Επίλυσης Επενδυτικών Διαφορών (International Centre for Settlement of Investment Disputes-ICSID)

#### **4.3.2.2 Πακέτο διαγραμματικής αναπαράστασης των δεδομένων**

Η αναπαράσταση των δεδομένων έγινε με την βοήθεια του πακέτου `ggplot2` της R. Το πακέτο αυτό δημιουργήθηκε από τον Hadley Wickham το 2005 και έκτοτε έχει γίνει ένα από τα πιο δημοφιλή πακέτα. Αυτό συμβαίνει διότι δίνει την δυνατότητα στον χρήστη για την γραφική απεικόνιση των δεδομένων σε πάρα πολλές μορφές και είναι πολύ εύκολο στην χρήση του.



#### **4.3.3.3 Πακέτα εφαρμογής προβλέψεων**

Για τις μεθόδους εκθετικής εξομάλυνσης, κινητού μέσου όρου και αφελή (Naive) χρησιμοποιήθηκε το πακέτο `forecast`. Αυτό το πακέτο δημιουργήθηκε από τον Rob Hyndman και χρησιμοποιείται ευρέως από τους χρήστες της R λόγω του μεγάλου πλήθους των μεθόδων που διαθέτει και λόγω των αρκετών διαθέσιμων επιλογών όσον αφορά τις μεθόδους. Το συγκεκριμένο πακέτο είναι διαθέσιμο μέσα από το CRAN και διατίθεται εντελώς δωρεάν. Επίσης ο δημιουργός του πακέτου έχει δημιουργήσει ένα διαδικτυακό βιβλίο, στο οποίο εξηγεί πολύ αναλυτικά τις μεθόδους που είναι υλοποιημένες στο πακέτο.

Για την εφαρμογή της απλής γραμμικής παλινδρόμησης χρησιμοποιήθηκε η συνάρτηση `lm` η οποία είναι διαθέσιμη στην R. Επίσης για την εφαρμογή της μεθόδου Theta χρησιμοποιήθηκαν οι συναρτήσεις της γραμμικής παλινδρόμησης και της απλής εκθετικής εξομάλυνσης που βρίσκεται στο πακέτο `forecast` που αναφέρθηκε παραπάνω.



# 5. Χρησιμότητα των Δεδομένων(Dataset)

## 5.1 Ανάλυση των Αγαθών (Commodities)

Στα οικονομικά, ένα αγαθό θεωρείται ως ένα εμπορεύσιμο προϊόν που παράγεται για να ικανοποιήσει τις επιθυμίες και τις ανάγκες. Σε αυτά μπορεί να περιλαμβάνονται αγαθά και υπηρεσίες.

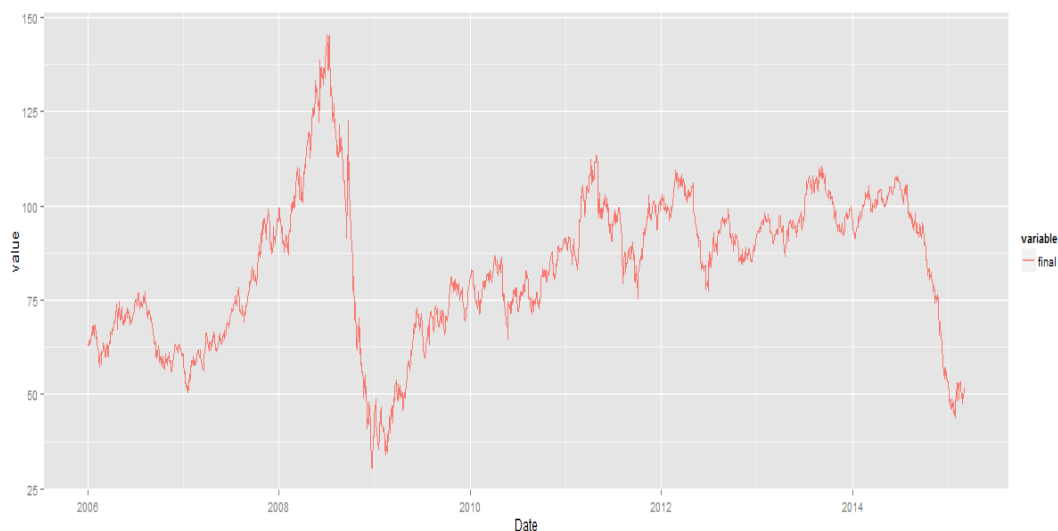
Πρόκειται για μία φυσική ουσία, όπως τα τρόφιμα, τα δημητριακά και τα μέταλλα, τα οποία είναι ανταλλάξιμα με άλλο προϊόν του ίδιου τύπου, και τα οποία οι επενδυτές αγοράζουν ή πωλούν, συνήθως μέσω συμβολαίων μελλοντικής εκπλήρωσης (future contracts). Η τιμή του εμπορεύματος εξαρτάται από την προσφορά και τη ζήτηση. Ο κίνδυνος συνήθως υπάρχει στα βασικά γεωργικά προϊόντα. Για παράδειγμα, ένας αγρότης παίρνει το ρίσκο παράγοντας ένα προϊόν με συγκεκριμένο κόστος, το οποίο θα τεθεί προς πώληση μελλοντικά, και αυτό διότι δεν γνωρίζει ποια θα είναι η τιμή πώλησης εκείνη την περίοδο.

### 5.1.1 Σημασία της τιμής των αγαθών

Συνήθως, με τον όρο τιμές των αγαθών εννοούνται οι τιμές της ενέργειας (κυρίως το αργό πετρέλαιο), οι τιμές των τροφίμων και οι τιμές ορισμένων από τα κύρια προϊόντα που χρησιμοποιούνται στην βιομηχανία. Οι τιμές της ενέργειας παρακολουθούνται πολύ στενά από τις οικονομικές μονάδες και τους φορείς χάραξης πολιτικής, διότι η τιμή του πετρελαίου υποδηλώνει το κόστος για μία από τις πιο σημαντικές εισροές για το σύνολο της οικονομίας. Η ενέργεια είναι ένας από τους κύριους παράγοντες κόστους για τις επιχειρήσεις. Είναι, επίσης, ένα σημαντικό προϊόν για τη θέρμανση των νοικοκυριών και για τις μετακινήσεις και επομένως η διαμόρφωση της τιμής της παίζει σημαντικό ρόλο. Από την άλλη πλευρά, οι τιμές των τροφίμων λόγω της υψηλής συμμετοχής τους στις δαπάνες, ιδίως για τα νοικοκυριά χαμηλού εισοδήματος, και της επίδρασης που έχουν στην αγοραστική δύναμη, δίνεται ιδιαίτερη προσοχή από τους ανθρώπους διαμόρφωσης πολιτικής. Εν τω μεταξύ, τα μέταλλα, όπως σιδηρομέταλλευμα, χαλκός και αλουμίνιο, είναι από τις σημαντικότερες πρώτες ύλες για τη βιομηχανία και τις κατασκευές.

Ο λόγος για τον οποίο οι τιμές των εμπορευμάτων έχουν προσελκύσει την προσοχή των μελετητών είναι οι τρέχουσες διακυμάνσεις στις τιμές των αγαθών ενός

μεγέθους που δεν είχε ιδιαίτερες διακυμάνσεις μέχρι το 2007. Η τιμή του αργού πετρελαίου τύπου Brent, η οποία ήταν 50 δολάρια το βαρέλι στις αρχές του 2007, ανέβηκε στα 145 δολάρια μέσα σε ενάμιση χρόνο, όπως φαίνεται και στην παρακάτω εικόνα:



Στη συνέχεια, παράλληλα με την εμβάθυνση της διεθνούς κρίσης στα τέλη του 2008, οι τιμές του αργού πετρελαίου επέστρεψαν στα 35 δολάρια το βαρέλι και τώρα κινείται γύρω στα 50 δολάρια το βαρέλι. Επίσης, αυτή η αβεβαιότητα σχετικά με τις τιμές του πετρελαίου, που προκύπτουν από την εξέλιξη της ζήτησης και της προσφοράς, και οι πιθανές επιπτώσεις τους, κρατάει το ενδιαφέρον για τη χάραξη της πολιτικής σχετικά με τις τιμές των εμπορευμάτων.

Δεδομένου, ότι οι τιμές των εμπορευμάτων, κυρίως του αργού πετρελαίου και των τροφίμων, είναι μεταξύ των παραγόντων κινδύνου που επηρεάζουν τις προοπτικές του πληθωρισμού. Έτσι, υπάρχει άμεση αναφορά στις τιμές των αγαθών, στις τελευταίες εκθέσεις του πληθωρισμού. Μάλιστα, σχεδόν το ήμισυ των αναθεωρήσεων σχετικά με τις επίσημες προβλέψεις του πληθωρισμού που δημοσιεύτηκαν τα τελευταία δύο χρόνια, προήλθαν λόγω αλλαγών των υποθέσεων που γινόντουσαν σχετικά με τις τιμές των τροφίμων και των βασικών εμπορευμάτων. Αυτό υποδηλώνει ότι η άνοδος των τιμών των βασικών αγαθών μπορεί να επιφέρει προκλήσεις για τις κεντρικές τράπεζες, όσον αφορά τις άμεσες πιέσεις στον πληθωρισμό. Επιπρόσθετα, οι αβεβαιότητες, που οφείλονται στις διακυμάνσεις των τιμών των βασικών εμπορευμάτων, ξεχωρίζουν ως ένα σημαντικό ζήτημα για την ανακοίνωση της νομισματικής πολιτικής και τις εκτιμήσεις της διοίκησης από τις κεντρικές τράπεζες.

Η δομή της δυναμικής των τιμών των εμπορευμάτων έχει επίσης μεγάλη σημασία και συμβάλλει στις αποφάσεις της νομισματικής πολιτικής από τις κεντρικές τράπεζες. Για παράδειγμα, οι πηγές των διακυμάνσεων των τιμών των βασικών εμπορευμάτων και η φύση των διαταραχών τους έχουν εξαιρετικά μεγάλη επιρροή για τον καθορισμό των πολιτικών αντιμετώπισης αυτών των διακυμάνσεων. Ομοίως, όχι

μόνο το επίπεδο, αλλά και η αστάθεια των τιμών των βασικών προϊόντων μπορεί να έχουν σημαντική επίπτωση στην μακροοικονομική δραστηριότητα. Για παράδειγμα, μία διαταραχή στο επίπεδο των τιμών του αργού πετρελαίου μπορεί να έχει διαφορετική επίδραση στην οικονομία, εάν παρατηρείται σε μια περίοδο όπου η αστάθεια των τιμών είναι επίσης μεγάλη.

Άρα είναι προφανές πως η διαμόρφωση των τιμών των αγαθών είναι αντικείμενο μελέτης και στενής παρακολούθησης, είτε από τις κυβερνήσεις για την διαμόρφωση της εσωτερικής και οικονομικής πολιτικής, είτε από τις επιχειρήσεις για τη χάραξη εσόδων-εξόδων, είτε από τους ερευνητές σαν αντικείμενο μελέτης, είτε ακόμα και από τους απλούς ανθρώπους για την διαμόρφωση του οικογενειακού προϋπολογισμού τους. Επομένως, η δημιουργία ενός συστήματος που αξιολογεί και κατατάσσει τις πιο γνωστές μεθόδους πρόβλεψης ανάλογα με τα σφάλματα που προκύπτουν σε σχέση με τα πραγματικά δεδομένα είναι πολύ σημαντική. Επίσης, αυτή η εφαρμογή μπορεί να αποτελέσει ένα εργαλείο για κάποιον που θέλει να μελετήσει τις τιμές των αγαθών και να προβλέψει την τιμή τους, διότι του παρέχει πολύτιμες πληροφορίες σε σχέση με την αξιολόγηση των μεθόδων πρόβλεψης που χρησιμοποιήθηκαν. Αυτό τον βοηθάει στην τελική επιλογή της μεθόδου πρόβλεψης για το αγαθό που θέλει με βάση την κατάταξη της.

### **5.1.2 Διαμόρφωση της τιμής των αγαθών**

Φυσικά υπάρχουν μερικοί παράγοντες που επηρεάζουν τις τιμές των εμπορευμάτων, οι οποίοι είναι κοινοί για όλα τα αγαθά. Δύο τέτοιοι παράγοντες είναι η προσφορά και η ζήτηση των αγαθών. Όταν υπάρχει μεγάλη προσφορά που υπερκαλύπτει την ζήτηση, τότε οι τιμές ωθούνται προς τα κάτω. Αντιθέτως, μεγάλη ζήτηση σημαίνει ώθηση των τιμών προς τα πάνω. Όπως είναι λογικό οι παράγοντες που επηρεάζουν την προσφορά και την ζήτηση είναι διαφορετικοί για κάθε αγαθό και για αυτό θα πρέπει να μελετιούνται ξεχωριστά, ώστε να λαμβάνονται υπόψιν στην πρόβλεψη των τιμών των αγαθών. Παρακάτω έχει γίνει μία τέτοια προσπάθεια αναγνώρισης και περιγραφικής ανάλυσης και των δεκαπέντε εμπορευμάτων που είναι διαθέσιμες στην εφαρμογή.

#### **5.1.2.1 Αργό πετρέλαιο (Crude Oil)**

Όπως είναι ήδη γνωστό η διαμόρφωση των τιμών του πετρελαίου επηρεάζουν με σημαντικό τρόπο την παγκόσμια οικονομία, είτε πρόκειται σε κρατικό επίπεδο, είτε σε ατομικό. Η τιμή του πετρελαίου επομένως επηρεάζεται συνήθως από τους εξής παράγοντες:

##### *OPEC (The Organization of Petroleum Exporting Countries)*

Ο OPEC, ένας οργανισμός ο οποίος αποτελείται από 13 χώρες: Αλγερία, Αγκόλα, Εκουαδόρ, Ινδονησία, Ιράν, Ιράκ, Κουβέιτ, Λιβύη, Νιγηρία, Κατάρ, Σαουδική Αραβία, Ηνωμένα Αραβικά Εμιράτα και Βενεζουέλα, είναι η μεγαλύτερη οικονομική οντότητα

η οποία έχει το μεγαλύτερο αντίκτυπο στις προμήθειες πετρελαίου παγκοσμίως. Ο ΟΡΕC είναι υπεύθυνος για το 40% της παγκόσμιας παραγωγής πετρελαίου, και καθορίζει τις πολιτικές μεταξύ των χωρών μελών, ώστε να μπορέσουν να ανταποκριθούν στην παγκόσμια κατανάλωση. Ο ΟΡΕC μπορεί να επηρεάσει την τιμή του πετρελαίου, αυξάνοντας ή μειώνοντας την παραγωγή πετρελαίου ανάμεσα στα κράτη-μέλη. Για παράδειγμα η αύξηση της τιμής του πετρελαίου το 2007 και το 2008 προκλήθηκε εξαιτίας της μείωσης της παραγωγής ανάμεσα στις χώρες του ΟΡΕC το 2006.

### Προσφορά, Ζήτηση και Αποθέματα

Τα αποθέματα πετρελαίου παγκοσμίως ισορροπούν την προσφορά και την ζήτηση. Αν η παραγωγή υπερβαίνει την ζήτηση η πλεονάζουσα προσφορά μπορεί να αποθηκευτεί. Όταν η κατανάλωση υπερβαίνει την ζήτηση, τα αποθέματα μπορούν να αξιοποιηθούν για την αντιμετώπιση της αύξησης της ζήτησης, και η σχέση μεταξύ των τιμών του πετρελαίου και αποθεμάτων πετρελαίου επιτρέπει διορθώσεις σε οποιαδήποτε κατεύθυνση. Προμηθευτές εκτός του ΟΡΕC παράγουν το 60% του πετρελαίου παγκοσμίως, και παρόλο που είναι 50% μεγαλύτεροι από τον ΟΡΕC, δεν διαθέτουν επαρκή αποθέματα για να είναι σε θέση να ελέγχουν τις τιμές και μπορούν να ανταποκριθούν μόνο στις διακυμάνσεις της αγοράς. Ο ΟΡΕC, ωστόσο μπορεί να επηρεάσει άμεσα τις τιμές της αγοράς, ειδικά όταν η προσφορά του πετρελαίου που παράγεται από χώρες εκτός του ΟΡΕC μειώνεται.

### Πολιτικές Αναταραχές

Εάν μία περιοχή πλούσια σε πετρέλαιο γίνεται πολιτικά ασταθής, οι αγορές αγαθών αντιδρούν ανεβάζοντας την τιμή πετρελαίου, έτσι ώστε οι προμήθειες να είναι διαθέσιμες στον υψηλότερο πλειοδότη. Σε αυτήν την περίπτωση, μόνο η αντίληψη της έλλειψης προσφοράς μπορεί να αυξήσει τις τιμές, ακόμα και όταν τα επίπεδα της παραγωγής παραμένουν σταθερά.

### Ο Καιρός

Όπως στα περισσότερα αγαθά, οι εποχιακές αλλαγές στον καιρό επηρεάζουν τη ζήτηση για πετρέλαιο. Κατά την διάρκεια του χειμώνα, καταναλώνεται περισσότερο πετρέλαιο θέρμανσης, και το καλοκαίρι, οι άνθρωποι οδηγούν περισσότερο και χρησιμοποιούν περισσότερη βενζίνη. Ακόμη και αν οι αγορές ξέρουν πότε να περιμένουμε αυτές τις αυξημένες περιόδους της ζήτησης, η τιμή του πετρελαίου ανεβαίνει και βγαίνει έξω από το επίπεδο με την εποχή κάθε χρόνο. Οι ακραίες καιρικές συνθήκες (τυφώνες, ανεμοστρόβιλοι, καταιγίδες) μπορούν να επηρεάσουν τις εγκαταστάσεις και υποδομές παραγωγής διακόπτοντας τον εφοδιασμό πετρελαίου κάτι που προκαλεί υπερτιμολόγηση.

### Συναλλαγματική Αξία του Δολαρίου

Το πετρέλαιο συναλλάσσεται και πωλείται διεθνώς σε αμερικάνικα δολάρια. Υποτίμηση του δολαρίου γενικά τείνει να αυξήσει τη ζήτηση του πετρελαίου και την αύξηση της τιμής του. Αντίθετα, η ενίσχυση του δολαρίου μειώνει το πραγματικό

εισόδημα στις χώρες κατανάλωσης, μειώνοντας τη ζήτηση για πετρέλαιο και μειώνοντας τις τιμές.

#### Ζήτηση χωρών εκτός ΟΟΣΑ

Η κατανάλωση πετρελαίου στις αναπτυσσόμενες χώρες που δεν είναι μέρος του Οργανισμού Οικονομικής Συνεργασίας και Ανάπτυξης (ΟΟΣΑ) έχει αυξηθεί σημαντικά τα τελευταία χρόνια. Ενώ η κατανάλωση πετρελαίου στις χώρες του ΟΟΣΑ μειώθηκε μεταξύ 2000 και 2010, η κατανάλωση πετρελαίου από χώρες εκτός ΟΟΣΑ έχει αυξηθεί περισσότερο από 40%. Η Κίνα, η Ινδία και η Σαουδική Αραβία είχαν την μεγαλύτερη αύξηση της κατανάλωσης του πετρελαίου ανάμεσα στις χώρες εκτός ΟΟΣΑ κατά τη διάρκεια αυτής της περιόδου.

#### **5.1.2.2 Ασήμι (Silver)**

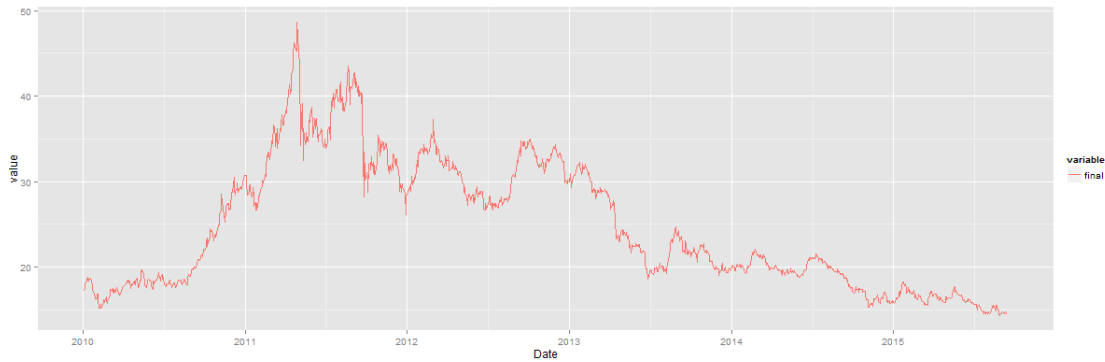
Το ασήμι θεωρείται ένα από τα πολύτιμα μέταλλα με χρήση και στην βιομηχανία. Επομένως η τιμή του μπορεί να επηρεάσει την τελική τιμή και άλλων βιομηχανικών προϊόντων. Η τιμή του διαμορφώνεται από τους εξής κύριους παράγοντες:

#### Βιομηχανική Ζήτηση

Οι νέες εφαρμογές για το ασήμι που διερευνώνται στις μπαταρίες, στους υπεραγωγούς και στα μικροκυκλώματα, μπορούν να αυξήσουν περαιτέρω την ζήτηση. Η επέκταση της μεσαίας τάξης στις αναδυόμενες κοινωνίες, που φιλοδοξούν να αποκτήσουν τον δυτικό τρόπο ζωής, μπορεί επίσης να συμβάλλει στην μακροπρόθεσμη αύξηση της βιομηχανικής ζήτησης για το ασήμι

#### Οι Τιμές του Χρυσού

Παρ' όλους τους θεμελιώδεις παράγοντες του ασημιού, ο χρυσός θεωρείται ως η κύρια κινητήρια δύναμη για τις τιμές του ασημιού. Σε ένα ανοδικό περιβάλλον, οι κερδοσκόποι έχουν την τάση να ενδιαφέρονται για τα περισσότερα από τα πολύτιμα μέταλλα. Έτσι αυτό οδηγεί σε μια αύξηση της ζήτησης για την επένδυση ασημιού. Το ασήμι έχει γενικά μία μικρότερη αγορά σε σχέση με αυτή του χρυσού. Την ίδια στιγμή αν το περιβάλλον είναι πτωτικό, οι επενδυτές χάνουν την εμπιστοσύνη τους στο ασήμι πολύ γρήγορα κάτι το οποίο προκαλεί τις τιμές να πέσουν. Από την ανάλυση της τάσης της αναλογίας χρυσού-ασημιού, μπορεί να φανεί καθαρά ότι το ασήμι έχει την τάση να ακολουθεί τις τιμές του χρυσού. Αυτό φαίνεται μάλιστα και από τα δύο διαγράμματα που βρίσκονται παρακάτω. Βλέπουμε ξεκάθαρα την τάση πως όταν οι τιμές του χρυσού αυξανόντουσαν τότε και οι τιμές του ασημιού σημείωναν άνοδο. Επίσης βλέπουμε πως κατά τη διάρκεια των ημερών που η τιμή του χρυσού μειωνόταν, τότε οι τιμή του ασημιού έπεφτε κατακόρυφα με ένα ακόμα μεγαλύτερο περιθώριο. Βασισμένοι σε αυτή την υπόθεση, προτείνεται η αγορά ασημιού κατά τη διάρκεια μίας ύφεσης και η πώλησή του κατά τη διάρκεια της οικονομικής άνθισης.



Χρονοσειρά τιμής Ασημιού



Χρονοσειρά τιμής Χρυσού

### Το Αμερικάνικο Δολάριο

Από την μελέτη της σχέσης μεταξύ του ασημιού και του δολαρίου μπορούμε να δούμε ξεκάθαρα ότι υπάρχει μία αντίστροφη σχέση μεταξύ των τιμών του ασημιού και του δολαρίου. Κατά τη διάρκεια της ύφεσης το δολάριο θεωρείται ασφαλές καταφύγιο, οι άνθρωποι σε όλο τον κόσμο έχουν την τάση να αποεπενδύουν από εμπορεύματα και να επενδύουν σε δολάρια.

### Μεγάλοι και Ιδιωτικοί Θεσμικοί Επενδυτές

Εκτός από την πιο σημαντική επίπτωση των Hun Brothers , το 1997, ο Warren Buffett αγόρασε 130 εκατομμύρια ουγκιές (4000 μετρικοί τόνοι) ασημιού με περίπου 4,5 δολάρια ανά ουγκιά (συνολική αξία 585 εκατομμύρια δολάρια). Επίσης, τον Απρίλιο του 2006, η iShares ξεκίνησε μια εταιρεία ανταλλαγών ασημιού, η οποία λέγεται iShares Silver Trust, όπου από τον Απρίλιο του 2008 κατέχει 180 εκατομμύρια Oz ασημιού ως αποθεματικά. Αυτό δείχνει σαφώς ότι οι μεγάλοι επενδυτές έχουν τη δύναμη να επηρεάζουν τις τιμές τις αγορές.

### **5.1.2.3 Αλουμίνιο (Aluminum)**

Η τιμή του αλουμινίου μπορεί να διαμορφωθεί από τους εξής παράγοντες:

#### Βιομηχανία Μεταφορών

Μεταξύ των ανεπτυγμένων χωρών, οι μεταφορές είναι η μεγαλύτερη αγορά αλουμινίου. Αλλαγές στη ζήτηση για αυτοκίνητα, αεροπλάνα, φορτηγά κλπ., μπορούν



να επηρεάσουν σημαντικά τις τιμές του αλουμινίου. Τα αυτοκίνητα συνήθως φτιάχνονταν σχεδόν εξ ολοκλήρου από χάλυβα, αλλά τα νέα επιβατηγά αυτοκίνητα και τα ελαφριά φορτηγά περιέχουν τώρα κατά μέσο όρο 300 λίβρες αλουμινίου, μία σημαντική αύξηση σε σχέση με πριν. Σε γενικές γραμμές, η ζήτηση στον κλάδο των αυτοκινήτων για το αλουμίνιο ανεβαίνει προς τα επάνω, κάτι οποίο θα μπορούσε να προκαλέσει μία αύξηση των τιμών του αλουμινίου.

Στην αεροναυπηγική βιομηχανία το αλουμίνιο χρησιμοποιείται εδώ και καιρό για την κατασκευή αεροπλάνων και ελικοπτέρων, αν και υπήρξε μία πρόσφατη αύξηση στη χρήση των σύνθετων υλικών από ίνες άνθρακα. Αυτά τα σύνθετα υλικά είναι ελαφρύτερα από το αλουμίνιο, ισχυρότερα από το ατσάλι και πολύ πιο εύκολο να συγκεντρωθούν, ένας συνδυασμός που συνιστά απειλή για την κυριαρχία του αλουμινίου στην αεροναυπηγική βιομηχανία. Κατασκευαστές αεροπλάνων έχουν ανακοινώσει σχέδια για την κατασκευή μοντέλων που αποτελούνται από μόλις 20% αλουμίνιο κατά βάρος, μία σημαντική μείωση από το 50% των προηγούμενων μοντέλων. Η άνοδος των σύνθετων υλικών με ίνες άνθρακα στον τομέα της κατασκευής της αεροναυπηγικής βιομηχανίας μπορεί να μειώσει την ζήτηση για αλουμίνιο και να προκαλέσει καθοδικές πιέσεις στις τιμές.

#### Κατασκευαστική Αγορά

Η αγορά κτιρίων και η κατασκευαστική αγορά αντιπροσωπεύουν ένα σημαντικό μέρος της συνολικής ζήτησης για αλουμίνιο. Στις ανεπτυγμένες χώρες, περίπου το 15-20% του συνολικού αλουμινίου που παράγεται διοχετεύεται στην κατασκευαστική αγορά, ενώ το ποσοστό αυτό είναι υψηλότερο στις αναπτυσσόμενες χώρες με περίπου 30%. Οποιαδήποτε αύξηση ανέγερσης κατοικιών ή εμπορικών κατασκευών μπορεί να αυξήσουν τη ζήτηση για αλουμίνιο με αποτέλεσμα υψηλότερες τιμές, ενώ μείωση της στέγασης μπορεί να οδηγήσουν σε μείωση των τιμών του αλουμινίου.

#### Αναδυόμενες Αγορές

Η ποσότητα του αλουμινίου που καταναλώνονται κατά κεφαλήν σχετίζεται σε μεγάλο βαθμό με την οικονομική ανάπτυξη. Οι άνθρωποι στις αναπτυγμένες χώρες καταναλώνουν κατά μέσο όρο 50-70 κιλά αλουμινίου ετησίως, ενώ οι άνθρωποι στις αναπτυσσόμενες χώρες (χώρες με κατά κεφαλήν ΑΕΠ μικρότερο από 10000 δολάρια) καταναλώνουν λιγότερο από 20 κιλά ανά έτος, κατά μέσο όρο. Όταν αναδυόμενες χώρες γίνονται πιο αναπτυγμένες, η κατανάλωση αλουμινίου σε αυτές τις χώρες αυξάνεται σημαντικά. Λαμβάνοντας υπόψη το μέγεθος των αναπτυσσόμενων χωρών όπως η Κίνα και η Ινδία, οι αναδυόμενες αγορές θα μπορούσαν να προκαλέσουν αυξήσεις τόσο της ζήτησης του αλουμινίου όσο και στις τιμές.

#### Το Κόστος της Ενέργειας

Το κόστος της ενέργειας μπορεί να επηρεάσει τη διαδικασία παραγωγής του αλουμινίου. Τήξη αλουμίνης σε αλουμίνιο απαιτεί μια συνεχή, μεγάλη παροχή ηλεκτρικού ρεύματος, η οποία αντιπροσωπεύει περίπου το 25% του κόστους της όλης

διαδικασίας τήξης. Αν το κόστος της ενέργειας γίνει πολύ υψηλό, τότε το κόστος αυτό μετατοπίζεται στην τιμή του αλουμινίου.

#### **5.1.2.4 Φυσικό Αέριο**

##### Αντίξοες Καιρικές Συνθήκες Μπορούν να Διαταράξουν την Παραγωγή

Τυφώνες και άλλες σοβαρές καιρικές συνθήκες μπορούν να επηρεάσουν την προσφορά φυσικού αερίου. Για παράδειγμα, το καλοκαίρι του 2005, οι τυφώνες κατά μήκος στις ακτές του Κόλπου των Η.Π.Α. έκλεισαν περίπου το 4% της συνολικής παραγωγής των Η.Π.Α. μεταξύ Αυγούστου 2005 και Ιουνίου 2006.

##### Ο χειμώνας

Ο χειμώνας επηρεάζει σε μεγάλο βαθμό την οικιστική και την εμπορική ζήτηση. Κατά τη διάρκεια των κρύων μηνών, η οικιστική, και η εμπορική χρήση καταναλώνει φυσικό αέριο για θέρμανση, το οποίο θέτει ανοδικές πιέσεις στις τιμές καθώς αυξάνεται η ζήτηση. Εάν παρουσιαστούν απροσδόκητες ή αντίξοες καιρικές συνθήκες, η επίδραση στις τιμές εντείνεται επειδή η παροχή είναι έτοιμη να αντιδράσει γρήγορα στις βραχυπρόθεσμες αυξήσεις της ζήτησης

##### Ο ζεστός καιρός

Ο ζεστός καιρός το καλοκαίρι μπορεί να αυξήσει την ζήτηση των μονάδων ηλεκτροπαραγωγής για φυσικό αέριο. Περίπου το 30% του ηλεκτρισμού των Η.Π.Α. παράγεται από φυσικό αέριο. Υψηλότερες από το κανονικό θερμοκρασίες μπορεί να αυξήσει τη ζήτηση για κλιματισμό το οποίο αυξάνει τη ζήτηση του τομέα της ηλεκτρικής ενέργειας για το φυσικό αέριο και μπορεί να οδηγήσει σε αύξηση των τιμών.

##### Ο Ανταγωνισμός με τα άλλα καύσιμα

Καταναλωτές φυσικού αερίου μεγάλου όγκου (κυρίως βιομηχανικοί καταναλωτές και παραγωγοί ηλεκτρισμού) μπορούν να εναλλάξουν μεταξύ του φυσικού αερίου, του άνθρακα και του πετρελαίου, ανάλογα με τις τιμές του κάθε είδους καυσίμου. Λόγω της αλληλεξάρτησης μεταξύ αυτών των αγορών καυσίμων, όταν οι τιμές των άλλων καυσίμων πέφτουν, κάθε μετατόπιση της ζήτησης από το φυσικό αέριο στον άνθρακα ή στο πετρέλαιο μειώνει τη ζήτηση του φυσικού αερίου και ρίχνει τη τιμή του προς τα κάτω.

#### **5.1.2.5 Χρυσός (Gold)**

Η τιμή του χρυσού εξαρτάται από την αξία του δολαρίου, όπως ακριβώς και στην περίπτωση του ασημιού. Επιπρόσθετα διαμορφώνεται και από τους εξής παράγοντες:

##### Παγκόσμια Κρίση

Επειδή οι τιμές του χρυσού έχουν την τάση να αυξάνονται όταν οι άνθρωποι δεν έχουν εμπιστοσύνη στις κυβερνήσεις ή στις χρηματοπιστωτικές αγορές, συχνά

καλείται ως αγαθό της κρίσης. Τα παγκόσμια γεγονότα έχουν συχνά αντίκτυπο στην τιμή του χρυσού, διότι θεωρείται ως πηγή ασφαλείας εν μέσω οικονομικής ή γεωπολιτικής αναταραχής. Για παράδειγμα η τιμή του χρυσού ανέβηκε αμέσως μετά την εισβολή των Ρώσων στην Ουκρανία, καθώς οι άνθρωποι ένιωθαν την αβεβαιότητα για την γεωπολιτική σταθερότητα στην περιοχή.

#### Επιτόκια

Ο χρυσός δεν καταβάλλει τόκους, όπως τα κρατικά ομόλογα ή οι λογαριασμοί καταθέσεων, αλλά οι τρέχουσες τιμές του χρυσού, συχνά, αντικατοπτρίζουν τις αυξήσεις και τις μειώσεις των επιτοκίων. Καθώς τα επιτόκια αυξάνονται, οι τιμές του χρυσού ενδέχεται να μειωθούν, διότι οι άνθρωποι πωλούν χρυσό για να απελευθερώσουν πόρους για άλλες επενδυτικές ευκαιρίες. Καθώς τα επιτόκια μειώνονται, η τιμή του χρυσού μπορεί να αυξηθεί και πάλι, επειδή υπάρχει ένα χαμηλότερο κόστος για την εκμετάλλευση χρυσού σε σύγκριση με άλλες επενδύσεις. Τα χαμηλά επιτόκια ισοδυναμούν με μεγαλύτερη έλξη για τον χρυσό.

#### Κοσμήματα και Βιομηχανία

Ο χρυσός δεν είναι μόνο πολύτιμος ως αντισταθμιστής κινδύνου αμοιβαίων κεφαλαίων και ως ένα ασφαλές επενδυτικό καταφύγιο. Ο χρυσός χρησιμοποιείται επίσης σε κοσμήματα και στη βιομηχανία. Πάνω από το ήμισυ η ζήτηση του χρυσού αφορά τα κοσμήματα, όπου η Κίνα, η Ινδία και οι Ηνωμένες Πολιτείες είναι τρεις χώρες με τις μεγαλύτερες απαιτήσεις.

Εκτός από κοσμήματα, ένα άλλο δώδεκα τοις εκατό της ζήτησης του χρυσού προέρχεται από τις βιομηχανικές εφαρμογές. Οι κατασκευαστές χρησιμοποιούν το χρυσό σε όλα τα είδη ηλεκτρονικών συσκευών, από υπολογιστές με συστήματα GPS και ιατρικές συσκευές.

#### **5.1.2.6 Χαλκός (Copper)**

Ο χαλκός είναι ένα από τα παλαιότερα μέταλλα που είναι γνωστό στον άνθρωπο. Η ιστορία δείχνει ότι αυτό το μέταλλο χρησιμοποιείται για την κατασκευή κερμάτων για πάνω από 10.000 χρόνια. Σήμερα, ο χαλκός χρησιμοποιείται σε μια ευρεία ποικιλία της βιομηχανίας. Βρίσκουμε χαλκό στα περισσότερα πράγματα από τα καλώδια μέχρι τους αγωγούς για τη θέρμανση και την ψύξη. Ο χαλκός έχει επίσης χρησιμοποιηθεί, ιστορικά, από τους εμπόρους οι οποίοι επιθυμούσαν να λάβουν κέρδος επενδύοντας σε αυτό το μέταλλο. Η τιμή του χαλκού διαμορφώνεται από την αύξηση της ζήτησης των αναδυόμενων χωρών, όπως στην περίπτωση του αλουμινίου. Εδώ, θα εξετάσουμε τους κορυφαίους παράγοντες που επηρεάζουν την τιμή του χαλκού:

#### Προμήθειες και Αποθέματα

Στις 28 Ιουλίου 2014 η Goldman Sachs κυκλοφόρησε ένα ερευνητικό σημείωμα για τα βασικά προϊόντα. Στο σημείωμα αυτό, η προέβλεψε ότι είναι πιθανό να γίνουμε μάρτυρες σε ένα πλεόνασμα της παγκόσμιας αγοράς του χαλκού της τάξης των 350.000 έως 500.000 mt πάνω από το 2014 έως το 2015. Ως αποτέλεσμα, προέβλεψε

ότι η τιμή του χαλκού θα πέσει κάτω από την τιμή εκείνη τη στιγμή, για μια περίοδο άνω των δώδεκα μηνών.

Πλεόνασμα στην αγορά του χαλκού σημαίνει ότι η προμήθεια του χαλκού υπερβαίνει τη ζήτηση του. Το ηθικό δίδαγμα αυτής της ιστορίας για τους εμπόρους είναι ότι τα υψηλά πλεονάσματα συχνά μεταφράζονται σε μια βουτιά στην τιμή του χαλκού, όπως ακριβώς συμβαίνει με τα περισσότερα αγαθά. Από την άλλη πλευρά, κάθε γεγονός που απειλεί την προμήθεια του χαλκού κατά ένα μεγάλο ποσοστό πρόκειται να προκαλέσει αύξηση των τιμών.

### Προοπτικές Ζήτησης του Χαλκού από την Κίνα

Φυσικά, έχουμε ήδη μιλήσει για την Κίνα παραπάνω. Ωστόσο, με το κύρος της οικονομίας της, είναι σημαντικό να εξετάσουμε την περιοχή ξεχωριστά, αντί να δώσουμε μια γενική εξήγηση, διότι για παράδειγμα ειδήσεις από τη χώρα αυτή μπορεί να επηρεάσουν την τιμή του χαλκού προς τα πάνω ή προς τα κάτω. Αυτό γίνεται κατανοητό αν αντιληφθεί κανείς ότι η Κίνα καταναλώνει περίπου το 40% του χαλκού στον κόσμο. Οι ΗΠΑ είναι η δεύτερη μετά την Κίνα στην παγκόσμια κατανάλωση του χαλκού.

Ένας απλός τρόπος για να κατανοήσουμε την επίδραση της Κίνας στην αγορά του χαλκού είναι συγκρίνοντας την Κίνα με το σύνολο της Βόρειας Αμερικής. Σύμφωνα με την Oracle Mining Corp, οι καταναλωτές της Βόρειας Αμερικής χρησιμοποιούν περίπου 10 κιλά χαλκού ανά κάτοικο. Ωστόσο, στην Κίνα, παρά την τεράστια οικονομική ανάπτυξη των τελευταίων ετών, η κατανάλωση χαλκού ανά κάτοικο εξακολουθεί να ανέρχεται σε περίπου 5.4kg. Αυτό δείχνει ότι υπάρχουν τεράστια περιθώρια αύξησης στην κατανάλωση του χαλκού από την Κίνα. Όσο ο λαός της Κίνας αστικοποιείται, χτίζει τις υποδομές του και γίνεται περισσότερο μια καταναλωτική κοινωνία, η κατανάλωση χαλκού κατά πάσα πιθανότητα θα πλησιάσει ή και θα ξεπεράσει τα επίπεδα της Βόρειας Αμερικής. Ο ρυθμός της κατανάλωσης του χαλκού στη χώρα αναμένεται να αυξηθεί κατά 9% ετησίως μέχρι το 2016. Αυτό και μόνο συνοψίζει την επίδραση που θα έχει η οικονομική κατάσταση στην Κίνα επί της τιμής του χαλκού. Δεύτερον, η Κίνα με πληθυσμό άνω του 1,3 δισεκατομμυρίων, σε σύγκριση με τα 530 εκατομμύρια ανθρώπους της Βόρειας Αμερικής, θα πρέπει να περιμένουμε ότι η οποιαδήποτε μικρή αύξηση της ζήτησης από την Κίνα θα μετακινήσει την τιμή του χαλκού σημαντικά.

### Διαταραχές στον Εφοδιασμό

Ως αποτέλεσμα του γεγονότος ότι ένα σημαντικό μέρος των παγκόσμιων αποθεμάτων χαλκού προέρχεται από τη Νότια Αμερική, οι διαταραχές του εφοδιασμού ή ο φόβος ενός τέτοιου ενδεχομένου στην περιοχή μπορεί να επηρεάσει τις παγκόσμιες τιμές σε μεγάλο βαθμό. Για παράδειγμα, γεγονότα όπως απεργία στα ορυχεία και οι φυσικές καταστροφές είναι αρκετά συχνές στην περιοχή. Όλα αυτά τα γεγονότα απειλούν τις προμήθειες από την περιοχή και μπορεί να ωθήσουν τις τιμές υψηλότερα, εάν εμφανιστούν. Θα πρέπει να αναφερθεί, ότι η Χιλή και το Περού μαζί

κατέχουν το περίπου 40% των αποδεδειγμένων αποθεμάτων χαλκού στον κόσμο και για το 42% περίπου της συνολικής παραγωγής χαλκού στον κόσμο.

#### Η Επίδραση των Υποκατάστατων του Χαλκού

Ο τελευταίος βασικός παράγοντας περιλαμβάνει τις δυνατότητες υποκατάστασης του χαλκού. Προκειμένου να μειωθεί το κόστος παραγωγής, οι βιομηχανίες που κάνουν χρήση του χαλκού έχουν επιδιώξει φθηνότερα μέταλλα που μπορούν να χρησιμοποιηθούν στη θέση του. Για παράδειγμα, το αλουμίνιο χρησιμοποιείται τώρα στα καλώδια ρεύματος, στις ηλεκτρικές συσκευές, στα ψυγεία αυτοκινήτων και στις σωλήνες ψύξης. Συνοπτικά, όσο μεγαλύτερη είναι η διαθεσιμότητα στα υποκατάστατα των μετάλλων, τόσο περισσότερη θα είναι η πίεση που ασκείται στη ζήτηση για χαλκό, το οποίο θα κάνει το μέταλλο φθηνότερο.

#### **5.1.2.7 Λευκόχρυσος ή Πλατίνα (Platinum)**

Η τιμή της πλατίνας, εκτός από την ζήτηση για τα κοσμήματα, διαμορφώνεται και από τους εξής παράγοντες:

#### Η Επίδραση της Νότιας Αφρικής

Όπως και σε πολλά άλλα προϊόντα, η παραγωγή της πλατίνας είναι βασισμένη κυρίως στις αναδυόμενες αγορές. Αλλά αυτό το μέταλλο είναι μια ειδική περίπτωση, διότι η παραγωγή της γίνεται σχεδόν αποκλειστικά στη Νότια Αφρική, μια μάλλον ασταθή αναδυόμενη αγορά. Το έθνος είναι σήμερα υπεύθυνο για το περίπου 75% της ετήσιας παραγωγής παγκοσμίως και είναι επίσης αποθηκευμένο το περίπου 95% των αποδεδειγμένων αποθεμάτων του πλανήτη, που τους καθιστά τον μεγαλύτερο παίκτη για αυτό το πολύτιμο μέταλλο. Με μια ασταθή αγορά που εξακολουθεί να προσπαθεί να μπει στο γκρουπ των ανεπτυγμένων χωρών, οι εξελίξεις στη Νότια Αφρική μπορεί να έχουν σημαντικό αντίκτυπο στις τιμές και τις τάσεις της πλατίνας. Οι απεργίες των ορυχείων, οι περίοδοι οικονομικής ύφεσης, και άλλες διαταραχές είναι κρίσιμοι παράγοντες για τους εμπόρους της πλατίνας.

#### Η Αυτοκινητοβιομηχανία

Περίπου το μισό της προσφοράς της πλατίνας πηγαίνει προς τις συσκευές ελέγχου των εκπομπών για τα αυτοκίνητα. Ως καταλύτης επιτρέπει την καύση των άκαυστων υδρογονανθράκων από την εξάτμιση σε διοξείδιο του άνθρακα και υδρατμούς. Επειδή η αυτοκινητοβιομηχανία αντιπροσωπεύει ένα σημαντικό τμήμα της παγκόσμιας ζήτησης πλατίνας, η τιμή για το μέταλλο αυτό μπορεί να επηρεάζεται από την πορεία του τομέα αυτού της οικονομίας. Μεγαλύτερη παραγωγή αυτοκινήτων ισοδυναμεί γενικά μεγαλύτερη ζήτηση για πλατίνα, ενώ οι δύσκολες στιγμές για τη βιομηχανία αυτοκινήτων μπορεί να μειώσει τις τιμές.

#### Υποκατάστατα της Πλατίνας στην Αυτοκινητοβιομηχανία

Για την μείωση του κόστους του αυτοκινήτου, ορισμένοι κατασκευαστές οχημάτων αντικαθιστούν τη πλατίνα με το παλλάδιο, καθώς το δεύτερο είναι φθηνότερο.

Ιστορικά, μόνο ο λευκόχρυσος χρησιμοποιείται σε καταλυτικούς μετατροπείς, αλλά οι νέες τεχνολογίες χρησιμοποιούν περίπου 25% παλλάδιο στις συσκευές. Πειράματα έχουν δείξει ότι το ποσοστό αυτό μπορεί να φθάσει έως και το 50% στο εγγύς μέλλον, περιορίζοντας περαιτέρω τη χρησιμοποίηση της πλατίνας. Θα πρέπει να σημειωθεί ότι ο λευκόχρυσος θεωρείται ευρέως ως το λιγότερο δραστικό μέταλλο για αυτό δεν μπορεί ποτέ να είναι σε θέση να αντικατασταθεί πλήρως, αλλά η μείωση της παρουσίας της στο χώρο της αυτοκινητοβιομηχανίας θα μπορούσε ασφαλώς να ασκήσει πιέσεις για τις τιμές προς τα κάτω.

#### **5.1.2.8 Ρύζι (Rice)**

Όταν συναντάμε τον όρο ακατέργαστο ρύζι, θα πρέπει να έχουμε στο μυαλό μας το ρύζι που προέρχεται από το χωράφι μετά τη συγκομιδή. Το ρύζι ανήκει στη λίστα των κορυφαίων τριών βασικών τροφίμων του κόσμου. Για να καταλάβουμε πόσο σημαντικό είναι το ρύζι για την ανθρωπότητα ακολουθούν ορισμένα στατιστικά στοιχεία. Το 2012, οι χώρες σε όλο τον κόσμο διέθεσαν 162.300.000 εκτάρια γης για καλλιέργεια ρυζιού, η οποία οδήγησε σε συνολική παραγωγή 738.100.000 τόνων ακατέργαστου ρυζιού. Η Κίνα και η Ινδία είναι οι δύο μεγαλύτεροι παραγωγοί ρυζιού στον κόσμο. Εδώ, θα εξετάσουμε τους κορυφαίους παράγοντες που επηρεάζουν την τιμή του ρυζιού.

##### Τα Αποθέματα

Λόγω της τεράστιας σημασίας του ρυζιού στις χώρες σε όλο τον κόσμο, τα αποθέματά του επηρεάζουν τις τιμές σημαντικά. Από το 2009, η συνολική κατανάλωση ρυζιού στον κόσμο, ανήλθε σε 531 600 000 μετρικούς τόνους αναποφλοιώτου ρυζιού. Τα αποθέματα πρέπει να είναι σε υψηλά επίπεδα όλη την ώρα για την ικανοποίηση μιας πιθανής τεράστιας ζήτησης. Ως εκ τούτου, αν σε οποιαδήποτε στιγμή, τα αποθέματα μειωθούν, οι έμποροι αρχίζουν να φοβούνται ότι θα υπάρξει έλλειψη ρυζιού. Όπως θα περίμενε κανείς, θα ακολουθήσει η αύξηση των τιμών. Από την άλλη πλευρά, τα υψηλά αποθέματα τείνουν να οδηγήσουν σε χαμηλότερες τιμές.

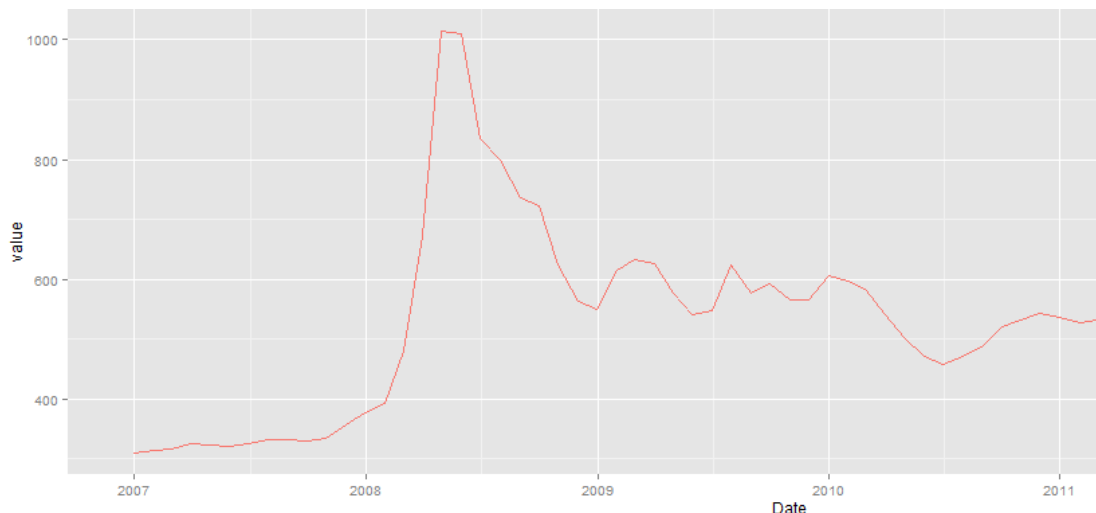
##### Η Ζήτηση στον Τομέα των Βιοκαυσίμων

Κατά τα τελευταία χρόνια, η χρήση του ρυζιού στον τομέα των βιοκαυσίμων έχει αυξηθεί. Εκτός από την ξηρασία, η οποία ήταν ο κύριος παράγοντας για την αύξηση των τιμών την περίοδο 2007-2009, οι αναλυτές πιστεύουν πως οι επιδοτήσεις των ΗΠΑ για το ρύζι της παραγωγής βιοκαυσίμων συνέβαλε στην αύξηση των τιμών. Με την αυξανόμενη ευαισθητοποίηση σχετικά με την υπερθέρμανση του πλανήτη, η ζήτηση για το ρύζι από αυτόν τον τομέα θα μπορούσε να πάει ακόμα υψηλότερα, η οποία θα οδηγήσει στην αύξηση των τιμών. Ως εκ τούτου, οι επενδυτές θα πρέπει να προσέχουν τις τάσεις στον τομέα των βιοκαυσίμων.

##### Τα Κόστη Καυσίμων και Λιπασμάτων

Η καλλιέργεια του ρυζιού απαιτεί, επίσης, σημαντικό ποσό λίπανσης του εδάφους, προκειμένου να μεγιστοποιηθεί η παραγωγή των γεωργικών εκμεταλλεύσεων. Αυτό σημαίνει ότι το κόστος των λιπασμάτων αποτελεί βασικό παράγοντα για την

τιμολόγηση του ρυζιού. Οι υψηλές τιμές των λιπασμάτων θα αύξανε το κόστος της παραγωγής, η οποία με τη σειρά της θα αύξανε την τιμή του ρυζιού. Για παράδειγμα, σύμφωνα με την Investopedia, οι τιμές του ρυζιού την περίοδο 2004-2008 επηρεάστηκαν από την αύξηση των εξόδων για τα καύσιμα και τα λιπάσματα, που οδήγησαν σε διπλασιασμό των τιμών σε μόλις τέσσερα χρόνια.

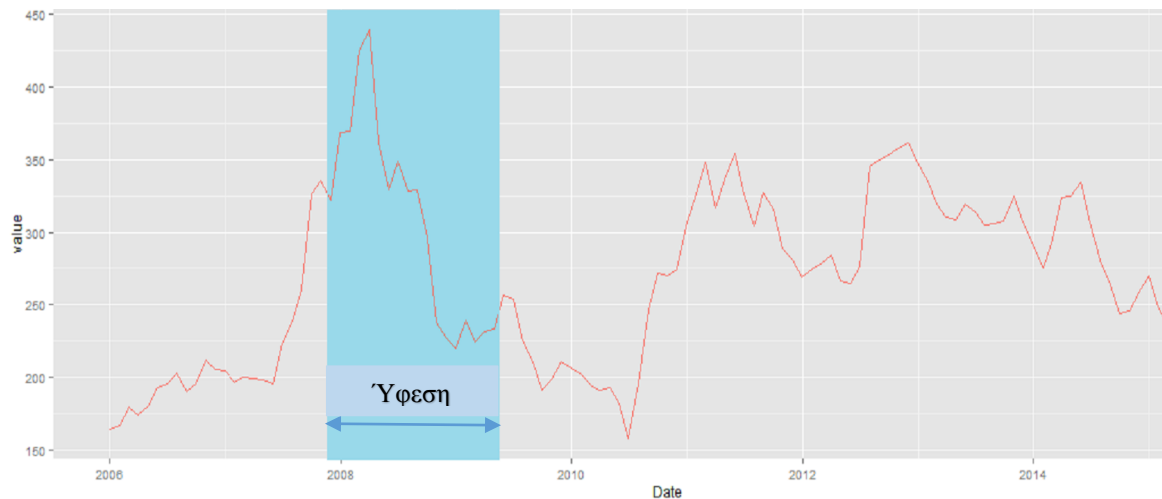


### 5.1.2.9 Σιτάρι (*Wheat*)

Το σιτάρι είναι ένα από τα πιο σημαντικά γεωργικά προϊόντα στον κόσμο. Το εμπόρευμα συνεχίζει να αυξάνει σε δημοτικότητα σε όλο τον κόσμο. Στην πραγματικότητα, το σιτάρι είναι η βασική τροφή στη Βόρεια Αφρική και τη Μέση Ανατολή. Αυτό δείχνει ότι το σιτάρι είναι σημαντικό τόσο στις ανεπτυγμένες χώρες όσο και στις αναδυόμενες αγορές. Λόγω του γεγονότος ότι χρησιμοποιείται ευρέως στην παραγωγή τροφίμων, το σιτάρι έχει τη δυνατότητα να χρησιμοποιηθεί ως μέσο αντιστάθμισης έναντι του πληθωρισμού. Αυτό σημαίνει ότι υπάρχουν μια σειρά από επιλογές που είναι διαθέσιμες στους επενδυτές που επιθυμούν να επωφεληθούν από το εμπόρευμα. Παρακάτω βλέπουμε τους κορυφαίους παράγοντες που επηρεάζουν την τιμή του σιταριού.

#### Η Οικονομία

Η κατάσταση της οικονομίας είναι ένας από τους μεγαλύτερους παράγοντες που επηρεάζουν την τιμή του σίτου. Η σχέση φαίνεται να είναι ανάλογη. Κατά τη διάρκεια αδύναμων οικονομιών, οι τιμές του σίτου έχουν την τάση να μειώνονται. Ωστόσο, όταν η οικονομία αρχίσει να ανακάμπτει, οι τιμές του σίτου ανεβαίνουν σχεδόν αμέσως. Θα μπορούσαν να μειωθούν αργότερα, αλλά πάντα υπήρχε μία άμεση αύξηση. Το παρακάτω διάγραμμα, το οποίο υπογραμμίζει την τελευταία ύφεση, το επιβεβαιώνει. Επιπλέον, για κάθε ύφεση που έχουν περάσει οι Η.Π.Α. τις τελευταίες τρεις δεκαετίες, η τάση αυτή έχει πραγματοποιηθεί. Έτσι, οι επενδυτές θα πρέπει να κρατούν πάντα στοιχεία για τα οικονομικά δεδομένα, όταν θέλουν να προβλέψουν τις τιμές του σιταριού.



### Κλιματικοί Παράγοντες

Όντας το σιτάρι ένα γεωργικό προϊόν, οι καιρικές προβλέψεις θα μπορούσαν ενδεχομένως να επηρεάσουν τις τιμές του σιταριού. Ιστορικά, σε αυτή την κατηγορία, η ξηρασία είχε την πιο σημαντική επίδραση στις τιμές του σιταριού. Όταν υπάρχουν προβλέψεις για επερχόμενη ξηρασία, οι τιμές του σίτου έχουν την τάση να ανεβαίνουν. Αυτό σημαίνει ότι η ξηρασία εμποδίζει την ανάπτυξη και, εν συνεχεία, την παραγωγή του σιταριού. Για παράδειγμα, το 2010 η ξηρασία στη Ρωσία είχε σαν αποτέλεσμα να απαγορευθούν οι εξαγωγές της χώρας, με αποτέλεσμα την αύξηση των τιμών.

#### **5.1.2.10 Καλαμπόκι (Corn)**

Ισχύει ότι την τιμή του καλαμποκιού επηρεάζουν το κλίμα, αλλά και γεωπολιτικά ζητήματα. Για παράδειγμα οι ταραχές στην Ουκρανία που είναι ο πέμπτος μεγαλύτερος παραγωγός ώθησε τις τιμές του καλαμποκιού προς τα πάνω. Άλλος ένας παράγοντας που διαμορφώνει την τιμή του καλαμποκιού είναι, η αιθανόλη. Με όλο και περισσότερη αιθανόλη που προέρχεται από το καλαμπόκι, η αγορά αιθανόλης είναι μια σημαντική μεταβλητή στην τιμολόγηση του καλαμποκιού. Βασικά, η αύξηση της ζήτησης για αιθανόλη πρόκειται να αυξήσει τη ζήτηση για το καλαμπόκι και άρα αύξηση των τιμών του καλαμποκιού.

#### **5.1.2.11 Σόγια (Soybean)**

Γενικά ότι ισχύει για το καλαμπόκι ισχύει και για τη σόγια. Αρχικά ως ένα καλλιεργήσιμο προϊόν επηρεάζεται πολύ από το κλίμα. Στη συνέχεια η σόγια χρησιμοποιείται και εκείνη για την παραγωγή βιοκαυσίμων και άρα γίνεται η ανάλογη συσχέτιση με την τιμή του πετρελαίου.

#### **5.1.2.12 Κακάο (Cocoa)**

Ισχύει ότι την τιμή του κακάο επηρεάζουν το κλίμα, αλλά και γεωπολιτικά ζητήματα. Τα δύο τρίτα της παγκόσμιας παραγωγής κακάο προέρχεται από τη Δυτική Αφρική, με την Ακτή Ελεφαντοστού να είναι ο μεγαλύτερος παραγωγός παγκοσμίως. Κατά τη



διάρκεια των τελευταίων ετών, η χώρα έπρεπε να παλέψει με τις πολιτικές ταραχές, τις ανταρσίες, το φόβο του Ebola, που είναι παράγοντες που μπορούν να μειώσουν την παραγωγή. Σε γενικές γραμμές, πρόκειται για μια περιοχή που η γεωπολιτική αβεβαιότητα είναι κάτι το συνηθισμένο. Επίσης οι υποδομές της χώρας αυτής είναι ελλιπέστατες και η μεταφορά επομένως γίνεται ακριβότερη, και οι πρόσθετες δαπάνες μετατοπίζονται πάνω στο εμπόρευμα.

#### **5.1.2.13 Καφές (Coffee)**

Παράγοντες όπως είναι η γεωπολιτική στις χώρες παραγωγής του καφέ, η έλλειψη υποδομών που παρατηρούνται σε αυτές τις χώρες, αλλά και η γενικότερη ζήτηση και παραγωγή των κόκκων καφέ είναι παράγοντες που επηρεάζουν την τιμή του.

#### **5.1.2.14 Βαμβάκι (Cotton)**

Και εδώ παράγοντες όπως είναι τα παγκόσμια αποθέματα, η ζήτηση των χωρών, αλλά και το κλίμα επηρεάζουν την διαμόρφωση των τιμών του βαμβακιού. Επειδή ο τρόπος είναι πανομοιότυπος με τα προηγούμενα καλλιεργήσιμα αγαθά, δεν θα αναλυθεί περαιτέρω.

#### **5.1.2.15 Κάρβουνο (Coal)**

Εδώ παράγοντες που θεωρούνται κυρίαρχοι στην διαμόρφωση των τιμών του κάρβουνου είναι η προσφορά και η ζήτηση. Ακόμη όμως ένας παράγοντας που φαίνεται να επηρεάζει πάρα πολύ την τιμή του κάρβουνου, είναι οι διάφοροι περιοριστικοί νόμοι που αφορούν τη χρήση του. Αυτό συμβαίνει διότι ως γνωστό η καύση του συγκεκριμένου ορυκτού δημιουργεί αρκετά περιβαλλοντικά προβλήματα και για αυτό γίνεται προσπάθεια χρήσης διαφορετικών μορφών ενέργειας, πιο φιλικές στο περιβάλλον.

## **5.2 Ανάλυση των δεδομένων των τιμών των μετοχών (Stocks)**

Μετοχή είναι ένα μερίδιο στην ιδιοκτησία μιας εταιρείας. Ουσιαστικά αντιπροσωπεύει την απαίτηση για τα περιουσιακά στοιχεία και τα κέρδη της εταιρείας. Όσο περισσότερες μετοχές κατέχει κάποιος, τόσο πολύ αυξάνει το ποσοστό ιδιοκτησίας στην εταιρεία. Είτε λέμε μετοχές, μετοχικό κεφάλαιο ή μερίσματα, όλα σημαίνει το ίδιο πράγμα.

### **Διαμόρφωση της τιμής των μετοχών**

Οι χρηματιστηριακές αγορές μπορεί να είναι ευμετάβλητες, και οι λόγοι, που οι μετοχές ανεβαίνουν και πέφτουν, είναι πολύπλοκοι. Επίσης, οι τιμές των μετοχών επηρεάζονται από διάφορους παράγοντες και γεγονότα, ορισμένοι από τους οποίους επηρεάζουν άμεσα τις τιμές των μετοχών και άλλοι που το κάνουν έμμεσα. Τέτοιοι παράγοντες συνήθως είναι:

### Παράγοντες στο εσωτερικό της εταιρίας

Οι εξελίξεις που μπορεί να προκύψουν στο πλαίσιο των επιχειρήσεων θα επηρεάσουν την τιμή της μετοχής της. Τέτοιες εξελίξεις μπορεί να είναι οι συγχωνεύσεις και εξαγορές, οι εκθέσεις των κερδών της, η ανάπτυξη ή η έγκριση ενός νέου καινοτόμου προϊόντος, η πρόσληψη ή η απόλυση των στελεχών της εταιρίας και τέλος, οι καταγγελίες για απάτη ή αμέλεια . Όπως είναι λογικό οι κινήσεις των τιμών των μετοχών είναι πιο έντονες όταν οι εξελίξεις στο εσωτερικό της εταιρίας είναι απροσδόκητες.

### Παγκόσμια Γεγονότα

Οι τιμές των μετοχών μίας εταιρίας και γενικά, η χρηματιστηριακή αγορά μπορεί να επηρεαστούν από παγκόσμια γεγονότα, όπως είναι ο πόλεμος, οι πολιτικές αναταραχές, οι φυσικές καταστροφές και η τρομοκρατία. Αυτές οι επιδράσεις μπορεί να είναι άμεσες ή έμμεσες, και συχνά εμφανίζονται με αλυσιδωτές αντιδράσεις. Η κοινωνική αβεβαιότητα και ο φόβος που δημιουργήθηκε από τις τρομοκρατικές επιθέσεις της 11ης Σεπτεμβρίου 2001, επηρέασαν άμεσα τις αγορές, δεδομένου ότι πολλοί επενδυτές στις Ηνωμένες Πολιτείες άρχισαν να συναλλάσσονται λιγότερο και επικεντρώθηκαν στις μετοχές και τα ομόλογα μικρότερου επενδυτικού κινδύνου. Ένα παράδειγμα έμμεσης επιρροής επί των αγορών είναι η ανακοίνωση μιας νέας στρατιωτικής επιχείρησης από μια χώρα σε απάντηση στο ξέσπασμα της εμφύλιας διαμάχης ή σύγκρουσης στο εξωτερικό. Η ανακοίνωση αυτή πιθανότατα θα προκαλέσει η τιμή των μετοχών του στρατιωτικού εξοπλισμού και των κατασκευαστριών όπλων να αυξηθεί λόγω της αναμενόμενης αύξησης των αμυντικών συμβάσεων, το οποίο με τη σειρά του μπορεί να αυξήσει την αξία των αποθεμάτων για τις εταιρείες που προμηθεύουν στρατιωτικό τμήμα και τεχνολογικό εξοπλισμό. Επομένως, είναι πιθανό να αυξηθεί η ζήτηση, και συνεπακόλουθα η τιμή, των φυσικών πόρων που χρησιμοποιούνται για την δημιουργία των στρατιωτικών υλικών, κάτι το οποίο μπορεί να αυξήσει την τιμή των μετοχών των εταιριών εξόρυξης και επεξεργασίας φυσικών πόρων.

### Πληθωρισμός και επιτόκια

Μία από τις πιο προβλέψιμες επιρροές της χρηματιστηριακής αγοράς είναι οι περιοδικές αναπροσαρμογές των επιτοκίων για την καταπολέμηση του πληθωρισμού. Όταν τα επιτόκια ανεβαίνουν, πολλοί επενδυτές πωλούν ή συναλλάσσονται αύξηση των μετοχών που ενέχουν μεγαλύτερο κίνδυνο για να αποκτήσουν τίτλους εξασφαλισμένους από την κυβέρνηση, όπως τα ομόλογα ώστε να επωφεληθούν από τα υψηλότερα επιτόκια που δίνουν και για να εξασφαλίσουν ότι οι επενδύσεις τους προστατεύονται.

### Συναλλαγματικές ισοτιμίες

Οι ισοτιμίες συναλλάγματος έχουν άμεση επίπτωση στην τιμή και την αξία των μετοχών σε χώρες του εξωτερικού, και οι μεταβολές στις συναλλαγματικές ισοτιμίες αυξάνει ή μειώνει το κόστος της επιχειρηματικής δραστηριότητας σε μια χώρα, η

οποία με τη σειρά της επηρεάζει την τιμή των μετοχών των επιχειρήσεων που δραστηριοποιούνται στο εξωτερικό. Ενώ οι μακροπρόθεσμες κινήσεις στις συναλλαγματικές ισοτιμίες επηρεάζονται από τις θεμελιώδεις δυνάμεις της αγοράς της προσφοράς και της ζήτησης και της τιμής αγοράς του συναλλάγματος, οι βραχυπρόθεσμες κινήσεις οδηγούνται από τα νέα, τις εξελίξεις και τις συναλλαγές συμβολαίων μελλοντικής εκπλήρωσης, τα οποία είναι δύσκολο να προβλεφθούν.

### Δημοσιότητα

Οι μετοχές και το χρηματιστήριο μπορούν να επηρεαστούν από τη διαφημιστική εκστρατεία για μια εταιρεία ή την κυκλοφορία νέων προϊόντων ή υπηρεσιών. Πολλοί άνθρωποι και οργανισμοί έχουν συμφέρον να προωθήσουν συγκεκριμένες μετοχές και οι βιομηχανίες να αυξήσουν την αξία των δικών τους μετοχών και των κερδών. Για αυτό το λόγο, οι θετικές οικονομικές εκθέσεις, τα ενημερωτικά δελτία του χρηματιστηρίου, τα blogs του Διαδικτύου, τα δελτία τύπου και τα δελτία ειδήσεων μπορούν να χτίσουν υψηλές προσδοκίες για τις επιδόσεις των εταιρειών, κάτι το οποίο οδηγεί στην αύξηση της τιμής των μετοχών τους.

Όπως είναι ήδη γνωστό, πάρα πολλοί επενδυτές ακόμα και απλοί άνθρωποι ασχολούνται με την αγορά του χρηματιστηρίου, διότι το βλέπουν σαν μία μορφή επένδυσης που μπορεί να τους αποφέρει κάποια κέρδη. Βεβαίως όπως φάνηκε και από τα παραπάνω, θα πρέπει να υπάρχει συστηματική παρακολούθηση πολλών γεγονότων που μπορεί να επηρεάσουν τις τιμές των μετοχών. Επιπρόσθετα, ίσως θα ήταν συνετό να υπάρχει η ιστορική παρακολούθηση των μετοχών μίας εταιρίας και να βλέπει ο επενδυτής πως αντιδράει η τιμή της μετοχής στις διακυμάνσεις της οικονομίας. Επομένως, η υλοποίηση της εφαρμογής βοηθάει στην καταγραφή των ιστορικών δεδομένων, αλλά και στην επιλογή οποιασδήποτε μετοχής θέλει ο χρήστης να μελετήσει. Βέβαια, το σημαντικότερο είναι πως δίνει την επιλογή της πρόβλεψης της τιμής της μετοχής παρακολουθώντας τα ιστορικά γεγονότα. Γενικά, προτείνεται ο επενδυτής να λαμβάνει υπόψιν του τις σημειακές προβλέψεις, αλλά και διάφορα ειδικά γεγονότα τα οποία μπορεί να επηρεάσουν την τιμή των μετοχών βραχυπρόθεσμα.

## **5.3 Ανάλυση των Δεδομένων του Τουρισμού (Tourism)**

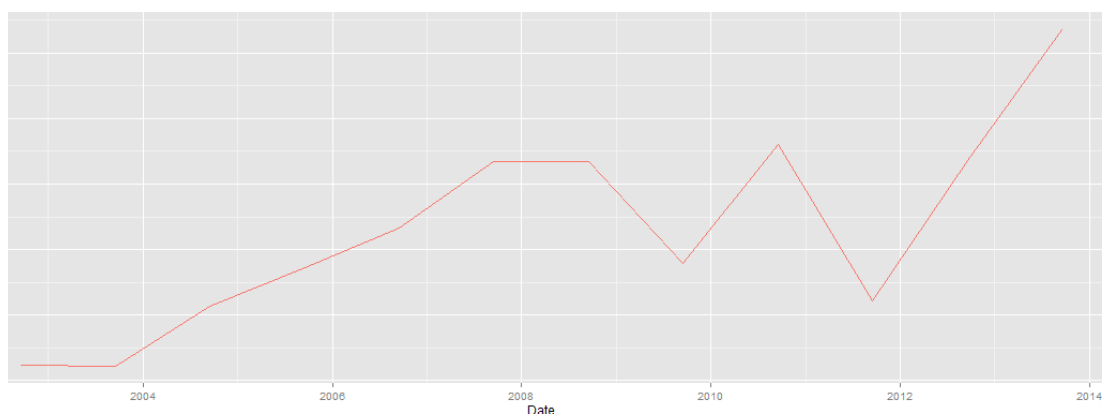
Η σημασία του τουρισμού αυξάνεται συνεχώς για την παγκόσμια οικονομία και τις διεθνείς σχέσεις. Παρά τους πολέμους, τις πολιτικές αναταραχές, τις φυσικές καταστροφές, τις τρομοκρατικές επιθέσεις, καθώς και την οικονομική και ενεργειακή κρίση σε διάφορα μέρη του κόσμου, το διεθνές εμπόριο υπηρεσιών στον τομέα του τουρισμού έχει αυξηθεί θεαματικά από το 1970. Το 2012, οι διεθνείς αφίξεις τουριστών ανήλθαν σε όλο τον κόσμο στους 1.035 δισεκατομμύρια. Ελαφρώς πάνω από τους μισούς ήταν για ταξίδια αναψυχής. Συγκριτικά, υπήρχαν μόνο 166 εκατομμύρια διεθνείς αφίξεις τουριστών σε όλο τον κόσμο το 1970. Οι επισκέπτες

του 2012 δαπάνησαν ένα τρισεκατομμύριο δολάρια για τα ταξίδια (εξαιρουμένων των διεθνών εξόδων μεταφοράς επιβατών τα οποία αποτιμώνται σε 213 δισεκατομμύρια δολάρια). Είτε με άμεσο είτε με έμμεσο τρόπο, αυτές οι δαπάνες αντιπροσωπεύουν το εννέα τοις εκατό του παγκόσμιου ΑΕΠ και το έξι τοις εκατό των εξαγωγών της.

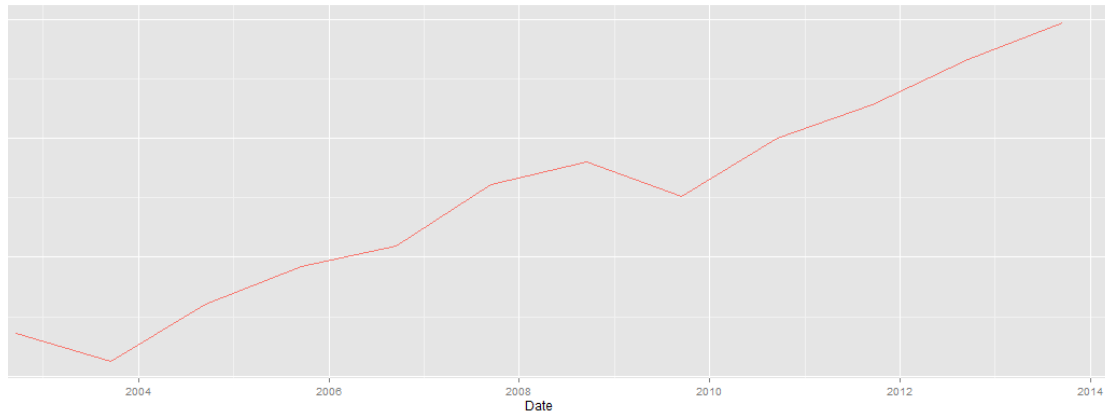
Οι χώρες και οι προορισμοί που εξαρτώνται από τον τουρισμό, το μερίδιο του τουρισμού στο ΑΕΠ μπορεί να υπερβαίνει το διπλάσιο του παγκόσμιου μέσου όρου. Σήμερα, οι διεθνείς τουριστικές εισπράξεις υπερβαίνουν το 1 δισεκατομμύριο δολάρια ετησίως σε περίπου 90 χώρες. Σε όλο τον κόσμο, ο εσωτερικός τουρισμός είναι συνήθως αρκετές φορές μεγαλύτερος. Ο τουρισμός έχει γίνει πραγματικά μια παγκόσμια οικονομική και κοινωνική δύναμη.

Κάποτε υπήρχε η υπόθεση ότι μόνο οι αναπτυσσόμενες χώρες επιδίωκαν να έχουν τον τουρισμό ως βασική αναπτυξιακή στρατηγική. Για παράδειγμα, η Ιαπωνία και οι Ηνωμένες Πολιτείες, έδιναν ιστορικά λίγη σημασία στην προσέλκυση των τουριστών. Αυτό φυσικά δεν ισχύει πλέον. Πρόσφατα, οι δύο χώρες εφάρμοσαν αλλαγές στην πολιτική τους και χαλάρωσαν το καθεστώς έκδοσης βίζας για την προώθηση των εισερχόμενων ταξιδιών από το εξωτερικό, τη δημιουργία θέσεων εργασίας, ώστε να τονωθούν οι υποτονικές οικονομίες τους.

Το 2013, η Ιαπωνία φιλοξένησε 10 εκατομμύρια ξένους επισκέπτες, το οποίο είναι ρεκόρ όλων των εποχών. Η Ιαπωνία ελπίζει να διπλασιάσει τον αριθμό αυτό μέχρι Ολυμπιακούς Αγώνες του Τόκιο το 2020, και να φθάσει του 30 εκατομμύρια επισκέπτες μέχρι το 2030. Οι Ηνωμένες Πολιτείες, αν και θεωρείται ως ένα έθνος εχθρικό προς τους ξένους επισκέπτες λόγω των αυστηρών κανονισμών εισόδου στη χώρα, προσπαθεί σκληρά για να βελτιώσει τη διεθνή εικόνα της. Το 2012 πέτυχε ένα ρεκόρ με 67 εκατομμύρια διεθνείς αφίξεις επισκεπτών, και ο Πρόεδρος Ομπάμα έχει θέσει ως στόχο να προσελκύσει 100 εκατομμύρια ξένους επισκέπτες μέχρι το 2021. Αυτή η αύξηση στις αφίξεις των τουριστών φαίνεται και από τα ακόλουθα διαγράμματα.



Αφίξεις τουριστών στην Ιαπωνία



### Αφίξεις Τουριστών στις Η.Π.Α.

Ως γνωστόν τα ταξίδια είναι δαπανηρά. Ιστορικά, μόνο οι πλούσιοι ιδιώτες μπορούσαν να αντέξουν οικονομικά ένα ταξίδι στο εξωτερικό, και έτειναν να ταξιδέψουν σε αναπτυγμένες χώρες με τουριστική υποδομή και υπηρεσίες ποιότητας. Δεν αποτελεί έκπληξη, πως η Ευρώπη και η Βόρεια Αμερική ήταν οι μεγαλύτερες πηγές και αποδέκτες των διεθνών τουριστών. Αλλά αυτό, επίσης, αλλάζει. Τις τελευταίες δεκαετίες, οι αφίξεις τουριστών στις αναπτυσσόμενες χώρες έχουν αυξηθεί πολύ πιο γρήγορα από ό, τι στις ανεπτυγμένες. Η περιοχή της Ασίας-Ειρηνικού έχει και θα συνεχίσει να έχει, την ταχύτερη ανάπτυξη. Ο Παγκόσμιος Οργανισμός Τουρισμού των Ηνωμένων Εθνών (United Nations World Tourism Organization-UNWTO) προβλέπει ότι οι διεθνείς τουριστικές αφίξεις θα αυξηθούν κατά 3,3 τοις εκατό ετησίως μεταξύ του 2010 και του 2030 και θα φτάσουν συνολικά οι αφίξεις τις 1,8 δισεκατομμύρια μέχρι το 2030. Η αύξηση στις αναπτυσσόμενες χώρες αναμένεται να είναι δύο φορές πιο γρήγορη από ότι στις προηγμένες. Το μερίδιο αγοράς του τουρισμού στις αναδυόμενες χώρες αναμένεται να ανέλθει στο 57 τοις εκατό μέχρι το 2030, σε σύγκριση με το 47 τοις εκατό του 2012. Ωστόσο, η πρόκληση για το πώς πρέπει να κατευθυνθούν και να κατανοηθούν τα οικονομικά οφέλη του τουρισμού στις φτωχότερες περιοχές και πληθυσμούς του κόσμου παραμένει.

Αν και η αυξανόμενη ευημερία και η πτώση του πραγματικού κόστους του ταξιδιού ήταν κύριοι λόγοι για την αύξηση των διεθνών ταξιδιών μετά το 1970, αυτό που έπαιξε καταλυτικό ρόλο ήταν οι αλλαγές στις πολιτικές των κυβερνήσεων σχετικά με τα ταξίδια. Η Ιαπωνία αποτελεί ένα χαρακτηριστικό παράδειγμα αυτού του φαινομένου. Οι πολίτες της Ιαπωνίας απαγορευόταν να ταξιδεύουν στο εξωτερικό για αναψυχή μέχρι τους Ολυμπιακούς Αγώνες του Τόκιο το 1964, και αργότερα υπήρχαν νομισματικοί περιορισμοί για τα ταξίδια στο εξωτερικό, κάτι το οποίο παρέμενε σε ισχύ μέχρι τα τέλη του 1970. Για πολλές αναπτυσσόμενες χώρες, συμπεριλαμβανομένου της Ιαπωνίας, που απαγορεύουν τα ταξίδια αναψυχής στο εξωτερικό είχε ως στόχο τη διατήρηση της ανεπάρκειας ξένου συναλλάγματος που είναι απαραίτητο για τη χρηματοδότηση της εκβιομηχάνισης. Τελικά, με μεγάλα και αυξανόμενα εμπορικά πλεονάσματα, η ελευθέρωση του εξερχόμενου ταξιδιού

συνέβαλε στην αποκλιμάκωση των διεθνών πολιτικών εντάσεων που προέκυψαν από τις εμπορικές ανισορροπίες.

Για αυτούς τους λόγους, ο τουρισμός έχει γνωρίσει τεράστια άνθιση και όπως αναφέρεται και παραπάνω, έχει ένα σημαντικό μερίδιο της παγκόσμιας οικονομίας. Ειδικά ορισμένες αναπτυσσόμενες χώρες βασίζουν την οικονομία τους πάνω στον τουρισμό. Επομένως, ο τουρισμός έχει γίνει αντικείμενο μελέτης πολλών ερευνών, όπου όπως σε κάθε έρευνα είναι απαραίτητη η μελέτη των ιστορικών δεδομένων σχετικά με τον τουρισμό. Άρα η εφαρμογή που υλοποιήθηκε μπορεί να αποτελέσει ένα πολύτιμο εργαλείο για κάποιον που θέλει να αναζητήσει τα ιστορικά δεδομένα των αφίξεων των τουριστών σε μία χώρα. Και αυτό διότι, παρέχει μία μεγάλη ευκολία σχετικά με την αναζήτηση, αφού το μόνο που χρειάζεται είναι η πληκτρολόγηση του κωδικού της χώρας που θέλει να μελετήσει, κάτι το οποίο γλιτώνει χρόνο από την όποια αναζήτηση σε βάσεις δεδομένων ή άλλους ιστότοπους. Πρέπει να τονιστεί πως με το που γίνουν διαθέσιμα καινούργια ιστορικά δεδομένα από την Παγκόσμια Τράπεζα, θα ανανεωθούν αυτόματα και στο πλαίσιο της εφαρμογής.

## **5.4 Ανάλυση της Ανεργίας (Unemployment)**

Το εθνικό ποσοστό ανεργίας ορίζεται ως το ποσοστό των ανέργων στο σύνολο του εργατικού δυναμικού. Είναι ευρέως αναγνωρισμένος ως ο βασικός δείκτης των επιδόσεων της αγοράς εργασίας και δίνεται από τον τύπο:

$$\text{Ποσοστό της Ανεργίας} = \frac{\text{Αριθμός Ανέργων}}{\text{Συνολικό Εργατικό Δυναμικό}} \cdot 100$$

Είναι ένας οικονομικός δείκτης ο οποίος παρακολουθείται στενά, και επίσης το ποσοστό ανεργίας προσελκύει τη μεγάλη προσοχή των μέσων ενημέρωσης, ειδικά σε περιόδους ύφεσης και σε δύσκολες οικονομικές περιόδους.

### **5.4.1 Προσδιορισμός της ανεργίας**

Αρχικά είναι πολύ σημαντικό να γίνει διάκριση μεταξύ του ποσοστού των ανθρώπων που είναι άνεργοι και όσων απλά δεν εργάζονται. Μερικοί άνθρωποι μπορεί να είναι στο σχολείο με πλήρες ωράριο, να εργάζονται από το σπίτι, να είναι άτομα με ειδικές ανάγκες, ή συνταξιούχοι. Αυτοί οι άνθρωποι δεν θεωρούνται μέρος του εργατικού δυναμικού και, επομένως, δεν περιλαμβάνονται στο ποσοστό ανεργίας. Μόνο αυτοί οι άνθρωποι, οι οποίοι αναζητούν ενεργά εργασία ή περιμένουν να επιστρέψουν στην εργασία θεωρούνται άνεργοι.

Οι οικονομολόγοι γενικά διακρίνουν την ανεργία μεταξύ τριών διαφορετικών τύπων της:

- Η Ανεργία τριβής (Frictional Unemployment) προκύπτει λόγω της ακατάπαυστης κίνησης των ανθρώπων μεταξύ περιοχών και θέσεων εργασίας ή διαφόρων σταδίων του κύκλου ζωής. Επίσης, περιλαμβάνονται τα άτομα που, λόγω φυσικών αιτιών, καθίστανται ανίκανα προς εργασία. Ουσιαστικά, αποτελεί το ελάχιστο επίπεδο ανεργίας που δεν μπορεί να μειωθεί σε μια δυναμική κοινωνία (εφ' όσον, ακόμη και αν η οικονομία είναι σε πλήρη απασχόληση, κάποιοι θα αποχωρούν από τις θέσεις τους, κάποιοι φοιτητές θα ψάχνουν για εργασία μετά την αποφοίτησή τους, κάποιοι γονείς θα επανεπεντάσσονται στο εργατικό δυναμικό, και κάποιοι Να θέλουν μια καλύτερη ή μια διαφορετική εργασία). Επίσης αυτός ο τύπος ανεργίας μπορεί να αυξηθεί όταν οι παροχές στους ανέργους είναι τόσο ελκυστικές που μερικοί προτιμούν να μην εργάζονται.
- Η Δομική ή Διαρθρωτική ανεργία (Structural Unemployment) οφείλεται στις αναντιστοιχίες μεταξύ προσφοράς και ζήτησης, καθώς κάποιοι κλάδοι (ή ειδικότητες, ή γεωγραφικές περιοχές) έχουν αυξανόμενη ζήτηση και κάποιοι άλλοι πτωτική ζήτηση. Επίσης συμβαίνει όταν μεταβαλλόμενες αγορές ή νέες τεχνολογίες κάνουν τις δεξιότητες των εργαζομένων να μοιάζουν με παρωχημένες.
- Η Κυκλική ανεργία (Cyclical Unemployment) οφείλεται στην μειωμένη συνολική ζήτηση προϊόντων και υπηρεσιών. Αυτή η μορφή ανεργίας είναι αποτέλεσμα της επιβράδυνσης της ανάπτυξης σε περιόδους οικονομικής ύφεσης, η οποία έχει ως επακόλουθο την αδυναμία απορρόφησης του εργατικού δυναμικού. Είναι αποτέλεσμα της κυκλικής φύσης της οικονομίας και εμφανίζεται κάθε φορά που υπάρχει μια γενικευμένη κάμψη της επιχειρηματικής δραστηριότητας.

Οι άνεργοι συνήθως εμπίπτουν σε μία από τις ακόλουθες τέσσερις κατηγορίες. Στη πρώτη κατηγορία ανήκουν οι άνθρωποι που απολύθηκαν ή πρόκειται να απολυθούν, είτε προσωρινά είτε μόνιμα. Από την άλλη πλευρά, στην δεύτερη κατηγορία ανήκουν άτομα που εγκατέλειψαν τη δουλειά τους ή εργαζόμενοι που έφυγαν εθελοντικά από την εργασία τους. Το μέγεθος αυτής της κατηγορίας είναι προφανές ότι μπορεί να αντανακλά την πραγματική εμπιστοσύνη στην κατάσταση της οικονομίας. Στην Τρίτη κατηγορία ανήκουν οι νεοεισερχόμενοι που αναζητούν εργασία για πρώτη φορά. Και, τέλος, στη τέταρτη κατηγορία ανήκουν οι επανεισερχόμενοι, οι οποίοι είναι οι εργαζόμενοι που εγκατέλειψαν το εργατικό δυναμικό για ένα χρόνο και τώρα επιστρέφουν, όπως είναι για παράδειγμα οι γονείς που επέλεξαν να επικεντρωθούν στις οικογένειες τους ή εκείνοι που έφυγαν για να αποκτήσουν μία πρόσθετη εκπαίδευση σε σχέση με αυτήν που έχουν.

Ως ένα επίπεδο η ανεργία θα είναι πάντα παρούσα σε μια οικονομία όπου οι βιομηχανίες μεγαλώνουν και συρρικνώνονται, καθώς συμβαίνουν τεχνολογικές εξελίξεις, καθώς οι νέες γενιές εισέρχονται στην αγορά εργασίας, και καθώς οι εργαζόμενοι φεύγουν εθελοντικά από την εργασία τους για την αναζήτηση καλύτερων ευκαιριών. Αυτός είναι ο λόγος που οι περισσότεροι οικονομολόγοι συμφωνούν ότι υπάρχει ένα φυσικό ποσοστό της ανεργίας στην οικονομία (συνήθως

4% -6%). Αυτό το φυσικό ποσοστό επηρεάζεται περισσότερο από τον αριθμό των νεαρών εργαζομένων στο εργατικό δυναμικό, καθώς επίσης και οι δημόσιες πολιτικές που αποθαρρύνουν την απασχόληση ή τη δημιουργία θέσεων εργασίας, όπως είναι ο υψηλός κατώτατος μισθός, τα γενναιόδωρα επιδόματα ανεργίας, και μερικά αντικίνητρα που συνδέονται με την απόλυση εργαζομένων.

#### **5.4.2 Σημασία της καταγραφής της ανεργίας**

Η απασχόληση είναι, συνήθως, η κύρια πηγή του εισοδήματος ενός ανθρώπου και έχει σημαντική επίδραση στις καταναλωτικές δαπάνες και γενικά στην οικονομική ανάπτυξη. Έτσι, το ποσοστό ανεργίας, το οποίο είναι ένας δείκτης υστέρησης της οικονομίας, μπορεί να παρέχει σημαντικές πληροφορίες σχετικά με την κατάσταση της οικονομίας ή την οικονομική υγεία των επιμέρους τομέων των επιχειρήσεων. Για παράδειγμα, η υψηλή ανεργία γενικά δείχνει ότι η οικονομία είναι υποτονική ή έχει μία πτώση το Ακαθάριστο Εγχώριο Προϊόν. Αντίθετα, χαμηλή ανεργία ή μείωση της ανεργίας μπορεί να αντικατοπτρίζει μια αναπτυσσόμενη οικονομία. Ταυτόχρονα, τα δεδομένα της ανεργίας μπορούν επίσης να αναφέρουν αλλαγές σε ορισμένες βιομηχανίες. Για παράδειγμα, η αύξηση των θέσεων εργασίας στο τομέα των κατασκευών θα μπορούσε να σηματοδοτήσει την έναρξη για ανέγερση νέων κατοικιών. Για τους λόγους αυτούς, το ποσοστό ανεργίας είναι ένας από τους οικονομικούς δείκτες που παρακολουθούνται στενά και ευρέως.

Παρόλο που υπάρχει μία διαμάχη όσον αφορά τις διαφορετικές μεθοδολογίες που χρησιμοποιούνται για τη μέτρηση της ανεργίας, σχεδόν όλοι αναγνωρίζουν τη σημασία της. Επειδή τα στατιστικά δεδομένα, τα οποία σχετίζονται με την ανεργία, παρακολουθούνται πολύ στενά και αρκετοί βασίζονται σε αυτά. Για αυτό το λόγο, οι διαφορές μεταξύ του προσδοκώμενου ρυθμού ανεργίας και του πραγματικού ποσοστού μπορεί να έχει ένα ευρύ αντίκτυπο στις αγορές κινητών αξιών. Μία απροσδόκητα χαμηλή ανεργία θα μπορούσε να παρακινήσει την Κεντρική Τράπεζα να αυξήσει τα επιτόκια, προκειμένου να αποφευχθεί μία «υπερθέρμανση» της οικονομίας και αντίστροφα.

Επομένως, η ανεργία, και συγκεκριμένα το ποσοστό της ανεργίας παίζει πολύ σημαντικό ρόλο για την οικονομία. Όπως αναφέρθηκε και παραπάνω το ποσοστό της ανεργίας αποτελεί ουσιαστικά έναν καθρέφτη, ο οποίος αντανακλάει την κατάσταση της οικονομίας. Άρα, όπως είναι λογικό η ανεργία είναι αντικείμενο μελέτης είτε για κυβερνητικούς λόγους, είτε για επιχειρηματικούς σκοπούς, είτε για ερευνητικούς σκοπούς και επομένως είναι απαραίτητη η μελέτη των ιστορικών δεδομένων σχετικά με την ανεργία. Αυτό συμβαίνει, διότι όπως αναφέρθηκε παραπάνω η ανεργία μπορεί να έχει μία κυκλική ροή και άρα η παρακολούθηση των ιστορικών δεδομένων μπορεί να δώσει πολύτιμες πληροφορίες. Συνεπώς, η εφαρμογή που υλοποιήθηκε μπορεί να αποτελέσει ένα πολύτιμο εργαλείο για κάποιον που θέλει να αναζητήσει τα ιστορικά δεδομένα του ποσοστού της ανεργίας σε μία χώρα. Και αυτό διότι, όπως και στην καρτέλα του τουρισμού, παρέχει μία μεγάλη ευκολία σχετικά με την



αναζήτηση των δεδομένων. Αυτό συμβαίνει, διότι το μόνο που απαιτείται είναι η πληκτρολόγηση του κωδικού της χώρας που κάποιος θέλει να μελετήσει. Θα πρέπει και εδώ να αναφερθεί πως με το που γίνουν διαθέσιμα καινούργια ιστορικά δεδομένα από την Παγκόσμια Τράπεζα, θα ανανεωθούν αυτόματα και στην εφαρμογή.



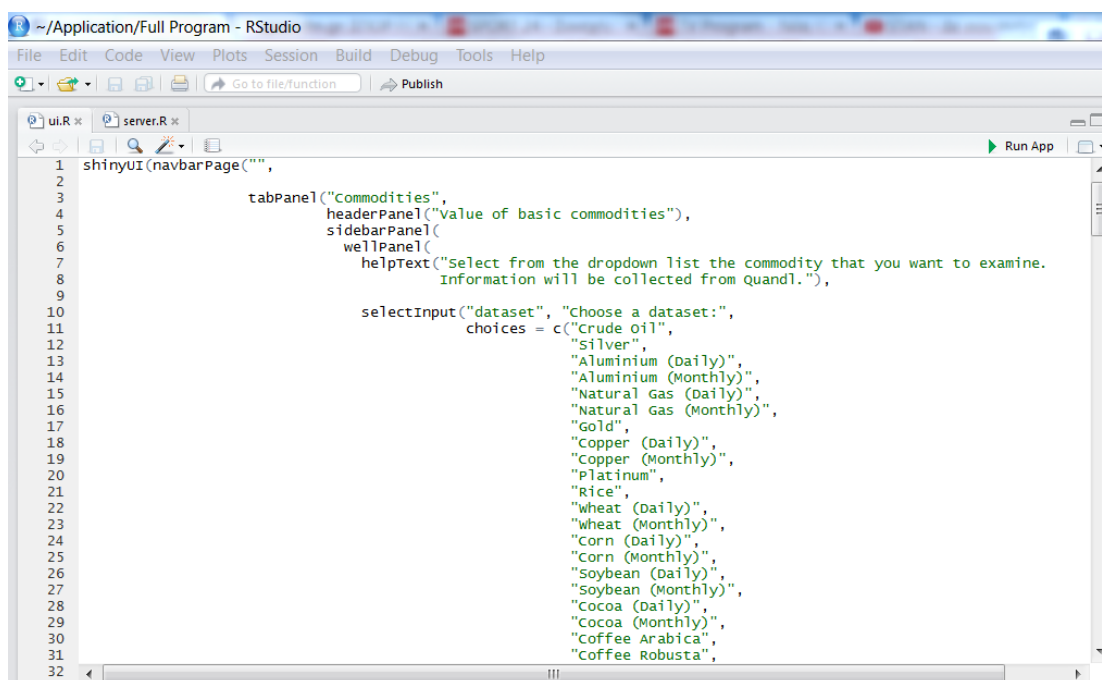
## 6. Εγχειρίδιο (Manual) Εφαρμογής

### 6.1 Είσοδος στην Εφαρμογή και Αρχική Σελίδα

#### 6.1.1 Είσοδος στην Εφαρμογή

Η είσοδος στο πρόγραμμα μπορεί να γίνει με τους εξής τρόπους:

- Ένας τρόπος είναι μέσω του περιβάλλοντος της *R*, το *R Studio* καταχωρώντας τον κώδικα στις ανάλογες επιλογές *ui.R* και *server.R* που έχουν αναλυθεί στο 4<sup>ο</sup> κεφάλαιο, όπως φαίνεται και στην παρακάτω εικόνα. Κατόπιν πατώντας την επιλογή Run App ανοίγει η εφαρμογή.



```
1 shinyUI(navbarPage("",
2
3     tabPanel("Commodities",
4         headerPanel("Value of basic commodities"),
5         sidebarPanel(
6             wellPanel(
7                 helpText("select from the dropdown list the commodity that you want to examine.
8                     Information will be collected from Quandl."),
9
10                selectInput("dataset", "Choose a dataset:",
11                    choices = c("Crude Oil",
12                        "Silver",
13                        "Aluminium (Daily)",
14                        "Aluminium (Monthly)",
15                        "Natural Gas (Daily)",
16                        "Natural Gas (Monthly)",
17                        "Gold",
18                        "Copper (Daily)",
19                        "Copper (Monthly)",
20                        "Platinum",
21                        "Rice",
22                        "Wheat (Daily)",
23                        "Wheat (Monthly)",
24                        "Corn (Daily)",
25                        "Corn (Monthly)",
26                        "Soybean (Daily)",
27                        "Soybean (Monthly)",
28                        "Cocoa (Daily)",
29                        "Cocoa (Monthly)",
30                        "Coffee Arabica",
31                        "Coffee Robusta",
32
```

- Ένας δεύτερος τρόπος είναι ανοίγοντας έναν πλοηγό διαδικτύου (*Web Browser*) και πληκτρολογώντας την κατάλληλη διεύθυνση, γίνεται απευθείας η είσοδος στην εφαρμογή.

#### 6.1.2 Αρχική Σελίδα

Αφού γίνει η επιτυχής είσοδος στην εφαρμογή και αρχίζει να λειτουργεί, τότε υπάρχει η αρχική σελίδα. Αυτή έχει την παρακάτω μορφή:

C:/Users/.../Documents/Application/Full Program - Shiny

http://127.0.0.1:5708 Open in Browser Publish

Commodities Stocks Tourism Unemployment

## Value of basic commodities

Select from the dropdown list the commodity that you want to examine. Information will be collected from Quandl.

Choose a dataset:  
Crude Oil

Date range  
2012-12-31 to 2015-09-16

Forecasting Methods:

- Real Values
- ses
- holt
- damped
- ets
- Moving Average
- hi
- Naive
- Theta

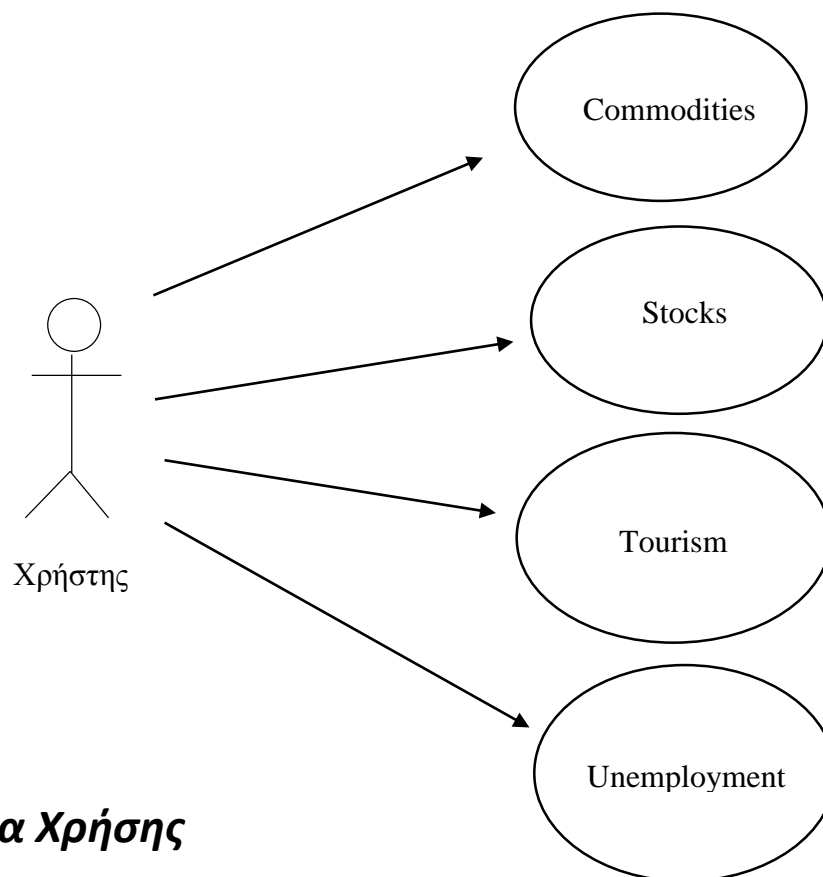
Forecasting methods ranking

Check the type of error coefficients:

Mean Error  
 Mean Absolute Error

	Historical Data	Graphic Representation	Ranking of Forecasting Methods	Method's Forecast
	Date	Value		
1	2015-09-14	44.07		
2	2015-09-11	44.75		
3	2015-09-10	45.85		
4	2015-09-09	44.13		
5	2015-09-08	45.92		
6	2015-09-04	46.02		
7	2015-09-03	46.75		
8	2015-09-02	46.30		
9	2015-09-01	45.38		
10	2015-08-31	49.20		
11	2015-08-28	45.29		
12	2015-08-27	42.47		
13	2015-08-26	38.50		
14	2015-08-25	39.15		
15	2015-08-24	38.22		
16	2015-08-21	40.45		
17	2015-08-20	41.00		
18	2015-08-19	40.75		
19	2015-08-18	42.58		
20	2015-08-17	41.93		
21	2015-08-14	42.45		
22	2015-08-13	42.27		
23	2015-08-12	43.22		
24	2015-08-11	43.11		
25	2015-08-10	44.94		

Όπως φαίνεται και παραπάνω σαν αρχική επιλογή είναι τα εμπορεύματα (*Commodities*). Γενικότερα ο χρήστης μπορεί να επιλέξει μέσα από τις επιλογές, εμπορεύματα (*Commodities*), μετοχές (*Stocks*), τουρισμός (*Tourism*) και ανεργία (*Unemployment*). Παρακάτω φαίνεται ένα διάγραμμα που κάνει πιο κατανοητό ποιες είναι οι αρχικές επιλογές του χρήστη και οι αρχικές εισοδοι που μπορούν να δοθούν:

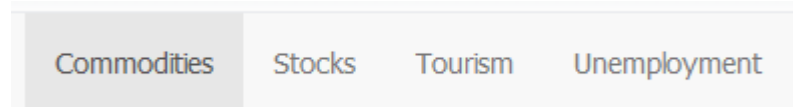


## 6.2 Σενάρια Χρήσης

Για να γίνουν πιο κατανοητές οι δυνατότητες που προσφέρει η εφαρμογή και οι επιλογές της, θα ήταν σκόπιμο να χωρίσουμε αυτές τις επιλογές σε τέσσερα σενάρια χρήσης. Αυτό διότι με αυτό τον τρόπο επιλέγονται τα σενάρια αναλόγως με τα δεδομένα που θέλει να μελετήσει ο χρήστης, δηλαδή τα αγαθά, τις μετοχές, ο τουρισμός και η ανεργία.

### 6.2.1 Επιλογή των Εμπορευμάτων (*Commodities*)

Ο χρήστης μπορεί να επιλέξει την συγκεκριμένη ομάδα δεδομένων, επιλέγοντας την ανάλογη καρτέλα με την συγκεκριμένη ονομασία, όπως φαίνεται παρακάτω.



Αφού γίνει η είσοδος σε αυτήν την καρτέλα, η πρώτη επιλογή που εμφανίζεται είναι η κατηγορία με την ονομασία *Historical Data*, όπου αναγράφονται τα δεδομένα που επιλέγουμε σε μορφή πίνακα. Κατόπιν, η επιλογή των δεδομένων γίνεται επιλέγοντας μέσα από τη drop-down λίστα που βρίσκεται στο χώρο *Choose a dataset*, όπως φαίνεται και στην εικόνα. Μέσω αυτής της λίστας μπορεί να γίνει η επιλογή των αγαθών. Η λίστα αυτή περιλαμβάνει δεκαπέντε αγαθά, τα οποία είναι είτε ημερήσιες, είτε μηνιαίες χρονοσειρές ή και τα δύο και για τα οποία έχει γίνει η περιγραφή τους στο κεφάλαιο 4. Προεπιλεγμένο αγαθό είναι το αργό πετρέλαιο (*Crude Oil*).

Όπως είναι εμφανές στην εικόνα βρίσκεται ένα πλαίσιο το οποίο ονομάζεται *Date Range* και με αυτό δίνεται η δυνατότητα στον χρήστη να επιλέξει εκείνος το εύρος των ημερομηνιών για οποιοδήποτε αγαθό της λίστας, που θέλει να μελετήσει. Σε περίπτωση που ο χρήστης επιλέξει ένα εύρος το οποίο ξεπερνάει τα διαθέσιμα ιστορικά δεδομένα, τότε στον πίνακα θα εμφανιστούν τα δεδομένα από την στιγμή που θα υπάρχουν καταγεγραμμένα ιστορικά δεδομένα. Για παράδειγμα αν η χρονοσειρά ξεκινάει από το 1980 και ο χρήστης επιλέξει το 1970, τότε στον πίνακα θα εμφανιστούν τα δεδομένα από το 1980 και μετά.

Select from the dropdown list the commodity that you want to examine.  
Information will be collected from Quandl.

**Choose a dataset:**

Crude Oil ▼

**Date range**

2012-12-31 to 2015-09-16

**Forecasting Methods:**

- Real Values
- ses
- holt
- damped
- ets
- Moving Average
- Irl
- Naive
- Theta

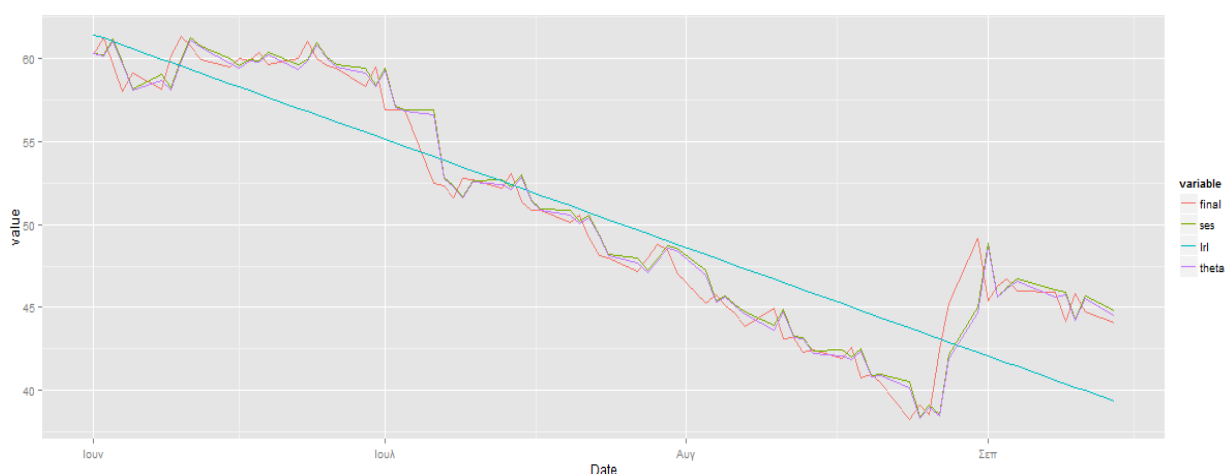
**Forecasting methods ranking**

**Check the type of error coefficients:**

- Mean Error
- Mean Absolute Error
- Symmetric Mean Absolute Percentage Error

Βλέπουμε πως ο χρήστης έχει τη δυνατότητα να επιλέξει και μία δεύτερη κατηγορία που ονομάζεται *Graphic Representation*. Σε αυτή την κατηγορία ο χρήστης έχει τη

δυνατότητα να δει τα δεδομένα των τιμών των αγαθών σε διαγραμματική απεικόνιση, όπου στον οριζόντιο άξονα είναι η ημερομηνία και στον κατακόρυφο η τιμή του αγαθού, όπως φαίνεται στην εικόνα:



Επίσης υπάρχει ένα πλαίσιο στην εφαρμογή, το οποίο ονομάζεται *Forecasting Methods* και από κάτω υπάρχουν οχτώ διαθέσιμες μέθοδοι πρόβλεψης και η επιλογή *Real Values*, που είναι οι πραγματική χρονοσειρά. Αυτές οι μέθοδοι οι εξής:

1. Απλή εκθετική εξομάλυνση – SES
2. Εκθετική εξομάλυνση γραμμικής τάσης – Holt
3. Εκθετική εξομάλυνση μη γραμμικής τάσης – Damped
4. Error-Trend-Seasonal ή Exponential Smoothing method – ETS
5. Κινητός μέσος όρος – KMO
6. Απλή γραμμική παλινδρόμηση – LRL
7. Απλοϊκή μέθοδος - Naive
8. Κλασική μέθοδος Theta

Για αυτές τις μεθόδους πρόβλεψης έχει γίνει εκτενής παρουσίαση σε προηγούμενο κεφάλαιο.

Ουσιαστικά εδώ υπάρχει ένα κουτί επιλογής (Check Box) δίπλα από κάθε μέθοδο πρόβλεψης. Ο χρήστης όταν πατήσει τη συγκεκριμένη επιλογή, θα εμφανιστεί στο διάγραμμα η χρονοσειρά με τις προσαρμοσμένες τιμές (fitted values) των προβλέψεων της επιλεγμένης μεθόδου. Επιλέχθηκε αυτός ο τρόπος, ούτως ώστε ο χρήστης να μπορεί να συγκρίνει απευθείας διαγραμματικά τις διαφορές των μεθόδων πρόβλεψης σε σχέση με τα πραγματικά δεδομένα και έτσι αποκτά μία γενική εικόνα σχετικά με αυτές. Επομένως μπορεί να γίνει ταυτόχρονη αναπαράσταση μέχρι και εννέα χρονοσειρών, όπου η μία χρονοσειρά έχει τα ιστορικά δεδομένα και οι άλλες οχτώ είναι για τις μεθόδους πρόβλεψης. Στο παράδειγμα της παραπάνω εικόνας έχουν επιλεγεί η πραγματική χρονοσειρά του αργού πετρελαίου και οι μέθοδοι ses, lrl και Theta.

Σε αυτή την καρτέλα των εμπορευμάτων υπάρχει μία διαφοροποίηση σε σχέση με τις άλλες καρτέλες. Αυτή η διαφοροποίηση έγκειται στο γεγονός πως υπάρχει μία επιπλέον κατηγορία με την ονομασία *Ranking of Forecasting Methods*.

Methods	errors
Holt	3.5748876
Theta	3.5812114
Ses	3.5881322
Naive	3.5881818
Damped	3.7391192
Ets	3.8013663
Moving_average	4.6320076
LRL	43.3697751

Όπως είναι εμφανές και από την ονομασία στην κατηγορία αυτή γίνεται μία σύγκριση των μεθόδων με βάση τρεις κατηγορίες σφαλμάτων και κατόπιν γίνεται αυτόματα η κατάταξή τους με βάση αυτά τα σφάλματα. Αυτές οι κατηγορίες σφαλμάτων είναι οι εξής:

1. Μέσο σφάλμα – *Mean Error*
2. Μέσο απόλυτο σφάλμα - *Mean Absolute Error*
3. Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα - *Symmetric Mean Absolute Percentage Error*

Η επιλογή των σφάλματος που θέλει ο χρήστης για να κάνει την σύγκριση των μεθόδων πρόβλεψης γίνεται μέσα από το πλαίσιο με την ονομασία *Forecasting Methods Ranking – Check the type of error coefficient* και εκεί ο χρήστης ενεργοποιεί ένα από τα κουμπιά (Radio Buttons) με τις ονομασίες *Mean Error*, *Mean Absolute Error* και *Symmetric Mean Absolute Percentage Error*.

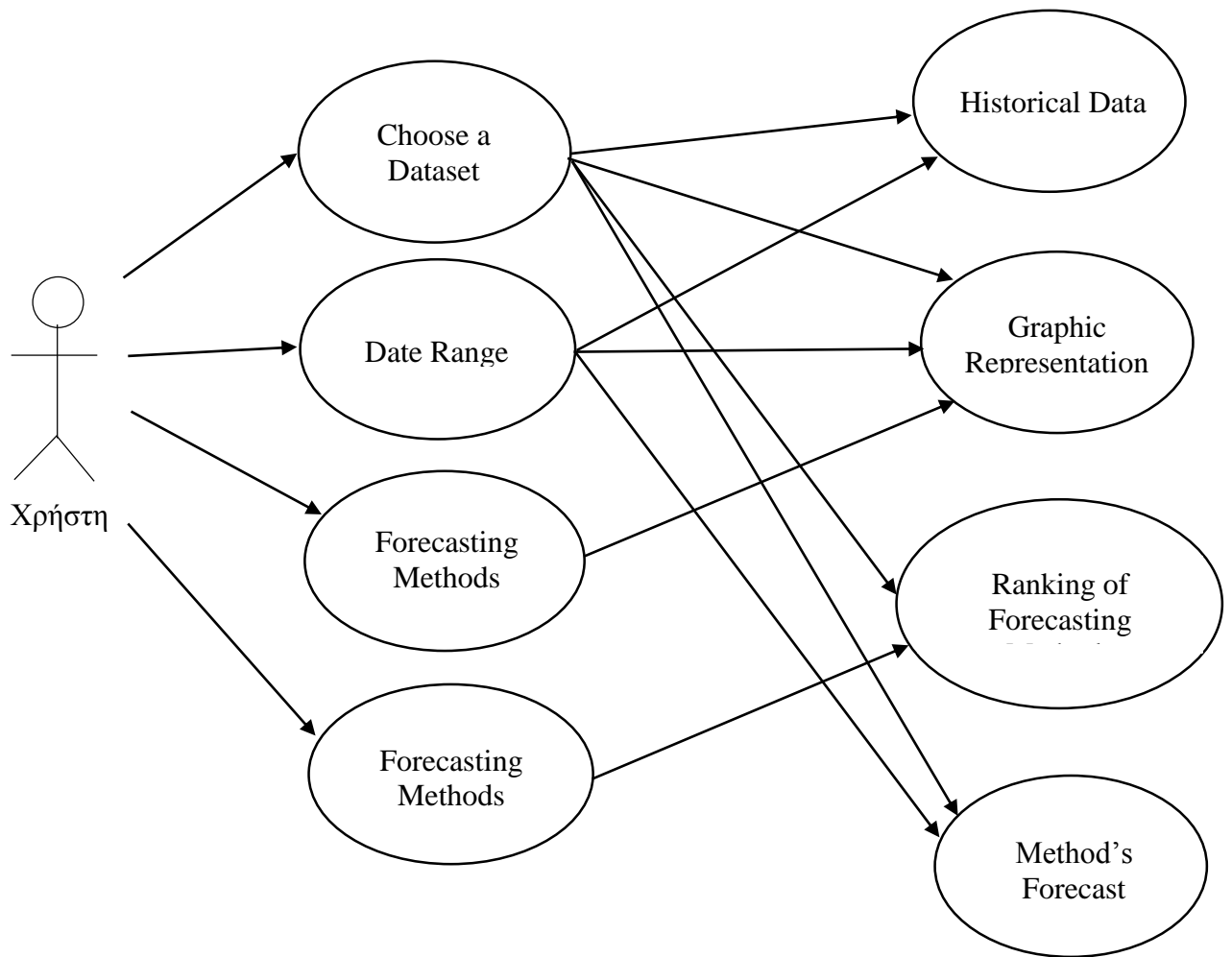
Είναι σημαντικό να αναφερθεί πως η σύγκριση των μεθόδων γίνεται διαρκώς και αλλάζει ανάλογα με τα καινούργια δεδομένα που έχουμε. Αυτό γίνεται, διότι κάθε φορά που έχουμε διαθέσιμα καινούργια δεδομένα, ανανεώνονται αυτόματα στην εφαρμογή. Η εφαρμογή έχει υλοποιηθεί με τέτοιο τρόπο ώστε να λαμβάνονται υπόψιν τα καινούργια δεδομένα και να ανανεώνεται αυτόματα η κατάταξη των μεθόδων ανάλογα με τον τύπο του σφάλματος που έχει επιλεγεί από τον χρήστη, ώστε να γίνει η σύγκριση των μεθόδων. Στην παραπάνω εικόνα φαίνεται ο τρόπος αναπαράστασης της κατάταξης.

Ακόμη μία επιλογή στη κατηγορία των εμπορευμάτων, είναι οι σημειακές προβλέψεις για την επόμενη περίοδο, οι οποίες αναγράφονται συνοπτικά σε έναν πίνακα, όπως φαίνεται και στην εικόνα:



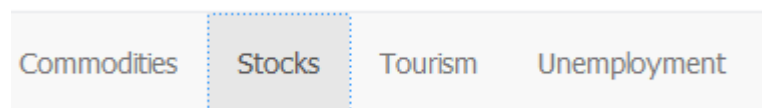
Forecast.Methods	Point Forecast
SES	44.1091864
Holt	44.1051435
Damped	44.0818625
ETS	44.1093589
Moving Average	44.8900000
LRL	89.8152341
Naive	44.0700000
Theta	44.1120055

Στο παρακάτω διάγραμμα αναπαρίσταται συνοπτικά, τις εισόδους που μπορεί να δώσει ο χρήστης και τις εξόδους που μπορεί να λάβει:

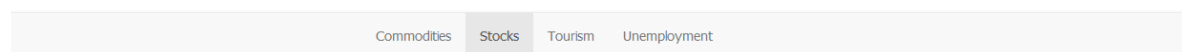


## 6.2.2 Επιλογή των Μετοχών (Stocks)

Για την ενεργοποίηση αυτής της επιλογής, επιλέγουμε την καρτέλα με την ονομασία *Stocks*:



Κατόπιν εμφανίζεται η αρχική σελίδα της καρτέλας που είναι η εξής:



### Stocks

Select a stock to examine. Information will be collected from Yahoo Finance.

**Symbol**

**Date range**

 to

**Forecasting Methods:**

- Real Values
- ses
- holt
- damped
- ets
- Moving Average
- lri
- Naive
- Theta

Historical Data    [Graphic Representation](#)    [Method's Forecast](#)

	SPY.Open	SPY.High	SPY.Low	SPY.Close	SPY.Volume	SPY.Adjusted
2012-12-31	139.66	142.56	139.54	142.41	243935200.00	135.65
2013-01-02	145.11	146.15	144.73	146.06	192059000.00	139.12
2013-01-03	145.99	146.37	145.34	145.73	144761800.00	138.81
2013-01-04	145.97	146.61	145.67	146.37	116817700.00	139.42
2013-01-07	145.85	146.11	145.43	145.97	110002500.00	139.04
2013-01-08	145.71	145.91	144.98	145.55	121265100.00	138.64
2013-01-09	145.87	146.32	145.64	145.92	90745600.00	138.99
2013-01-10	146.73	147.09	145.97	147.08	130735400.00	140.10
2013-01-11	147.04	147.15	146.61	147.07	113917300.00	140.09
2013-01-14	146.89	147.07	146.43	146.97	89567200.00	139.99
2013-01-15	146.29	147.21	146.20	147.07	93172600.00	140.09
2013-01-16	146.77	147.28	146.61	147.05	104849500.00	140.07
2013-01-17	147.70	148.42	147.15	148.00	133833500.00	140.97
2013-01-18	147.97	148.49	147.43	148.33	169906000.00	141.29
2013-01-22	148.33	149.13	147.98	149.13	111797300.00	142.05
2013-01-23	149.13	149.50	148.86	149.37	104596100.00	142.28
2013-01-24	149.15	150.14	149.01	149.41	146426400.00	142.31
2013-01-25	149.88	150.25	149.37	150.25	147211600.00	143.12
2013-01-28	150.29	150.33	149.51	150.07	113357700.00	142.94
2013-01-29	149.77	150.85	149.67	150.66	105694400.00	143.51
2013-01-30	150.64	150.94	149.93	150.07	137447700.00	142.94
2013-01-31	149.89	150.38	149.60	149.70	108975800.00	142.59
2013-02-01	150.65	151.42	150.39	151.24	131173000.00	144.06
2013-02-04	150.32	151.27	149.43	149.54	159073600.00	142.44
2013-02-05	150.35	151.48	150.29	151.05	113912400.00	143.88

Όπως είναι εμφανές από την παραπάνω εικόνα, η πρώτη επιλογή που εμφανίζεται είναι η κατηγορία με την ονομασία *Historical Data*, όπου αναγράφονται τα δεδομένα σε μορφή πίνακα. Στη συνέχεια, η επιλογή των δεδομένων γίνεται από τον χρήστη, απλά πληκτρολογώντας, στο πλαίσιο με την ονομασία *Symbol* όπως φαίνεται και στην εικόνα, το σύμβολο της μετοχής που θέλει να μελετήσει. Σαν προεπιλεγμένη επιλογή έχει οριστεί η μετοχή Standard & Poors 500 (S&P 500). Αντίστοιχα άλλα παραδείγματα είναι η μετοχή της Apple έχει το σύμβολο AAPL, της Yahoo το σύμβολο YHOO, των Starbucks είναι SBUX και ούτε κάθε εξής.

Όπως φαίνεται στην παρακάτω εικόνα βρίσκεται ένα πλαίσιο το οποίο ονομάζεται *Date Range* και με αυτό ο χρήστης έχει την δυνατότητα να επιλέξει εκείνος το εύρος

των ημερομηνιών για οποιαδήποτε μετοχή θέλει να μελετήσει. Σε περίπτωση που ο χρήστης επιλέξει ένα εύρος το οποίο ξεπερνάει τα διαθέσιμα ιστορικά δεδομένα, τότε στον πίνακα θα εμφανιστούν τα καταγεγραμμένα. Για παράδειγμα αν η χρονοσειρά ξεκινάει από το 1990 και ο χρήστης επιλέξει το 1980, τότε στον πίνακα θα εμφανιστούν τα δεδομένα από το 1990 και μετά.

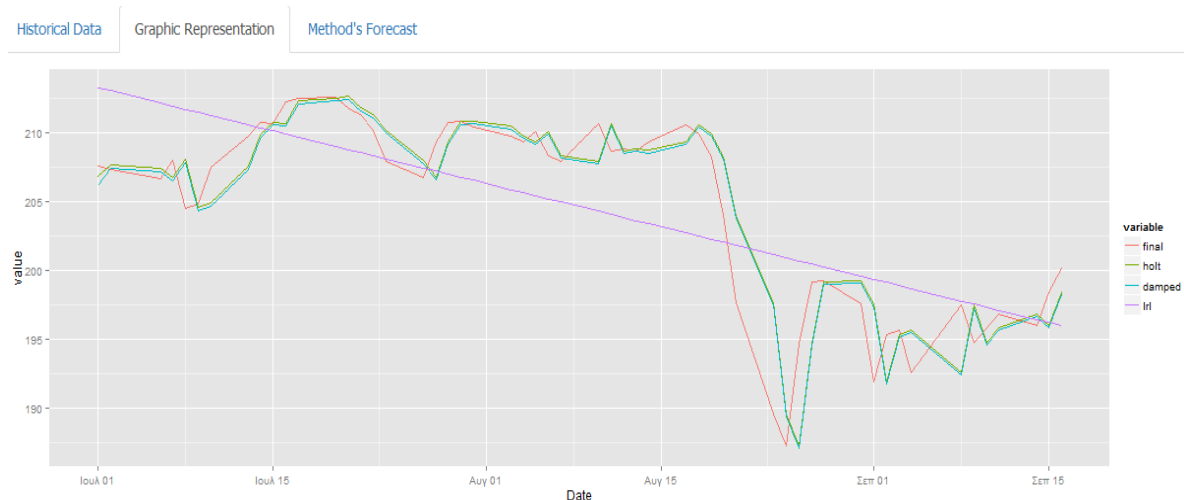
Select a stock to examine. Information will be collected from Yahoo Finance.

**Symbol**

**Date range**

 to

Επιπρόσθετα, ο χρήστης έχει τη δυνατότητα να επιλέξει και μία δεύτερη κατηγορία που ονομάζεται *Graphic Representation*. Σε αυτή την κατηγορία τα δεδομένα των τιμών κλεισίματος των μετοχών αναπαριστούνται σε διαγραμματικά, όπου στον οριζόντιο άξονα είναι η ημερομηνία και στον κατακόρυφο η τιμή κλεισίματος της μετοχής, όπως φαίνεται στην εικόνα:



Και εδώ υπάρχει ένα πλαίσιο, το οποίο ονομάζεται *Forecasting Methods* και από κάτω υπάρχουν οι οχτώ διαθέσιμες μέθοδοι πρόβλεψης και η επιλογή *Real Values*,

που είναι οι πραγματική χρονοσειρά, όπως έχει αναφερθεί και προηγουμένως. Στο παράδειγμα της παραπάνω εικόνας έχουν επιλεγεί η πραγματική χρονοσειρά του αργού πετρελαίου και οι μέθοδοι holt, damped και lrl.

### Forecasting Methods:

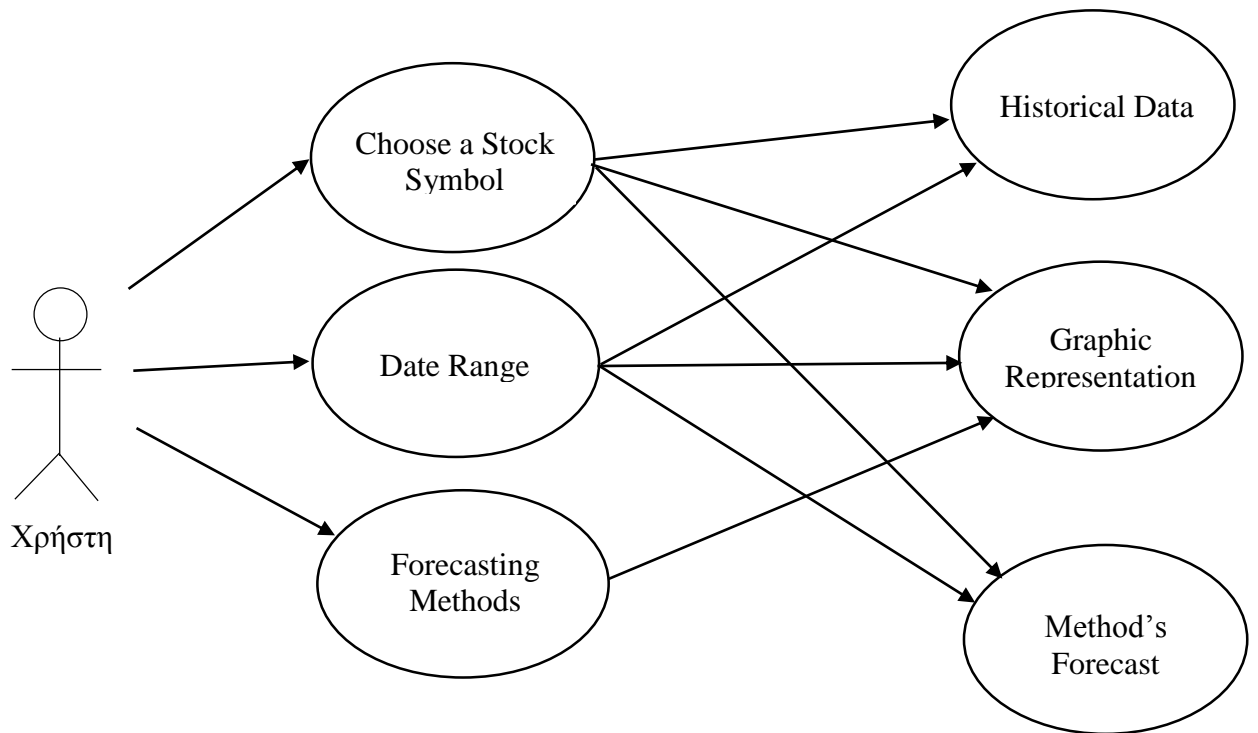
- Real Values
- ses
- holt
- damped
- ets
- Moving Average
- lrl
- Naive
- Theta

Ακόμη μία επιλογή στη κατηγορία των μετοχών, είναι όπως και στην περίπτωση των αγαθών, οι σημειακές προβλέψεις για την επόμενη περίοδο, οι οποίες αναγράφονται συνοπτικά σε έναν πίνακα, όπως φαίνεται και στην εικόνα:

Historical Data    Graphic Representation    Method's Forecast

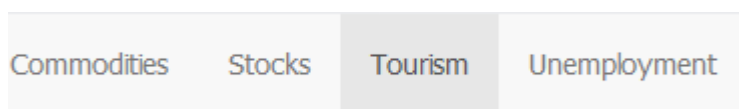
Forecast.Methods	Point Forecast
SES	200.0363439
Holt	200.0506905
Damped	200.0360002
ETS	200.0361342
Moving Average	198.1899973
LRL	166.3996028
Naive	200.1399990
Theta	200.0430169

Στο παρακάτω διάγραμμα αναπαρίσταται συνοπτικά, τις εισόδους που μπορεί να δώσει ο χρήστης και τις εξόδους που μπορεί να λάβει:



### 6.2.3 Επιλογή του Τουρισμού (Tourism)

Για να γίνει αυτή η επιλογή, ο χρήστης επιλέγει την καρτέλα με την ονομασία *Tourism*:



Κατόπιν εμφανίζεται η αρχική σελίδα της καρτέλας που είναι η εξής:

## Tourist's Arrival

Select a country to examine with the appropriate country symbol. Information will be collected from World Bank.

**Country Symbol**

Select the starting and the ending year.

**Starting Year**

**Ending Year**

**Forecasting Methods:**

Real Values

ses

holt

damped

ets

Moving Average

lr

Naive

Theta

	iso2c	country	ST.INT.ARVL	year
1	US	United States	60010000.00	2010.00
2	US	United States	55103000.00	2009.00
3	US	United States	58007000.00	2008.00
4	US	United States	56135000.00	2007.00
5	US	United States	50977000.00	2006.00
6	US	United States	49206000.00	2005.00
7	US	United States	46086000.00	2004.00
8	US	United States	41218000.00	2003.00
9	US	United States	43581000.00	2002.00
10	US	United States	46927000.00	2001.00
11	US	United States	51238000.00	2000.00

Εξ ορισμού επιλέγοντας αυτήν την καρτέλα, θα εμφανίσει την κατηγορία *Historical Data*, όπου αναγράφονται τα δεδομένα σε μορφή πίνακα. Στην πρώτη στήλη αναγράφεται το σύμβολο της χώρας, στη δεύτερη στήλη αναγράφεται η χώρα, στην τρίτη στήλη ο αριθμός των επισκεπτών της χώρας και στην τέταρτη στήλη η χρονιά που αφίχθησαν οι τουρίστες.

Στο πεδίο *Country Symbol* ο χρήστης πληκτρολογεί με κεφαλαία γράμματα το σύμβολο της χώρας, που θέλει να μελετήσει τις αφίξεις των τουριστών. Εξ ορισμού έχει τοποθετηθεί το σύμβολο *US*, το οποίο αναφέρεται στις Ηνωμένες Πολιτείες της Αμερικής (*USA*). Για παράδειγμα για την Ελλάδα είναι το σύμβολο *GR* (Greece), για την Ιταλία το σύμβολο *IT* (Italy), για την Κίνα το σύμβολο *CN* (China), για την Γαλλία το σύμβολο *FR* (France) και ούτω καθεξής.

Όπως φαίνεται στην παρακάτω εικόνα βρίσκονται δύο πλαίσια τα οποία ονομάζονται *Starting Year* και *Ending Year* και με αυτά ο χρήστης έχει την δυνατότητα να επιλέξει εκείνος το εύρος των ετών για οποιαδήποτε χώρα θέλει να μελετήσει. Σε περίπτωση που ο χρήστης επιλέξει ένα εύρος το οποίο ξεπερνάει τα διαθέσιμα ιστορικά δεδομένα, τότε στον πίνακα θα εμφανιστούν τα καταγεγραμμένα.

# Tourist's Arrival

Select a country to examine with the appropriate country symbol. Information will be collected from World Bank.

**Country Symbol**

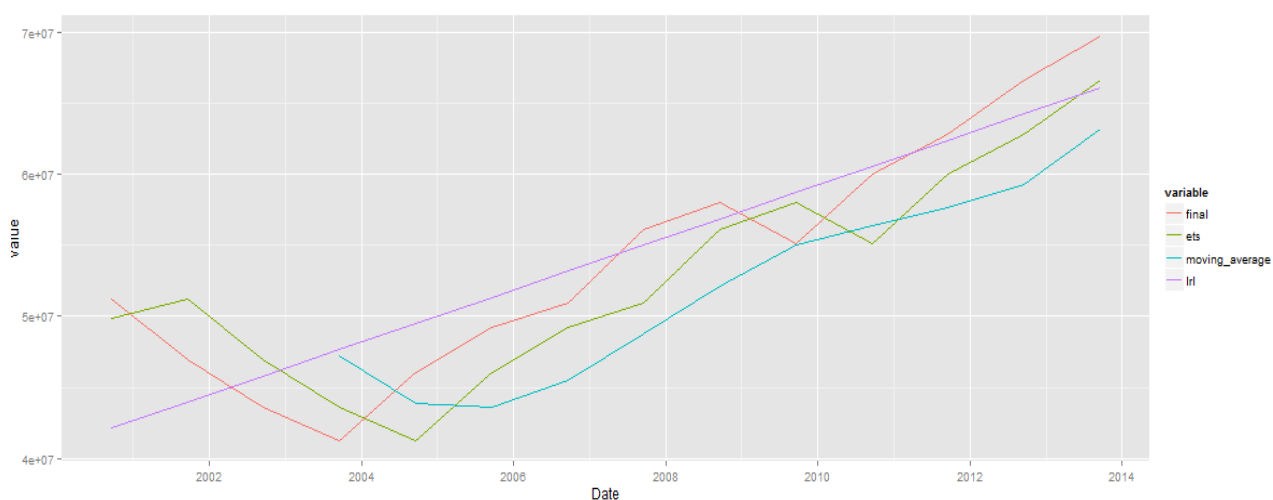
Select the starting and the ending year.

**Starting Year**

**Ending Year**

Επίσης, δίνεται η δυνατότητα στον χρήστη να επιλέξει την κατηγορία *Graphic Representation*, όπου απεικονίζονται διαγραμματικά ο αριθμός των αφίξεων ανά έτος για τη χώρα που θα επιλέξει ο χρήστης με τον τρόπο που αναφέρθηκε προηγουμένως.

[Historical Data](#) [Graphic Representation](#) [Method's Forecast](#)



Και εδώ υπάρχει το γνωστό πλέον πλαίσιο, το οποίο ονομάζεται *Forecasting Methods* και από κάτω υπάρχουν οι οχτώ διαθέσιμες μέθοδοι πρόβλεψης και η επιλογή *Real Values*, που είναι οι πραγματική χρονοσειρά, όπως έχει αναφερθεί και προηγουμένως. Στο παράδειγμα της παραπάνω εικόνας έχουν επιλεγεί η πραγματική χρονοσειρά των αφίξεων στις Η.Π.Α. και οι μέθοδοι ets, moving average και lrl.

**Forecasting Methods:**

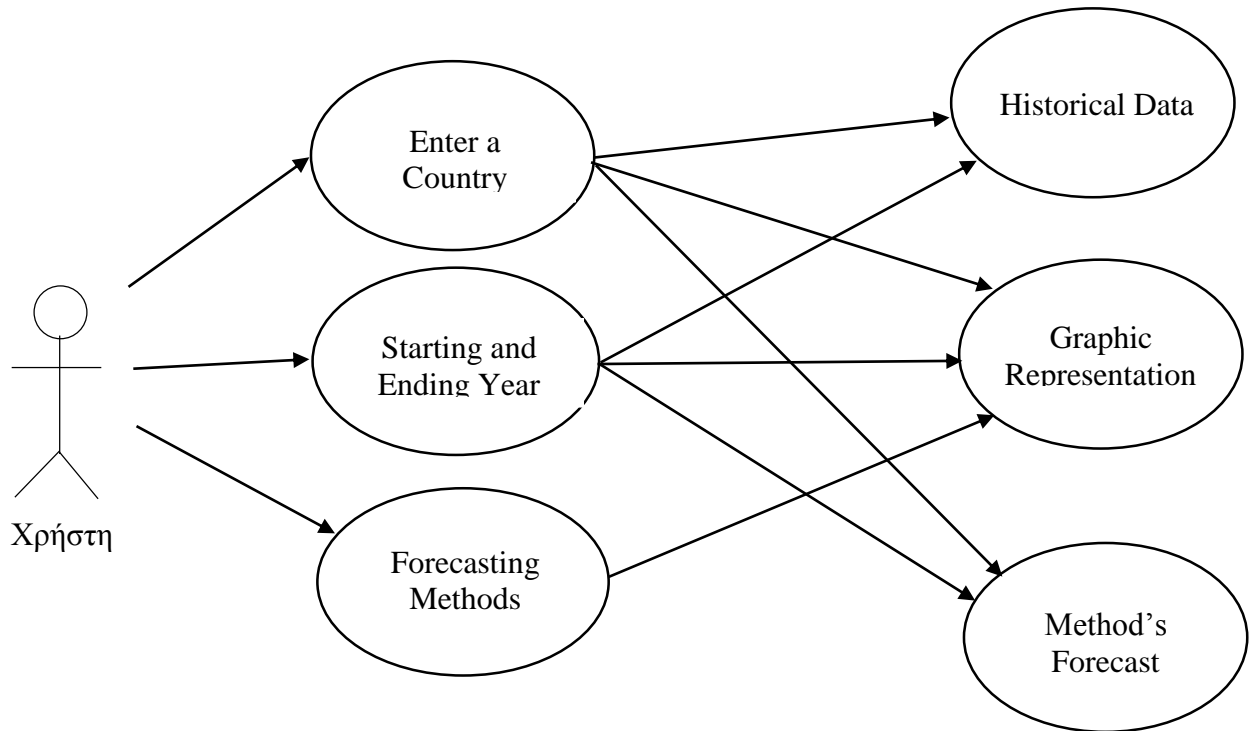
- Real Values
- ses
- holt
- damped
- ets
- Moving Average
- lrl
- Naive
- Theta

Ακόμη μία επιλογή στη κατηγορία του τουρισμού, είναι όπως και στην περίπτωση των αγαθών, οι σημειακές προβλέψεις για την επόμενη περίοδο, οι οποίες αναγράφονται συνοπτικά σε έναν πίνακα, όπως φαίνεται και στην εικόνα:

Historical Data		Graphic Representation		Method's Forecast	
Forecast.Methods	Point Forecast				
SES	69767688.8442385				
Holt	71026134.6492869				
Damped	70697539.2187883				
ETS	69767688.7914125				
Moving Average	66415333.3333333				
LRL	66119527.9136425				
Naive	69768000.0000000				
Theta	69770306.0953680				

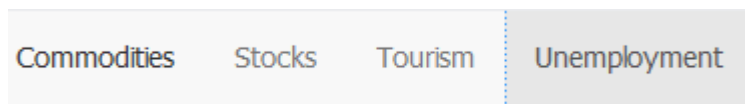


Στο παρακάτω διάγραμμα αναγράφονται συνοπτικά, οι εισόδους που μπορεί να δώσει ο χρήστης και τις εξόδους που μπορεί να λάβει:



#### 6.2.4 Επιλογή της Ανεργίας (Unemployment)

Για την ανεργία, ο χρήστης επιλέγει την καρτέλα με την ονομασία *Unemployment*:



Η αρχική σελίδα της καρτέλας είναι:

## Rates of Unemployment

Select a country to examine with the appropriate country symbol. Information will be collected from World Bank.

**Country Symbol**

Select the starting and the ending year.

**Starting Year**

**Ending Year**

**Forecasting Methods:**

Real Values

ses

holt

damped

ets

Moving Average

lr

Naive

Theta

	iso2c	country	SL.UEM.TOTL.NE.ZS	year
1	US	United States	9.60	2010.00
2	US	United States	9.30	2009.00
3	US	United States	5.80	2008.00
4	US	United States	4.60	2007.00
5	US	United States	4.60	2006.00
6	US	United States	5.10	2005.00
7	US	United States	5.50	2004.00
8	US	United States	6.00	2003.00
9	US	United States	5.80	2002.00
10	US	United States	4.70	2001.00
11	US	United States	4.00	2000.00

Από την αρχή επιλέγοντας αυτήν την καρτέλα, εμφανίζει την κατηγορία *Historical Data*, όπου αναγράφονται τα δεδομένα σε μορφή πίνακα. Στην πρώτη στήλη αναγράφεται το σύμβολο της χώρας που χρησιμοποιείται διεθνώς, στη δεύτερη στήλη αναγράφεται το όνομα της χώρας, στην τρίτη στήλη το ποσοστό της ανεργίας της χώρας και στην τέταρτη στήλη η χρονιά που καταμετρήθηκε αυτό το ποσοστό.

Στο πεδίο *Country Symbol* ο χρήστης πληκτρολογεί με κεφαλαία γράμματα τον κωδικό της χώρας που χρησιμοποιείται διεθνώς και που θέλει να μελετήσει τα ποσοστά ανεργίας των εκάστοτε χωρών. Σαν προεπιλογή έχει τοποθετηθεί το σύμβολο *US*, το οποίο αναφέρεται στις Ηνωμένες Πολιτείες της Αμερικής (*USA*). Για παράδειγμα για την Ελλάδα είναι το σύμβολο *GR* (Greece), για την Ιταλία το σύμβολο *IT* (Italy), για την Κίνα το σύμβολο *CN* (China), για την Γαλλία το σύμβολο *FR* (France) και ούτω καθεξής.

Όπως φαίνεται στην παρακάτω εικόνα βρίσκονται δύο πλαίσια τα οποία ονομάζονται *Starting Year* και *Ending Year*, όπου δίνεται η δυνατότητα στον χρήστη να επιλέξει εκείνος το εύρος των ετών για οποιαδήποτε χώρα θέλει να μελετήσει. Σε περίπτωση που ο χρήστης επιλέξει ένα εύρος το οποίο ξεπερνάει τα διαθέσιμα ιστορικά δεδομένα, τότε στον πίνακα θα εμφανιστούν μόνο τα διαθέσιμα.

# Rates of Unemployment

Select a country to examine with the appropriate country symbol. Information will be collected from World Bank.

**Country Symbol**

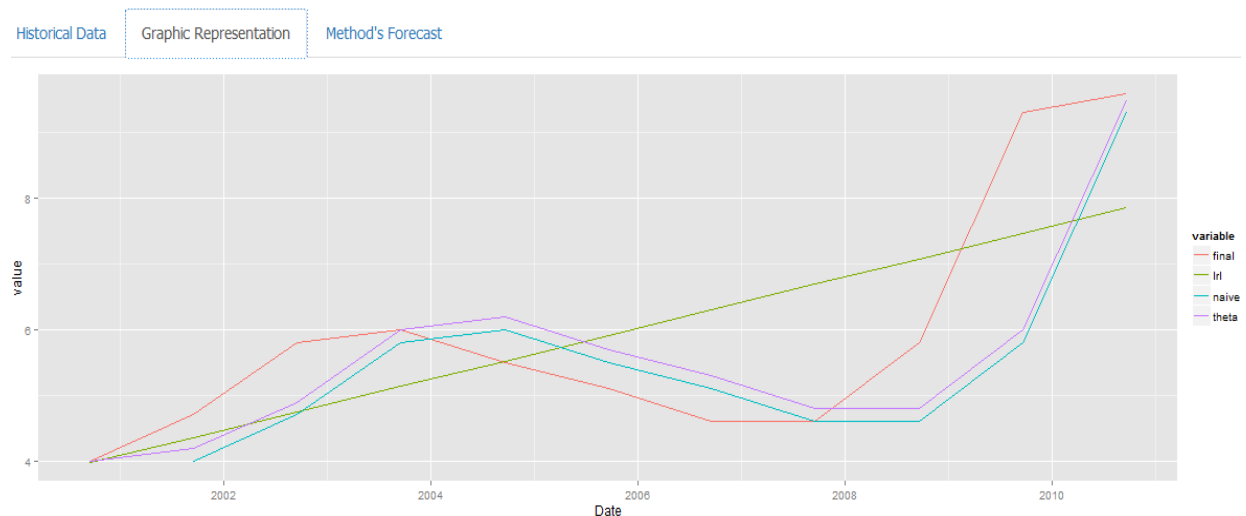
Select the starting and the ending year.

**Starting Year**

**Ending Year**

Επίσης, ο χρήστης έχει την δυνατότητα να επιλέξει την κατηγορία *Graphic Representation*, όπου υπάρχει ένα διάγραμμα στο οποίο στον κάθετο άξονα τοποθετούνται τα καταγεγραμμένα ποσοστά της ανεργίας κάθε χώρας και στον οριζόντιο άξονα οι χρονιές που καταγράφηκαν αυτά τα ποσοστά. Η επιλογή της χώρας γίνεται με τον τρόπο που περιεγράφηκε στο παραπάνω βήμα.



Εδώ βρίσκεται το γνωστό πλέον πλαίσιο, το οποίο ονομάζεται *Forecasting Methods* και από κάτω υπάρχουν οι διαθέσιμες μέθοδοι πρόβλεψης

και η επιλογή *Real Values*, που είναι οι πραγματική χρονοσειρά, όπως έχει αναφερθεί και προηγουμένως. Στο παράδειγμα της παραπάνω εικόνας έχουν επιλεγεί η πραγματική χρονοσειρά του ποσοστού της ανεργίας στις Η.Π.Α. και οι μέθοδοι lrl, naive και theta.

### Forecasting Methods:

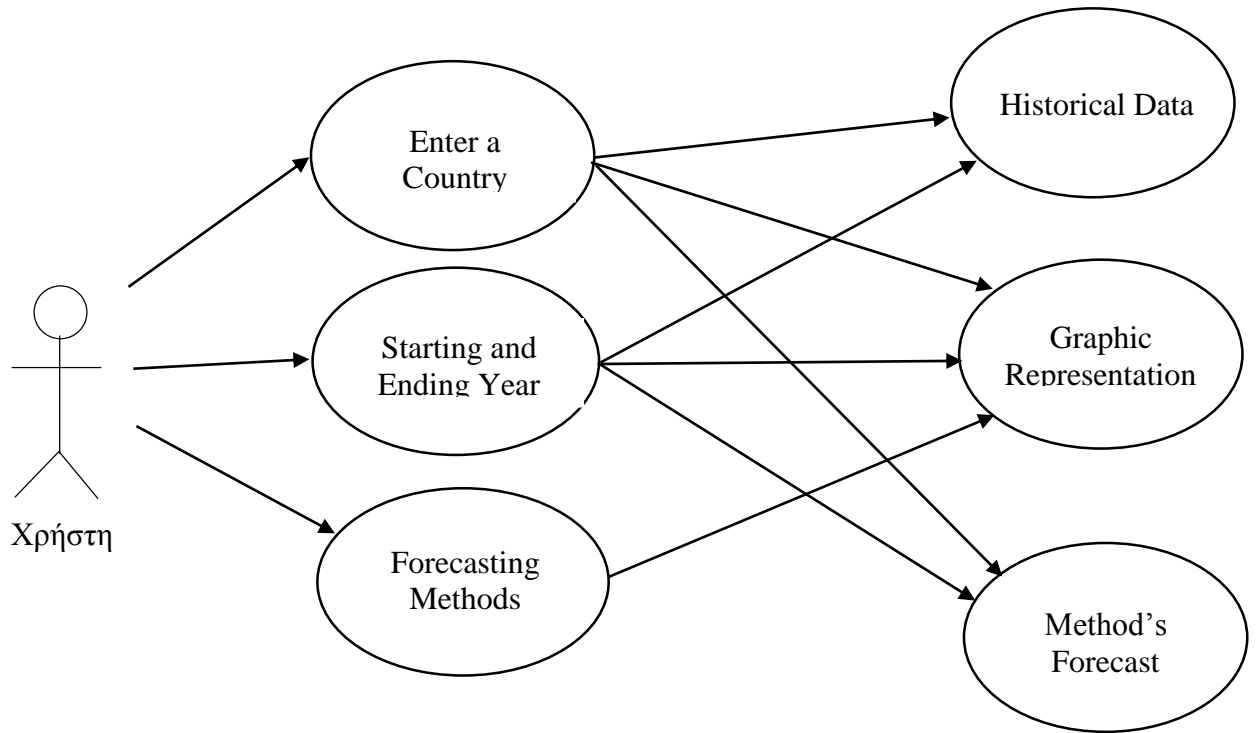
- Real Values
- ses
- holt
- damped
- ets
- Moving Average
- lrl
- Naive
- Theta

Στην κατηγορία της ανεργίας υπάρχει ακόμη μία επιλογή, που όπως και στις προηγούμενες κατηγορίες, δίνονται οι σημειακές προβλέψεις για την επόμενη περίοδο, οι οποίες αναγράφονται συνοπτικά σε έναν πίνακα, όπως φαίνεται και στην εικόνα:

Historical Data
Graphic Representation
Method's Forecast

Forecast.Methods	Point Forecast
SES	8.1000804
Holt	8.4584154
Damped	8.3222594
ETS	8.1000804
Moving Average	8.8666668
LRL	8.6432057
Naive	8.1000004
Theta	8.1006324

Στο παρακάτω διάγραμμα αναγράφονται συνοπτικά, οι εισοδοι που μπορεί να δώσει ο χρήστης και τις εξόδους που μπορεί να λάβει:





# 7. Σύγκριση Μεθόδων Πρόβλεψης

Στη παρούσα διπλωματική εργασία, εκτός από την υλοποίηση της εφαρμογής, υπήρξε η σκέψη για την υλοποίηση ενός πειράματος. Αυτό είχε σαν αντικείμενο την μελέτη των χρονοσειρών των αγαθών, που συμπεριλαμβάνονται στην εφαρμογή εξάγοντας σημειακές προβλέψεις για την πρώτη και την τρίτη περίοδο. Ο σκοπός ήταν η αξιολόγηση των μεθόδων πρόβλεψης ανά αγαθό, ανά χρονικό ορίζοντα των χρονοσειρών (ημερήσιες και μηνιαίες), ανά ορίζοντα πρόβλεψης (πρώτη και τρίτη περίοδος πρόβλεψης) και ανά είδος σφάλματος. Τέλος, έγινε μία συνολική βαθμολόγηση όλων των μεθόδων σε όλα τα αγαθά, με βάση την κατάταξη των μεθόδων. Η βαθμολόγηση αυτή έγινε με διαχωρισμό των χρονοσειρών με βάση τον χρονικό ορίζοντά τους, τον ορίζοντα πρόβλεψης και το είδος σφάλματος.

## 7.1 Γενική Μεθοδολογία

Στα πλαίσια της διπλωματικής, εκτός από την υλοποίηση της εφαρμογής, πραγματοποιήθηκε ένα πείραμα που αφορά τις χρονοσειρές των αγαθών. Πιο συγκεκριμένα, πραγματοποιήθηκαν οι σημειακές προβλέψεις για τις τιμές των εμπορευμάτων, οι οποίες έχουν αναφερθεί σε προηγούμενο κεφάλαιο, για χρονικό ορίζοντα μίας περιόδου, αλλά και για χρονικό ορίζοντα τριών περιόδων. Κατόπιν η σύγκριση των μεθόδων γίνεται υπολογίζοντας πέντε σφάλματα, τα οποία είναι:

- Μέσο σφάλμα – *Mean Error*
- Μέσο απόλυτο σφάλμα - *Mean Absolute Error*
- Μέσο απόλυτο ποσοστιαίο σφάλμα - *Mean Absolute Percentage Error*
- Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα - *symmetric Mean Absolute Percentage Error*
- Μέσο απόλυτο κανονικοποιημένο σφάλμα - *Mean Absolute scaled Error*

Προφανώς οι μέθοδοι που υλοποιήθηκαν και αξιολογήθηκαν είναι οι οχτώ που αναφέρθηκαν στα προηγούμενα κεφάλαια και που χρησιμοποιούνται στην εφαρμογή. Οι μέθοδοι αυτοί είναι οι εξής:

- Απλή εκθετική εξομάλυνση – SES
- Εκθετική εξομάλυνση γραμμικής τάσης – Holt
- Εκθετική εξομάλυνση μη γραμμικής τάσης – Damped
- Error-Trend-Seasonal ή Exponential Smoothing method – ETS
- Κινητός μέσος όρος – KMO
- Απλή γραμμική παλινδρόμηση – LRL

- Απλοϊκή μέθοδος - Naive
- Κλασική μέθοδος Theta

Για την εξαγωγή προβλέψεων, οι οποίες οδηγούν στην αξιολόγηση των μεθόδων, χρησιμοποιήθηκε η τεχνική της κυλιόμενης πρόβλεψης (*rolling forecasting*). Η τεχνική αυτή χρησιμοποιεί συγκεκριμένο πλήθος παρατηρήσεων που λαμβάνονται υπόψη για την εξαγωγή της πρόβλεψης. Στη δική μας περίπτωση θεωρήσαμε το μεγαλύτερο ποσοστό των παρατηρήσεων (περίπου 80%) γνωστό και το υπόλοιπο ποσοστό έμπαινε αφού γινόταν η εξαγωγή της πρόβλεψης, σαν μελλοντική τιμή. Κάθε φορά δηλαδή που πραγματοποιείται η πρόβλεψη, έρχεται για την ίδια περίοδο η πραγματική τιμή της χρονοσειράς που στην ουσία θεωρείται σαν μελλοντική τιμή. Με αυτό τον τρόπο γίνεται μία καλύτερη αξιολόγηση των μεθόδων πρόβλεψης και επίσης δεν είναι απαραίτητο να περιμένει κανείς να περάσει το χρονικό διάστημα που έχει οριστεί εξαρχής για να μάθει τις καινούργιες πραγματικές τιμές. Το συγκεκριμένο πλήθος των παρατηρήσεων που θα ληφθεί υπόψη, ορίζεται από τον ερευνητή και εξαρτάται από το πλήθος των δεδομένων και από τη φύση τους.

Στη περίπτωση τώρα των αγαθών το πλήθος των δεδομένων που χρησιμοποιείται για την εξαγωγή της πρόβλεψης δεν είναι το ίδιο για κάθε χρονοσειρά, και αυτό διότι όπως έχει ήδη αναφερθεί στην περιγραφή των αγαθών, υπάρχουν και ημερήσιες και μηνιαίες χρονοσειρές διαφορετικού μεγέθους.

Παρακάτω βρίσκονται συνοπτικά τα βήματα που ακολουθήθηκαν για την συγκέντρωση των χρονοσειρών και για την εξαγωγή των προβλέψεων:

#### **Βήμα 1°: Συγκέντρωση πραγματικών τιμών**

Όπως έχει ήδη αναφερθεί και σε προηγούμενο κεφάλαιο, οι χρονοσειρές αντλήθηκαν μέσα από τον ιστότοπο Quandl χρησιμοποιώντας την γλώσσα προγραμματισμού R και συγκεκριμένα το πακέτο της με την ονομασία Quandl.

#### **Βήμα 2°: Παραγωγή προβλέψεων**

Η τεχνική που ακολουθήθηκε για την εξαγωγή των προβλέψεων είναι η τεχνική της κυλιόμενης πρόβλεψης (*rolling forecasting*), η οποία αναλύθηκε στην προηγούμενη παράγραφο. Για την εφαρμογή των μεθόδων πρόβλεψης χρησιμοποιήθηκε το πακέτο *forecast* της R και παρήχθησαν προβλέψεις για την επόμενη περίοδο και για την τρίτη χρονική περίοδο.

#### **Βήμα 3°: Αποθήκευση δεδομένων**

Η αποθήκευση των δεδομένων και των προβλέψεων ανά μέθοδο πρόβλεψης έγινε μέσω της R σε αρχεία *csv* και *excel*.

#### **Βήμα 4°: Υπολογισμός σφαλμάτων**

Τα σφάλματα που υπολογίστηκαν και που θα αναφερθούν παρακάτω, είναι ο μέσος όρος των σφαλμάτων ανά αγαθό και ανά μέθοδο πρόβλεψης.



## **Βήμα 5<sup>ο</sup>: Αξιολόγηση και σύγκριση των μεθόδων με βάση τα σφάλματα**

Σε αυτό το βήμα έγινε η αξιολόγηση και η σύγκριση των μεθόδων με βάση τα σφάλματα που υπολογίσαμε στο προηγούμενο βήμα. Η σύγκριση έγινε διαχωρίζοντας τις χρονοσειρές με βάση τον χρονικό ορίζοντά τους, δηλαδή σε ημερήσιες και σε μηνιαίες χρονοσειρές. Επίσης ο διαχωρισμός σε συνεχίστηκε με βάση το είδος των σφαλμάτων, αλλά και του ορίζοντα πρόβλεψης, δηλαδή για πρόβλεψη της επόμενης περιόδου ( $h=1$ ) και για την τρίτη περίοδο ( $h=3$ ).

## **7.2 Αξιολόγηση Μεθόδων για Πρόβλεψη Πρώτης και Τρίτης Χρονικής Περιόδου**

Στη περίπτωση αυτή χρησιμοποιώντας την παραπάνω τεχνική που περιγράφηκε, πραγματοποιήθηκαν οι προβλέψεις για την επόμενη χρονική περίοδο και για την τρίτη επόμενη περίοδο. Στη συνέχεια, υπολογίστηκαν όλα τα είδη σφαλμάτων, που αναφέρθηκαν πιο πάνω. Θα πρέπει να τονιστεί πως το πλήθος  $n$  που χρησιμοποιείται στον υπολογισμό των σφαλμάτων, δεν αντιπροσωπεύει το μήκος της χρονοσειράς, αλλά το μήκος των προβλέψεων που έχουν εξαχθεί. Εν προκειμένω, στο συγκεκριμένο πείραμα σαν μήκος θεωρείται το περίπου 20% της συνολικής χρονοσειράς, το οποίο θεωρήσαμε άγνωστο αρχικά, αλλά στη συνέχεια αποτέλεσε τις μελλοντικές τιμές της χρονοσειράς που θεωρήσαμε αρχικά, δηλαδή το 80%.

Πρέπει να αναφερθεί εδώ πως ο διαχωρισμός της αξιολόγησης των μεθόδων έγινε με βάση το αγαθό και τοποθετήθηκαν στα ίδια διαγράμματα τα ανάλογα σφάλματα των οχτώ μεθόδων πρόβλεψης, για παραγωγή προβλέψεων της επόμενης περιόδου και για των τριών επόμενων περιόδων. Αυτό γίνεται για να γίνει μία απευθείας σύγκριση των μεθόδων στα ίδια δεδομένα, αλλά για διαφορετική περίοδο πρόβλεψης.

Παρακάτω παρατίθενται συνοπτικά οι μέθοδοι πρόβλεψης και τα αντίστοιχα σφάλματά τους.

### **7.2.1 Αργό πετρέλαιο (Crude Oil)**

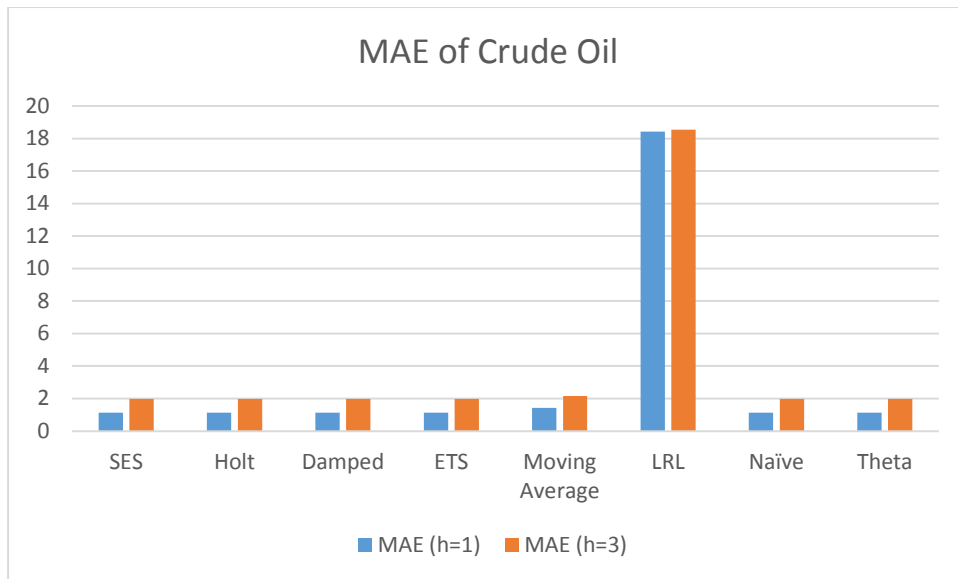
Όπως έχει ήδη αναφερθεί σε προηγούμενο κεφάλαιο πρόκειται για μία ημερήσια χρονοσειρά με μεγάλο πλήθος παρατηρήσεων. Ο πίνακας σφαλμάτων για το μέσο σφάλμα (ME) και του μέσου απόλυτου σφάλματος του αργού πετρελαίου για προβλέψεις της επόμενης περιόδου ( $h=1$ ) και για την τρίτη επόμενη περίοδο ( $h=3$ ) είναι:

<b>Crude Oil</b>	ME (h=1)	MAE (h=1)	ME (h=3)	MAE (h=3)
SES	<b>-0,00469</b>	1,137069	<b>-0,01416</b>	1,985946
Holt	-0,01456	1,136881	-0,04291	<b>1,98505</b>
Damped	-0,00529	1,137184	-0,01534	1,986295
ETS	-0,00497	<b>1,13604</b>	-0,01437	1,985853
Moving Average	-0,00715	1,421457	-0,0198	2,157089
LRL	11,09653	18,43264	10,98262	18,54222
Naïve	<i>-0,00493</i>	<i>1,136131</i>	<i>-0,01422</i>	1,986
Theta	-0,01061	1,137056	-0,02618	<i>1,985581</i>

Σε αυτή την περίπτωση βλέπουμε, όσον αφορά το μέσο σφάλμα και για τις δύο περιπτώσεις, πως υπάρχουν τεράστιες διαφορές όσον αφορά την μέθοδο LRL σε σχέση με τις υπόλοιπες μεθόδους πρόβλεψης. Εξαιτίας αυτού είναι αρκετά δύσκολη η αναπαράστασή της σε ένα διάγραμμα και για αυτό προτιμήθηκε να μην γίνει.

Πριν αναλυθούν τα αποτελέσματα του παραπάνω πίνακα, πρέπει να αναφερθεί και η σημασία του αντίστοιχου δείκτη. Ο δείκτης του μέσου σφάλματος (ME) δείχνει τον βαθμό της προκατάληψης που υπάρχει στην αντίστοιχη μέθοδο για την οποία υπολογίζεται. Μία θετική τιμή του δείκτη αυτού, σημαίνει απαισιοδοξία στην αντίστοιχη μέθοδο διότι εννοείται πως κατά μέσο όρο οι τιμές της πρόβλεψης που παράγονται είναι μικρότερες από τις πραγματικές. Αντιθέτως, αρνητικές τιμές του δείκτη δίνουν το χαρακτηριστικό της αισιοδοξίας στην αντίστοιχη μέθοδο πρόβλεψης καθώς συνεπάγεται υψηλότερες προβλέψεις από τις πραγματικές. Σαν θετικό μπορεί να θεωρηθεί πως δεν παρουσιάζονται υψηλές τιμές, εκτός από την περίπτωση της LRL, στις τιμές του σφάλματος. Είναι θετικό το φαινόμενο αρχικά, ότι δεν παρουσιάζονται ιδιαίτερα υψηλές τιμές του δείκτη σε κανένα από τους συνδυασμούς των μεθόδων. Γενικά στην παραγωγή προβλέψεων μίας χρονικής περιόδου αποδείχτηκε πως η SES και ακολούθως η Naïve έχουν την μικρότερη, κατά απόλυτη τιμή του σφάλματος, συνεπώς είναι και οι καλύτερες όσον αφορά την απόδοση των μεθόδων σχετικά με την προκατάληψη. Ακριβώς το ίδιο ισχύει και για την παραγωγή προβλέψεων τριών χρονικών περιόδων μπροστά.

Τώρα όσον αφορά το Μέσο Απόλυτο Σφάλμα (MAE) είναι ένας δείκτης σφάλματος που είναι συνδεδεμένος με την ακρίβεια της μεθόδου. Από τις τιμές των σφαλμάτων έχουμε το αντίστοιχο διάγραμμα:



Παρατηρώντας πιο προσεκτικά το παραπάνω διάγραμμα βλέπουμε, πως στην πρώτη περίπτωση η μέθοδος με την μεγαλύτερη ακρίβεια είναι η ETS και ακολουθεί η Naïve. Στη περίπτωση παραγωγής προβλέψεων τριών περιόδων οι αντίστοιχοι μέθοδοι είναι οι Holt και Theta, ενώ η Naïve βρίσκεται στην πέμπτη θέση σε σχέση με τις υπόλοιπες μεθόδους. Επίσης τα σφάλματα των προβλέψεων για την επόμενη περίοδο είναι μικρότερα σε σχέση με τα υπόλοιπα, κάτι το οποίο είναι λογικό σε σχέση με την ακρίβεια των μεθόδων όταν μεγαλώνει η περίοδος πρόβλεψης.

Στον παρακάτω πίνακα βρίσκονται τα ποσοστιαία σφάλματα των μεθόδων πρόβλεψης του αργού πετρελαίου:

Crude Oil	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	1,3824%	1,3811%	1,000303	2,4105%	2,4062%	1,749885
Holt	1,3822%	1,3808%	1,000138	2,4100%	<b>2,4049%</b>	<b>1,749095</b>
Damped	1,3819%	1,3809%	1,000405	<b>2,4098%</b>	2,4062%	1,750192
ETS	<b>1,3809%</b>	<b>1,3796%</b>	<b>0,999398</b>	2,4110%	2,4068%	1,749803
Moving Average	1,7225%	1,7204%	1,250486	2,6153%	2,6084%	1,900685
LRL	22,7127%	22,9075%	16,21558	22,9248%	23,0676%	16,33818
Naïve	1,3810%	1,3797%	0,999478	2,4113%	2,4072%	1,749932
Theta	1,3825%	1,3811%	1,000292	2,4104%	2,4057%	1,749564

Τα παραπάνω σφάλματα και πιο συγκεκριμένα τα MAPE και sMAPE επιλέχθηκαν για να μας δείξουν σε ποσοστιαία μορφή τις τιμές των σφαλμάτων. Όπως φαίνεται από τον παραπάνω πίνακα οι τιμές των σφαλμάτων σε σχέση με τις πραγματικές τιμές, είναι πολύ μικρές. Αυτό σημαίνει ότι οι μέθοδοι πρόβλεψης δίνουν αρκετά καλές προβλέψεις, ειδικά για την επόμενη περίοδο. Φυσικά αυτό δεν ισχύει για την μέθοδο της LRL και για του κινητού μέσου όρου. Όπως είναι λογικό τα σφάλματα για την παραγωγή προβλέψεων της τρίτης περιόδου είναι μεγαλύτερα.

Στη πρώτη περίπτωση καλύτερα αποτελέσματα δίνει η μέθοδος ETS και ακολουθεί η Naïve και για τα τρία ήδη σφάλματων. Στην δεύτερη περίπτωση με τον αυξημένο ορίζοντα πρόβλεψης, δεν πρέπει να αμεληθεί το γεγονός πως και εδώ η ακρίβεια της μεθόδου Theta είναι καλύτερη σε σχέση με πριν κατέχοντας την δεύτερη θέση, ενώ κυρίαρχη μέθοδος αποδείχτηκε η Holt σύμφωνα με αυτά τα σφάλματα και τον συγκεκριμένο ορίζοντα πρόβλεψης.

Τα παραπάνω σφάλματα είναι ιδανικά για την σύγκριση των μεθόδων σε διαφορετικές χρονοσειρές. Για αυτό το λόγο θα ακολουθήσουν σε επόμενη παράγραφο και διαχωρισμένα ανά βραχυπρόθεσμο (ημερήσιες χρονοσειρές) και μεσοπρόθεσμο (μηνιαίες χρονοσειρές) ορίζοντα, τα σχετικά διαγράμματα συμπεριλαμβανομένων όλων των αγαθών, για να φανεί σε ποια αγαθά ανταποκρίνονται καλύτερα οι μέθοδοι πρόβλεψης.

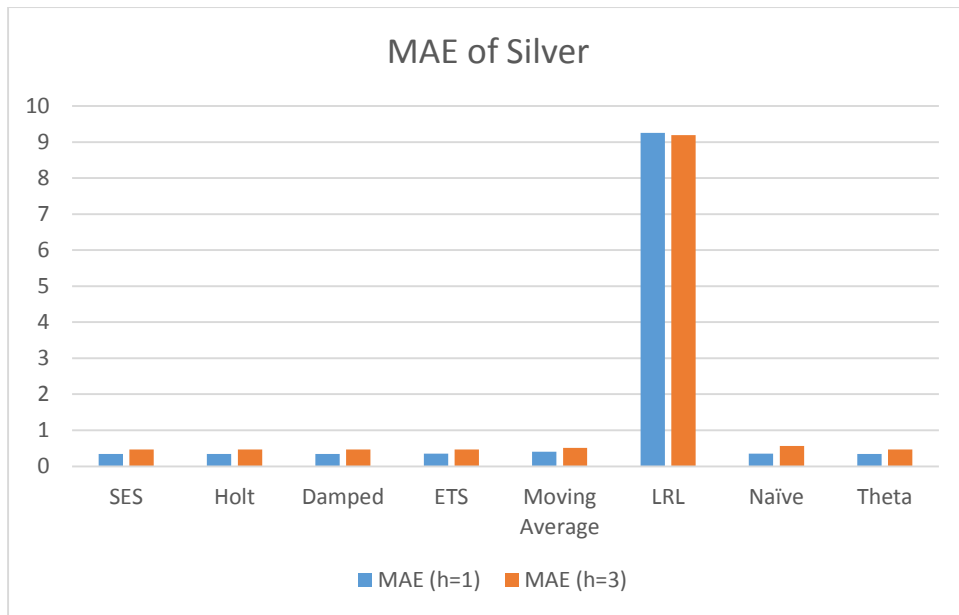
### 7.2.2 Ασήμι (Silver)

Και εδώ πρόκειται για μία χρονοσειρά με πολλά ημερήσια δεδομένα, για την οποία ακολουθήσαμε την ίδια διαδικασία με το αργό πετρέλαιο. Δηλαδή υπολογίσαμε τα σφάλματα προκατάληψης (ME), ακρίβειας (MAE) και ποσοστιαία (MAPE, sMAPE) με παραγωγή προβλέψεων για την επόμενη περίοδο και για τις επόμενες τρεις. Τα σφάλματα προκατάληψης και ακρίβειας φαίνονται συγκεντρωμένα παρακάτω:

<b>Silver</b>	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	0,002866	0,345857	0,005362	0,463996
Holt	<b>0,000524</b>	0,345795	<b>0,000985</b>	0,46384
Damped	0,001537	0,346209	0,002915	0,464003
ETS	-0,0037	0,347007	-0,00642	<b>0,463569</b>
Moving Average	0,005351	0,405568	0,007721	0,514417
LRL	9,234182	9,254111	9,144729	9,190223
Naïve	0,002605	0,348132	0,007783	0,563682
Theta	0,002471	<b>0,345701</b>	0,004935	0,463996

Και σε αυτή την περίπτωση η μέθοδος LRL παρουσιάζει τα μεγαλύτερα σφάλματα είτε πρόκειται για το μέσο σφάλμα, είτε για το μέσο απόλυτο σφάλμα. Στην περίπτωση της προκατάληψης φαίνεται πως σε σχέση με το πετρέλαιο οι μέθοδοι πρόβλεψης παρουσιάζουν απαισιοδοξία, λόγω των θετικών τιμών του μέσου σφάλματος. Η μόνη μέθοδος που διαφοροποιείται σχετικά είναι η ETS. Η μέθοδος με την μεγαλύτερη ακρίβεια παρουσιάζεται η Holt και για την παραγωγή προβλέψεων της επόμενης περιόδου και για την τρίτη περίοδο. Στη συνέχεια ακολουθούν η Theta για την πρώτη περίπτωση και η Damped για τη δεύτερη.

Για την ακρίβεια θα είναι χρήσιμο να δώσουμε βάση στο επόμενο διάγραμμα:



Και εδώ βλέπουμε πως η LRL έχει μεγάλη διαφορά σε σχέση με τις υπόλοιπες μεθόδους. Όπως και πριν οι ακρίβεια όλων των μεθόδων με εξαίρεση την LRL είναι αρκετά καλή. Παρατηρώντας προσεκτικά το διάγραμμα και τον παραπάνω πίνακα βλέπουμε πως η μέθοδος με την καλύτερη ακρίβεια είναι η Theta και ακολουθεί η Holt για την εξαγωγή προβλέψεων της επόμενης περιόδου. Στην άλλη περίπτωση παρόλο που έχει μειωθεί γενικά η ακρίβεια πρόβλεψης, όπως είναι λογικό, σαν καταλληλότερη μέθοδος τώρα παρουσιάζεται η ETS και ακολουθούν η Holt και η Theta. Δηλαδή σε σχέση με πριν η ETS παρουσίασε βελτίωση στην ακρίβειά της.

Τώρα όσον αφορά τα ποσοστιαία σφάλματα που εξαγάγαμε, παρουσιάζονται συνοπτικά στον ακόλουθο πίνακα που συγκεντρώνει και τις δύο περιπτώσεις παραγωγής προβλέψεων:

Silver	MAPE (h=1)	sMAPE (h=1)	MAE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAE (h=3)
SES	1,6887%	1,6848%	0,993357	2,3015%	2,2924%	1,340152
Holt	1,6885%	1,6844%	<b>0,993178</b>	<b>2,3009%</b>	<b>2,2913%</b>	1,339701
Damped	1,6911%	1,6871%	0,994368	2,3025%	2,2933%	1,340172
ETS	1,6995%	1,6944%	0,99666	2,3043%	2,2935%	<b>1,338917</b>
Moving Average	2,0153%	2,0078%	1,164855	2,5782%	2,5672%	1,485783
LRL	41,8680%	56,0941%	26,57927	41,6008%	55,7122%	26,54397
Naïve	1,6975%	1,6943%	0,99989	2,8163%	2,8040%	1,628073
Theta	<b>1,6880%</b>	<b>1,6841%</b>	0,993282	2,3016%	2,2924%	1,340153

Και εδώ βλέπουμε πως η LRL έχει τεράστια διαφορά σε σχέση με τις υπόλοιπες μεθόδους. Επίσης όλες οι μέθοδοι διαφέρουν στην ακρίβεια τους στο δεύτερο δεκαδικό πράγμα που σημαίνει πως τα δεδομένα του ασημιού ανταποκρίνονται πολύ καλά στις μεθόδους προβλέψεων. Στη περίπτωση του σφάλματος MAPE και sMAPE και για παραγωγή προβλέψεων της επόμενης περιόδου, η μέθοδος Theta έχει την

μικρότερη τιμή και ακολουθεί η Holt. Αυτό σημαίνει πως το ποσοστό του σφάλματος των προβλέψεων της μεθόδου Theta σε σχέση με τις πραγματικές τιμές είναι αρκετά μικρό. Στη περίπτωση του κανονικοποιημένου σφάλματος MAE η Holt παρουσιάζει μικρή τιμή και ακολουθεί με πολύ μικρή διαφορά η μέθοδος Theta.

Τώρα στην περίπτωση παραγωγής προβλέψεων για την τρίτη περίοδο και ο υπολογισμός των αντίστοιχων σφαλμάτων δίνει περίπου τα ίδια αποτελέσματα όσον αφορά τη μέθοδο με την μικρότερη τιμή. Για τα σφάλματα MAPE και sMAPE την μικρότερη τιμή έχει η μέθοδος Holt και αμέσως μετά ακολουθεί η Theta. Όσον αφορά το κανονικοποιημένο σφάλμα η ETS έχει βελτιωθεί σε σχέση με την πρώτη περίπτωση και ακολουθεί η Holt.

### 7.2.3 Αλουμίνιο (Aluminum)

Στη περίπτωση αυτή διατίθενται μία ημερήσια, αλλά και μία μηνιαία χρονοσειρά για τις τιμές του αλουμινίου. Για αυτό το λόγο η παρουσίαση των αποτελεσμάτων θα γίνει ξεχωριστά, για τα ημερήσια δεδομένα και για τα μηνιαία, όπως φαίνεται παρακάτω:

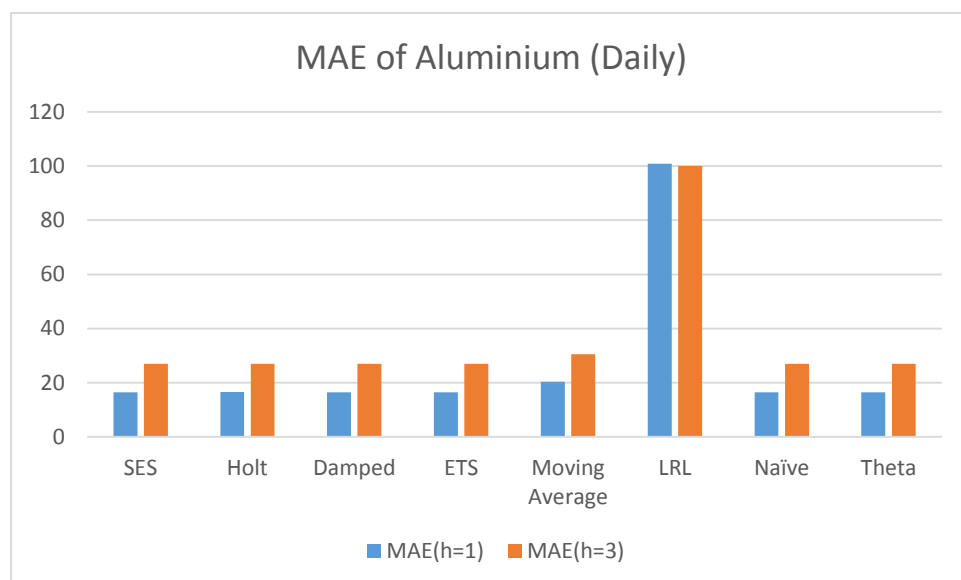
#### 7.2.3.1 Ημερήσια Δεδομένα

Για τα ημερήσια δεδομένα ακολουθήθηκε η ίδια ακριβώς μεθοδολογία με προηγουμένως και εξαγάγαμε τα εξής σφάλματα:

<b>Aluminum (Daily)</b>	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-1,64303	16,47886	-4,65928	27,04673
Holt	-1,54791	16,5319	<b>-4,27846</b>	27,03586
Damped	-1,64099	16,47856	-4,65359	27,0452
ETS	-1,64303	16,47886	-4,65928	27,04674
Moving Average	-3,39727	20,34171	-6,59626	30,57665
LRL	82,7864	100,824	81,53814	100,0794
Naïve	-1,64286	16,47884	-4,65909	27,04679
Theta	<b>-1,48205</b>	<b>16,46213</b>	-4,32455	<b>26,9965</b>

Σε αυτή την περίπτωση βλέπουμε πως υπάρχει αισιοδοξία στις προβλέψεις, σύμφωνα με το δείκτη προκατάληψης ME, λόγω του αρνητικού προσήμου για όλες τις μεθόδους, πλην της LRL η οποία παρουσιάζει και εδώ μεγάλα σφάλματα. Τη μικρότερη τιμή κατά απόλυτη τιμή παρουσιάζει η μέθοδος Theta και ακολουθεί η μέθοδος Damped, στην περίπτωση προβλέψεων της επόμενης περιόδου. Στην άλλη περίπτωση η μέθοδος Holt και η μέθοδος Theta έχουν την μικρότερη απόλυτη τιμή του μέσου σφάλματος εν συγκρίσει με τις άλλες μεθόδους. Φυσικά και εδώ οι απόλυτες τιμές των μέσων σφαλμάτων είναι μεγαλύτερες στην περίπτωση της περιόδου πρόβλεψης για την τρίτη περίοδο, εν αντιθέσει με τις προβλέψεις που έγιναν για την επόμενη περίοδο.

Παρακάτω βρίσκεται το διάγραμμα του μέσου απόλυτου σφάλματος για τα ημερήσια δεδομένα του αλουμινίου, συμπεριλαμβανομένου και των δύο περιπτώσεων σχετικά με τον ορίζοντα πρόβλεψης:



Αρχικά είναι πολύ εύκολο να αναγνωρίσει κανείς πως η μέθοδος με τα μεγαλύτερα σφάλματα είναι η LRL και ακολουθεί με σχετικά μεγάλη απόκλιση σε σχέση με τις υπόλοιπες μεθόδους η μέθοδος κινητού μέσου όρου. Επίσης είναι εμφανές πως η ακρίβεια της Theta είναι καλύτερη σε σχέση με τις υπόλοιπες και για τις δύο περιπτώσεις. Πολύ καλά αποτελέσματα δίνουν επίσης η Holt και η Damped.

Τώρα όσον αφορά την επιλογή της καταλληλότερης μεθόδου για τα ημερήσια δεδομένα του αλουμινίου λαμβάνουμε υπόψιν τα ποσοστιαία σφάλματα MAPE, sMAPE και MASe που βρίσκονται συγκεντρωμένα στον παρακάτω πίνακα και για τις δύο περιπτώσεις εξαγωγής προβλέψεων:

Aluminum (Daily)	MAPE (h=1)	sMAPE (h=1)	MASe (h=1)	MAPE (h=3)	sMAPE (h=3)	MASe (h=3)
SES	0,8780%	0,8772%	0,997914	1,4411%	1,4381%	1,650625
Holt	0,8808%	0,8801%	1,001127	1,4402%	1,4375%	1,649961
Damped	0,8780%	0,8772%	0,997897	1,4410%	1,4380%	1,650531
ETS	0,8780%	0,8772%	0,997915	1,4411%	1,4381%	1,650625
Moving Average	1,0832%	1,0814%	1,231838	1,6283%	1,6236%	1,866051
LRL	5,1891%	5,4187%	6,105626	5,1553%	5,3821%	6,107709
Naïve	0,8780%	0,8772%	0,997913	1,4411%	1,4381%	1,650629
Theta	<b>0,8771%</b>	<b>0,8764%</b>	<b>0,996901</b>	<b>1,4382%</b>	<b>1,4355%</b>	<b>1,647559</b>

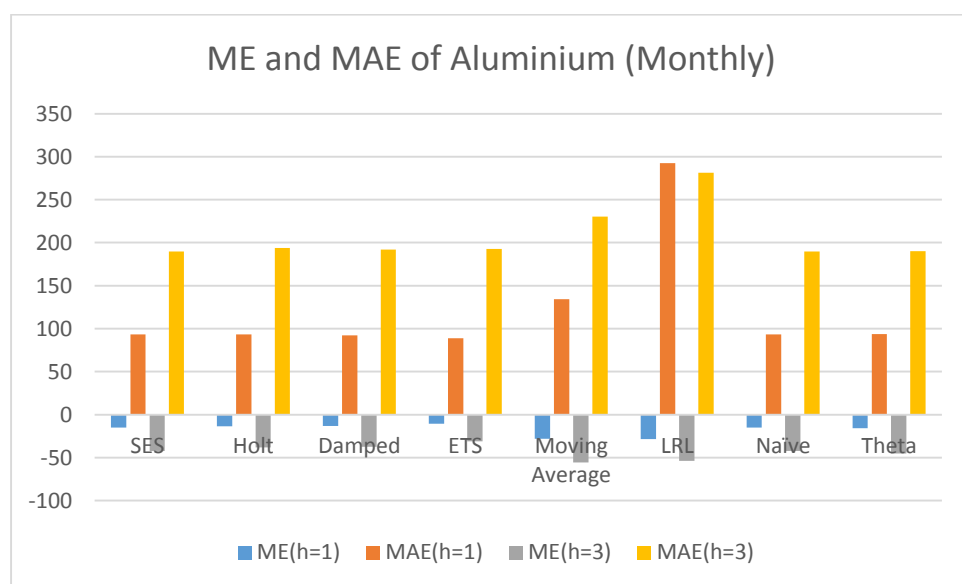
Εδώ βλέπουμε πως για την ημερήσια χρονοσειρά του αλουμινίου η καταλληλότερη μέθοδος πρόβλεψης με βάση τα ποσοστιαία σφάλματα είναι η Theta και ακολουθούν η Damped και η Holt. Τελευταίες είναι πάλι οι LRL και κινητού μέσου όρου.

### 7.2.3.2 Μηνιαία Δεδομένα

Στα μηνιαία δεδομένα της χρονοσειράς του αλουμινίου λάβαμε τα εξής αποτελέσματα σχετικά με τα σφάλματα ME και MAE για τις προβλέψεις που εξαγάγαμε και για την πρόβλεψη της επόμενης περιόδου και για την τρίτη επόμενη:

Aluminum (Monthly)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-14,9076	93,31287	-42,0831	189,8623
Holt	-13,5312	93,45493	-38,0119	193,7038
Damped	-13,0618	92,1689	-37,5221	191,9686
ETS	<b>-10,4128</b>	<b>88,96161</b>	<b>-31,0364</b>	192,6015
Moving Average	-27,9947	134,1389	-55,7266	230,2739
LRL	-28,6295	292,5786	-53,6084	281,2735
Naïve	-14,8872	93,29093	-42,0621	<b>189,8395</b>
Theta	-15,9861	93,48438	-45,3463	189,9362

Στο παρακάτω διάγραμμα βρίσκονται οι τιμές των παραπάνω σφαλμάτων και για τις δύο περιπτώσεις και για τις οχτώ μεθόδους:



Παρατηρώντας το διάγραμμα διαπιστώνουμε πως και στις δύο περιπτώσεις οι μέθοδοι πρόβλεψης έχουν έντονο το στοιχείο της αισιοδοξίας σχετικά με τις προβλέψεις τους. Αυτό γίνεται αντιληπτό από την αρνητική τιμή του μέσου σφάλματος που έχει κάθε μέθοδος. Επίσης εύκολα διαπιστώνεται πως το μικρότερο μέσο σφάλμα κατά απόλυτη τιμή το έχει η μέθοδος ETS και για τις δύο περιπτώσεις περιόδων πρόβλεψης και ακολουθεί η μέθοδος Damped.

Τώρα όσον αφορά την ακρίβεια σφάλματος των μεθόδων, στην περίπτωση εξαγωγής προβλέψεων για την επόμενη περίοδο η μέθοδος με την καλύτερη ακρίβεια εμφανίζεται η ETS ακολουθούμενη από την Damped. Στην περίπτωση εξαγωγής προβλέψεων για την τρίτη περίοδο υπάρχει μία διαφοροποίηση και πιο συγκεκριμένα καλύτερη μέθοδος αποδεικνύεται η Naïve ακολουθούμενη από την



SES. Επίσης η μέθοδος Theta βελτιώθηκε σε σχέση με την πρώτη περίπτωση όπου είναι στην τρίτη θέση από την έκτη. Αντιθέτως η ETS από την μέθοδο με τη μεγαλύτερη ακρίβεια που ήταν στη πρώτη περίπτωση, εδώ έχασε αρκετά από την ακρίβειά της.

Για να γίνει η επιλογή της μεθόδου για αυτά τα δεδομένα θα πρέπει να μελετήσουμε τα ποσοστιαία σφάλματα, όπως έχουμε κάνει και στις προηγούμενες περιπτώσεις. Πιο συγκεκριμένα οι τιμές των σφαλμάτων είναι:

<b>Aluminum (Monthly)</b>	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	4,6470%	4,5886%	0,996022	9,9337%	9,3636%	2,01927
Holt	4,6513%	4,5986%	0,997539	10,1749%	9,6354%	2,060126
Damped	4,5894%	4,5354%	0,983812	10,0738%	9,5505%	2,041673
ETS	<b>4,4282%</b>	<b>4,3956%</b>	<b>0,949577</b>	10,0421%	9,7055%	2,048403
Moving Average	6,8926%	6,6502%	1,4318	12,2649%	11,3228%	2,449067
LRL	14,6726%	14,3022%	3,122986	14,4930%	13,9343%	2,99147
Naïve	4,6461%	4,5878%	0,995788	<b>9,9325%</b>	9,3627%	<b>2,019029</b>
Theta	4,6570%	4,5962%	0,997853	9,9448%	<b>9,3586%</b>	2,020056

Όπως φαίνεται από τα παραπάνω η επιλογή της μεθόδου με βάση τα ποσοστιαία σφάλματα είναι πολύ εμφανής, παρόλο που όλες οι μέθοδοι δίνουν αρκετά καλά αποτελέσματα. Στην περίπτωση παραγωγής προβλέψεων για την επόμενη περίοδο καλύτερα ανταποκρίνεται η ETS και ακολουθεί η Damped. Αντιθέτως στην περίπτωση εξαγωγής προβλέψεων για τις επόμενες τρεις περιόδους η μέθοδος Naïve έχει καλά αποτελέσματα και ακολουθεί η SES, ενώ στη περίπτωση του συμμετρικού σφάλματος την μικρότερη τιμή την έχει η Theta εξαιτίας της πολύ καλής ανταπόκρισης της μεθόδου στην αύξηση του ορίζοντα πρόβλεψης κάτι που έχει παρατηρηθεί και στα προηγούμενα δεδομένα.

## **7.2.4 Φυσικό Αέριο (Natural Gas)**

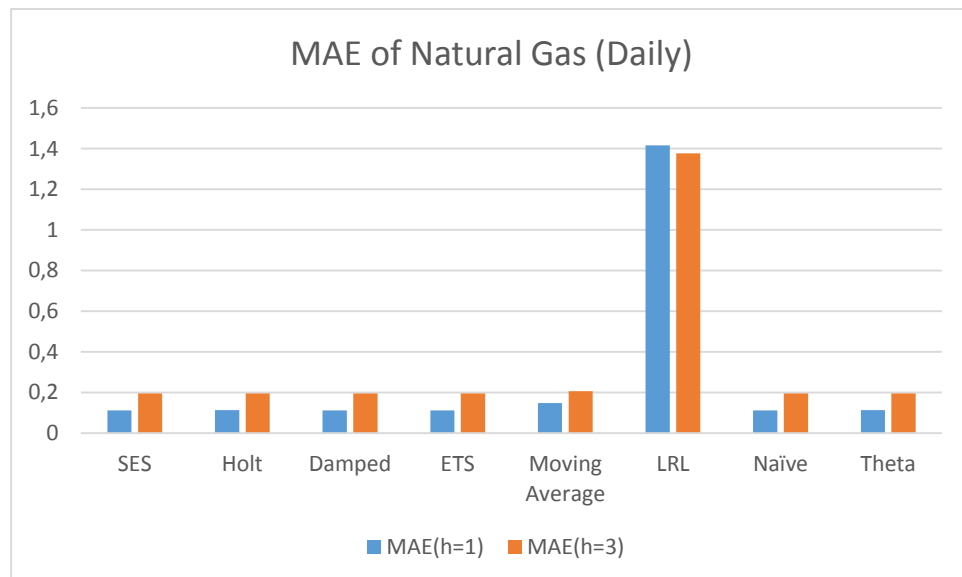
Για το Φυσικό Αέριο όπως και για την περίπτωση του αλουμινίου υπάρχουν διαθέσιμες προς επεξεργασία μία ημερήσια και μία μηνιαία χρονοσειρά, οι οποίες αναλύονται παρακάτω.

### **7.2.4.1 Ημερήσια Δεδομένα**

Με βάση τον συγκεκριμένο χρονικό ορίζοντα της χρονοσειράς βγάλαμε τα εξής αποτελέσματα για το μέσο σφάλμα και για το μέσο απόλυτο σφάλμα των μεθόδων πρόβλεψης που εφαρμόσαμε και πιο συγκεκριμένα για προβλέψεις που αφορούν την επόμενη περίοδο και για την τρίτη περίοδο:

Natural Gas (Daily)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,001525	0,111958	-0,00511	0,194486
Holt	-0,000481	0,112625	<b>-0,00058</b>	0,194959
Damped	-0,001858	0,11203	-0,00609	0,194652
ETS	-0,001563	0,111676	-0,00515	<b>0,194458</b>
Moving Average	-0,003236	0,148248	-0,00675	0,205264
LRL	1,4162515	1,416252	1,376186	1,376186
Naïve	-0,001563	<b>0,111672</b>	-0,00515	0,194459
Theta	<b>-0.0000359</b>	0,112275	-0,00239	0,194462

Είναι εμφανές πως για ακόμη μία φορά οι μέθοδοι πρόβλεψης στο σύνολό τους έχουν έντονο το στοιχείο της αισιοδοξίας, λόγω του αρνητικού προσήμου των μέσων σφαλμάτων. Η μέθοδος με την μικρότερη κατά απόλυτη τιμή του μέσου σφάλματος είναι η Theta και ακολουθεί η Holt. Όσον αφορά την ακρίβεια των μεθόδων προβλέψεων και πιο συγκεκριμένα των μέσων απόλυτων σφαλμάτων κρίνεται σκόπιμη η παρουσίαση του επόμενου διαγράμματος:



Αυτό που είναι γεγονός σχετικά με τα απόλυτα σφάλματα είναι πως οι διαφορές των σφαλμάτων των προβλέψεων κάθε μεθόδου, εκτός της LRL και του κινητού μέσου όρου, είναι ελάχιστες και δύσκολα ανιχνεύσιμες από το διάγραμμα. Παρόλα αυτά η μέθοδος με το μικρότερο απόλυτο σφάλμα για την περίπτωση προβλέψεων της επόμενης περιόδου είναι η Naïve και με ελάχιστη διαφορά είναι η ETS. Στην περίπτωση των προβλέψεων για την τρίτη επόμενη περίοδο καταλληλότερη είναι η ETS ακολουθούμενη από την Naïve. Ο λόγος που οι διαφορές είναι τόσο μικρές είναι πιθανότατα στις μικρές διακυμάνσεις των τιμών διαχρονικά και στην εν γένει διατήρηση συγκεκριμένης τιμής για κάποιο υπολογίσιμο διάστημα.

Για να φανεί πόσο μικρή είναι η διαφορά των σφαλμάτων των προβλέψεων σε σχέση με τις πραγματικές τιμές θα πρέπει να ανατρέξουμε στα αποτελέσματα των ποσοστιαίων σφαλμάτων για κάθε μέθοδο. Επομένως:

<b>Natural Gas (Daily)</b>	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	2,6377%	2,6285%	1,000622	4,7114%	4,6544%	1,791248
Holt	2,6532%	2,6428%	1,006582	4,7176%	4,6651%	1,7956
Damped	2,6393%	2,6296%	1,001269	4,7151%	4,6562%	1,792774
ETS	2,6329%	2,6240%	0,998104	4,7111%	4,6536%	<b>1,790989</b>
Moving Average	3,5337%	3,5091%	1,324968	4,9400%	4,8871%	1,890513
LRL	34,2205%	43,3833%	12,65775	33,5531%	42,4097%	12,67489
Naïve	<b>2,6329%</b>	<b>2,6239%</b>	<b>0,998073</b>	4,7111%	<b>4,6536%</b>	1,790994
Theta	2,6446%	2,6365%	1,00346	<b>4,7083%</b>	4,6546%	1,791023

Εδώ είναι εμφανές αυτό που αναφέρθηκε προηγουμένως, δηλαδή η μικρή τιμή που παρουσιάζουν τα σφάλματα και τα οποία κυμαίνονται στο 2,62%-2,65%. Μικρότερα ποσοστιαία σφάλματα έχουν οι μέθοδοι ETS και Naïve για την περίπτωση πρόβλεψης για την επόμενη περίοδο και για τα τρία είδη σφαλμάτων. Τώρα αυτό που χρήζει προσοχής είναι η βελτίωση της μεθόδου Theta στην περίπτωση αύξησης του ορίζοντα πρόβλεψης σε σημείο που αποδείχτηκε η καλύτερη μέθοδος με βάση το MAPE. Και εδώ οι Naïve και ETS έχουν πολύ καλά αποτελέσματα.

#### **7.2.4.2 Μηνιαία Δεδομένα**

Στα μηνιαία δεδομένα της χρονοσειράς του φυσικού αερίου λάβαμε τα εξής αποτελέσματα σχετικά με τα σφάλματα ME και MAE για τις προβλέψεις που εξαγάγαμε και για την πρόβλεψη της επόμενης περιόδου και για την τρίτη επόμενη:

<b>Natural Gas (Monthly)</b>	ME(h=1)	MAE(h=1)	ME(h=1)	MAE(h=3)
SES	-0,031573	0,283472	-0,07574	0,559485
Holt	-0,04061	0,2863	-0,10255	0,565762
Damped	-0,031562	0,283414	-0,07568	<b>0,559395</b>
ETS	-0,031529	0,283401	-0,07569	0,559487
Moving Average	-0,062542	0,396554	-0,10897	0,607126
LRL	-2,898256	2,898256	-2,92175	2,921753
Naïve	<b>-0,031525</b>	<b>0,28339</b>	<b>-0,07569</b>	0,559483
Theta	-0,041216	0,285854	-0,0767	0,559569

Αυτό που είναι ορατό είναι η αισιοδοξία των μεθόδων προβλέψεων και για τις δύο περιπτώσεις εξαγωγής προβλέψεων ανάλογα με την περίοδο πρόβλεψης. Αυτό γίνεται κατανοητό από την εξέταση του προσήμου του μέσου σφάλματος το οποίο είναι αρνητικό για όλες τις μεθόδους και για τις δύο περιπτώσεις. Γενικά, η μέθοδος Naïve όπως επίσης και η μέθοδος ETS, είναι οι δύο μέθοδοι που παρουσιάζουν τις μικρότερες κατά απόλυτη τιμή, τιμές του μέσου σφάλματος, συνεπώς είναι και οι καλύτερες όσον αφορά την απόδοση των μεθόδων σχετικά με την προκατάληψη και για την παραγωγή προβλέψεων για την επόμενη περίοδο, αλλά και για την τρίτη περίοδο.

Σχετικά με την ακρίβεια των προβλέψεων για κάθε μέθοδο θα πρέπει να ανατρέξουμε στις τιμές των μέσων απόλυτων σφαλμάτων κάθε μεθόδου. Στην

περίπτωση παραγωγής προβλέψεων για την επόμενη περίοδο η μέθοδος Naïve και η μέθοδος ETS υποδηλώνουν την μεγαλύτερη ακρίβεια με βάση τα αποτελέσματά τους. Στην περίπτωση πρόβλεψης για την τρίτη περίοδο η πιο αξιόπιστη μέθοδος για τα μηνιαία δεδομένα με βάση την ακρίβεια αποδείχτηκε η Damped ακολουθούμενη από την Naïve.

Τώρα για την επιλογή της καταλληλότερης μεθόδου πρόβλεψης θα πρέπει να ελέγξουμε τα ποσοστιαία σφάλματα δίνοντας παραπάνω έμφαση στο MAPE. Συγκεκριμένα έχουμε:

<b>Natural Gas (Monthly)</b>	MAPE (h=1)	sMAPE (h=1)	MAE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAE (h=3)
SES	7,7800%	7,7499%	1,002523	15,9227%	15,4785%	2,046898
Holt	7,8652%	7,8162%	1,012524	16,1431%	15,5840%	2,069861
Damped	7,7784%	7,7484%	1,002318	<b>15,9203%</b>	<b>15,4761%</b>	<b>2,046568</b>
ETS	7,7782%	7,7482%	1,002273	15,9225%	15,4786%	2,046904
Moving Average	11,1749%	10,9739%	1,402446	17,8347%	16,9554%	2,221194
LRL	88,1950%	58,1577%	10,24993	89,9043%	58,8427%	10,68934
Naïve	<b>7,7779%</b>	<b>7,7479%</b>	<b>1,002232</b>	15,9224%	15,4785%	2,046888
Theta	7,8585%	7,8063%	1,010948	15,9271%	15,4784%	2,047203

Παρατηρώντας το συνολικό ποσοστό των σφαλμάτων διαπιστώνουμε πως τα σφάλματα είναι μεγαλύτερα σε σχέση με τις προηγούμενες εξεταζόμενες χρονοσειρές για την περίπτωση της μηνιαίας χρονοσειράς του φυσικού αερίου. Πιο αναλυτικά η μέθοδος Naïve και ακολούθως η μέθοδος ETS είχαν τις μικρότερες τιμές σφαλμάτων σχετικά με τις υπόλοιπες για την περίπτωση πρόβλεψης για την επόμενη περίοδο.

Στη περίπτωση της πρόβλεψης με ορίζοντα πρόβλεψης τρεις περιόδους η καλύτερη μέθοδος που προκύπτει με βάση τα σφάλματα είναι η Damped και εν συνεχεία η Naïve.

### 7.2.5 Χρυσός (Gold)

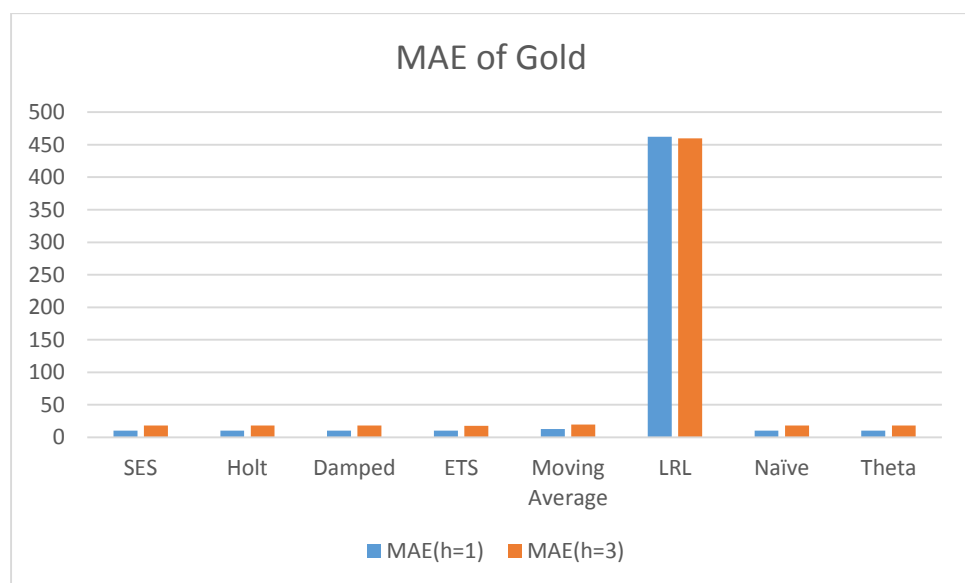
Για τον χρυσό διαθέτουμε και εξετάζουμε μία ημερήσια χρονοσειρά για την οποία έχουμε πάρει τα εξής αποτελέσματα σχετικά με την προκατάληψη και την ακρίβεια των μεθόδων πρόβλεψης. Εξετάσαμε δύο περιπτώσεις, η πρώτη περίπτωση αναφέρεται στην εξαγωγή προβλέψεων για την επόμενη περίοδο και η δεύτερη περίπτωση για την εξαγωγή προβλέψεων για την τρίτη περίοδο. Συγκεκριμένα:

<b>Gold</b>	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	0,267768	10,22064	0,711062	17,95227
Holt	<b>0,016866</b>	10,21392	<b>-0,01014</b>	17,91765
Damped	0,257135	10,22264	0,680627	17,9621
ETS	-0,13153	<b>10,20269</b>	-0,41912	<b>17,86774</b>

Moving Average	0,507559	12,79476	0,92874	19,48698
LRL	462,3323	462,3323	459,6781	459,6781
Naïve	0,252584	10,23137	0,696777	17,95116
Theta	0,237406	10,21908	0,649487	17,94873

Στην πρώτη περίπτωση και όσον αφορά την προκατάληψη όλες οι μέθοδοι, εκτός από την ETS έχουν έντονο το στοιχείο της απαισιοδοξίας, λόγω του θετικού προσήμου του ME. Επίσης κατά απόλυτη τιμή οι μέθοδοι με την μικρότερη τιμή είναι η Holt και η ETS. Στην δεύτερη περίπτωση όλες οι μέθοδοι, πλην Holt και ETS έχουν το στοιχείο της απαισιοδοξίας στις προβλέψεις τους, λόγω πάλι του θετικού προσήμου του ME.

Στο παρακάτω διάγραμμα φαίνεται αμέσως η ακρίβεια των μεθόδων με βάση το μέσο απόλυτο σφάλμα και για τις δύο περιπτώσεις:



Όπως φαίνεται και από τον πίνακα, αλλά και από το διάγραμμα οι μέθοδοι με την μεγαλύτερη ακρίβεια και για τις δύο περιπτώσεις είναι οι ETS και Holt, ενώ θα πρέπει να αναφερθεί και η βελτίωση της μεθόδου Theta για την δεύτερη περίπτωση, δηλαδή για την παραγωγή προβλέψεων μεγαλύτερου προβλεπτικού ορίζοντα.

Για την περαιτέρω αξιολόγηση των μεθόδων θα πρέπει να ανατρέξουμε στα ποσοστιαία σφάλματα τα οποία είναι συγκεντρωμένα στον παρακάτω πίνακα:

Gold	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	0,9208%	0,9203%	0,99903	1,6228%	1,6215%	1,758864
Holt	0,9204%	0,9197%	0,998373	1,6200%	1,6177%	1,755473
Damped	0,9210%	0,9205%	0,999225	1,6236%	1,6221%	1,759827
ETS	<b>0,9195%</b>	<b>0,9186%</b>	<b>0,997275</b>	<b>1,6152%</b>	<b>1,6119%</b>	<b>1,750583</b>
Moving Average	1,1556%	1,1547%	1,250641	1,7687%	1,7674%	1,909227
LRL	37,9562%	48,2211%	45,19127	37,7386%	47,9375%	45,03673
Naïve	0,9219%	0,9214%	1,000078	1,6224%	1,6210%	1,758755
Theta	0,9207%	0,9201%	0,998877	1,6225%	1,6210%	1,758518

Χωρίς αμφιβολία περί της κατάλληλης επιλογής της μεθόδου για την περίπτωση της ημερήσιας χρονοσειράς του χρυσού είναι η ETS κι ακολούθως η Damped. Βεβαίως τα σφάλματα των μεθόδων είναι πάρα πολύ μικρά σε σχέση με τα πραγματικά δεδομένα, εκτός φυσικά της LRL.

## 7.2.6 Χαλκός (Copper)

Όσον αφορά αυτό το αγαθό υπάρχουν διαθέσιμες μία ημερήσια και μία μηνιαία χρονοσειρά, όπου παρουσιάζονται ξεχωριστά παρακάτω:

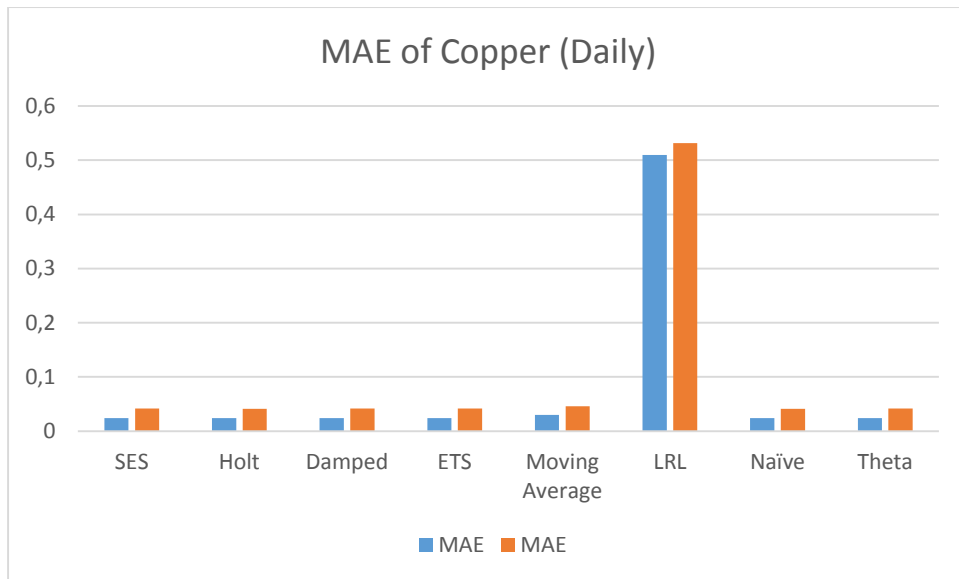
### 7.2.6.1 Ημερήσια Δεδομένα

Για τα ημερήσια δεδομένα ακολουθήθηκε η ίδια ακριβώς μεθοδολογία με προηγουμένως και εξαγάγαμε τα εξής σφάλματα:

<b>Copper (Daily)</b>	ME	MAE	ME	MAE
SES	-0,00191	0,023893	-0,00627	0,041396
Holt	<b>-0,00156</b>	<b>0,023883</b>	<b>-0,00527</b>	0,041362
Damped	-0,00174	0,023887	-0,00576	0,04152
ETS	-0,00191	0,023893	-0,00627	0,041396
Moving Average	-0,00355	0,029842	-0,00823	0,045769
LRL	-0,50944	0,509445	-0,53137	0,531365
Naïve	-0,00177	0,023922	-0,00611	<b>0,041287</b>
Theta	-0,00204	0,023898	-0,00652	0,041405

Όπως σε όλες τις περιπτώσεις έτσι και εδώ, οι μέθοδοι παρουσιάζονται αισιόδοξες στις προβλέψεις τους και για τις δύο περιόδους πρόβλεψης, εξαιτίας του αρνητικού προσήμου του ME που παρουσιάζεται σε όλους τις μεθόδους. Κατά απόλυτη τιμή η μέθοδος Holt έχει την μικρότερη τιμή σε σχέση με τις άλλες μεθόδους και ακολουθεί η Damped.

Στη συνέχεια χρησιμοποιώντας το παρακάτω διάγραμμα γίνεται μία ανάλυση της ακρίβειας των μεθόδων για τις δύο περιπτώσεις. Στην πρώτη περίπτωση αποδείχτηκε πως η μέθοδοι με την μεγαλύτερη ακρίβεια είναι οι Holt και Damped. Για την περίπτωση παραγωγής προβλέψεων της τρίτης χρονικής περιόδου ακριβέστερη μέθοδος αποδεικνύεται η Naïve και ακολούθως η Damped. Γενικά οι μέθοδοι δίνουν πολύ μικρές τιμές σφαλμάτων στο σύνολό τους.



Για την αξιολόγηση των μεθόδων θα πρέπει να ληφθούν υπόψη τα ποσοστιαία σφάλματα των μεθόδων σε σχέση με τις πραγματικές τιμές. Αυτά τα σφάλματα συγκεντρώνονται στον παρακάτω πίνακα:

Copper (Daily)	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	0,8108%	0,8097%	1,000997	1,4194%	1,4150%	1,736788
Holt	<b>0,8103%</b>	<b>0,8093%</b>	<b>1,000554</b>	1,4176%	1,4138%	1,735355
Damped	0,8105%	0,8095%	1,000741	1,4236%	1,4195%	1,74198
ETS	0,8108%	0,8097%	1,000998	1,4194%	1,4150%	1,736795
Moving Average	1,0113%	1,0091%	1,250223	1,5649%	1,5587%	1,920256
LRL	17,4717%	15,8870%	21,34305	18,5725%	16,7535%	22,29363
Naïve	0,8118%	0,8108%	1,002197	<b>1,4159%</b>	<b>1,4117%</b>	<b>1,732199</b>
Theta	0,8110%	0,8099%	1,001214	1,4198%	1,4153%	1,73716

Παρατηρώντας κανείς τα παραπάνω σφάλματα φαίνεται πως όλοι οι μέθοδοι πρόβλεψης δίνουν αξιόπιστα αποτελέσματα Στην περίπτωση πρόβλεψης για την επόμενη περίοδο τα καλύτερα αποτελέσματα δίνει η μέθοδος Holt και ακολουθεί η Damped. Στην περίπτωση προβλέψεων για την τρίτη περίοδο καλύτερη μέθοδος εμφανίζεται η Naïve.

### 7.2.6.2 Μηνιαία Δεδομένα

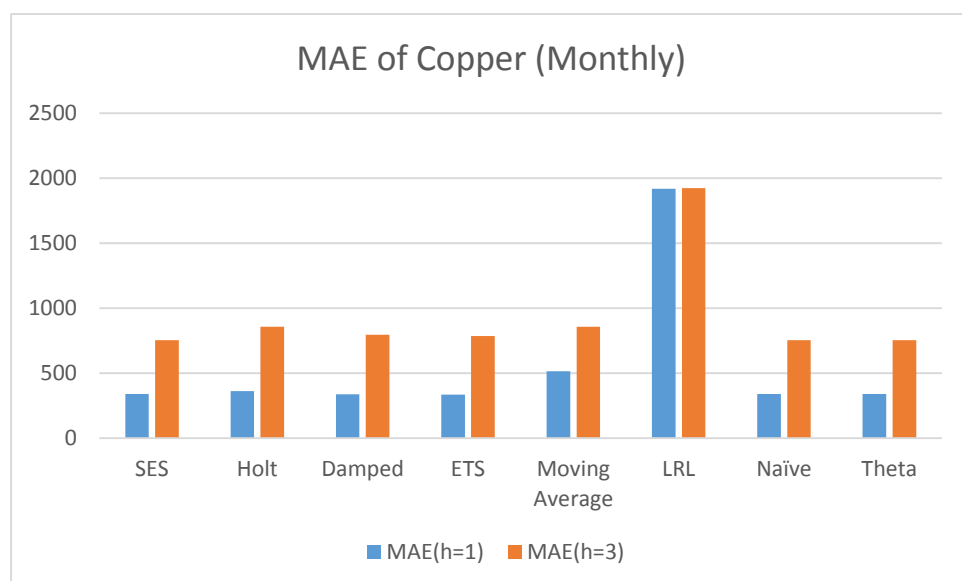
Στην περίπτωση των μηνιαίων δεδομένων έχουμε τα εξής σφάλματα που εξετάζουν την προκατάληψη και την ακρίβεια των μεθόδων:

Copper (Monthly)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-33,5045	340,2575	-85,627	752,4454
Holt	-48,6975	361,2225	-131,973	856,8677
Damped	-27,0681	338,2194	-66,8279	795,6463
ETS	<b>-21,5756</b>	<b>335,2794</b>	<b>-54,2106</b>	786,2879
Moving Average	-57,4482	514,3752	-114,2	856,9351
LRL	1743,227	1916,995	1753,918	1922,985

Naïve	-33,502	340,2455	-85,624	<b>752,4376</b>
Theta	-39,6807	340,0413	-86,2332	752,449

Όσον αφορά την προκατάληψη ισχύει ότι και για τα ημερήσια δεδομένα, σε σχέση με την τιμή του προσήμου του ME. Τώρα κατά απόλυτη τιμή την μικρότερη τιμή την έχει η ETS και για τις δύο περιπτώσεις εξαγωγής προβλέψεων ανάλογα με την περίοδο πρόβλεψης.

Τώρα όσον αφορά την ακρίβεια εξετάζοντας το παρακάτω διάγραμμα φαίνεται εκ πρώτης όψεως η βελτίωση της μεθόδου Theta για την περίπτωση εξαγωγής προβλέψεων για την τρίτη περίοδο, σε σχέση με την παραγωγή προβλέψεων για την επόμενη περίοδο. Επίσης στην πρώτη περίπτωση συνολικά μεγαλύτερη ακρίβεια έχει η ETS και για την δεύτερη περίπτωση η Naïve.



Σχετικά με την αξιολόγηση των μεθόδων με βάση τα ποσοστιαία σφάλματα έχουμε τα εξής:

Copper (Monthly)	MAPE (h=1)	sMAPE (h=1)	MAE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAE (h=3)
SES	5,3023%	5,1603%	1,000652	12,8626%	11,7108%	2,17425
Holt	5,7949%	5,7323%	1,062307	15,2643%	15,2104%	2,475987
Damped	5,2109%	5,1609%	0,994658	13,4683%	12,8689%	2,299083
ETS	<b>5,1642%</b>	<b>5,1246%</b>	<b>0,986012</b>	13,3010%	12,7958%	2,272041
Moving Average	8,5412%	8,0215%	1,512709	15,0800%	13,3978%	2,476182
LRL	25,3943%	30,2009%	5,637626	25,5693%	30,4182%	5,556617
Naïve	5,3020%	5,1601%	1,000616	<b>12,8624%</b>	11,7106%	<b>2,174228</b>
Theta	5,2986%	5,1519%	1,000016	12,8630%	<b>11,7102%</b>	2,174261

Και εδώ η μέθοδος που επιλέγεται εν τέλει και αξιολογείται καλύτερα για την περίπτωση προβλέψεων της επόμενης περιόδου είναι η ETS. Τώρα όσον αφορά την δεύτερη περίπτωση η μέθοδος που αξιολογείται ως η καταλληλότερη είναι η Naïve με πολύ καλά αποτελέσματα της Theta και της SES.



### 7.2.7 Πλατίνα (Platinum)

Σε αυτή την περίπτωση υπάρχει διαθέσιμη μία ημερήσια χρονοσειρά, όπου ο πίνακας είναι ο εξής:

Platinum	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,61353	13,45931	-1,89099	24,58042
Holt	-0,807	13,48067	-2,4663	24,70797
Damped	<b>-0,48365</b>	<b>13,45354</b>	<b>-1,48404</b>	24,80291
ETS	-0,61353	13,45931	-1,89099	24,58042
Moving Average	-1,19267	17,5266	-2,52589	27,29494
LRL	-90,7852	239,4427	-101,701	247,3519
Naïve	-0,61348	13,45922	-1,89092	<b>24,58028</b>
Theta	-0,76785	13,45825	-2,19434	24,59558

Για την εκτίμηση της προκατάληψης ισχύει ότι όλες οι μέθοδοι έχουν το στοιχείο της αισιοδοξίας λόγω του αρνητικού προσήμου του ME για όλες τις μεθόδους. Τώρα κατά απόλυτη τιμή η μέθοδος με την μικρότερη τιμή είναι η Damped και για τις δύο περιπτώσεις.

Για την εκτίμηση της ακρίβειας βλέπουμε χαρακτηριστικά πως η Naïve παρουσιάζει πολύ καλά αποτελέσματα και στις δύο περιπτώσεις, όπως και η Damped για την πρώτη περίπτωση και η SES για την δεύτερη.

Για την συνολική αξιολόγηση των μεθόδων θα μελετήσουμε και τα ποσοστιαία σφάλματα τα οποία είναι τα εξής:

Platinum	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	0,8930%	0,8922%	0,999251	1,6359%	1,6321%	1,83436
Holt	0,8945%	0,8936%	1,000838	1,6449%	1,6405%	1,843879
Damped	<b>0,8924%</b>	<b>0,8918%</b>	<b>0,998823</b>	1,6501%	1,6471%	1,850963
ETS	0,8930%	0,8922%	0,999252	1,6359%	1,6321%	1,83436
Moving Average	1,1583%	1,1563%	1,301217	1,8165%	1,8109%	2,036936
LRL	17,0783%	15,7629%	17,7768	18,0139%	16,3998%	18,4591
Naïve	0,8930%	0,8922%	0,999245	<b>1,6359%</b>	<b>1,6321%</b>	<b>1,834349</b>
Theta	0,8931%	0,8922%	0,999173	1,6373%	1,6331%	1,835491

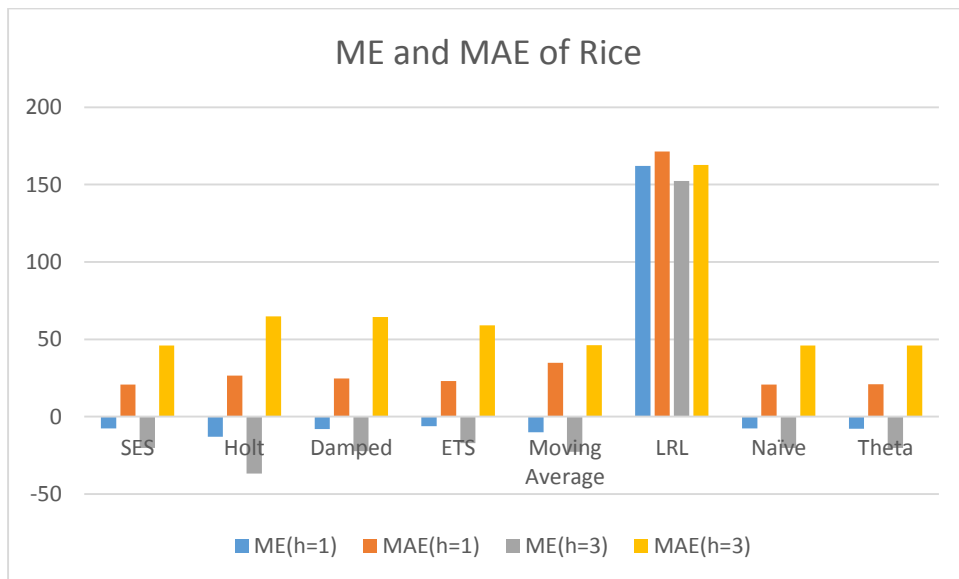
Στην περίπτωση των προβλέψεων μίας περιόδου καταλληλότερη μέθοδος φαίνεται η Damped και πολύ καλά αποτελέσματα δίνει η Naïve. Στη περίπτωση εξαγωγής προβλέψεων για την τρίτη περίοδο καλύτερη αξιολογήθηκε η Naïve και η Holt με πολύ καλά αποτελέσματα.

### 7.2.8 Ρύζι (Rice)

Αντιθέτως με την περίπτωση της πλατίνας, στη περίπτωση των τιμών του ρυζιού έχουμε διαθέσιμη μία μηνιαία χρονοσειρά. Επομένως ο πίνακας των σφαλμάτων αναφέρεται σε μηνιαίες προβλέψεις και είναι ο εξής:

Rice	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-7,49658	20,85556	-20,2171	46,02281
Holt	-12,938	26,68112	-36,7179	64,76675
Damped	-7,96924	24,73192	-22,2246	64,4078
ETS	<b>-6,15159</b>	23,04418	<b>-17,0678</b>	59,03405
Moving Average	-10,0897	34,76539	-22,7298	46,29012
LRL	162,0224	171,4103	152,2543	162,6104
Naïve	-7,49605	<b>20,85488</b>	-20,2166	<b>46,022</b>
Theta	-7,66078	20,90537	-20,2336	46,02907

Για την περίπτωση της προκατάληψης βλέπουμε πως όλες οι μέθοδοι έχουν το στοιχείο της αισιοδοξίας, λόγω των αρνητικών προσήμων των μεθόδων και για τις δύο περιπτώσεις. Στο παρακάτω διάγραμμα είναι συγκεντρωμένες οι τιμές των σφαλμάτων ME και MAE όλων των μεθόδων και για τις δύο περιπτώσεις.



Βλέπουμε πως η μέθοδος ETS και κατόπιν η Naïve έχουν την μικρότερη απόλυτη τιμή του σφάλματος ME. Στη περίπτωση της ακρίβειας των προβλέψεων πολύ καλά αποτελέσματα δίνουν η Naïve και κατόπιν η SES με την Theta.

Για την τελική αξιολόγηση των μεθόδων και της επιλογής της καταλληλότερης μεθόδου για την χρονοσειρά της τιμής του ρυζιού θα εξετάσουμε τα ποσοστιαία σφάλματα, όπως φαίνονται παρακάτω:

Rice	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	3,6602%	3,6036%	0,991661	8,3317%	7,9700%	2,424532
Holt	4,3147%	4,1649%	1,26866	10,8567%	9,5908%	3,411983
Damped	4,0042%	3,9108%	1,175978	10,9615%	10,1938%	3,393073
ETS	3,8232%	3,7593%	1,095728	10,2632%	9,7877%	3,109978
Moving Average	5,9719%	5,8677%	1,653059	8,5886%	8,1265%	2,438614
LRL	28,8791%	36,1189%	8,150387	28,4041%	35,1868%	8,566495
Naïve	<b>3,6601%</b>	<b>3,6035%</b>	<b>0,991629</b>	<b>8,3315%</b>	<b>7,9699%</b>	<b>2,424489</b>
Theta	3,6721%	3,6146%	0,994029	8,3332%	7,9713%	2,424862

Όπως φαίνεται από τον παραπάνω πίνακα πέρα πάσης αμφιβολίας η καταλληλότερη μέθοδος για την τιμή του ρυζιού αποδεικνύεται πως είναι η Naïve. Πολύ καλά αποτελέσματα δίνει επίσης η SES και η Theta.

## 7.2.9 Σιτάρι (Wheat)

Σε αντίθεση με τις χρονοσειρές της πλατίνας και του ρυζιού, το σιτάρι διαθέτει και ημερήσια και μηνιαία χρονοσειρά, οι οποίες αναλύονται ξεχωριστά παρακάτω:

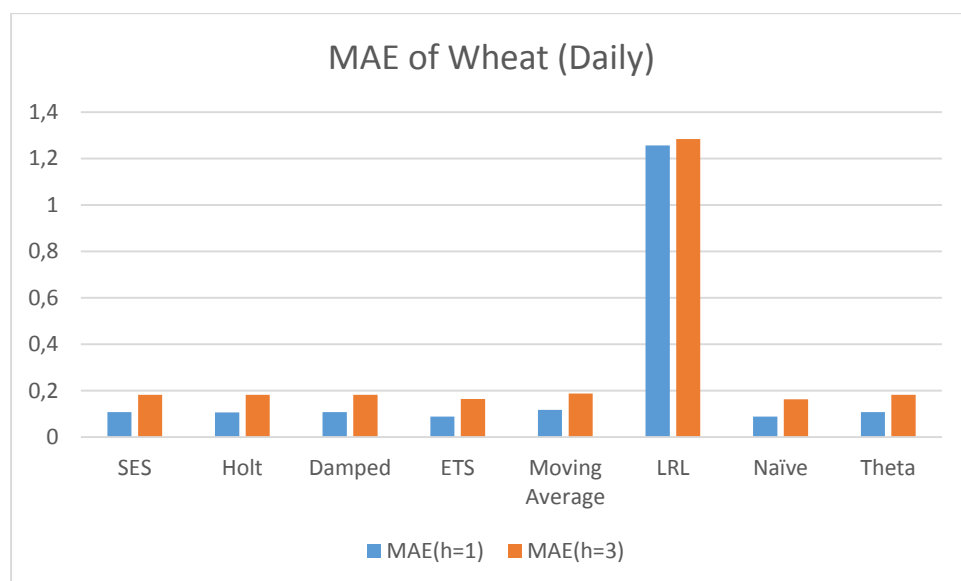
### 7.2.9.1 Ημερήσια Δεδομένα

Με βάση τον συγκεκριμένο χρονικό ορίζοντα της χρονοσειράς βγάλαμε τα εξής αποτελέσματα για το μέσο σφάλμα και για το μέσο απόλυτο σφάλμα των μεθόδων πρόβλεψης που εφαρμόσαμε και πιο συγκεκριμένα για προβλέψεις που αφορούν την επόμενη και την τρίτη περίοδο:

Wheat (Daily)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,00642	0,107097	-0,01999	0,181672
Holt	-0,00755	0,106761	-0,02213	0,18158
Damped	-0,00641	0,107037	-0,01996	0,181614
ETS	-0,00355	0,088798	<b>-0,01539</b>	0,164752
Moving Average	-0,00659	0,117224	-0,02031	0,188339
LRL	-1,23991	1,255664	-1,26888	1,284051
Naïve	<b>-0,00348</b>	<b>0,088192</b>	-0,01569	<b>0,163522</b>
Theta	-0,00714	0,107187	-0,02103	0,181782

Είναι εμφανές πως για ακόμη μία φορά οι μέθοδοι πρόβλεψης στο σύνολό τους έχουν έντονο το στοιχείο της αισιοδοξίας, λόγω του αρνητικού προσήμου των μέσων σφαλμάτων. Η μέθοδος με την μικρότερη κατά απόλυτη τιμή του μέσου σφάλματος είναι η Naive και ακολουθεί η ETS. Όσον αφορά την ακρίβεια των μεθόδων προβλέψεων και πιο συγκεκριμένα των μέσων απόλυτων σφαλμάτων κρίνεται σκόπιμη η παρουσίαση του επόμενου διαγράμματος. Αυτό που είναι γεγονός σχετικά με τα απόλυτα σφάλματα είναι πως οι διαφορές των σφαλμάτων των προβλέψεων κάθε μεθόδου, εκτός της LRL και του κινητού μέσου όρου, είναι ελάχιστες και

δύσκολα ανιχνεύσιμες από το διάγραμμα. Παρόλα αυτά η μέθοδος με το μικρότερο απόλυτο σφάλμα για την περίπτωση προβλέψεων της επόμενης περιόδου είναι η Naïve και με ελάχιστη διαφορά είναι η ETS. Στην περίπτωση των προβλέψεων της τρίτης περιόδου καταλληλότερη είναι η Naive ακολουθούμενη από την ETS.



Για την αξιολόγηση των μεθόδων θα πρέπει να ληφθούν υπόψη τα ποσοστιαία σφάλματα των μεθόδων σε σχέση με τις πραγματικές τιμές. Αυτά τα σφάλματα συγκεντρώνονται στον παρακάτω πίνακα:

Wheat (Daily)	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	1,9864%	1,9872%	1,212511	0,034121	0,034037	2,056342
Holt	1,9799%	1,9805%	1,208711	0,034098	0,034008	2,0553
Damped	1,9853%	1,9860%	1,211828	0,03411	0,034026	2,055689
ETS	1,6343%	1,6367%	1,005338	0,03089	0,030869	1,864828
Moving Average	2,1766%	2,1777%	1,327165	0,035444	0,035365	2,131809
LRL	25,1217%	21,1354%	14,21615	0,258118	0,217155	14,53418
Naïve	<b>1,6231%</b>	<b>1,6252%</b>	<b>0,998475</b>	<b>0,030642</b>	<b>0,030608</b>	<b>1,850902</b>
Theta	1,9881%	1,9886%	1,213536	0,034144	0,034054	2,057591

Όπως φαίνεται από τους παραπάνω πίνακες η καταλληλότερη μέθοδος πρόβλεψης αποδεικνύεται πως είναι η Naïve και ακολούθως η ETS και για την εξαγωγή προβλέψεων της επόμενης περιόδου, αλλά και για την τρίτη επόμενη όπως φαίνεται από τα αποτελέσματα.

### 7.2.9.2 Μηνιαία Δεδομένα

Με βάση τον συγκεκριμένο χρονικό ορίζοντα της χρονοσειράς βγάλαμε τα εξής αποτελέσματα για το μέσο σφάλμα και για το μέσο απόλυτο σφάλμα των μεθόδων πρόβλεψης που εφαρμόσαμε:

Wheat (Monthly)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-1,77189	15,12378	-4,90731	29,1395
Holt	-2,11068	15,16104	-5,92506	29,46359
Damped	-1,10006	<b>14,95987</b>	-3,32891	30,40916
ETS	<b>-0,81241</b>	14,97565	<b>-2,62589</b>	30,50428
Moving Average	-4,29411	21,96589	-7,36561	34,33439
LRL	61,84503	66,44159	60,09648	65,79859
Naïve	-1,77163	15,12326	-4,90706	<b>29,13929</b>
Theta	-1,89158	15,13636	-4,91922	29,14228

Αυτό που είναι ορατό είναι η αισιοδοξία των μεθόδων προβλέψεων και για τις δύο περιπτώσεις εξαγωγής προβλέψεων ανάλογα με την περίοδο πρόβλεψης. Αυτό γίνεται κατανοητό από την εξέταση του προσήμου του μέσου σφάλματος το οποίο είναι αρνητικό για όλες τις μεθόδους και για τις δύο περιπτώσεις. Γενικά, η μέθοδος ETS όπως επίσης και η μέθοδος Damped, είναι οι δύο μέθοδοι που παρουσιάζουν τις μικρότερες κατά απόλυτη τιμή, τιμές του μέσου σφάλματος, συνεπώς είναι και οι καλύτερες όσον αφορά την απόδοση των μεθόδων σχετικά με την προκατάληψη και για την παραγωγή προβλέψεων για την επόμενη και την τρίτη περίοδο.

Σχετικά με την ακρίβεια των προβλέψεων για κάθε μέθοδο θα πρέπει να ανατρέξουμε στις τιμές των μέσων απόλυτων σφαλμάτων κάθε μεθόδου. Στην περίπτωση παραγωγής προβλέψεων για την επόμενη περίοδο η μέθοδος Damped και η μέθοδος ETS υποδηλώνουν την μεγαλύτερη ακρίβεια με βάση τα αποτελέσματά τους. Στην περίπτωση πρόβλεψης για την τρίτη περίοδο η πιο αξιόπιστη μέθοδος για τα μηνιαία δεδομένα με βάση την ακρίβεια αποδείχτηκε η Naive ακολουθούμενη από την SES.

Τώρα για την επιλογή της καταλληλότερης μεθόδου πρόβλεψης θα πρέπει να ελέγξουμε τα ποσοστιαία σφάλματα δίνοντας παραπάνω έμφαση στο MAPE. Συγκεκριμένα έχουμε:

Wheat (Monthly)	MAPE (h=1)	sMAPE (h=1)	MAE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAE (h=3)
SES	5,5895%	5,5812%	1,014514	10,9400%	10,7820%	1,97489
Holt	5,6064%	5,5918%	1,017013	11,0694%	10,8753%	1,996855
Damped	<b>5,5152%</b>	<b>5,5308%</b>	<b>1,003519</b>	11,3869%	11,3464%	2,060939
ETS	5,5202%	5,5447%	1,004578	11,4185%	11,4191%	2,067386
Moving Average	8,1362%	8,0307%	1,473488	12,7962%	12,5078%	2,326967
LRL	22,1751%	25,9143%	4,45695	22,2211%	25,8314%	4,45941
Naïve	5,5894%	5,5810%	1,014479	<b>10,9399%</b>	<b>10,7819%</b>	<b>1,974876</b>
Theta	5,5967%	5,5860%	1,015358	10,9413%	10,7828%	1,975078

Παρατηρώντας το συνολικό ποσοστό των σφαλμάτων διαπιστώνουμε πως τα σφάλματα είναι μεγαλύτερα σε σχέση με τις προηγούμενες εξεταζόμενες χρονοσειρές για την περίπτωση της μηνιαίας χρονοσειράς της τιμής του σιταριού. Πιο

αναλυτικά η μέθοδος Damped είχε τις μικρότερες τιμές σφαλμάτων σχετικά με τις υπόλοιπες για την περίπτωση πρόβλεψης για την επόμενη περίοδο.

Στη περίπτωση της πρόβλεψης με ορίζοντα πρόβλεψης τρεις περιόδους η καλύτερη μέθοδος που προκύπτει με βάση τα σφάλματα είναι η Naïve και για αυτό θεωρείται καταλληλότερη.

## 7.2.10 Καλαμπόκι (Corn)

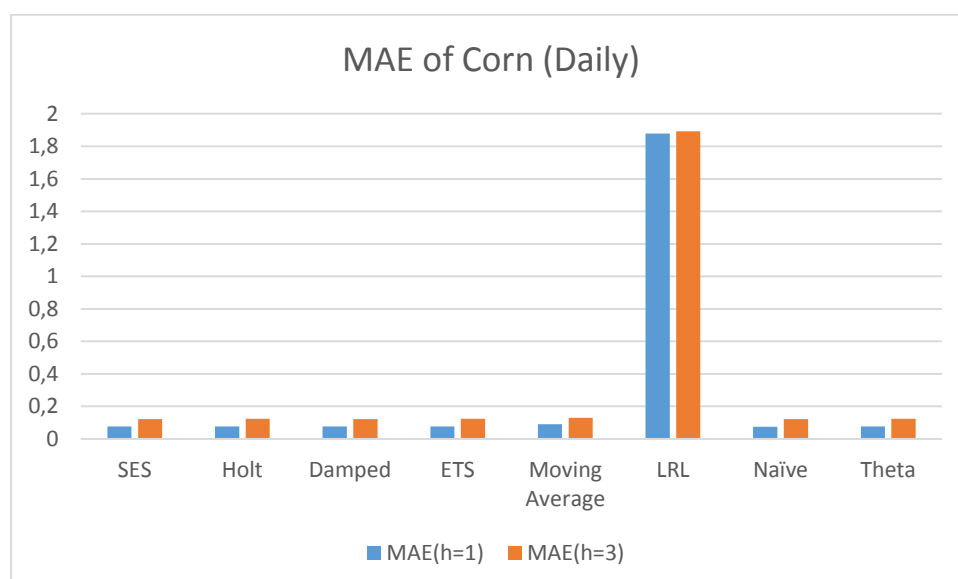
Για το συγκεκριμένο εμπόρευμα διαθέτουμε, επίσης, δύο χρονοσειρές μία ημερήσια και μία μηνιαία, οι οποίες αναλύονται ως εξής.

### 7.2.10.1 Ημερήσια Δεδομένα

Και σε αυτό το αγαθό ακολουθούμε την ίδια διαδικασία παρουσίασης των αποτελεσμάτων και αξιολόγησης μεθόδων. Αρχικά οι τιμές των σφαλμάτων σχετικά με την προκατάληψη και την ακρίβεια είναι οι εξής:

Corn (Daily)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,00533	0,075836	-0,01629	0,122793
Holt	-0,00659	0,075907	-0,0195	0,123112
Damped	-0,00517	0,075858	-0,01586	0,122803
ETS	-0,00571	0,076302	-0,01711	0,123416
Moving Average	-0,00921	0,090407	-0,02063	0,129965
LRL	-1,10419	1,879064	-1,13886	1,892644
Naïve	<b>-0,00489</b>	<b>0,073594</b>	<b>-0,0156</b>	<b>0,121248</b>
Theta	-0,00634	0,075912	-0,01813	0,122941

Και εδώ ισχύει η περίπτωση της αισιοδοξίας όσον αφορά τις μεθόδους πρόβλεψης. Η απόλυτη τιμή της μεθόδου που είναι μικρότερη και για τις δύο περιπτώσεις είναι η Naïve. Για την εξέταση της ακρίβειας των μεθόδων εξετάζουμε τον δείκτη MAE όπου οι τιμές βρίσκονται στο παρακάτω διάγραμμα:



Στην περίπτωση ακρίβειας των μεθόδων, η μέθοδος η οποία έχει την μικρότερη τιμή του δείκτη ακριβείας MAE είναι η μέθοδος Naïve και έπεται η Holt με την δεύτερη μικρότερη τιμή. Και εδώ οι τιμές των μέσων σφαλμάτων και των μέσων απόλυτων σφαλμάτων είναι σχετικά μικρές. Για να εξακριβώσουμε βέβαια πόσο αντιστοιχεί σε σχέση με τις πραγματικές τιμές χρειαζόμαστε τα ποσοστιαία σφάλματα, με τα οποία θα γίνει και η επιλογή της καταλληλότερης μεθόδου. Αρχικά παρατίθενται τα σφάλματα:

<b>Corn (Daily)</b>	MAPE (h=1)	sMAPE (h=1)	MAeE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAeE (h=3)
SES	1,4694%	1,4508%	1,03151	2,3992%	2,3697%	1,69309
Holt	1,4710%	1,4517%	1,032481	2,4056%	2,3744%	1,697488
Damped	1,4694%	1,4503%	1,031811	2,3988%	2,3691%	1,693225
ETS	1,4789%	1,4602%	1,037844	2,4103%	2,3804%	1,701674
Moving Average	1,7536%	1,7204%	1,229701	2,5481%	2,5042%	1,791981
LRL	43,9778%	34,9112%	25,5588	44,5220%	35,2699%	26,09606
Naïve	<b>1,4289%</b>	<b>1,4137%</b>	<b>1,001019</b>	<b>2,3714%</b>	<b>2,3452%</b>	<b>1,671791</b>
Theta	1,4711%	1,4522%	1,032552	2,4028%	2,3724%	1,695132

Με βάση τις παραπάνω τιμές των ποσοστιαίων σφαλμάτων είναι εμφανής η υπεροχή της Naïve σε σχέση με τις υπόλοιπες μεθόδους πρόβλεψης.

### 7.2.10.2 Μηνιαία δεδομένα

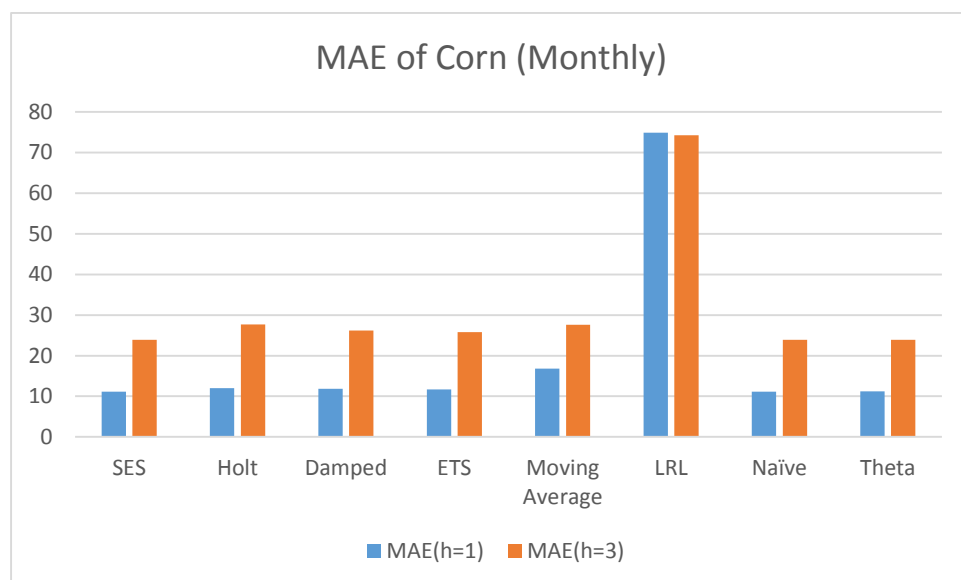
Με βάση τον συγκεκριμένο χρονικό ορίζοντα της χρονοσειράς βγάλαμε τα εξής αποτελέσματα για το μέσο σφάλμα και για το μέσο απόλυτο σφάλμα των μεθόδων πρόβλεψης που εφαρμόσαμε:

<b>Corn (Monthly)</b>	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,92973	11,16913	-3,11325	23,88206
Holt	-1,72504	11,97576	-5,51739	27,71666
Damped	-0,79613	11,84674	-2,82292	26,16225
ETS	<b>-0,69386</b>	11,70148	<b>-2,56924</b>	25,80384
Moving Average	-1,63895	16,84446	-3,77718	27,62706
LRL	69,22523	74,90939	68,33208	74,25669
Naïve	-0,92965	<b>11,16895</b>	-3,11318	<b>23,88188</b>
Theta	-1,01219	11,18785	-3,12133	23,88399

Σε αυτή την περίπτωση βλέπουμε, όσον αφορά το μέσο σφάλμα και για τις δύο περιπτώσεις, πως υπάρχουν τεράστιες διαφορές όσον αφορά την μέθοδο LRL σε σχέση με τις υπόλοιπες μεθόδους πρόβλεψης.

Είναι θετικό το φαινόμενο αρχικά, ότι δεν παρουσιάζονται ιδιαίτερα υψηλές τιμές του δείκτη σε κανένα από τους συνδυασμούς των μεθόδων. Γενικά στην παραγωγή προβλέψεων μίας χρονικής περιόδου αποδείχτηκε πως η ETS και ακολούθως η Damped έχουν την μικρότερη, κατά απόλυτη τιμή του σφάλματος. Ακριβώς το ίδιο ισχύει και για την παραγωγή προβλέψεων της τρίτης χρονικής περιόδου.

Τώρα όσον αφορά το Μέσο Απόλυτο Σφάλμα (MAE) είναι ένας δείκτης σφάλματος που είναι συνδεδεμένος με την ακρίβεια της μεθόδου. Από τις τιμές των σφαλμάτων έχουμε το αντίστοιχο διάγραμμα:



Παρατηρώντας πιο προσεκτικά το παραπάνω διάγραμμα βλέπουμε, πως στην πρώτη περίπτωση η μέθοδος με την μεγαλύτερη ακρίβεια είναι η Naive και ακολουθεί η SES. Στη περίπτωση παραγωγής προβλέψεων τριών περιόδων οι αντίστοιχοι μέθοδοι είναι οι ίδιες. Επίσης τα σφάλματα των προβλέψεων για την επόμενη περίοδο είναι μικρότερα σε σχέση με τα υπόλοιπα, κάτι το οποίο είναι λογικό σε σχέση με την ακρίβεια των μεθόδων όταν μεγαλώνει η περίοδος πρόβλεψης.

Στον παρακάτω πίνακα βρίσκονται τα ποσοστιαία σφάλματα των μεθόδων πρόβλεψης του καλαμποκιού:

Corn (Monthly)	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	4,9683%	4,9300%	0,991702	10,9678%	10,5947%	2,213302
Holt	5,2761%	5,2333%	1,063322	12,7848%	12,4140%	2,568679
Damped	5,2196%	5,1933%	1,051866	12,1430%	11,8390%	2,424622
ETS	5,1585%	5,1350%	1,038969	11,9461%	11,6381%	2,391406
Moving Average	7,6460%	7,4894%	1,495612	12,6673%	12,1481%	2,560375
LRL	29,6338%	36,6681%	6,65117	29,4725%	36,4733%	6,88184
Naive	<b>4,9682%</b>	<b>4,9299%</b>	<b>0,991686</b>	<b>10,9677%</b>	<b>10,5946%</b>	<b>2,213286</b>
Theta	4,9773%	4,9373%	0,993364	10,9688%	10,5953%	2,213481

Όπως φαίνεται από τον παραπάνω πίνακα οι τιμές των σφαλμάτων για την μέθοδο της LRL και του κινητού μέσου όρου είναι μεγαλύτερα σε σχέση με τις άλλες μεθόδους. Όπως είναι λογικό τα σφάλματα για την παραγωγή προβλέψεων για την τρίτη περίοδο είναι μεγαλύτερα.



Στη πρώτη περίπτωση καλύτερα αποτελέσματα δίνει η μέθοδος Naive και ακολουθεί η SES και για τα τρία ήδη σφάλματων. Στην δεύτερη περίπτωση με τον αυξημένο ορίζοντα πρόβλεψης, δεν πρέπει να αμεληθεί το γεγονός πως και εδώ η ακρίβεια της μεθόδου Theta είναι καλύτερη σε σχέση με πριν κατέχοντας την δεύτερη θέση, ενώ κυρίαρχη μέθοδος αποδείχτηκε η Naive σύμφωνα με αυτά τα σφάλματα και τον συγκεκριμένο ορίζοντα πρόβλεψης.

### 7.2.11 Σόγια (Soybean)

Η σόγια, όπως και τα προηγούμενα καλλιεργήσιμα αγαθά, διαθέτει δύο χρονοσειρές οι οποίες χρησιμοποιήθηκαν για την πραγματοποίηση προβλέψεων. Οι χρονοσειρές είναι μία ημερήσια και μία μηνιαία.

#### 7.2.11.1 Ημερήσια δεδομένα

Κατά τα γνωστά ακολουθείται και εδώ η ίδια διαδικασία με τις προηγούμενες περιπτώσεις. Αρχικά θα παραθέσουμε τα σφάλματα που αφορούν την προκατάληψη και την ακρίβεια των προβλέψεων των μεθόδων πρόβλεψης για την επόμενη και για την τρίτη περίοδο:

<b>Soybean (Daily)</b>	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,00985	0,173718	-0,02606	0,29312
Holt	-0,01308	0,173805	-0,03429	0,293229
Damped	-0,00913	0,174035	-0,0244	0,292884
ETS	-0,00962	0,173735	-0,02578	0,293398
Moving Average	-0,01487	0,21298	-0,03135	0,313242
LRL	-1,2179	2,351471	-1,29678	2,39721
Naïve	<b>-0,00768</b>	<b>0,16752</b>	<b>-0,02372</b>	<b>0,289857</b>
Theta	-0,01204	0,173729	-0,02998	0,293061

Και εδώ όσον αφορά την προκατάληψη των μεθόδων ισχύουν ότι ακριβώς και στη περίπτωση του καλαμποκιού, λόγω του αρνητικού προσήμου του ME για όλες τις μεθόδους.

Τώρα όσον αφορά την αξιολόγηση της ακρίβειας ισχύουν τα ίδια με την περίπτωση του καλαμποκιού, δηλαδή πιο αναλυτικά φαίνεται πως την καλύτερη ακρίβεια την έχει η μέθοδος Naive και για τις προβλέψεις του επόμενου χρονικού ορίζοντα, αλλά και για την τρίτη περίοδο. Επίσης πολύ καλά αποτελέσματα δίνουν οι μέθοδοι Damped και SES, ενώ αρκετά αξιόπιστη διακρίνεται η Theta.

Στον παρακάτω πίνακα παρουσιάζονται τα ποσοστιαία σφάλματα των προβλέψεων των μεθόδων πρόβλεψης, οι οποίοι δείκτες μας παρέχουν χρήσιμες πληροφορίες σχετικά με την επιλογή της καταλληλότερης επιλογής μεθόδου. Συγκεκριμένα έχουμε:

<b>Soybean (Daily)</b>	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	1,3485%	1,3313%	1,035868	2,2814%	2,2455%	1,763954
Holt	1,3495%	1,3320%	1,036391	2,2839%	2,2465%	1,764616
Damped	1,3505%	1,3333%	1,037763	2,2798%	2,2447%	1,762538
ETS	1,3482%	1,3312%	1,035974	2,2832%	2,2474%	1,765631
Moving Average	1,6495%	1,6231%	1,269985	2,4415%	2,3954%	1,88505
LRL	20,7957%	18,0295%	14,02168	21,3159%	18,4279%	14,42609
Naïve	<b>1,2991%</b>	<b>1,2880%</b>	<b>0,998911</b>	<b>2,2508%</b>	<b>2,2217%</b>	<b>1,74432</b>
Theta	1,3488%	1,3313%	1,035935	2,2814%	2,2447%	1,763601

Στην περίπτωση εξαγωγής προβλέψεων για την επόμενη περίοδο προκύπτει πως η καταλληλότερη μέθοδος πρόβλεψης, δηλαδή με τα μικρότερα ποσοστιαία σφάλματα είναι η Naïve και ακολουθούν οι SES και ETS. Τώρα στην δεύτερη περίπτωση ισχύει επίσης πως η μέθοδος με τα μικρότερα σφάλματα είναι η Naïve, αλλά εδώ υπάρχει μία διαφοροποίηση. Αυτή έγκειται στο γεγονός πως η Theta παρουσίασε πολύ μικρές τιμές σφαλμάτων σε σχέση με τις υπόλοιπες μεθόδους, πλην της Naïve.

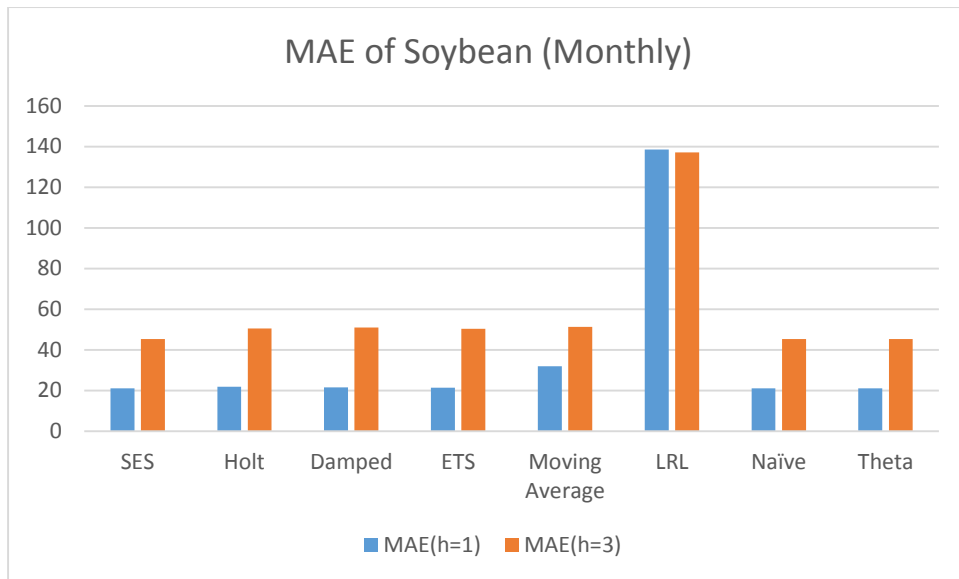
### 7.2.11.2 Μηνιαία Δεδομένα

Και εδώ πρόκειται για μία χρονοσειρά, για την οποία ακολουθήσαμε την ίδια διαδικασία με το αργό πετρέλαιο. Δηλαδή υπολογίσαμε τα σφάλματα προκατάληψης (ME), ακρίβειας (MAE) και ποσοστιαία (MAPE, sMAPE) για παραγωγή προβλέψεων για την επόμενη, αλλά και για την τρίτη περίοδο. Τα σφάλματα προκατάληψης και ακρίβειας φαίνονται συγκεντρωμένα παρακάτω:

<b>Soybean (Monthly)</b>	ME(h=1)	MAE(h=1)	ME(h=1)	MAE(h=3)
SES	-1,48819	20,97314	-5,23816	45,31287
Holt	-1,99841	21,89192	-6,76776	50,51072
Damped	-0,6984	21,49474	-3,3561	50,97647
ETS	<b>-0,6336</b>	21,28334	<b>-3,19351</b>	50,38236
Moving Average	-3,18868	31,98713	-6,90914	51,22631
LRL	133,3725	138,5556	131,9031	137,1656
Naïve	-1,48802	<b>20,97244</b>	-5,238	45,31282
Theta	-1,64379	20,98824	-5,25341	<b>45,31277</b>

Και σε αυτή την περίπτωση η μέθοδος LRL παρουσιάζει τα μεγαλύτερα σφάλματα είτε πρόκειται για το μέσο σφάλμα, είτε για το μέσο απόλυτο σφάλμα. Στην περίπτωση της προκατάληψης φαίνεται πως οι μέθοδοι πρόβλεψης παρουσιάζουν αισιοδοξία, λόγω των αρνητικών τιμών του μέσου σφάλματος. Η μέθοδος με την μεγαλύτερη ακρίβεια παρουσιάζεται η ETS και για την παραγωγή προβλέψεων της επόμενης, αλλά και της τρίτης περιόδου. Στη συνέχεια ακολουθεί η Damped.

Για την ακρίβεια θα είναι χρήσιμο να δώσουμε βάση στο επόμενο διάγραμμα:



Και εδώ βλέπουμε πως η LRL έχει μεγάλη διαφορά σε σχέση με τις υπόλοιπες μεθόδους. Όπως και πριν οι ακρίβεια όλων των μεθόδων με εξαίρεση την LRL είναι αρκετά καλή. Παρατηρώντας προσεκτικά το διάγραμμα βλέπουμε πως η μέθοδος με την καλύτερη ακρίβεια είναι η Naive και ακολουθεί η SES για την εξαγωγή προβλέψεων της επόμενης περιόδου. Στην άλλη περίπτωση παρόλο που έχει μειωθεί γενικά η ακρίβεια πρόβλεψης, όπως είναι λογικό, σαν καταλληλότερη μέθοδος τώρα παρουσιάζεται η Theta και ακολουθούν η Naive και η SES.

Τώρα όσον αφορά τα ποσοστιαία σφάλματα που εξαγάγαμε, παρουσιάζονται συνοπτικά στον ακόλουθο πίνακα που συγκεντρώνει και τις δύο περιπτώσεις παραγωγής προβλέψεων:

Soybean (Monthly)	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	4,7801%	4,7113%	0,99187	10,4691%	10,0147%	2,175267
Holt	5,0271%	4,9617%	1,035321	11,9097%	11,5022%	2,424792
Damped	4,8937%	4,8494%	1,016537	11,8620%	11,5023%	2,44715
ETS	4,8381%	4,7947%	1,00654	11,7051%	11,3424%	2,418629
Moving Average	7,2873%	7,0933%	1,512747	11,9059%	11,3509%	2,459144
LRL	29,2490%	35,3434%	6,552622	29,0942%	35,1391%	6,5847
Naive	<b>4,7799%</b>	<b>4,7111%</b>	<b>0,991786</b>	<b>10,4691%</b>	<b>10,0147%</b>	<b>2,175264</b>
Theta	4,7848%	4,7146%	0,992584	10,4693%	<b>10,0146%</b>	<b>2,175262</b>

Και εδώ βλέπουμε πως η LRL έχει τεράστια διαφορά σε σχέση με τις υπόλοιπες μεθόδους. Στη περίπτωση του σφάλματος MAPE και sMAPE και για παραγωγή προβλέψεων της επόμενης περιόδου, η μέθοδος Naive έχει την μικρότερη τιμή και ακολουθεί η SES. Στη περίπτωση του κανονικοποιημένου σφάλματος MAsE η Naive παρουσιάζει μικρή τιμή και ακολουθεί με πολύ μικρή διαφορά η μέθοδος SES.

Τώρα στην περίπτωση παραγωγής προβλέψεων για την τρίτη περίοδο, ο υπολογισμός των αντίστοιχων σφαλμάτων δίνει περίπου τα ίδια αποτελέσματα όσον

αφορά τη μέθοδο με την μικρότερη τιμή. Για τα σφάλματα MAPE την μικρότερη τιμή έχει η μέθοδος Naive και αμέσως μετά ακολουθεί η SES. Όσον αφορά τα σφάλματα sMAPE το κανονικοποιημένο σφάλμα η Theta έχει βελτιωθεί σε σχέση με την πρώτη περίπτωση και ακολουθεί η Naive.

### 7.2.12 Κακάο (Cocoa)

Υπάρχουν διαθέσιμες μία ημερήσια και μία μηνιαία χρονοσειρά για το συγκεκριμένο εμπόρευμα. Αυτές παρουσιάζονται αναλυτικά παρακάτω.

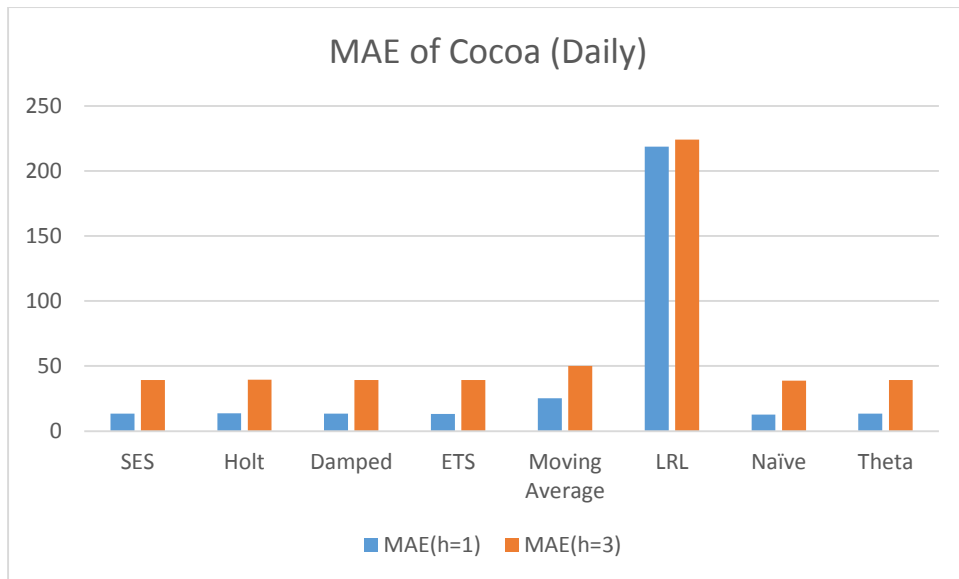
#### 7.2.12.1 Ημερήσια δεδομένα

Η ίδια διαδικασία με προηγουμένως ακολουθείται και εδώ. Άρα με βάση αυτή τη λογική ακολουθούν τα σφάλματα προκατάληψης και ακρίβειας των μεθόδων πρόβλεψης για τους δύο χρονικούς ορίζοντες πρόβλεψης:

Cocoa (Daily)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	1,912268	13,31658	3,735436	39,25483
Holt	<b>1,18301</b>	13,66862	<b>1,601375</b>	39,48728
Damped	1,911618	13,31655	3,733468	39,25121
ETS	1,900961	13,24289	3,727209	39,18954
Moving Average	3,601272	25,2525	4,932292	50,08333
LRL	191,5311	218,7365	198,7291	224,242
Naive	1,831063	<b>12,75204</b>	3,679688	<b>38,74219</b>
Theta	1,827747	13,32701	3,585944	39,25328

Στην περίπτωση της προκατάληψης φαίνεται πως σε σχέση με τη Σόγια οι μέθοδοι πρόβλεψης παρουσιάζουν απαισιοδοξία, λόγω των θετικών τιμών του μέσου σφάλματος. Η μέθοδος με την μεγαλύτερη ακρίβεια παρουσιάζεται η Holt και για την παραγωγή προβλέψεων της επόμενης περιόδου και για την τρίτη και έπεται η μέθοδος Theta.

Για τον υπολογισμό, την διερεύνηση και την αξιολόγηση της ακρίβειας είναι απαραίτητος ο έλεγχος των δεικτών MAE για κάθε μέθοδο. Στο διάγραμμα που ακολουθεί με μία πιο προσεκτική ματιά φαίνονται οι διαφορές των μεθόδων. Βλέπουμε λοιπόν πως η μέθοδος με την καλύτερη ακρίβεια και συνεπώς με το μικρότερο μέσο απόλυτο σφάλμα είναι η Naive, και για την περίπτωση πρόβλεψης για την επόμενη περίοδο, αλλά και για την τρίτη. Κατόπιν η δεύτερη μέθοδος με την καλύτερη ακρίβεια αποδεικνύεται η ETS και για τις δύο περιπτώσεις.



Σύμφωνα με την διαδικασία που έχει ήδη ακολουθηθεί, θα πρέπει να γίνει μία περαιτέρω διερεύνηση σχετικά με τις μεθόδους πρόβλεψης και την αξιολόγησή του. Αυτό αφορά στην σύγκριση των ποσοστιαίων σφαλμάτων ανά μέθοδο πρόβλεψης κάτι το οποίο οδηγεί με την σειρά του στην επιλογή της κατάλληλης μεθόδου για τα ημερήσια δεδομένα της τιμής του κακάο. Άρα τα σφάλματα αυτά είναι:

Cocoa (Daily)	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	0,4028%	0,4031%	1,041425	1,1832%	1,1829%	3,000966
Holt	0,4135%	0,4137%	1,068956	1,1906%	1,1896%	3,018737
Damped	0,4028%	0,4031%	1,041422	1,1831%	1,1828%	3,00069
ETS	0,4006%	0,4009%	1,035662	1,1813%	1,1809%	2,995975
Moving Average	0,7643%	0,7647%	1,974875	1,5096%	1,5093%	3,828788
LRL	6,4088%	6,6984%	17,10631	6,5522%	6,8508%	17,14293
Naïve	<b>0,3857%</b>	<b>0,3860%</b>	<b>0,997275</b>	<b>1,1678%</b>	<b>1,1675%</b>	<b>2,961776</b>
Theta	0,4032%	0,4034%	1,04224	1,1832%	1,1828%	3,000848

Είναι απόλυτα εμφανές λαμβάνοντας υπόψιν ακόμα και τα τρία ήδη σφαλμάτων, πως η καταλληλότερη μέθοδος πρόβλεψης είναι η Naïve και ακολουθεί η ETS.

### 7.2.12.2 Μηνιαία δεδομένα

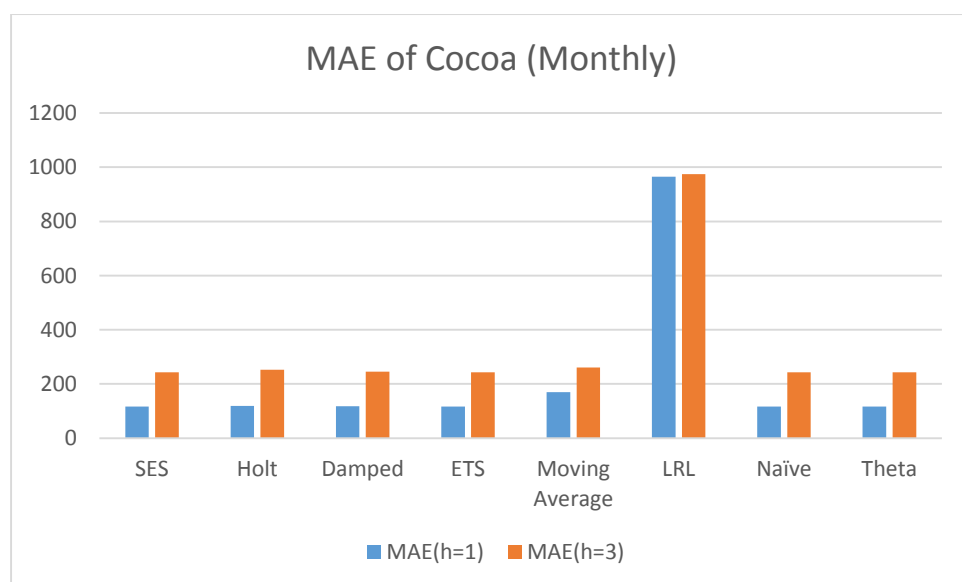
Όπως και για τα ημερήσια, έτσι και εδώ παραθέτουμε τα σφάλματα ME και MAE, όπως φαίνονται στο παρακάτω σχήμα:

Cocoa (Monthly)	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	7,077847	117,0086	15,47602	243,2548
Holt	<b>-2,00815</b>	119,2312	<b>-11,8575</b>	252,2031
Damped	5,826815	117,6152	11,97242	245,7132
ETS	7,054451	117,032	15,45235	<b>243,2421</b>

Moving Average	11,8919	169,776	18,59914	260,3208
LRL	965,3244	965,3244	975,0966	975,0966
Naïve	7,077326	<b>117,005</b>	15,47576	243,2544
Theta	6,914385	117,0152	15,46231	243,2499

Και σε αυτή την περίπτωση η μέθοδος LRL παρουσιάζει τα μεγαλύτερα σφάλματα είτε πρόκειται για το μέσο σφάλμα, είτε για το μέσο απόλυτο σφάλμα. Στην περίπτωση της προκατάληψης φαίνεται πως οι μέθοδοι πρόβλεψης παρουσιάζουν απαισιοδοξία, λόγω των θετικών τιμών του μέσου σφάλματος. Η μόνη μέθοδος που διαφοροποιείται σχετικά είναι η Holt. Η μέθοδος με την μεγαλύτερη ακρίβεια παρουσιάζεται η Holt και για την παραγωγή προβλέψεων της επόμενης, αλλά και της τρίτης περιόδου. Στη συνέχεια ακολουθεί η Damped και για τις δύο περιπτώσεις.

Με βάση το παρακάτω διάγραμμα μπορεί να γίνει μία αξιολόγηση της ακρίβειας των μεθόδων:



Με βάση τον πίνακα και το παραπάνω διάγραμμα προκύπτει πως στην περίπτωση εξαγωγής προβλέψεων για την επόμενη περίοδο, η μέθοδος με την μεγαλύτερη ακρίβεια είναι η Naïve και ακολουθεί η SES. Στην περίπτωση εξαγωγής προβλέψεων για την τρίτη περίοδο, η μέθοδος ETS θεωρείται η μέθοδος με την μεγαλύτερη ακρίβεια, και ακολουθεί η Theta η οποία βελτίωσε αρκετά την θέση της σε σχέση με την πρώτη περίπτωση.

Τώρα για να μπορέσουν να αξιολογηθούν περαιτέρω οι μέθοδοι πρόβλεψης με σκοπό την επιλογή της κατάλληλης μεθόδου παρουσιάζονται παρακάτω τα ποσοστιαία σφάλματα:

Cocoa (Monthly)	MAPE (h=1)	sMAPE (h=1)	MAE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAE (h=3)
SES	4,2893%	4,2814%	0,994548	8,8730%	8,7957%	2,108735
Holt	4,3676%	4,3466%	1,01344	9,2415%	9,0633%	2,186306

Damped	4,3122%	4,3029%	0,999704	8,9697%	8,8768%	2,130046
ETS	4,2904%	4,2824%	0,994747	<b>8,8725%</b>	<b>8,7951%</b>	<b>2,108624</b>
Moving Average	6,1931%	6,1465%	1,44306	9,4741%	9,3685%	2,256676
LRL	33,1047%	41,2415%	8,205052	33,3713%	41,6161%	8,452946
Naïve	<b>4,2891%</b>	<b>4,2812%</b>	<b>0,994518</b>	8,8730%	8,7957%	2,10873
Theta	4,2894%	4,2815%	0,994604	<b>8,8728%</b>	<b>8,7955%</b>	<b>2,108692</b>

Για την πρώτη περίπτωση η καταλληλότερη μέθοδος πρόβλεψης είναι η Naïve και ακολουθεί η SES για όλα τα είδη σφαλμάτων. Επομένως η Naïve αποδείχτηκε η καταλληλότερη μέθοδος πρόβλεψης για παραγωγή προβλέψεων της επόμενης περιόδου. Αντίστοιχα, για την παραγωγή προβλέψεων της τρίτης περιόδου η ETS αξιολογήθηκε ως η κατάλληλη μέθοδος πρόβλεψης και έπεται η Theta.

### 7.2.13 Καφές (Coffee)

Στην περίπτωση του καφέ έχουμε δύο είδη καφέ τον Arabica και τον Robusta, όπου υπάρχουν διαθέσιμες δύο μηνιαίες χρονοσειρές, ανάλογα με την ποικιλία καφέ. Επομένως θα μελετηθούν ξεχωριστά παρακάτω.

#### 7.2.13.1 Ποικιλία Arabica

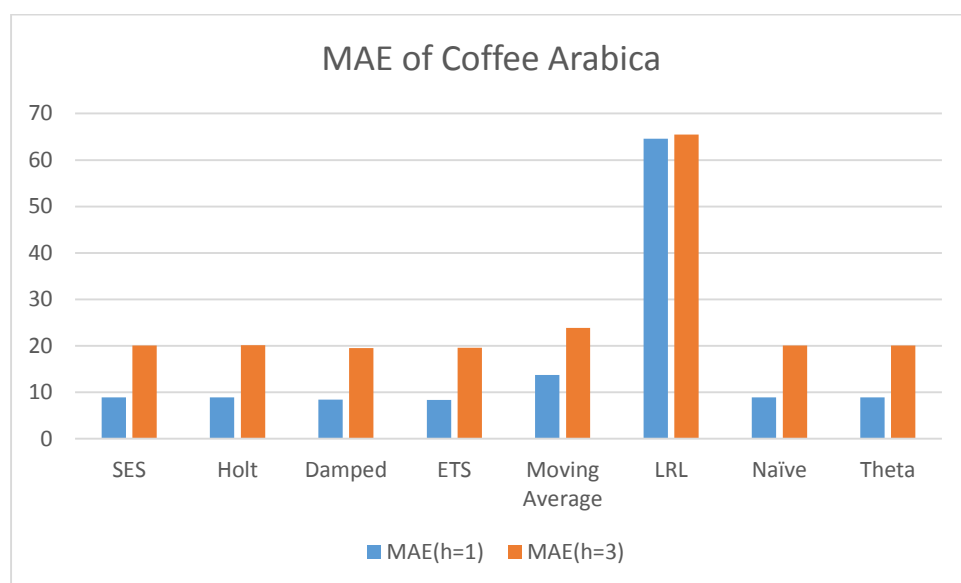
Στην περίπτωση αυτής της ποικιλίας πραγματοποιήθηκαν προβλέψεις με την τεχνική της κυλιόμενης πρόβλεψης με εξαγωγή προβλέψεων για την επόμενη περίοδο και για την τρίτη περίοδο. Αρχικά, όπως και στα προηγούμενα εμπορεύματα, η παρουσίαση και η ανάλυση των εμπορευμάτων θα γίνει με τον ίδιο τρόπο με πριν. Ως εκ τούτου, θα πρέπει πρωτίστως να μελετηθεί η προκατάληψη των μεθόδων, εξετάζοντας τον δείκτη ME και δευτερευόντως η ακρίβεια των μεθόδων μελετώντας τον δείκτη MAE. Ο πίνακας των σφαλμάτων και για τις δύο περιπτώσεις προβλέψεων είναι ο εξής:

Coffee Arabica	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	0,256175	8,883625	0,584373	20,05571
Holt	0,236655	8,870556	0,525497	20,1501
Damped	0,181897	8,38266	0,313134	<b>19,49748</b>
ETS	<b>0,137201</b>	<b>8,359129</b>	<b>0,192856</b>	<b>19,54245</b>
Moving Average	0,360659	13,73694	0,723922	23,84816
LRL	63,64847	64,5585	64,56312	65,47735
Naïve	0,256163	8,883372	0,584353	20,05541
Theta	0,266258	8,894988	0,585422	20,05783

Από τα παραπάνω αποτελέσματα φαίνεται με την πρώτη ματιά, πως οι μέθοδοι παρουσιάζουν το στοιχείο της απαισιοδοξίας στις προβλέψεις τους. Αυτό συμπεραίνεται και εδώ, από το θετικό πρόσημο του σφάλματος ME των προβλέψεων όλων των μεθόδων. Και εδώ η LRL παρουσιάζει το μεγαλύτερο σφάλμα κατά απόλυτη τιμή, και ακολούθως η μέθοδος κινητού μέσου όρου. Στον αντίποδα η μέθοδος ETS

παρουσιάζει τα μικρότερα σφάλματα κατά απόλυτη τιμή, όπου υπάρχει μία σχετική διαφορά σε σχέση με τις άλλες μεθόδους. Η επόμενη μέθοδος με τις αμέσως μικρότερες τιμές είναι η μέθοδος Damped.

Εξετάζοντας το παρακάτω διάγραμμα για τα σφάλματα ακρίβειας καταλήγουμε στα εξής συμπεράσματα. Πρώτον είναι αμέσως εμφανές πως την μικρότερη ακρίβεια έχει η μέθοδος LRL και ακολούθως η μέθοδος κινητού μέσου όρου. Στη συνέχεια η μέθοδος ETS διαπιστώνεται πως είναι η μέθοδος με την καλύτερη ακρίβεια όσον αφορά την περίπτωση εξαγωγής προβλέψεων για την επόμενη περίοδο. Μετά από άποψη ακρίβειας, πολύ καλά αποτελέσματα έδωσε η Damped, η οποία παρουσίασε την μεγαλύτερη ακρίβεια στην περίπτωση παραγωγής προβλέψεων για την τρίτη περίοδο και ακολούθησε η ETS.



Τώρα για την αξιολόγηση και την επιλογή της κατάλληλης μεθόδου πρόβλεψης για τα μηνιαία δεδομένα της τιμής του καφέ Arabica χρησιμοποιούμε τα ποσοστιαία σφάλματα, τα οποία είναι συγκεντρωμένα στον παρακάτω πίνακα:

Coffee Arabica	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	4,7478%	4,7825%	0,990995	10,5405%	10,7170%	2,208432
Holt	4,7406%	4,7755%	0,989537	10,5897%	10,7679%	2,218825
Damped	<b>4,5141%</b>	<b>4,5482%</b>	<i>0,93511</i>	<b>10,2168%</b>	<b>10,3825%</b>	<b>2,146962</b>
ETS	4,5199%	4,5524%	<b>0,932486</b>	10,2933%	10,4580%	2,15191
Moving Average	7,2581%	7,3395%	1,532396	12,4811%	12,6749%	2,626036
LRL	31,8813%	40,2066%	7,201691	32,0789%	40,6859%	7,210027
Naïve	4,7476%	4,7823%	0,990966	10,5403%	10,7168%	2,208398
Theta	4,7540%	4,7889%	0,992262	10,5417%	10,7183%	2,208665

Το πλέον σίγουρο είναι πως αν ληφθούν υπόψιν και τα τρία είδη σφαλμάτων, τότε η καταλληλότερη μέθοδος πρόβλεψης για τα συγκεκριμένα δεδομένα είναι η Damped και ακολουθεί η ETS.



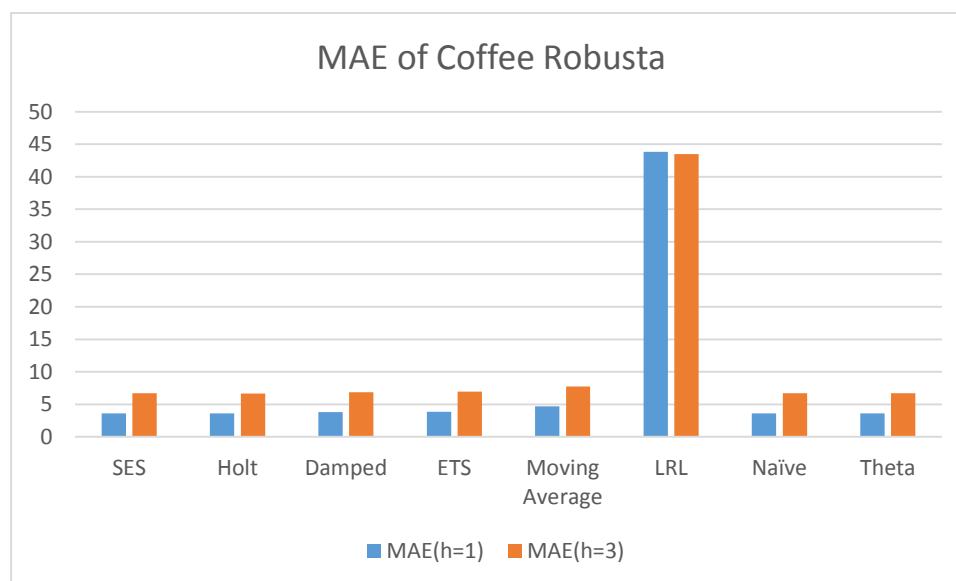
### 7.2.13.2 Ποικιλία Robusta

Με το ίδιο σκεπτικό με πριν ακολουθούμε την ίδια ροή για την παρουσίαση των αποτελεσμάτων. Επομένως τα ανάλογα σφάλματα είναι τα εξής:

Coffee Robusta	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,1741	3,614017	-0,5852	6,683084
Holt	<b>0,03135</b>	<b>3,605042</b>	<b>0,030573</b>	<b>6,643705</b>
Damped	-0,10952	3,779867	-0,45858	6,834825
ETS	-0,07236	3,841018	-0,37072	6,972764
Moving Average	-0,46349	4,662326	-0,84624	7,757686
LRL	43,83029	43,83029	43,49314	43,49314
Naïve	-0,17407	3,61407	-0,58518	6,683059
Theta	-0,09703	3,607914	-0,57756	6,682269

Σε αυτή την περίπτωση η μέθοδος LRL παρουσιάζει τα μεγαλύτερα σφάλματα είτε πρόκειται για το μέσο σφάλμα, είτε για το μέσο απόλυτο σφάλμα. Στην περίπτωση της προκατάληψης φαίνεται πως οι μέθοδοι πρόβλεψης παρουσιάζουν αισιοδοξία, λόγω των αρνητικών τιμών του μέσου σφάλματος. Οι μόνες μέθοδοι που διαφοροποιούνται σχετικά είναι η Holt και η LRL. Γενικά, η μέθοδος Holt όπως επίσης και η μέθοδος ETS, είναι οι δύο μέθοδοι που παρουσιάζουν τις μικρότερες κατά απόλυτη τιμή, τιμές του μέσου σφάλματος, συνεπώς είναι και οι καλύτερες όσον αφορά την απόδοση των μεθόδων σχετικά με την προκατάληψη.

Για την μελέτη της ακρίβειας και συνεπώς του δείκτη MAE έχουμε το ακόλουθο διάγραμμα:



Αν γίνει μία σύγκριση με το ύψος των σφαλμάτων του καφέ Arabica, παρατηρείται πως το ύψος σφαλμάτων της ποικιλίας Robusta είναι μικρότερο συγκριτικά με τις τιμές των σφαλμάτων της ποικιλίας Arabica, όσον αφορά το MAE.

Οι μέθοδοι που παρουσιάζουν την μεγαλύτερη ακρίβεια είναι η μέθοδος Holt και για τις δύο περιπτώσεις προβλέψεων. Στη συνέχεια, η Theta παρουσιάζει επίσης πολύ καλά αποτελέσματα.

Τώρα κατά τα γνωστά παίρνουμε τα ποσοστιαία σφάλματα για την περαιτέρω αξιολόγηση των μεθόδων πρόβλεψης:

<b>Coffee Robusta</b>	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	3,6489%	3,6365%	0,994984	6,7872%	6,7383%	1,824555
Holt	<b>3,6356%</b>	<b>3,6311%</b>	<b>0,992513</b>	<b>6,7216%</b>	<b>6,7168%</b>	<b>1,813804</b>
Damped	3,8053%	3,8010%	1,040645	6,8725%	6,8198%	1,865982
ETS	3,8683%	3,8676%	1,057481	7,0241%	6,9770%	1,903641
Moving Average	4,7239%	4,6967%	1,283597	7,9130%	7,8625%	2,117933
LRL	43,3721%	56,1817%	12,06703	43,1851%	55,8615%	11,8741
Naïve	3,6490%	3,6365%	0,994999	6,7872%	6,7383%	1,824548
Theta	<b>3,6410%</b>	<b>3,6315%</b>	<b>0,993304</b>	<b>6,7861%</b>	<b>6,7378%</b>	<b>1,824333</b>

Άρα στην περίπτωση του καφέ ποικιλίας Robusta η καταλληλότερη μέθοδος επιλέγεται η μέθοδος Holt και ακολουθεί η Theta με μικρή διαφορά. Και εδώ τα χειρότερα αποτελέσματα τα είχε η LRL και ακολούθως η μέθοδος κινητού μέσου όρου.

## **7.2.14 Βαμβάκι (Cotton)**

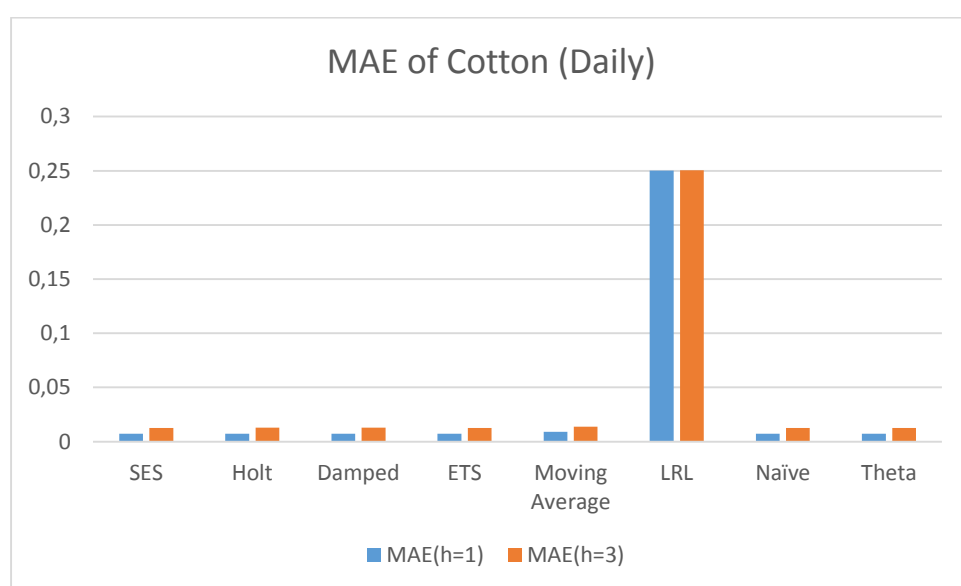
Στην περίπτωση του βαμβακιού, όπως και στις περιπτώσεις των προηγούμενων καλλιεργήσιμων αγαθών, διατίθενται μία ημερήσια και μία μηνιαία χρονοσειρά.

### **7.2.14.1 Ημερήσια Δεδομένα**

Με βάση τον συγκεκριμένο χρονικό ορίζοντα της χρονοσειράς βγάλαμε τα εξής αποτελέσματα για το μέσο σφάλμα και για το μέσο απόλυτο σφάλμα των μεθόδων πρόβλεψης που εφαρμόσαμε και πιο συγκεκριμένα για προβλέψεις που αφορούν την επόμενη και την τρίτη περίοδο:

<b>Cotton (Daily)</b>	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,00028	0,007334	-0,00075	0,012757
Holt	<b>0,000112</b>	0,007412	<b>0,000405</b>	0,012893
Damped	-0,00023	0,00736	-0,00059	0,012885
ETS	-0,00028	0,007334	-0,00075	0,012757
Moving Average	-0,00056	0,009216	-0,00107	0,013789
LRL	-0,25011	0,250106	-0,25057	0,250567
Naïve	-0,00028	<b>0,007334</b>	-0,00075	<b>0,012757</b>
Theta	-0,00038	0,007338	-0,00094	0,012766

Είναι εμφανές πως για ακόμη μία φορά οι μέθοδοι πρόβλεψης στο σύνολό τους έχουν έντονο το στοιχείο της αισιοδοξίας, λόγω του αρνητικού προσήμου των μέσων σφαλμάτων. Η μέθοδος που διαφοροποιείται είναι η μέθοδος Holt, της οποίας το πρόσημο είναι θετικό. Η μέθοδος με την μικρότερη κατά απόλυτη τιμή του μέσου σφάλματος είναι η Holt και ακολουθεί η Damped. Όσον αφορά την ακρίβεια των μεθόδων προβλέψεων και πιο συγκεκριμένα των μέσων απόλυτων σφαλμάτων κρίνεται σκόπιμη η παρουσίαση του επόμενου διαγράμματος. Αυτό που είναι γεγονός σχετικά με τα απόλυτα σφάλματα είναι πως οι διαφορές των σφαλμάτων των προβλέψεων κάθε μεθόδου, εκτός της LRL και του κινητού μέσου όρου, είναι ελάχιστες και δύσκολα ανιχνεύσιμες από το διάγραμμα. Παρόλα αυτά η μέθοδος με το μικρότερο απόλυτο σφάλμα για την περίπτωση προβλέψεων της επόμενης περιόδου είναι η Naïve, η SES και η ETS. Στην περίπτωση των προβλέψεων της τρίτης περιόδου καταλληλότερη είναι η Naïve ακολουθούμενη από την ETS και SES.



Για την αξιολόγηση των μεθόδων θα πρέπει να ληφθούν υπόψη τα ποσοστιαία σφάλματα των μεθόδων σε σχέση με τις πραγματικές τιμές. Αυτά τα σφάλματα συγκεντρώνονται στον παρακάτω πίνακα:

<b>Cotton (Daily)</b>	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	1,0323%	1,0312%	<b>0,997964</b>	1,8232%	1,8214%	<b>1,752847</b>
Holt	1,0431%	1,0427%	1,008523	1,8430%	1,8452%	1,771565
Damped	1,0358%	1,0348%	1,001534	1,8416%	1,8404%	1,770432
ETS	1,0323%	1,0312%	0,997964	1,8232%	1,8214%	1,752847
Moving Average	1,3057%	1,3047%	1,254036	1,9774%	1,9744%	1,894683
LRL	37,2240%	30,7148%	34,03195	37,4435%	30,8947%	34,42849
Naïve	<b>1,0323%</b>	<b>1,0312%</b>	0,997969	<b>1,8232%</b>	<b>1,8214%</b>	1,75285
Theta	1,0329%	1,0317%	0,998469	1,8248%	1,8226%	1,754038

Όπως φαίνεται από τους παραπάνω πίνακες η καταλληλότερη μέθοδος πρόβλεψης αποδεικνύεται πως είναι η Naïve και ακολούθως η SES και η ETS και για την εξαγωγή προβλέψεων και με τους δύο ορίζοντες πρόβλεψης.

### 7.2.14.2 Μηνιαία Δεδομένα

Με βάση τον συγκεκριμένο χρονικό ορίζοντα της χρονοσειράς βγάλαμε τα εξής αποτελέσματα για το μέσο σφάλμα και για το μέσο απόλυτο σφάλμα των μεθόδων πρόβλεψης που εφαρμόσαμε:

<b>Cotton (Monthly)</b>	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,03559	5,567919	-0,10601	13,83753
Holt	0,302536	5,841505	0,961871	15,40465
Damped	0,155514	5,259527	0,392717	13,0087
ETS	0,091642	<b>5,117885</b>	0,230204	<b>12,74667</b>
Moving Average	-0,10531	9,239651	-0,19486	16,606
LRL	24,11514	26,36444	24,3731	26,8151
Naïve	-0,03558	5,567674	-0,106	13,83729
Theta	<b>-0,02757</b>	5,572731	<b>-0,1052</b>	13,83809

Αυτό που είναι ορατό είναι η αισιοδοξία των μεθόδων προβλέψεων SES, Naïve και κινητού μέσου όρου και για τις δύο περιπτώσεις εξαγωγής προβλέψεων ανάλογα με την περίοδο πρόβλεψης. Αυτό γίνεται κατανοητό από την εξέταση του προσήμου του μέσου σφάλματος το οποίο είναι αρνητικό για όλες τις μεθόδους και για τις δύο περιπτώσεις. Οι υπόλοιπες μέθοδοι αντίστοιχα έχουν έντονο το στοιχείο της απαισιοδοξίας, λόγω του θετικού δείκτη σφάλματος ME. Γενικά, η μέθοδος Theta όπως επίσης και η μέθοδος Naïve, είναι οι δύο μέθοδοι που παρουσιάζουν τις μικρότερες κατά απόλυτη τιμή, τιμές του μέσου σφάλματος, συνεπώς είναι και οι καλύτερες όσον αφορά την απόδοση των μεθόδων σχετικά με την προκατάληψη για την παραγωγή προβλέψεων και για τους δύο ορίζοντες πρόβλεψης.

Σχετικά με την ακρίβεια των προβλέψεων για κάθε μέθοδο θα πρέπει να ανατρέξουμε στις τιμές των μέσων απόλυτων σφαλμάτων κάθε μεθόδου. Στην περίπτωση παραγωγής προβλέψεων για την επόμενη περίοδο η μέθοδος Damped και η μέθοδος ETS υποδηλώνουν την μεγαλύτερη ακρίβεια με βάση τα αποτελέσματά τους. Το ίδιο συμβαίνει και για την περίπτωση προβλέψεων για την τρίτη περίοδο.

Τώρα για την επιλογή της καταλληλότερης μεθόδου πρόβλεψης θα πρέπει να ελέγξουμε τα ποσοστιαία σφάλματα δίνοντας παραπάνω έμφαση στο MAPE. Συγκεκριμένα έχουμε:

<b>Cotton (Monthly)</b>	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	5,2385%	5,1726%	0,991086	12,5520%	12,3563%	2,450359
Holt	5,4713%	5,5236%	1,039784	15,2118%	16,2869%	2,727866
Damped	5,0186%	5,0324%	0,936192	12,4415%	12,4845%	2,30359
ETS	<b>4,9134%</b>	<b>4,9141%</b>	<b>0,91098</b>	<b>12,2689%</b>	<b>12,2306%</b>	<b>2,257189</b>

Moving Average	8,4463%	8,3340%	1,644651	15,3113%	14,9711%	2,940602
LRL	22,5702%	27,7398%	4,692851	22,7745%	28,1587%	4,748436
Naïve	5,2383%	5,1725%	0,991042	12,5518%	12,3561%	2,450318
Theta	5,2434%	5,1783%	0,991942	12,5526%	12,3570%	2,450459

Παρατηρώντας το συνολικό ποσοστό των σφαλμάτων διαπιστώνουμε πως τα σφάλματα είναι μεγαλύτερα σε σχέση με τις προηγούμενες εξεταζόμενες χρονοσειρές για την περίπτωση της μηνιαίας χρονοσειράς του βαμβακιού. Πιο αναλυτικά η μέθοδος ETS είχε τις μικρότερες τιμές σφαλμάτων σχετικά με τις υπόλοιπες για την περίπτωση πρόβλεψης για την επόμενη περίοδο.

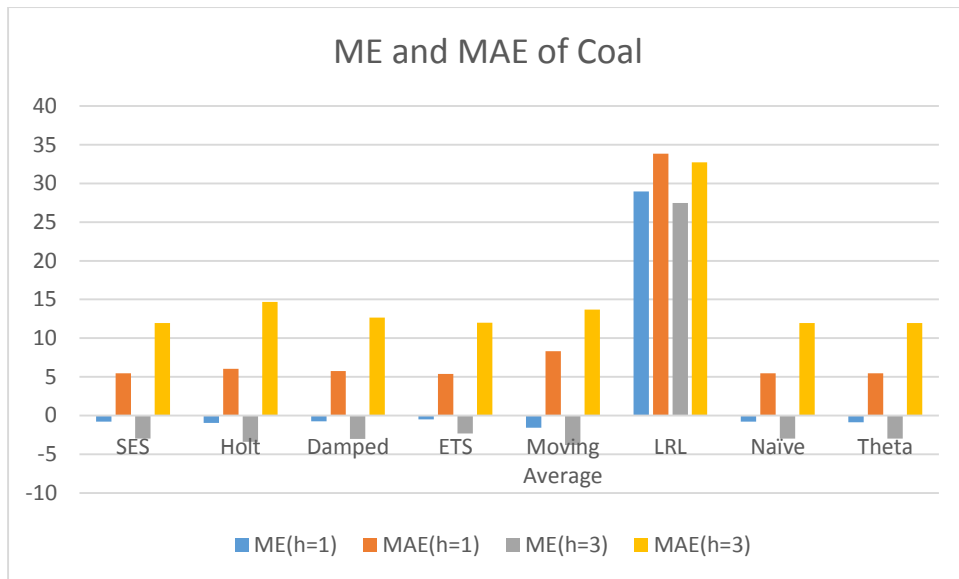
Στη περίπτωση της πρόβλεψης με ορίζοντα πρόβλεψης τρεις περιόδους η καλύτερη μέθοδος που προκύπτει με βάση τα σφάλματα είναι και πάλι η ETS και για αυτό θεωρείται καταλληλότερη.

### 7.2.15 Κάρβουνο (Coal)

Αντιθέτως με την περίπτωση του βαμβακιού, στη περίπτωση του κάρβουνου έχουμε διαθέσιμη μία μηνιαία χρονοσειρά. Επομένως ο πίνακας των σφαλμάτων αναφέρεται σε μηνιαίες προβλέψεις και είναι ο εξής:

Coal	ME(h=1)	MAE(h=1)	ME(h=3)	MAE(h=3)
SES	-0,79539	5,453699	-2,99075	11,94686
Holt	-0,9567	6,057534	-3,41061	14,66887
Damped	-0,73266	5,731308	-3,00335	12,6492
ETS	<b>-0,48544</b>	<b>5,393908</b>	<b>-2,30088</b>	11,99074
Moving Average	-1,58806	8,300233	-3,80478	13,70831
LRL	28,97263	33,82776	27,44925	32,72011
Naïve	-0,79942	5,449419	-2,99482	<b>11,94259</b>
Theta	-0,85474	5,467194	-2,99646	11,94903

Για την περίπτωση της προκατάληψης βλέπουμε πως όλες οι μέθοδοι έχουν το στοιχείο της αισιοδοξίας, λόγω των αρνητικών προσήμων των μεθόδων και για τις δύο περιπτώσεις, πλην της μεθόδου LRL. Στο παρακάτω διάγραμμα είναι συγκεντρωμένες οι τιμές των σφαλμάτων ME και MAE όλων των μεθόδων και για τις δύο περιπτώσεις.



Βλέπουμε πως η μέθοδος ETS και κατόπιν η Naïve έχουν την μικρότερη απόλυτη τιμή του σφάλματος ME στην περίπτωση εξαγωγής προβλέψεων της επόμενης περιόδου. Στη περίπτωση της ακρίβειας των προβλέψεων για χρονικό ορίζοντα τριών περιόδων πολύ καλά αποτελέσματα δίνουν η Naïve και κατόπιν η SES με την Theta.

Για την τελική αξιολόγηση των μεθόδων και της επιλογής της καταλληλότερης μεθόδου για την χρονοσειρά της τιμής του κάρβουνου θα εξετάσουμε τα ποσοστιαία σφάλματα, όπως φαίνονται παρακάτω:

Coal	MAPE (h=1)	sMAPE (h=1)	MAsE (h=1)	MAPE (h=3)	sMAPE (h=3)	MAsE (h=3)
SES	5,1874%	5,0837%	1,012746	12,0365%	11,2579%	2,460311
Holt	5,7707%	5,7037%	1,124878	14,9589%	14,1606%	3,020876
Damped	5,3830%	5,2988%	1,064298	12,4255%	11,4007%	2,60495
ETS	<b>5,1486%</b>	5,0869%	<b>1,001643</b>	<b>11,7355%</b>	<b>11,0189%</b>	2,469348
Moving Average	8,1127%	7,7843%	1,541345	14,1011%	12,8658%	2,823061
LRL	30,4166%	37,7221%	6,281781	30,2620%	37,1803%	6,73831
Naïve	5,1847%	<b>5,0806%</b>	1,011952	12,0341%	11,2548%	<b>2,459431</b>
Theta	5,2049%	5,0985%	1,015252	12,0394%	11,2602%	2,460757

Όπως φαίνεται από τον παραπάνω πίνακα πέρα πάσης αμφιβολίας η καταλληλότερη μέθοδος για την τιμή του κάρβουνου αποδεικνύεται πως είναι η ETS και ακολούθως η Naïve για την περίπτωση προβλέψεων της επόμενης, αλλά και της τρίτης περιόδου. Πολύ καλά αποτελέσματα δίνουν επίσης η SES και η Theta.

### 7.3 Σύγκριση Μεθόδων σύμφωνα με τον δείκτη MAPE

Το σφάλμα MAPE είναι ιδανικό για την σύγκριση των μεθόδων σε διαφορετικές χρονοσειρές. Για αυτό το λόγο θα ακολουθήσουν παρακάτω, διαχωρισμένα ανά

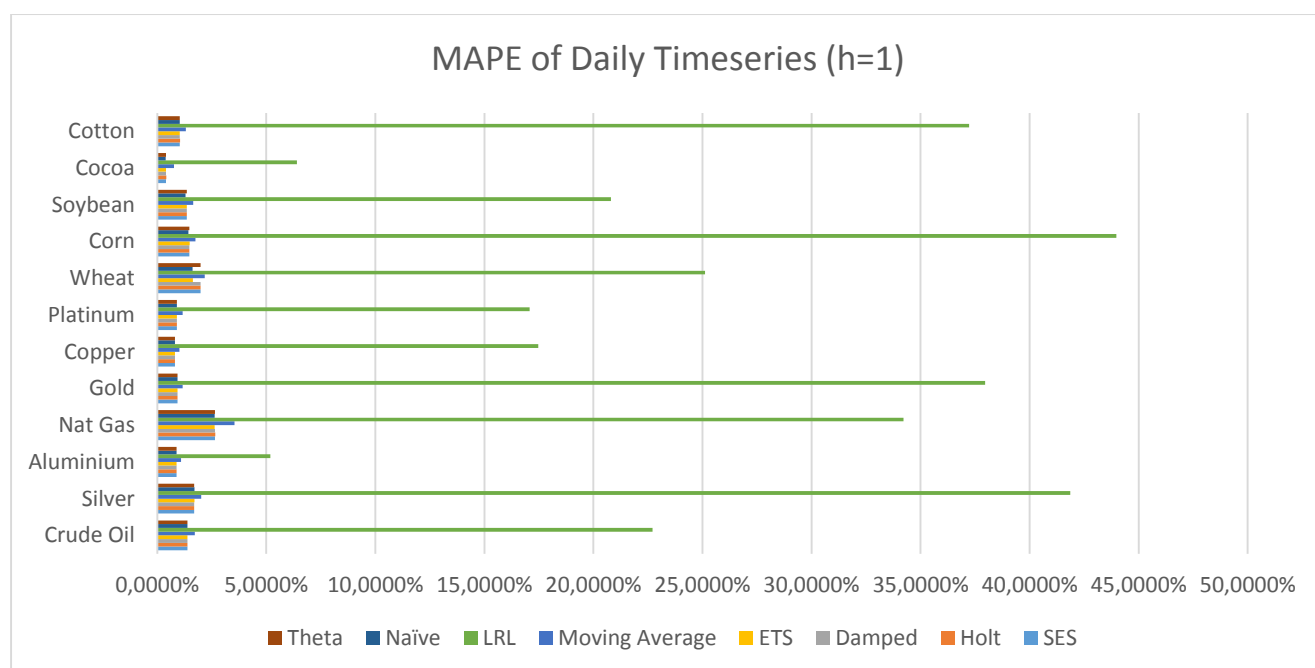
βραχυπρόθεσμο (ημερήσιες χρονοσειρές) και μεσοπρόθεσμο (μηνιαίες χρονοσειρές) ορίζοντα, τα σχετικά διαγράμματα συμπεριλαμβανομένων όλων των αγαθών, για να φανεί σε ποια αγαθά ανταποκρίνονται καλύτερα οι μέθοδοι πρόβλεψης. Κατόπιν θα υπάρχει ένας συγκεντρωτικός πίνακας επιλογής της καταλληλότερης μεθόδου σύμφωνα με τον δείκτη MAPE.

### 7.3.1 Προβλέψεις για την επόμενη περίοδο ( $h=1$ )

Εδώ μελετάμε την περίπτωση των σφαλμάτων MAPE για την περίπτωση παραγωγής προβλέψεων της επόμενης περιόδου. Φυσικά η σύγκριση είναι πιο σωστό να γίνει στα δεδομένα του ίδιου χρονικού ορίζοντα. Επομένως θα αναλυθεί περαιτέρω σε ημερήσιες και μηνιαίες χρονοσειρές

#### 7.3.1.1 Ημερήσιες Χρονοσειρές

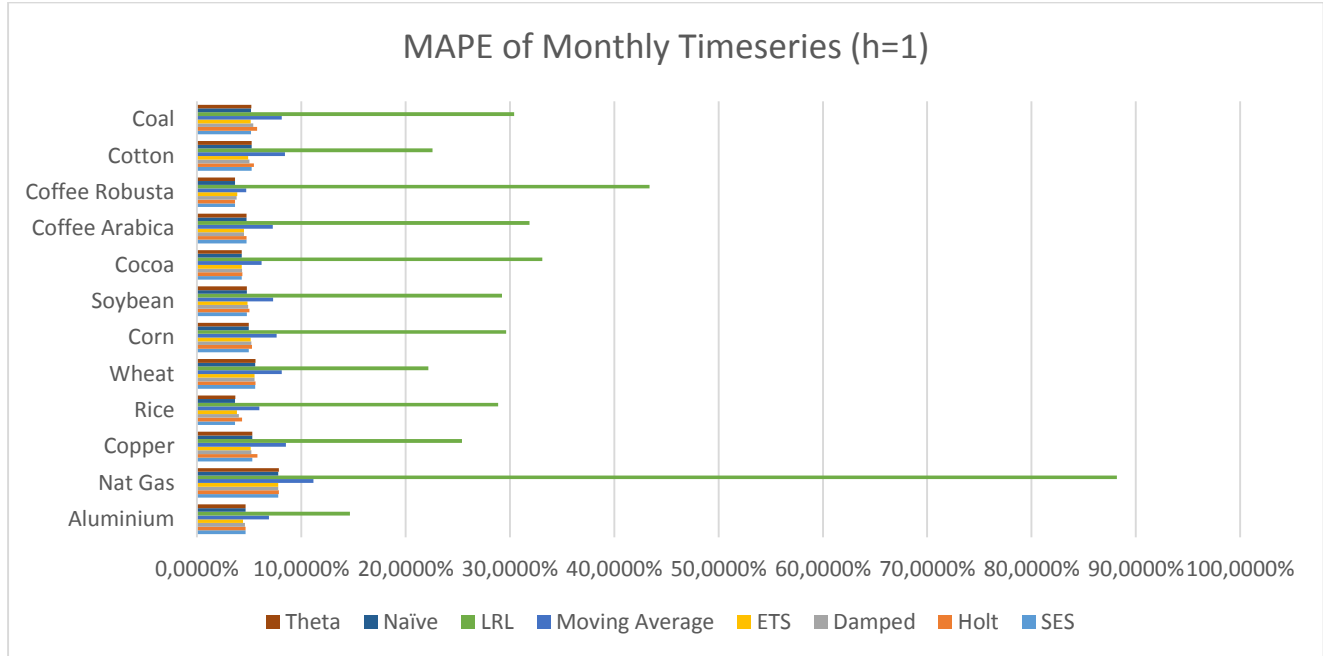
Στο διάγραμμα που ακολουθεί αναγράφονται όλες οι τιμές των σφαλμάτων MAPE για κάθε αγαθό και για τις οχτώ μεθόδους πρόβλεψης:



Από το παραπάνω διάγραμμα μπορούμε να βγάλουμε το συμπέρασμα, πως τα ημερήσια δεδομένα που έχουν τα μικρότερα σφάλματα από την εφαρμογή των μεθόδων πρόβλεψης για παραγωγή προβλέψεων της επόμενης περιόδου είναι το βαμβάκι (Cotton) και στη συνέχεια ο χαλκός (Copper). Τα χειρότερα αποτελέσματα συγκριτικά με τα υπόλοιπα αγαθά, τα έχει το φυσικό αέριο (Natural Gas) και ακολουθεί το σιτάρι με το ασήμι.

### 7.3.1.2 Μηνιαίες Χρονοσειρές

Στο διάγραμμα που ακολουθεί αναγράφονται όλες οι τιμές των σφαλμάτων MAPE για κάθε αγαθό και για τις οχτώ μεθόδους πρόβλεψης για τις μηνιαίες τιμές των αγαθών:



Στην περίπτωση των μηνιαίων χρονοσειρών τα δεδομένα που έχουν τα μικρότερα σφάλματα είναι οι τιμές της ποικιλίας καφέ Robusta και με τα μεγαλύτερα σφάλματα οι τιμές του φυσικού αερίου.

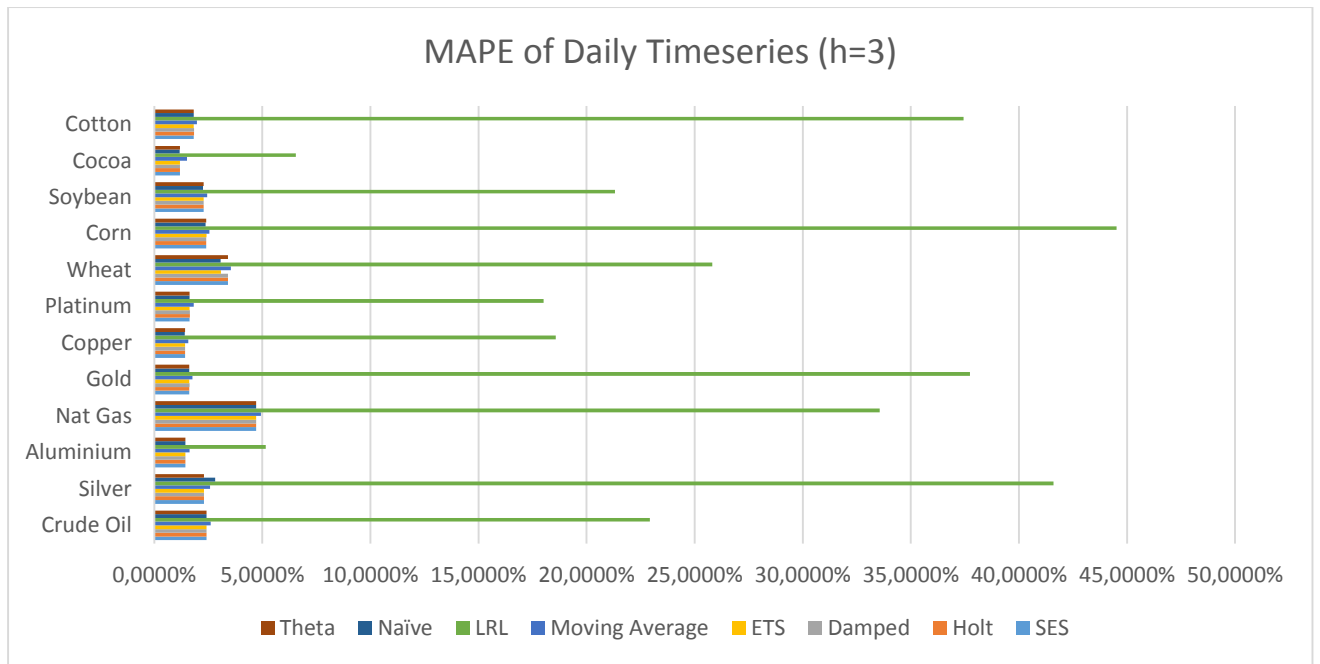
### 7.3.2 Προβλέψεις για την τρίτη περίοδο (h=3)

Εδώ μελετάμε την περίπτωση των σφαλμάτων MAPE για την περίπτωση παραγωγής προβλέψεων της επόμενης περιόδου. Φυσικά η σύγκριση είναι πιο σωστό να γίνει στα δεδομένα του ίδιου χρονικού ορίζοντα. Επομένως θα αναλυθεί περαιτέρω σε ημερήσιες και μηνιαίες χρονοσειρές.

#### 7.3.2.1 Ημερήσιες Χρονοσειρές

Στο διάγραμμα που ακολουθεί αναγράφονται όλες οι τιμές των σφαλμάτων MAPE για κάθε αγαθό και για τις οχτώ μεθόδους πρόβλεψης στην περίπτωση παραγωγής προβλέψεων για την τρίτη περίοδο:

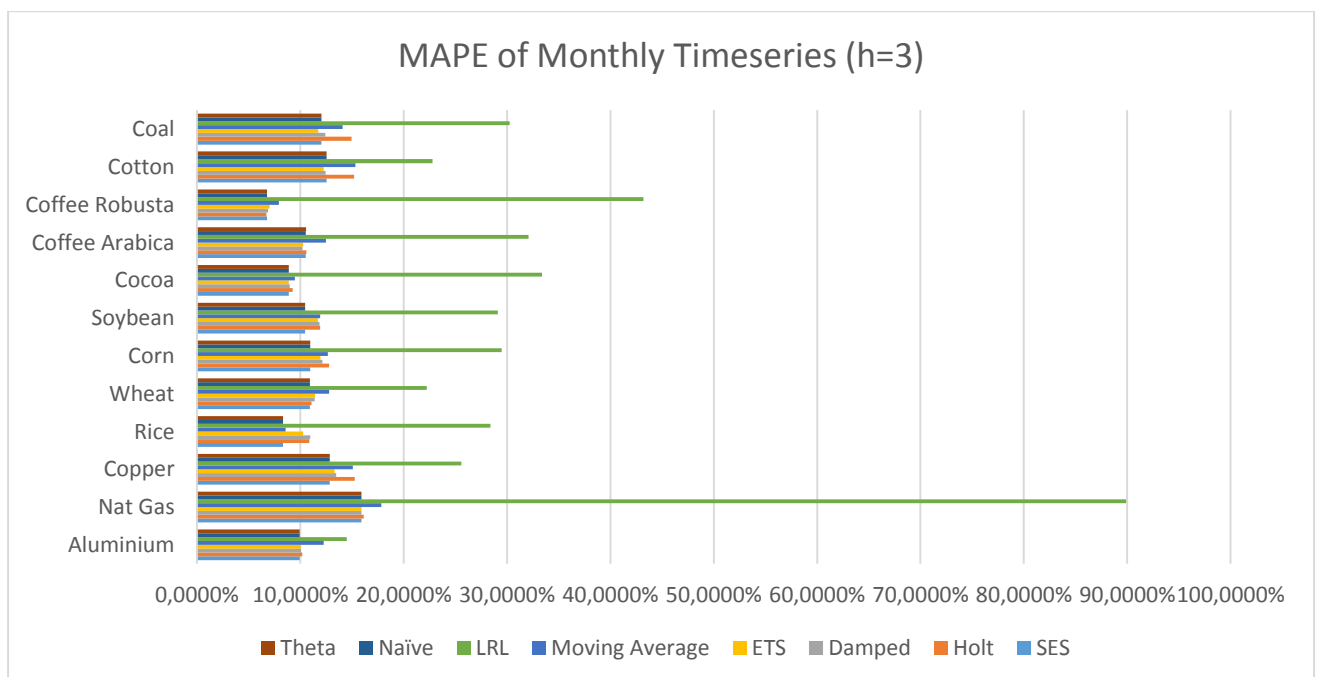




Από το παραπάνω διάγραμμα μπορούμε να βγάλουμε το συμπέρασμα, πως τα ημερήσια δεδομένα που έχουν τα μικρότερα σφάλματα από την εφαρμογή των μεθόδων πρόβλεψης για παραγωγή προβλέψεων της επόμενης περιόδου είναι το βαμβάκι (Cotton) και στη συνέχεια ο χαλκός (Copper). Τα χειρότερα αποτελέσματα συγκριτικά με τα υπόλοιπα αγαθά, τα έχει το φυσικό αέριο (Natural Gas) και ακολουθεί το σιτάρι (Wheat) και το ασήμι(Silver).

### 7.3.2.2 Μηνιαίες Χρονοσειρές

Στο διάγραμμα που ακολουθεί αναγράφονται όλες οι τιμές των σφαλμάτων MAPE για κάθε αγαθό και για τις οχτώ μεθόδους πρόβλεψης για τις μηνιαίες τιμές των αγαθών:



Στην περίπτωση των μηνιαίων χρονοσειρών τα δεδομένα που έχουν τα μικρότερα σφάλματα είναι οι τιμές της ποικιλίας καφέ Robusta και ακολουθούν οι τιμές του ρυζιού (Rice) και του κακάο (Cocoa). Οι τιμές με τα μεγαλύτερα σφάλματα είναι οι τιμές του φυσικού αερίου (Natural Gas) και του χαλκού (Copper).

## 7.4 Επιλογή Καταλληλότερου Μοντέλου Πρόβλεψης για Κάθε Αγαθό

Όσον αφορά τον προσδιορισμό του καταλληλότερου μοντέλου πρόβλεψης για την εκάστοτε κατηγορία αγαθών αυτός θα πραγματοποιηθεί μέσω της διαδικασίας της παραμετρικής κατάταξης των μεθόδων. Αυτό εν γένει συμβαίνει από την επιλογή της μεθόδου με την μικρότερη τιμή του σφάλματος MAPE. Ο διαχωρισμός που έγινε και εδώ αφορά τον χρονικό ορίζοντα των χρονοσειρών και τον ορίζοντα εξαγωγής προβλέψεων.

### 7.4.1 Προβλέψεις για την επόμενη περίοδο ( $h=1$ )

Παρακάτω βρίσκεται ένα συγκεντρωτικός πίνακας με βάση τα ελάχιστα MAPE που παρατηρήθηκαν στη παράγραφο 7.1 και επισήμανση του μοντέλου με την χαμηλότερη τιμή σφάλματος MAPE.

Αγαθά	Είδος Χρονοσειράς	Επιλογή Μεθόδου	Τιμή σφάλματος MAPE( $h=1$ )
1) Πετρέλαιο	Ημερήσια	ETS	1,3796%
2) Ασήμι	Ημερήσια	Theta	1,6880%
3) Αλουμίνιο	Ημερήσια	Theta	0,8771%
	Μηνιαία	ETS	4,4282%
4) Φυσικό Αέριο	Ημερήσια	Naïve	2,6329%
	Μηνιαία	Naïve	7,7779%
5) Χρυσός	Ημερήσια	ETS	0,9195%
6) Χαλκός	Ημερήσια	Holt	0,8103%
	Μηνιαία	ETS	5,1642%
7) Λευκόχρυσος	Ημερήσια	Damped	0,8924%
8) Ρύζι	Μηνιαία	Naïve	3,6601%
9) Σιτάρι	Ημερήσια	Naïve	1,6231%
	Μηνιαία	Damped	5,5102%
10) Καλαμπόκι	Ημερήσια	Naïve	1,4289%
	Μηνιαία	Naïve	4,9682%
11) Σόγια	Ημερήσια	Naïve	1,2991%
	Μηνιαία	Naïve	4,7799%
12) Κακάο	Ημερήσια	Naïve	0,3857%
	Μηνιαία	Naïve	4,2891%
13α) Καφές Arabica	Μηνιαία	Damped	4,5141%
13β) Καφές Robusta	Μηνιαία	Holt	3,6356%
14) Βαμβάκι	Ημερήσια	Naïve	1,0323%
	Μηνιαία	ETS	4,9134%
15) Κάρβουνο	Μηνιαία	ETS	5,1486%

### 7.4.2 Προβλέψεις για την τρίτη περίοδο (h=3)

Παρακάτω βρίσκεται ένα συγκεντρωτικός πίνακας με βάση τα ελάχιστα MAPE που παρατηρήθηκαν στη παράγραφο 7.1 και επισήμανση του μοντέλου με την χαμηλότερη τιμή σφάλματος MAPE.

<b>Αγαθά</b>	<b>Είδος Χρονοσειράς</b>	<b>Επιλογή μεθόδου</b>	<b>Τιμή σφάλματος MAPE(h=3)</b>
1) Πετρέλαιο	Ημερήσια	Damped	2,4098%
2) Ασήμι	Ημερήσια	Holt	2,3009%
3) Αλουμίνιο	Ημερήσια	Theta	1,4382%
	Μηνιαία	Naïve	9,9325%
4) Φυσικό Αέριο	Ημερήσια	Theta	4,7083%
	Μηνιαία	Damped	15,9203%
5) Χρυσός	Ημερήσια	ETS	1,6152%
6) Χαλκός	Ημερήσια	Naïve	1,4159%
	Μηνιαία	Naïve	12,8624%
7) Λευκόχρυσος	Ημερήσια	Naïve	1,6359%
8) Ρύζι	Μηνιαία	Naïve	8,3315%
9) Σιτάρι	Ημερήσια	Naïve	3,0642%
	Μηνιαία	Naïve	10,9399%
10) Καλαμπόκι	Ημερήσια	Naïve	2,3714%
	Μηνιαία	Naïve	10,9677%
11) Σόγια	Ημερήσια	Naïve	2,2508%
	Μηνιαία	Naïve	10,4691%
12) Κακάο	Ημερήσια	Naïve	1,1678%
	Μηνιαία	ETS	8,8725%
13α) Καφές Arabica	Μηνιαία	Damped	10,2168%
13β) Καφές Robusta	Μηνιαία	Holt	6,7216%
14) Βαμβάκι	Ημερήσια	Naïve	1,8232%
	Μηνιαία	ETS	12,2689%
15) Κάρβουνο	Μηνιαία	ETS	11,7355%

## **7.5 Κατάταξη και Βαθμολόγηση των Μεθόδων Πρόβλεψης Συνολικά στα Αγαθά**

Ένας άλλος τρόπος αξιολόγησης των μεθόδων γίνεται με βάση την κατάταξή τους στα αποτελέσματα που δίνει για κάθε χρονοσειρά. Συγκεκριμένα, παραθέτουμε πίνακα με τις θέσεις που έχει καταλάβει κάθε μοντέλο στο σύνολο των χρονοσειρών των αγαθών. Με βάση αυτόν τον πίνακα, προκύπτει και το επόμενο είδος πίνακα που θα παρουσιάσουμε στο κεφάλαιο αυτό και δίνει ακόμα καλύτερη εικόνα της αποτελεσματικότητας κάθε χρονοσειράς. Συγκεκριμένα, για τις θέσεις από 1 έως 8 (κατάταξη των οχτώ μεθόδων) υπάρχει ο αντίστοιχος παράγοντας κόστους, που είναι η ίδια η θέση. Για μία χρονοσειρά δηλαδή που έχει καταλάβει 3 φορές την 2η θέση το αποτέλεσμα είναι  $3 \times 2 = 6$  μονάδες κόστους. Με βάση αυτή τη διαδικασία και διαιρώντας με τον αριθμό των συνολικών χρονοσειρών προκύπτει και η κατάταξη αποτελεσματικότητας κάθε μεθόδου. Με παρόμοιο τρόπο θα παρουσιαστούν τα αποτελέσματα και ανά είδος χρονοσειράς ανάλογα με τον ορίζοντα της χρονοσειράς, δηλαδή για ημερήσιες και μηνιαίες χρονοσειρές. Επίσης θα γίνει κατηγοριοποίηση σύμφωνα με τον ορίζοντα πρόβλεψης, δηλαδή της επόμενης περιόδου και των επόμενων τριών.

Η παραπάνω διαδικασία γίνεται με σκοπό την γενική αξιολόγηση των μεθόδων, στα πλαίσια ενός ευρύτερου συνόλου δεδομένων, εν προκειμένω τα αγαθά. Δηλαδή να προκύψει μία γενική εκτίμηση των μεθόδων που ανταποκρίνονται καλύτερα στις χρονοσειρές του dataset των εμπορευμάτων σε σχέση με την προκατάληψη, την ακρίβεια και τη ποσοστιαία απόκλιση των προβλέψεων σε σχέση με τις πραγματικές τιμές.

Θα πρέπει να τονιστεί πως οι πίνακες κατάταξης και βαθμολόγησης των μεθόδων έγινε και για τα πέντε είδη σφαλμάτων που έχουμε υπολογίσει. Επομένως παρακάτω παρουσιάζονται οι πίνακες με τις ανάλογες κατηγορίες.

### **7.5.1 Κατάταξη Μεθόδων Παραγωγής Προβλέψεων για την Επόμενη Περίοδο ( $h=1$ )**

Σε αυτή την περίπτωση θα αναφερθούν σε αναλυτικούς πίνακες και ανά σφάλμα η κατάταξη των μεθόδων για την παραγωγή προβλέψεων της επόμενης περιόδου, κατηγοριοποιημένα ανάλογα με τον χρονικό ορίζοντα ης εκάστοτε χρονοσειράς. Άρα έχουμε τους παρακάτω πίνακες που αναγράφεται αναλυτικά η κατάταξη των μεθόδων ανά σφάλμα.

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	5	1	0	0	0	3	2
2	0	2	5	1	0	0	2	1
3	3	0	2	3	0	0	3	2
4	4	0	1	3	0	0	4	0
5	2	0	2	4	2	0	0	3
6	2	3	1	1	0	0	0	4
7	0	2	0	0	10	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη ME – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	5	1	0	0	0	3	2
2	0	4	10	2	0	0	4	2
3	9	0	6	9	0	0	9	6
4	16	0	4	12	0	0	16	0
5	10	0	10	20	10	0	0	15
6	12	18	6	6	0	0	0	24
7	0	14	0	0	70	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>48</b>	<b>41</b>	<b>37</b>	<b>49</b>	<b>80</b>	<b>96</b>	<b>32</b>	<b>49</b>
<b>Γενική Κατάταξη</b>	<b>4</b>	<b>3,42</b>	<b>3,08</b>	<b>4,08</b>	<b>6,67</b>	<b>8</b>	<b>2,67</b>	<b>4,08</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη ME – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	2	0	8	0	0	1	1
2	0	0	8	2	0	0	2	0
3	3	2	1	0	0	0	4	2
4	5	0	1	2	0	0	3	1
5	2	1	1	0	1	0	2	5
6	2	4	1	0	2	0	0	3
7	0	3	0	0	9	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη ME – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	2	0	8	0	0	1	1
2	0	0	16	4	0	0	4	0
3	9	6	3	0	0	0	12	6
4	20	0	4	8	0	0	12	4
5	10	5	5	0	5	0	10	25
6	12	24	6	0	12	0	0	18
7	0	21	0	0	63	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>51</b>	<b>58</b>	<b>34</b>	<b>20</b>	<b>80</b>	<b>96</b>	<b>0</b>	<b>54</b>
<b>Γενική Κατάταξη</b>	<b>4,25</b>	<b>4,83</b>	<b>2,83</b>	<b>1,67</b>	<b>6,67</b>	<b>8</b>	<b>3,25</b>	<b>4,5</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη ME – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	2	0	0	6	2
2	3	2	2	3	0	0	1	1
3	3	2	2	1	0	0	2	2
4	4	1	3	2	0	0	0	2
5	2	1	2	3	0	0	0	4
6	0	5	2	1	0	0	3	1
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	2	0	0	6	2
2	6	4	4	6	0	0	2	2
3	9	6	6	3	0	0	6	6
4	16	4	12	8	0	0	0	8
5	10	5	10	15	0	0	0	20
6	0	30	12	6	0	0	18	6
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>41</b>	<b>50</b>	<b>45</b>	<b>40</b>	<b>84</b>	<b>96</b>	<b>32</b>	<b>44</b>
<b>Γενική Κατάταξη</b>	<b>3,42</b>	<b>4,17</b>	<b>3,75</b>	<b>3,33</b>	<b>7</b>	<b>8</b>	<b>2,67</b>	<b>3,67</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	5	0	0	5	0
2	4	0	4	2	0	0	1	1
3	2	1	1	0	0	0	3	5
4	4	0	0	4	0	0	3	1
5	2	1	6	0	0	0	0	3
6	0	9	0	1	0	0	0	2
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	5	0	0	5	0
2	8	0	8	4	0	0	2	2
3	6	3	3	0	0	0	9	15
4	16	0	0	16	0	0	12	4
5	10	5	30	0	0	0	0	15
6	0	54	0	6	0	0	0	12
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>40</b>	<b>63</b>	<b>42</b>	<b>31</b>	<b>84</b>	<b>96</b>	<b>28</b>	<b>48</b>
<b>Γενική Κατάταξη</b>	<b>3,33</b>	<b>5,25</b>	<b>3,5</b>	<b>2,58</b>	<b>7</b>	<b>8</b>	<b>2,33</b>	<b>4</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	1	1	2	0	0	5	2
2	0	2	3	4	0	0	3	0
3	7	1	1	1	0	0	1	1
4	2	2	4	2	0	0	0	2
5	2	1	2	1	0	0	1	5
6	0	5	1	2	0	0	2	2
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	1	1	2	0	0	5	2
2	0	4	6	8	0	0	6	0
3	21	3	3	3	0	0	3	3
4	8	8	16	8	0	0	0	8
5	10	5	10	5	0	0	5	25
6	0	30	6	12	0	0	12	12
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>40</b>	<b>51</b>	<b>42</b>	<b>38</b>	<b>84</b>	<b>96</b>	<b>31</b>	<b>50</b>
<b>Γενική Κατάταξη</b>	<b>3.33</b>	<b>4.25</b>	<b>4.2</b>	<b>3.17</b>	<b>7</b>	<b>8</b>	<b>2.58</b>	<b>4.17</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	4	0	0	5	0
2	4	0	3	3	0	0	1	1
3	2	1	1	0	0	0	3	5
4	4	0	0	4	0	0	3	1
5	2	1	6	0	0	0	0	3
6	0	9	0	1	0	0	0	2
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	2	2	4	0	0	5	0
2	8	0	6	6	0	0	2	2
3	6	3	3	0	0	0	9	15
4	16	0	0	16	0	0	12	4
5	10	5	30	0	0	0	0	15
6	0	54	0	6	0	0	0	12
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>40</b>	<b>64</b>	<b>41</b>	<b>32</b>	<b>84</b>	<b>96</b>	<b>28</b>	<b>48</b>
<b>Γενική Κατάταξη</b>	<b>3.33</b>	<b>5.33</b>	<b>3.42</b>	<b>2.67</b>	<b>7</b>	<b>8</b>	<b>2.33</b>	<b>4</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	2	0	0	6	2
2	1	2	3	4	0	0	1	1
3	5	2	1	1	0	0	1	2
4	4	1	4	1	0	0	1	1
5	1	1	2	2	0	0	1	5
6	1	5	1	2	0	0	2	1
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	2	0	0	6	2
2	2	4	6	8	0	0	2	2
3	15	6	3	3	0	0	3	6
4	16	4	16	4	0	0	4	4
5	5	5	10	10	0	0	5	25
6	6	30	6	12	0	0	12	6
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>44</b>	<b>50</b>	<b>42</b>	<b>39</b>	<b>84</b>	<b>96</b>	<b>32</b>	<b>45</b>
<b>Γενική Κατάταξη</b>	<b>3,67</b>	<b>4,17</b>	<b>3,5</b>	<b>3,25</b>	<b>7</b>	<b>8</b>	<b>2,67</b>	<b>3,75</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	3	0	0	6	0
2	5	0	2	3	0	0	0	2
3	1	1	1	1	0	0	4	4
4	5	0	0	4	0	0	2	1
5	1	0	7	0	0	0	0	4
6	0	10	0	1	0	0	0	1
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	3	0	0	6	0
2	10	0	4	6	0	0	0	4
3	3	3	3	3	0	0	12	12
4	20	0	0	16	0	0	8	4
5	5	0	35	0	0	0	0	20
6	0	60	0	6	0	0	0	6
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>38</b>	<b>64</b>	<b>44</b>	<b>34</b>	<b>84</b>	<b>96</b>	<b>26</b>	<b>46</b>
<b>Γενική Κατάταξη</b>	<b>3,17</b>	<b>5,33</b>	<b>3,67</b>	<b>2,83</b>	<b>7</b>	<b>8</b>	<b>2,17</b>	<b>3,83</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Μηνιαίων Χρονοσειρών



Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	2	1	2	0	0	5	1
2	2	1	2	4	0	0	1	2
3	3	2	2	0	0	0	3	2
4	4	1	3	2	0	0	0	2
5	2	1	2	3	0	0	0	4
6	0	5	2	1	0	0	3	1
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	2	1	2	0	0	5	1
2	4	2	4	8	0	0	2	4
3	9	6	6	0	0	0	9	6
4	16	4	12	8	0	0	0	8
5	10	5	10	15	0	0	0	20
6	0	30	12	6	0	0	18	6
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>40</b>	<b>49</b>	<b>45</b>	<b>39</b>	<b>84</b>	<b>96</b>	<b>34</b>	<b>45</b>
<b>Γενική Κατάταξη</b>	<b>3,33</b>	<b>4,08</b>	<b>3,75</b>	<b>3,25</b>	<b>7</b>	<b>8</b>	<b>2,83</b>	<b>3,75</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	5	0	0	5	0
2	4	0	4	2	0	0	1	1
3	2	1	1	0	0	0	3	5
4	4	0	0	4	0	0	3	1
5	2	1	6	0	0	0	0	3
6	0	9	0	1	0	0	0	2
7	0	0	0	0	12	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	5	0	0	5	0
2	8	0	8	4	0	0	2	2
3	6	3	3	0	0	0	9	15
4	16	0	0	16	0	0	12	4
5	10	5	30	0	0	0	0	15
6	0	54	0	6	0	0	0	12
7	0	0	0	0	84	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>40</b>	<b>63</b>	<b>42</b>	<b>31</b>	<b>84</b>	<b>96</b>	<b>28</b>	<b>48</b>
<b>Γενική Κατάταξη</b>	<b>3,33</b>	<b>5,25</b>	<b>3,5</b>	<b>2,58</b>	<b>7</b>	<b>8</b>	<b>2,33</b>	<b>4</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Μηνιαίων Χρονοσειρών

Με βάση τα παραπάνω είναι εμφανές πως για τα ημερήσια δεδομένα σε σχέση με την προκατάληψη, οι μέθοδοι Naïve και ακολούθως η Damped ανταποκρίνονται καλύτερα με βάση το μέτρο της απόλυτης τιμής των σφαλμάτων ME. Αντιθέτως, οι μέθοδοι LRL και κινητού μέσου όρου πήραν την χειρότερη βαθμολογία με μεγάλη διαφορά από τις υπόλοιπες μεθόδους. Όσον αφορά τις μηνιαίες χρονοσειρές, η μέθοδος που ανταποκρίνεται καλύτερα σε σχέση με την προκατάληψη είναι η ETS και ακολουθεί η Damped, ενώ η Naïve βρίσκεται τώρα στην τρίτη θέση. Τελευταίες είναι και πάλι οι μέθοδοι LRL και κινητού μέσου όρου.

Τώρα η βαθμολογία με βάση τον δείκτη της ακρίβειας (MAE) για τα ημερήσια δεδομένα εμφάνισε την Naïve ως την πλέον κατάλληλη μέθοδο πρόβλεψης. Στην δεύτερη θέση εμφανίζεται η μέθοδος ETS και οι μέθοδοι της απλής γραμμικής παλινδρόμησης και του κινητού μέσου όρου ακολουθούν στις τελευταίες θέσεις. Ακριβώς τα ίδια ισχύουν και για την ακρίβεια των μηνιαίων χρονοσειρών των αγαθών.

Στην περίπτωση των ποσοστιαίων σφαλμάτων των προβλέψεων σε σχέση με τις πραγματικές τιμές, δεν υπήρξε κάποια διαφοροποίηση σε σχέση με πριν, είτε αφορά τις ημερήσιες χρονοσειρές, είτε τις μηνιαίες. Επομένως η Naïve πήρε την πρώτη θέση στη σχετική βαθμολόγηση και η ETS ακολουθεί. Αντιστοίχως ως η πλέον χειρότερη αποδείχτηκε η LRL και προτελευταία η μέθοδος κινητού μέσου όρου.

Ακριβώς η ίδια διαπίστωση βγήκε για τα συμμετρικά σφάλματα, αλλά και για τα κανονικοποιημένα. Δηλαδή για το sMAPE και για το MAsE.

### **7.5.2 Κατάταξη Μεθόδων Παραγωγής Προβλέψεων για την Τρίτη Περίοδο**

Σε αυτή την περίπτωση θα αναφερθούν σε αναλυτικούς πίνακες και ανά σφάλμα η κατάταξη των μεθόδων για την παραγωγή προβλέψεων της τρίτης περιόδου, κατηγοριοποιημένα ανάλογα με τον χρονικό ορίζοντα της εκάστοτε χρονοσειράς. Άρα οι πίνακες σχετικά με την βαθμολόγηση των μεθόδων ανά είδος σφάλματος και ανά χρονικό ορίζοντα των χρονοσειρών βρίσκεται παρακάτω.

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	7	1	1	0	0	2	0
2	0	0	5	1	0	0	3	3
3	4	0	2	2	0	0	2	2
4	4	0	2	3	0	0	3	0
5	0	0	1	5	2	0	1	3
6	3	2	1	0	2	0	0	4
7	0	3	0	0	8	0	1	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη ME – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	7	1	1	0	0	2	0
2	0	0	10	2	0	0	6	6
3	12	0	6	6	0	0	6	6
4	16	0	8	12	0	0	12	0
5	0	0	5	25	10	0	5	15
6	18	12	5	0	12	0	0	24
7	0	21	0	0	56	0	7	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>47</b>	<b>40</b>	<b>35</b>	<b>46</b>	<b>78</b>	<b>96</b>	<b>38</b>	<b>51</b>
<b>Γενική Κατάταξη</b>	<b>3,92</b>	<b>3,33</b>	<b>2,92</b>	<b>3,83</b>	<b>6,5</b>	<b>8</b>	<b>3,17</b>	<b>4,25</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη ME – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	2	1	8	0	0	0	1
2	1	0	7	1	0	0	3	0
3	2	2	1	2	0	0	5	0
4	5	0	0	0	1	0	2	4
5	2	0	2	1	0	0	2	5
6	2	4	1	0	3	0	0	2
7	0	4	0	0	7	1	0	0
8	0	0	0	0	1	11	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη ME – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	2	1	8	0	0	0	1
2	2	0	14	2	0	0	6	0
3	6	6	3	6	0	0	15	0
4	20	0	0	0	4	0	8	16
5	10	0	10	5	0	0	10	25
6	12	24	6	0	18	0	0	12
7	0	28	0	0	49	7	0	0
8	0	0	0	0	8	88	0	0
<b>Σύνολο</b>	<b>50</b>	<b>60</b>	<b>34</b>	<b>21</b>	<b>79</b>	<b>95</b>	<b>39</b>	<b>54</b>
<b>Γενική Κατάταξη</b>	<b>4,17</b>	<b>5</b>	<b>2,83</b>	<b>1,75</b>	<b>6,58</b>	<b>7,92</b>	<b>3,25</b>	<b>4,5</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη ME – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	1	0	3	0	0	6	1
2	2	4	1	2	0	0	2	1
3	1	1	3	3	0	0	0	4
4	5	0	1	1	0	0	1	4
5	3	3	3	1	0	0	1	1
6	0	3	4	2	1	0	1	1
7	0	0	0	0	11	0	1	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	2	1	0	3	0	0	6	1
2	4	8	2	4	0	0	4	2
3	3	3	9	9	0	0	0	12
4	20	0	4	4	0	0	4	16
5	15	15	15	5	0	0	5	5
6	0	18	24	12	6	0	6	6
7	0	0	0	0	77	0	7	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>44</b>	<b>45</b>	<b>54</b>	<b>37</b>	<b>83</b>	<b>96</b>	<b>32</b>	<b>42</b>
<b>Γενική Κατάταξη</b>	<b>3,67</b>	<b>3,75</b>	<b>4,5</b>	<b>3,08</b>	<b>6,92</b>	<b>8</b>	<b>2,67</b>	<b>3,5</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	2	0	0	6	1
2	6	0	1	1	0	0	2	2
3	2	0	0	0	0	0	4	6
4	4	1	1	5	1	0	0	0
5	0	1	6	2	0	0	0	3
6	0	6	2	2	2	0	0	0
7	0	3	0	0	9	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	2	0	0	6	1
2	12	0	2	2	0	0	4	4
3	6	0	0	0	0	0	12	18
4	16	4	4	20	4	0	0	0
5	0	5	30	10	0	0	0	15
6	0	36	12	12	12	0	0	0
7	0	21	0	0	63	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>34</b>	<b>67</b>	<b>50</b>	<b>46</b>	<b>79</b>	<b>96</b>	<b>20</b>	<b>38</b>
<b>Γενική Κατάταξη</b>	<b>2,83</b>	<b>5,58</b>	<b>4,17</b>	<b>3,83</b>	<b>6,58</b>	<b>8</b>	<b>1,67</b>	<b>3,17</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	1	0	0	7	2
2	3	4	2	2	0	0	1	0
3	2	1	2	3	0	0	1	3
4	3	0	2	1	0	0	1	5
5	4	2	2	3	0	0	0	1
6	0	4	3	2	1	0	1	1
7	0	0	0	0	11	0	1	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	1	1	0	0	7	2
2	6	8	4	4	0	0	2	0
3	6	3	6	9	0	0	3	9
4	12	0	8	4	0	0	4	20
5	20	10	10	15	0	0	0	5
6	0	24	18	12	6	0	6	6
7	0	0	0	0	77	0	7	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>44</b>	<b>46</b>	<b>47</b>	<b>45</b>	<b>83</b>	<b>96</b>	<b>29</b>	<b>42</b>
<b>Γενική Κατάταξη</b>	<b>3,67</b>	<b>3,83</b>	<b>3,92</b>	<b>3,75</b>	<b>6,92</b>	<b>8</b>	<b>2,42</b>	<b>3,5</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	3	0	0	6	0
2	6	0	1	1	0	0	2	2
3	1	0	0	1	0	0	4	6
4	5	1	0	4	1	0	0	1
5	0	0	8	1	0	0	0	3
6	0	6	0	2	4	0	0	0
7	0	4	1	0	7	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	3	0	0	6	0
2	12	0	2	2	0	0	4	4
3	3	0	0	3	0	0	12	18
4	20	4	0	16	4	0	0	4
5	0	0	40	5	0	0	0	15
6	0	36	0	12	24	0	0	0
7	0	28	7	0	49	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>35</b>	<b>69</b>	<b>51</b>	<b>41</b>	<b>77</b>	<b>96</b>	<b>22</b>	<b>41</b>
<b>Γενική Κατάταξη</b>	<b>2,92</b>	<b>5,75</b>	<b>4,25</b>	<b>3,42</b>	<b>6,42</b>	<b>8</b>	<b>1,83</b>	<b>3,42</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAPE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	2	0	2	0	0	7	1
2	1	3	1	3	0	0	1	3
3	5	1	4	1	0	0	0	1
4	3	0	2	1	0	0	1	5
5	3	3	2	3	0	0	0	1
6	0	3	3	2	1	0	2	1
7	0	0	0	0	11	0	1	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	2	0	2	0	0	7	1
2	2	6	2	6	0	0	2	6
3	15	3	12	3	0	0	0	3
4	12	0	8	4	0	0	4	20
5	15	15	10	15	0	0	0	5
6	0	18	18	12	6	0	12	6
7	0	0	0	0	77	0	7	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>44</b>	<b>44</b>	<b>50</b>	<b>42</b>	<b>83</b>	<b>96</b>	<b>32</b>	<b>35</b>
<b>Γενική Κατάταξη</b>	<b>3,67</b>	<b>3,67</b>	<b>4,17</b>	<b>3,5</b>	<b>6,92</b>	<b>8</b>	<b>2,67</b>	<b>2,92</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	3	0	0	3	3
2	3	0	0	1	0	0	5	3
3	8	0	0	0	0	0	1	3
4	1	1	1	3	1	0	3	2
5	0	2	7	1	1	0	0	1
6	0	4	0	4	4	0	0	0
7	0	4	2	0	6	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	3	0	0	3	3
2	6	0	0	2	0	0	10	6
3	24	0	0	0	0	0	3	9
4	4	4	4	12	4	0	12	8
5	0	10	35	5	5	0	0	5
6	0	24	0	24	24	0	0	0
7	0	28	14	0	42	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>34</b>	<b>67</b>	<b>55</b>	<b>46</b>	<b>75</b>	<b>96</b>	<b>28</b>	<b>31</b>
<b>Γενική Κατάταξη</b>	<b>2,83</b>	<b>5,58</b>	<b>4,58</b>	<b>3,83</b>	<b>6,25</b>	<b>8</b>	<b>2,33</b>	<b>2,58</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη sMAPE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	1	0	3	0	0	6	1
2	2	4	1	3	0	0	1	1
3	2	1	3	2	0	0	1	3
4	4	0	1	1	0	0	1	5
5	3	3	3	1	0	0	1	1
6	0	3	4	2	1	0	1	1
7	0	0	0	0	11	0	1	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	1	1	0	3	0	0	6	1
2	4	8	2	6	0	0	2	2
3	6	3	9	6	0	0	3	9
4	16	0	4	4	0	0	4	20
5	15	15	15	5	0	0	5	5
6	0	18	24	12	6	0	6	6
7	0	0	0	0	77	0	7	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>42</b>	<b>45</b>	<b>54</b>	<b>36</b>	<b>83</b>	<b>96</b>	<b>33</b>	<b>43</b>
<b>Γενική Κατάταξη</b>	<b>3,5</b>	<b>3,75</b>	<b>4,5</b>	<b>3,00</b>	<b>6,92</b>	<b>8</b>	<b>2,75</b>	<b>3,58</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Ημερήσιων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	2	0	0	6	1
2	6	0	1	1	0	0	2	2
3	2	0	0	0	0	0	4	6
4	4	1	1	5	1	0	0	0
5	0	1	6	2	0	0	0	3
6	0	6	2	2	2	0	0	0
7	0	3	0	0	9	0	0	0
8	0	0	0	0	0	12	0	0

Πίνακας Κατάταξης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Μηνιαίων Χρονοσειρών

Θέση	SES	Holt	Damped	ETS	Moving Average	LRL	Naive	Theta
1	0	1	2	2	0	0	6	1
2	12	0	2	2	0	0	4	4
3	6	0	0	0	0	0	12	18
4	16	4	4	20	4	0	0	0
5	0	5	30	10	0	0	0	15
6	0	36	12	12	12	0	0	0
7	0	21	0	0	63	0	0	0
8	0	0	0	0	0	96	0	0
<b>Σύνολο</b>	<b>34</b>	<b>67</b>	<b>50</b>	<b>46</b>	<b>79</b>	<b>96</b>	<b>22</b>	<b>38</b>
<b>Γενική Κατάταξη</b>	<b>2,83</b>	<b>5,58</b>	<b>4,17</b>	<b>3,83</b>	<b>6,58</b>	<b>8</b>	<b>1,83</b>	<b>3,17</b>

Πίνακας Βαθμολόγησης Μεθόδων με βάση το δείκτη MAsE – Σύνολο Μηνιαίων Χρονοσειρών

Σύμφωνα με τους παραπάνω πίνακες είναι εμφανές πως για τα ημερήσια δεδομένα σε σχέση με την προκατάληψη, οι μέθοδοι Damped και ακολούθως η Naive ανταποκρίνονται καλύτερα με βάση το μέτρο της απόλυτης τιμής των σφαλμάτων ME για τις σημειακές προβλέψεις της τρίτης περιόδου. Αντιθέτως, οι μέθοδοι LRL και κινητού μέσου όρου αποδείχτηκαν πως είναι οι χειρότερες μέθοδοι με μεγάλη διαφορά από τις υπόλοιπες μεθόδους. Όσον αφορά τις μηνιαίες χρονοσειρές, η μέθοδος που ανταποκρίνεται καλύτερα σε σχέση με την προκατάληψη είναι η ETS και ακολουθεί η Damped, ενώ η Naive βρίσκεται τώρα στην τρίτη θέση. Τελευταίες είναι και πάλι οι μέθοδοι LRL και κινητού μέσου όρου.

Τώρα η βαθμολογία με βάση τον δείκτη της ακρίβειας (MAE) για τα ημερήσια δεδομένα εμφάνισε την Naive ως την πλέον κατάλληλη μέθοδο πρόβλεψης. Στην δεύτερη θέση εμφανίζεται η μέθοδος ETS και στην τρίτη θέση η Theta. Οι μέθοδοι της απλής γραμμικής παλινδρόμησης και του κινητού μέσου όρου ακολουθούν στις τελευταίες θέσεις. Για τα μηνιαία δεδομένα υπάρχει μία διαφοροποίηση σε σχέση με πριν και αφορά την μέθοδο SES που κατατάσσεται στη δεύτερη θέση με βάση την βαθμολόγηση. Αντίστοιχα η μέθοδος ETS παρατηρήθηκε στην τέταρτη θέση της σχετικής βαθμολογίας. Οι υπόλοιπες μέθοδοι δεν διαφοροποιούνται σε σχέση με τα ημερήσια δεδομένα.

Στην περίπτωση των ποσοσטיαίων σφαλμάτων των προβλέψεων σε σχέση με τις πραγματικές τιμές, η βαθμολογία του δείκτη MAPE για τα ημερήσια δεδομένα, απέδειξε τη μέθοδο Naive ως την καλύτερη, με την μέθοδο Theta να ακολουθεί. Όσον αφορά τα μηνιαία δεδομένα πρώτη εμφανίζεται πάλι η Naive, αλλά εδώ ακολουθεί η μέθοδος SES και έπειτα η Theta.

Ακριβώς η ίδια διαπίστωση βγήκε για τον δείκτη sMAPE με την μόνη διαφορά πως στα μηνιαία δεδομένα την δεύτερη θέση καταλαμβάνει η Theta και στην τρίτη βρίσκεται η Naive. Στη τελευταία θέση παρατηρείται για ακόμη μία φορά η μέθοδος LRL και στην προτελευταία η μέθοδος κινητού μέσου όρου.

Ο δείκτης MAE επιβεβαίωσε και εδώ την κυριαρχία της μεθόδου Naive για τα ημερήσια δεδομένα, με την μέθοδο ETS να ακολουθεί. Στην περίπτωση των μηνιαίων χρονοσειρών πάλι η Naive εμφανίζεται πρώτη με την μέθοδο SES και Theta να ακολουθούν. Εδώ η μέθοδος ETS δεν ανταποκρίνεται τόσο καλά με πριν και για αυτό παρατηρείται στην τέταρτη θέση.



# 8. Συμπεράσματα-Προοπτικές

## 8.1 Συμπεράσματα

Η παρούσα διπλωματική χωρίζεται σε δύο στάδια. Το ένα αφορά στην υλοποίηση της εφαρμογής, όπου ο χρήστης μπορεί πολύ εύκολα να αναζητήσει και να επεξεργαστεί, με τους τρόπους που έχουν ήδη αναφερθεί σε προηγούμενο κεφάλαιο, ένα τεράστιο πλήθος χρονοσειρών. Οι οποίες χρονοσειρές μπορούν να αναπαρασταθούν είτε σε μορφή πίνακα, είτε σε διαγραμματική απεικόνιση. Επίσης, μπορούν να εξαχθούν προβλέψεις για την επόμενη χρονική περίοδο χρησιμοποιώντας αυτόματα τις οχτώ μεθόδους πρόβλεψης που έχουν ήδη αναφερθεί.

Η καινοτομία της εφαρμογής, έγκειται στον τρόπο άντλησης των δεδομένων μέσα από διάφορες βάσεις δεδομένων που διαθέτουν οι ιστότοποι από τους οποίους κάνουμε την συγκεκριμένη ενέργεια. Κάτι τέτοιο έχει σαν αποτέλεσμα την αυτόματη ανανέωση των δεδομένων της εφαρμογής, την ίδια στιγμή που γίνεται η ενημέρωση των δεδομένων του αντίστοιχου ιστότου.

Το παραπάνω βοήθησε στην υλοποίηση της ιδέας για την αυτόματη αξιολόγηση των μεθόδων πρόβλεψης, ανάλογα με τα σφάλματα που μελετάμε, τα οποία στην εφαρμογή είναι το μέσο σφάλμα, το μέσο απόλυτο ποσοστιαίο σφάλμα και το συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα. Όπως αναφέρθηκε η αξιολόγηση και η κατάταξη των μεθόδων γίνεται αυτόματα και ανανεώνεται συνεχώς με την προσθήκη καινούργιων δεδομένων. Η τεχνική που χρησιμοποιείται για την εξαγωγή της κατάταξης των μεθόδων είναι η τεχνική της κυλιόμενης πρόβλεψης (rolling forecasting) και η οποία έχει αναλυθεί περαιτέρω σε προηγούμενο κεφάλαιο.

Η χρησιμότητα της εφαρμογής είναι μεγάλη, ειδικά αν αναλογιστεί κανείς τα διαθέσιμα dataset που υπάρχουν. Αυτό διότι οι συγκεκριμένοι τομείς παρουσιάζουν μεγάλο ενδιαφέρον και είναι αρκετά δημοφιλείς στους ερευνητές, στους επιχειρηματίες και στους επενδυτές. Επίσης το σημαντικό είναι η άμεση ανατροφοδότηση των δεδομένων την στιγμή που γίνονται διαθέσιμα.

Το δεύτερο στάδιο της διπλωματικής αφορά στη περαιτέρω διερεύνηση της αξιολόγησης των μεθόδων πρόβλεψης στην κατηγορία των αγαθών. Η τεχνική που χρησιμοποιήθηκε και εδώ είναι η τεχνική της κυλιόμενης πρόβλεψης (rolling forecasting), όπου θεωρήσαμε το 80% της πραγματικής χρονοσειράς γνωστό και το υπόλοιπο 20% θεωρήθηκαν οι μελλοντικές τιμές της χρονοσειράς που ερχόταν κάθε χρονική περίοδο δίπλα από την πρόβλεψη της χρονοσειράς.

Βέβαια, στην αξιολόγηση των μεθόδων λήφθηκαν υπόψιν πέντε σφάλματα, σε αντίθεση με την εφαρμογή που λήφθηκαν τρία. Τα σφάλματα αυτά είναι το μέσο σφάλμα, το μέσο απόλυτο σφάλμα, το μέσο απόλυτο ποσοστιαίο σφάλμα, το συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα και το μέσο απόλυτο κανονικοποιημένο σφάλμα. Επίσης πραγματοποιήθηκαν οι προβλέψεις για την επόμενη περίοδο, αλλά επίσης και για τις τρεις επόμενες περιόδους, ώστε να δούμε πως ανταποκρίνονται οι μέθοδοι πρόβλεψης σε μεγαλύτερο ορίζοντα πρόβλεψης.

Από τα πειράματα που διενεργήθηκαν βγήκαν τα εξής συμπεράσματα:

- Γενικά η εφαρμογή των μεθόδων πρόβλεψης πάνω στις χρονοσειρές των αγαθών είτε των ημερήσιων δεδομένων, είτε των μηνιαίων είχαν πολύ καλή απόδοση, εκτός των μεθόδων της LRL και του κινητού μέσου όρου που ήταν σταθερά οι μέθοδοι με τα μεγαλύτερα σφάλματα.
- Όπως είναι λογικό τα σφάλματα των προβλέψεων για τις επόμενες τρεις περιόδους ήταν αυξημένα σχετικά με τις προβλέψεις της επόμενης περιόδου. Αυτό είναι συνηθισμένο όταν αυξάνουμε τον ορίζοντα πρόβλεψης μεγαλώνουν και τα σφάλματα.
- Ακόμη ένα συμπέρασμα αφορά την προκατάληψη των μεθόδων πρόβλεψης. Από τα μέσα σφάλματα που υπολογίστηκαν από τις προβλέψεις που παρήχθησαν με εφαρμογή των μεθόδων, προέκυψε πως υπάρχει έντονο το στοιχείο της αισιοδοξίας στις προβλέψεις, κατά ένα μεγάλο ποσοστό και σχεδόν σε όλες τις χρονοσειρές τιμών των αγαθών, ανεξαρτήτου χρονικού ορίζοντα. Η LRL διαφοροποιείται σχετικά με την παραπάνω διαπίστωση.
- Σύμφωνα με την παράγραφο 7.3, αν λάβουμε υπόψη τον δείκτη MAPE των σφαλμάτων των προβλέψεων κάθε μεθόδου για κάθε ορίζοντα πρόβλεψης, βγαίνει το συμπέρασμα πως τα αγαθά που ανταποκρίνονται καλύτερα στις προβλέψεις είναι για τις ημερήσιες χρονοσειρές του Κακάο (Cocoa), του βαμβακιού (Cotton) και στη συνέχεια ο χαλκός (Copper). Στον αντίποδα οι χρονοσειρές με τα μεγαλύτερα σφάλματα στις προβλέψεις τους είναι οι χρονοσειρές του Φυσικού Αερίου (Natural Gas) και ακολουθούν το σιτάρι (Wheat) και το ασήμι (Silver).  
Στην περίπτωση των μηνιαίων δεδομένων τα αγαθά που είχαν τα καλύτερα αποτελέσματα είναι οι τιμές της ποικιλίας καφέ Robusta και ακολουθούν οι τιμές του ρυζιού (Rice) και του κακάο (Cocoa). Σε αντίθεση βρίσκονται οι τιμές με τα μεγαλύτερα ποσοστιαία σφάλματα (MAPE), οι οποίες είναι οι τιμές του φυσικού αερίου (Natural Gas) και του χαλκού (Copper).
- Στην παράγραφο 7.4 έγινε μία καταγραφή των μεθόδων πρόβλεψης που έχουν τα μικρότερα ποσοστιαία σφάλματα ανά αγαθό. Αυτό έγινε διότι το ποσοστιαίο σφάλμα MAPE λειτουργεί σαν κριτήριο επιλογής της κατάλληλης μεθόδου για κάθε αγαθό. Από τους δύο συγκεντρωτικούς πίνακες, φάνηκε πως

η Naïve εμφανίζεται πιο συχνά από τις άλλες μεθόδους, ως η πλέον κατάλληλη μέθοδος για την παραγωγή προβλέψεων των αγαθών, είτε αυτό αφορά τις ημερήσιες ή τις μηνιαίες τιμές, είτε ακόμα και τον ορίζοντα πρόβλεψης.

- Συγκεκριμένα η Naïve για την παραγωγή προβλέψεων της επόμενης περιόδου ( $h=1$ ) ανεξαρτήτως του χρονικού ορίζοντα της χρονοσειράς, σε 24 ημερήσιες και μηνιαίες χρονοσειρές είχε το μικρότερο δείκτη σφάλματος MAPE συγκριτικά με όλες τις μεθόδους, 11 φορές, που πρακτικά σημαίνει στο 45,83% των χρονοσειρών.

Μετά την Naïve, η μέθοδος που είχε τα πιο μικρά σφάλματα MAPE είναι η ETS για τα ημερήσια και για τα μηνιαία δεδομένα 6 φορές που μεταφράζεται στο 25% των χρονοσειρών.

Κατόπιν είναι η Damped που εμφανίζεται με τα μικρότερα σφάλματα MAPE 3 φορές, δηλαδή στο 12,5% των χρονοσειρών.

Τέλος οι μέθοδοι Theta και Holt εμφανίζονται από 2 φορές, δηλαδή στο 8,33% των χρονοσειρών.

- Στην περίπτωση παραγωγής προβλέψεων της τρίτης περιόδου, η μέθοδος Naïve έχει και πάλι τα μικρότερα σφάλματα MAPE 13 φορές, συγκριτικά με τις υπόλοιπες μεθόδους, το οποίο μεταφράζεται στο 54,17% των συνολικών χρονοσειρών. Αυτό σημαίνει πως αυξάνοντας τον ορίζοντα πρόβλεψης η Naïve ανταποκρίνεται καλύτερα σε σχέση με τις υπόλοιπες μεθόδους σε αυξημένο ορίζοντα πρόβλεψης.

Στη συνέχεια η μέθοδος που ακολουθεί είναι η ETS, διότι έχει το μικρότερο MAPE σε σχέση με τις άλλες μεθόδους, 4 φορές, δηλαδή στο 16,67%. Αυτό σημαίνει πως η ETS αυξάνοντας τον ορίζοντα πρόβλεψης δεν ανταποκρίνεται τόσο καλά όπως προηγουμένως.

Επίσης φαίνεται πως δεν υπάρχει κάποια μεταβολή για τις μεθόδους Damped (12,5%), Theta (8,33%) και Holt (8,33%).

- Συνολικά ο μέσος όρος του ποσοστού του σφάλματος των MAPE κάθε μεθόδου για την παραγωγή προβλέψεων της πρώτης περιόδου, προκύπτει από το μέσο όρο των MAPE των 24 διαθέσιμων χρονοσειρών, ανεξαρτήτου χρονικού ορίζοντα. Επομένως προκύπτει ο εξής συγκεντρωτικός πίνακας:

1) ETS	3,1067%
2) Naïve	3,1174%
3) SES	3,1371%
4) Theta	3,1433%
5) Damped	3,1543%
6) Holt	3,2504%
7) Moving Average	4,5714%
8) LRL	29,5653%

Άρα με βάση τον μέσο όρο των σφαλμάτων MAPE, η μέθοδος με τον μικρότερο δείκτη είναι η ETS και δεύτερη η Naive. Στη τελευταία θέση βρίσκεται η LRL με συγκριτικά μεγάλη διαφορά ανάμεσα στις υπόλοιπες μεθόδους που είναι της τάξης του 30%. Γενικά οι διαφορές ανάμεσα στις μεθόδους είναι πολύ μικρές όπως φαίνεται στον παραπάνω πίνακα.

- Παρακάτω φαίνεται ο μέσος όρος του ποσοστού του σφάλματος των MAPE κάθε μεθόδου για την παραγωγή προβλέψεων της τρίτης περιόδου, προκύπτει από το μέσο όρο των MAPE των 24 διαθέσιμων χρονοσειρών, ανεξαρτήτου χρονικού οριζοντα. Επομένως προκύπτει ο εξής συγκεντρωτικός πίνακας με τον μέσο όρο των MAPE:

1) SES	6,5358%
2) Theta	6,5369%
3) Naive	6,5393%
4) ETS	6,7132%
5) Damped	6,8092%
6) Holt	7,1504%
7) Moving Average	7,4730%
8) LRL	29,7515%

Δηλαδή όπως φαίνεται οι μέθοδοι με τον μικρότερο μέσο όρο σφαλμάτων είναι η SES και δεύτερη η Theta, η οποία φαίνεται να βελτιώνει την ακρίβειά της συγκριτικά με τις υπόλοιπες μεθόδους, καθώς αυξάνεται ο οριζοντας πρόβλεψης. Δηλαδή υπάρχει μία διαφοροποίηση σε σχέση με πριν. Ειδικά η μέθοδος ETS τώρα έχει τον τέταρτο μικρότερο μέσο όρο σε σχέση με τις υπόλοιπες μεθόδους. Επίσης γεγονός αποτελεί ο διπλασιασμός των ποσοστών των σφαλμάτων κάθε μεθόδου, εκτός από την μέθοδο LRL που διατηρείται στα ίδια επίπεδα του 30%.

- Βέβαια τα παραπάνω αποτελούν την επιλογή της κατάλληλης μεθόδου ξεχωριστά για κάθε αγαθό και δεν αποτελούν κριτήριο για όλα τα αγαθά, διότι σε κάποιο αγαθό η Naive μπορεί να έχει το καλύτερο αποτέλεσμα, ενώ για κάποιο άλλο να εμφανίζεται με μεγαλύτερη τιμή σφαλμάτων από τις άλλες μεθόδους.

Για αυτό το λόγο έγινε η προσπάθεια κατάταξης και βαθμολόγησης των μεθόδων στο σύνολο των αγαθών, με διαχωρισμό του χρονικού οριζοντα των αγαθών, αλλά και του οριζοντα πρόβλεψης των αγαθών. Η κατάταξη και η βαθμολόγηση φαίνονται στην παράγραφο 7.5 και έγινε με βάση κάθε είδους σφάλματος.

Για την περίπτωση παραγωγής προβλέψεων της επόμενης περιόδου, είναι εμφανές πως οι μέθοδοι που ανταποκρίνονται καλύτερα με βάση την βαθμολόγηση και των πέντε σφαλμάτων είναι, πρώτη από όλες η μέθοδος

Ναΐνε και ακολουθεί η μέθοδος ETS. Εδώ η μέθοδος Theta συνολικά δεν παρουσιάζει πολύ καλά αποτελέσματα.

Αυτό βέβαια αλλάζει στην περίπτωση παραγωγής προβλέψεων για την τρίτη περίοδο όπου η αποτελεσματικότητα της μεθόδου Theta έχει ανέβει αισθητά και σε σχέση με τα σφάλματα καταλαμβάνει την δεύτερη και την τρίτη θέση, ανάλογα με το είδος του σφάλματος. Αντίστοιχα μία αισθητή μείωση της αποτελεσματικότητας εμφανίζεται στην περίπτωση της μεθόδου ETS, ενώ η μέθοδος SES έχει σταθερά καλά αποτελέσματα. Και στην περίπτωση αυτή η μέθοδος που βαθμολογείται καλύτερα είναι η Ναΐνε, τόσο για τα ημερήσια, όσο και για τα μηνιαία δεδομένα.

- Επίσης, από τα παραπάνω συμπεράσματα προκύπτει πως καλύτερη μέθοδος είναι η Ναΐνε. Επομένως το μέσο απόλυτο κανονικοποιημένο σφάλμα (MAE) αποκτά ιδιαίτερη σημασία και προτείνεται ως μέτρο σφάλματος. Πράγματι αυτό φαίνεται και από τα αποτελέσματα του σφάλματος MAE, τα οποία για τις μεθόδους της εκθετικής εξομάλυνσης, την Ναΐνε και την Theta παίρνουν τιμές κοντά στη μονάδα για την περίπτωση παραγωγής προβλέψεων της επόμενης περιόδου, κάτι το οποίο επιβεβαιώνει τον προηγούμενο ισχυρισμό. Πρακτικά αν το σφάλμα MAE μία μεθόδου είναι μικρότερο της μονάδας αυτό σημαίνει πως η μέθοδος αυτή είναι καταλληλότερη από την αφελή μέθοδο σύμφωνα με τους Hyndman και Koehler (2006). Από την άλλη για τιμές μεγαλύτερη της μονάδας η αφελής μέθοδος θεωρείται καλύτερη.
- Επίσης παρατηρώντας κανείς την κατάταξη των μεθόδων για τις μηνιαίες χρονοσειρές των τιμών του ρυζιού, του καλαμποκιού και τις σόγιας υπάρχουν πολλές ομοιότητες σχετικά με την κατάταξη των μεθόδων πρόβλεψης.

Γενικά για πιο μακροπρόθεσμες προβλέψεις θα πρέπει να λαμβάνονται υπόψιν διάφοροι παράγοντες που διαμορφώνουν και επηρεάζουν την τιμή των αγαθών. Αυτοί οι παράγοντες περιεγράφηκαν αναλυτικά στο 5<sup>ο</sup> κεφάλαιο της διπλωματικής. Αυτό που προτείνεται σε γενική βάση είναι να λαμβάνονται υπόψιν και οι μέθοδοι προβλέψεων για τις τιμές των αγαθών, αλλά επίσης και το οικονομικό περιβάλλον ή όποιο περιβάλλον δύναται να επηρεάσει την τιμή πώλησης ενός αγαθού μακροοικονομικά. Αυτό μπορεί να φανεί χρήσιμο στους επενδυτές που ασχολούνται με τις αγοραπωλησίες των εμπορευμάτων, είτε με τον κλασικό τρόπο αγοραπωλησίας, είτε μέσω των συμβολαίων μελλοντικής εκπλήρωσης (future contracts).

## 8.2 Προοπτικές

### 8.2.1 Προοπτικές Εφαρμογής

Αρχικά, θα αναφερθούμε στις προοπτικές και τις δυνατότητες της εφαρμογής οι οποίες είναι τεράστιες. Αυτό συμβαίνει διότι το περιβάλλον της  $R$  και συγκεκριμένα του  $r - shiny$  δίνει απεριόριστες δυνατότητες και στην στατιστική επεξεργασία των χρονοσειρών, αλλά και της διαδραστικότητας που προσφέρεται.

Πρώτα από όλα, η εφαρμογή έχει την δυνατότητα να εμπλουτιστεί με ορισμένα καινούργια χαρακτηριστικά. Αρχικά, θα μπορούσε να εμπλουτιστεί με περισσότερα dataset. Ο ιστότοπος Quandl, αλλά και η Παγκόσμια Τράπεζα διαθέτουν στη βάση τους πάρα πολλά δεδομένα, τα οποία τα έχουν διαθέσιμα μέσα από το API τους όπου γίνεται η άντληση των δεδομένων με τον τρόπο που έχει εξηγηθεί σε προηγούμενο κεφάλαιο. Για να γίνει αυτό αρκούσε να γνωρίζει ο χρήστης την κωδική ονομασία του dataset, την οποία μπορεί να δει μέσα από τον ιστότοπο. Άρα θα μπορούσε να υπήρχε σαν επιλογή να πληκτρολογεί ο χρήστης την κωδική ονομασία και να του εμφανίζει τις χρονοσειρές που θέλει. Το σημαντικό πλεονέκτημα αυτού είναι ότι διευρύνεται κατά πολύ το πλήθος των διαθέσιμων χρονοσειρών. Το μειονέκτημα είναι πως επαφίεται στον χρήστη η εύρεση της κωδικής ονομασίας κάτι το οποίο δεν ενδείκνυται στο μέσο χρήστη, από θέμα οργάνωσης.

Άλλη μία επέκταση σχετικά με την εφαρμογή είναι πως επηρεάζει και σχετίζεται ο αριθμός αναζητήσεων μίας λέξης την τιμή κάποιου αγαθού. Για παράδειγμα αν η αναζήτηση της πρότασης “Πόλεμος στο Ιράκ” - “War in Iraq” επηρεάζει ή αν υπάρχει κάποια συσχέτιση με την τιμή του αργού πετρελαίου. Για τον σκοπό αυτό δημιουργήθηκε το ανάλογο κομμάτι κώδικα που εξετάζει την συσχέτιση, χρησιμοποιώντας τον συντελεστή συσχέτισης του Pearson, ανάμεσα στον αριθμό αναζητήσεων μίας λέξης με την τιμή κλεισίματος μίας μετοχής, όπως φαίνεται και στην παρακάτω εικόνα, που έχουμε επιλέξει την τιμή της μετοχής της Apple από το 2004 (ημερομηνία ίδρυσης της Google) μέχρι και σήμερα:

Select a stock to examine. Information will be collected from yahoo finance.

**Symbol**

Κατόπιν επιλέξαμε τον αριθμό των αναζητήσεων με τη λέξη apple:

Select a keyword to examine. Information will be collected from Google Trends.

**Key Word**

Και βγάλαμε την εξής συσχέτιση:

## Correlation

**Check the type of correlation coefficients:**

- Pearson product-moment correlation coefficient
- Spearman's rank correlation coefficient (Spearman's rho)
- Kendall tau rank correlation coefficient (Kendall's tau)

[1] 0.5590598

Δηλαδή υπάρχει μέση γραμμική συσχέτιση σύμφωνα με τον συντελεστή του Pearson. Πάλι λόγω της πολυπλοκότητας αυτού του χαρακτηριστικού για τον μέσο χρήστη επιλέχθηκε να μην τοποθετηθεί το συγκεκριμένο κομμάτι κώδικα στην κύρια εφαρμογή.

Επίσης ακόμη ένα χαρακτηριστικό που θα μπορούσε να προστεθεί στην εφαρμογή είναι να επεκταθεί η αξιολόγηση των μεθόδων πρόβλεψης και για τα υπόλοιπα dataset με βάση την τεχνική της κυλιόμενης πρόβλεψης. Δυστυχώς όμως οι μέθοδοι Damped και ETS απαιτούν μεγάλη υπολογιστική ισχύ και κατά συνέπεια ο χρόνος εξαγωγής των προβλέψεων σε μία μεγάλη χρονοσειρά μπορεί να είναι πραγματικά τεράστιος. Για αυτό το λόγο το συγκεκριμένο χαρακτηριστικό δεν επεκτάθηκε και στα υπόλοιπα dataset.

Επίσης άλλη μία προσθήκη που θα μπορούσε να γίνει, αφορά τις μεθόδους πρόβλεψης. Υπάρχει η δυνατότητα προσθήκης και άλλων μεθόδων πρόβλεψης, ακόμα και συνδυασμός μεθόδων υλοποιώντας σε μορφή κώδικα την υπάρχουσα μεθοδολογία που αφορούν τις μεθόδους. Ακόμα ένα χαρακτηριστικό που προτιμήθηκε να μην προστεθεί αφορά τις ήδη υπάρχουσες μεθόδους πρόβλεψης και πιο συγκεκριμένα τις μεθόδους εξομάλυνσης SES, Holt, Damped και ETS. Όπως έχει ήδη αναφερθεί σε προηγούμενο κεφάλαιο η επιλογή των αρχικών συνθηκών της μεθόδου, όπως είναι το αρχικό επίπεδο, γίνεται με βάση την ελαχιστοποίηση του αθροίσματος των τετραγωνικών σφαλμάτων. Για λόγους πληρότητας θα μπορούσε να δοθεί η δυνατότητα στον χρήστη να επιλέγει εκείνος, για παράδειγμα το αρχικό επίπεδο, κάτι το οποίο θα τον βοηθούσε ερευνητικά. Αυτό όμως έρχεται σε σύγκρουση με το σκεπτικό δημιουργίας της εφαρμογής, το οποίο αναφέρεται στη χρησιμοποίησή του από έναν μέσο χρήστη, ο οποίος δεν χρειάζεται να έχει γνώσεις στατιστικής και τεχνικών προβλέψεων.

Σίγουρα ένα ακόμα στοιχείο το οποίο χρήζει βελτίωσης είναι το αισθητικό κομμάτι της εφαρμογής. Η βελτίωση αυτή μπορεί να γίνει με γνώσεις της γλώσσας HTML και πραγματικά μπορεί να κάνει πολύ ενδιαφέρουσα την εφαρμογή από οπτικής πλευράς. Κάτι τέτοιο θα βελτιώνε την διαδραστικότητα της εφαρμογής με τον χρήστη, διότι θα έκανε πιο φιλική την χρήση της.

Μία περαιτέρω ανάπτυξη της εφαρμογής θα μπορούσε να είναι η υλοποίηση της σε εφαρμογή για κινητά και tablet, δηλαδή να γίνει application και να δίνεται δωρεάν μέσα από τα play store κάθε λογισμικού. Αυτό που χρειάζεται είναι γνώσεις Android ή IOS, αναλόγως σε ποια πλατφόρμα γίνει η υλοποίηση του Application.

Συμπερασματικά, οι όποιες αλλαγές μπορούν να γίνουν που αφορούν τη φιλοσοφία της εφαρμογής, αλλά και οι όποιες προσθήκες και βελτιώσεις που μπορεί να υπάρξουν, σίγουρα μπορεί να οδηγήσει στη βελτίωση της χρήσης της εφαρμογής, καθώς επίσης και την επέκταση των δυνατοτήτων της ούτως ώστε να ανταποκρίνονται σε χρήστες κάθε τύπου.

## **8.2.2 Προοπτικές Ακρίβειας Μεθόδων**

Η πρόβλεψη των τιμών των εμπορευμάτων είναι ένα θέμα που χρίζει περαιτέρω ανάλυσης. Η έρευνα γύρω από τις τιμές των αγαθών αφορά όλο και περισσότερο κόσμο, διότι πλέον ενδιαφέρουν επενδυτικά περισσότερο κόσμο, αναφορικά με τις



τιμές του χρυσού, του ασημιού, κλπ. Επίσης αφορά τα οικονομικά σε παγκόσμιο επίπεδο, όσον αφορά τις τιμές των αγαθών που συνδέονται με την ενέργεια και τις μετακινήσεις, όπως είναι το αργό πετρέλαιο, το φυσικό αέριο, το κάρβουνο κλπ. Τα σημεία στα οποία θα μπορούσαν να επικεντρωθούν επόμενες μελέτες ή έρευνες αφορούν:

- Μελέτη απόδοσης κι άλλων μοντέλων ή μεθόδων πρόβλεψης για τα δεδομένα των τιμών των αγαθών.
- Παραγωγή Κριτικών Προβλέψεων και συνδυασμός τους με τα αποτελέσματα κλασσικών μεθόδων πρόβλεψης. Οι κριτικές προβλέψεις μπορούν να βασίζονται σε σχέση με τους παράγοντες διαμόρφωσης των τιμών των αγαθών που αναφέρθηκαν στο 5<sup>ο</sup> κεφάλαιο, όπως επίσης και το οικονομικό περιβάλλον γενικότερα.
- Πρόβλεψη των τιμών των αγαθών για εβδομαδιαία και για ετήσια δεδομένα.
- Χρησιμοποίηση διαφορετικών συντελεστών βαρύτητας στους συνδυασμούς των μεθόδων πρόβλεψης. Οι μέθοδοι με τα περισσότερο ακριβή αποτελέσματα θα μπορούσαν να έχουν μεγαλύτερους συντελεστές βαρύτητας σε σχέση με τις υπόλοιπες.
- Αναζήτηση και εφαρμογή προβλέψεων και σε άλλες χρονοσειρές αγαθών, όπως είναι το παλλάδιο, ο σίδηρος κλπ.
- Χρησιμοποίηση περισσότερων μεθόδων, όπως είναι τα *arima models*, τα νευρωνικά κλπ.
- Αξιολόγηση των μεθόδων με περισσότερους στατιστικούς δείκτες, όπως είναι το μέσο γεωμετρικό σφάλμα.



## 9. Βιβλιογραφία

1. Hyndman, Rob, et al. Forecasting with exponential smoothing: the state space approach. Springer Science & Business Media, 2008.
2. Hyndman, Rob J., and Yeasmin Khandakar. Automatic time series for forecasting: the forecast package for R. No. 6/07. Monash University, Department of Econometrics and Business Statistics, 2007.
3. Hyndman, Rob J., and Yanan Fan. "Sample quantiles in statistical packages." *The American Statistician* 50.4 (1996): 361-365.
4. Jeffrey Jarret Gutenberg, Μέθοδοι Πρόβλεψης
5. Makridakis, Spyros, Steven C. Wheelwright, and Rob J. Hyndman. Forecasting methods and applications. John Wiley & Sons, 2008.
6. Piesse, Jenifer, and Colin Thirtle. "Three bubbles and a panic: An explanatory review of recent food commodity price events." *Food policy* 34.2 (2009): 119-129.
7. Sargent, Thomas J. "Commodity price expectations and the interest rate." *The Quarterly Journal of Economics* (1969): 127-140.
8. Hyndman, Rob J., et al. "A state space framework for automatic forecasting using exponential smoothing methods." *International Journal of Forecasting* 18.3 (2002): 439-454.
9. Mackey, Michael C. "Commodity price fluctuations: price dependent delays and nonlinearities as explanatory factors." *Journal of Economic Theory* 48.2 (1989): 497-509.
10. Eric Zivot, Working with Financial Time Series Data in R, 2014
11. Frankel, Jeffrey A. "Expectations and commodity price dynamics: The overshooting model." *American Journal of Agricultural Economics* (1986): 344-348.
12. Cooper, Chris. *Tourism: Principles and practice*. Pearson education, 2005.
13. Claessens, Stijn, and Ronald C. Duncan. *Managing commodity price risk in developing countries*. Johns Hopkins University Press, 1993.
14. Hyndman, Rob J., and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2014.
15. Gandrud, Christopher. *Reproducible Research with R and R Studio*. CRC Press, 2013.
16. Newbery, David M., M. Scott, and D. Lal. *Commodity price stabilization*. Oxford University Press, 1990.

17. Hyndman, Rob J., and Anne B. Koehler. "Another look at measures of forecast accuracy." *International journal of forecasting* 22.4 (2006): 679-688.
18. Blomberg, S. Brock, and Ethan S. Harris. "The commodity-consumer price connection: fact or fable?." *Economic Policy Review* 1.3 (1995).
19. Just, Richard E., and Gordon C. Rausser. "Commodity price forecasting with large-scale econometric models and the futures market." *American Journal of Agricultural Economics* 63.2 (1981): 197-208.
20. Fildes R., Hibon M., Makridakis S., Meade N. , (1998), Generalising about univariate forecasting methods: further empirical evidence, *International journal of Forecasting* 14 339-358
21. Assimakopoulos, V. And Nikolopoulos ,N. (2000), The Theta modelQ a decomposition approach to forecasting, *International Journal of Forecasting*, Vol.16, No4, pp. 521-530
22. Πετρόπουλος Φ., Ασημακόπουλος Β., (2011), Επιχειρησιακές Προβλέψεις, Αθήνα
23. Crude Oil <<https://c1wsolutions.wordpress.com/2012/04/30/factors-affect-price-of-oil/>>
24. Silver <<http://tejas.iimb.ac.in/articles/80.php>>
25. Wikinvest.com <<http://www.wikinvest.com/commodity/Aluminum>>
26. U.S. Energy Information Administration (EIA) <[http://www.eia.gov/energyexplained/index.cfm?page=natural\\_gas\\_factors\\_affecting\\_prices](http://www.eia.gov/energyexplained/index.cfm?page=natural_gas_factors_affecting_prices)>
27. Scottsdale Bullion and Coin <<http://www.sbcgold.com/blog/10-factors-regularly-influence-gold-prices/>>
28. Futures Knowledge <<http://www.futuresknowledge.com/news-and-analysis/metals/the-top-factors-that-move-the-price-of-copper/>>
29. Investopedia.com<<http://www.investopedia.com/university/commodities/commodities15.asp>>
30. Commodityhq.com<<http://commodityhq.com/2012/four-little-known-factors-driving-the-price-of-platinum/>>
31. Futures Knowledge <<http://www.futuresknowledge.com/news-and-analysis/grains/the-top-factors-that-move-the-price-of-rough-rice/>>
32. Futures Knowledge <<http://www.futuresknowledge.com/news-and-analysis/grains/the-top-factors-that-move-the-price-wheat/>>
33. Futures Knowledge <<http://www.futuresknowledge.com/news-and-analysis/grains/the-top-factors-that-move-the-price-of-corn/>>
34. Futures Knowledge <<http://www.futuresknowledge.com/news-and-analysis/grains/the-top-factors-that-move-the-price-of-soybeans/>>
35. Quora.com <<http://www.quora.com/What-factors-affect-the-price-of-cocoa>>
36. Investopedia<<http://www.investopedia.com/university/commodities/commodities2.asp>>

37. Futures Knowledge <<http://www.futuresknowledge.com/news-and-analysis/softs/top-factors-that-move-the-price-of-cotton/>>
38. Wikipedia <[www.wikipedia.com](http://www.wikipedia.com)>
39. Stack Overflow <[www.stackoverflow.com/](http://www.stackoverflow.com/)>
40. Μαυροειδής Νικόλαος, Διπλωματική Εργασία (2012) Πρόβλεψη Ζήτησης Φυσικού Αερίου: Ανάλυση και Μοντελοποίηση
41. R-shiny <<http://shiny.rstudio.com/>>
42. Package forecast <<https://cran.r-project.org/web/packages/forecast/index.html>>
43. Package WDI <<https://cran.r-project.org/web/packages/WDI/index.html>>
44. Package Quandl <<https://cran.r-project.org/web/packages/Quandl/index.html>>
45. Package Quantmod <<https://cran.r-project.org/web/packages/quantmod/index.html>>
46. Package ggplot2 <<https://cran.r-project.org/web/packages/ggplot2/index.html>>
47. Yahoo Finance <<http://finance.yahoo.com/>>
48. Quandl <<https://www.quandl.com/>>
49. World Bank <<http://www.worldbank.org/>>