



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
Μονάδα Προβλέψεων & Στρατηγικής  
Forecasting & Strategy Unit

# *Τεχνικές Προβλέψεων*

## Προβλέψεις

# Το Μοντέλο $\Theta$ (1 από 3)

Η Μέθοδος  $\Theta$  : Μια μέθοδος παραγωγής προβλέψεων βασισμένη σε μια διαφορετική προσέγγιση της αποσύνθεσης

- Η μέθοδος  $\Theta$  (Assimakopoulos et. al. 1999, 2000) είναι μια μονοδιάστατη μέθοδος πρόβλεψης.
- Η μέθοδος βασίζεται στην μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς μέσα από την παράμετρο  $\theta$  (Theta), η οποία εφαρμόζεται απευθείας (πολλαπλασιαστικά) στις διαφορές δεύτερης τάξης των δεδομένων.
- Η καινούργια χρονοσειρά που δημιουργείται διατηρεί την μέση τιμή και κλίση (παλινδρόμησης) της αρχικής χρονοσειράς αλλά όχι και τις τοπικές καμπυλότητες. Οι χρονοσειρές που παράγονται με αυτή την διαδικασία ονομάζονται γραμμές  $\Theta$  (Theta Lines).
- Βασικό ποιοτικό χαρακτηριστικό αυτών των γραμμών είναι η καλύτερη προσέγγιση της μακροπρόθεσμης συμπεριφοράς-τάσης των δεδομένων ή ανάδειξη-τονισμός των βραχυπρόθεσμων χαρακτηριστικών, ανάλογα με την τιμή της παραμέτρου  $\theta$  ( $<, >1$ ).

# Το Μοντέλο $\Theta$ (2 από 3)

Η Μέθοδος  $\Theta$  : Μια μέθοδος παραγωγής προβλέψεων βασισμένη σε μια διαφορετική προσέγγιση της αποσύνθεσης

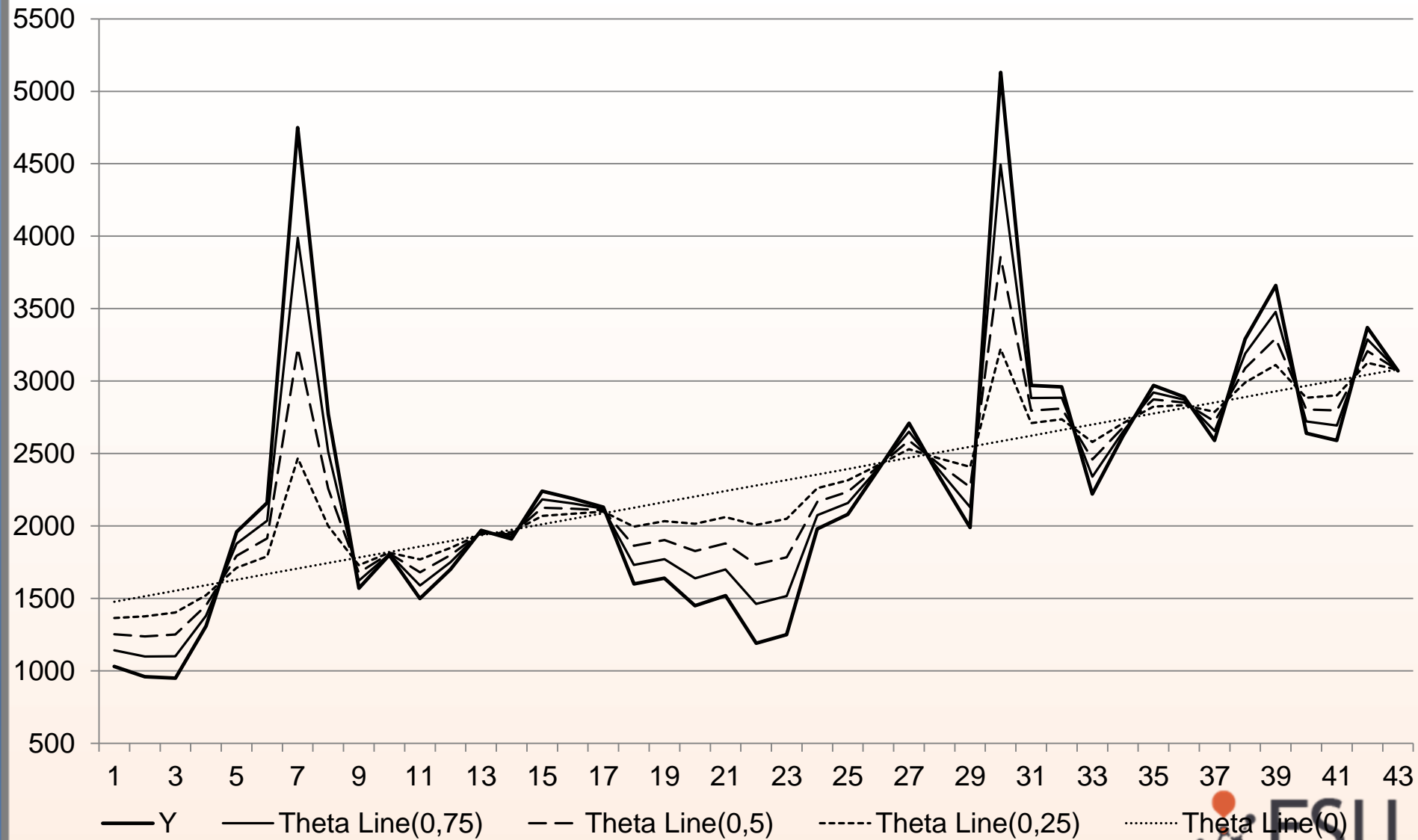
- Η προτεινόμενη μέθοδος αποσυνθέτει (διαχωρίζει) την αρχική χρονοσειρά σε δύο ή περισσότερες γραμμές Theta.
- Η κάθε γραμμή Theta προεκτείνεται στο μέλλον ξεχωριστά (με την ίδια ή και με διαφορετικές μεθόδους πρόβλεψης) και οι παραγόμενες προβλέψεις συνδυάζονται για να προκύψει η τελική πρόβλεψη.
- Ο απλός συνδυασμός δύο γραμμών Theta, για  $\theta=0$  (ευθεία γραμμή) και  $\theta=2$  (διπλασιασμός των τοπικών καμπυλοτήτων) χρησιμοποιήθηκε για την παραγωγή προβλέψεων για τις 3003 χρονοσειρές του διεθνούς διαγωνισμού προβλέψεων M3 (Makridakis et al., 2000).
- Η μέθοδος παρήγαγε πολύ καλά αποτελέσματα, με μικρά σφάλματα προβλεπτικής ακρίβειας, ιδιαίτερα για τις μηνιαίες χρονοσειρές και τα μικροοικονομικά δεδομένα.

# Το Μοντέλο $\Theta$ (3 από 3)

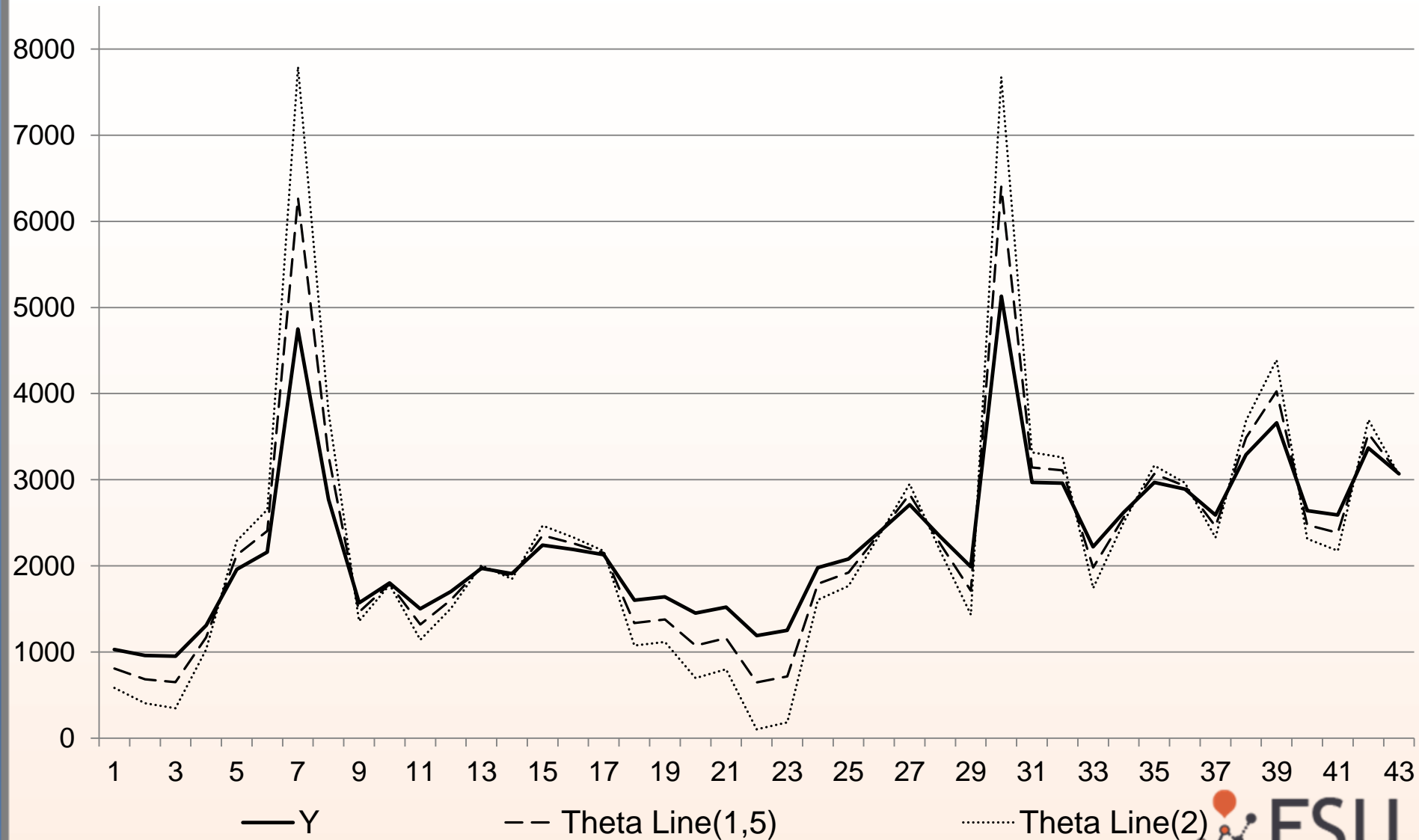
Η Μέθοδος  $\Theta$  : Μια μέθοδος παραγωγής προβλέψεων βασισμένη σε μια διαφορετική προσέγγιση της αποσύνθεσης

- Η μέθοδος Theta εισήγαγε μια διαφορετική προσέγγιση της αποσύνθεσης. Ο διαχωρισμός των αποεποχικοποιημένων δεδομένων γίνεται σε συνιστώσες (γραμμές Theta) μακροπρόθεσμης και βραχυπρόθεσμης τάσης.
- Η πρόκληση για τη συγκεκριμένη μέθοδο ήταν να αυξήσει το βαθμό αξιοποίησης της χρήσιμης πληροφορίας που είναι κρυμμένη μέσα στα δεδομένα, πριν την εφαρμογή ενός μοντέλου επέκτασης των δεδομένων στο μέλλον (extrapolation model).
- 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, μέσα από αυτήν την διαδικασία, πιο αποδοτικός.

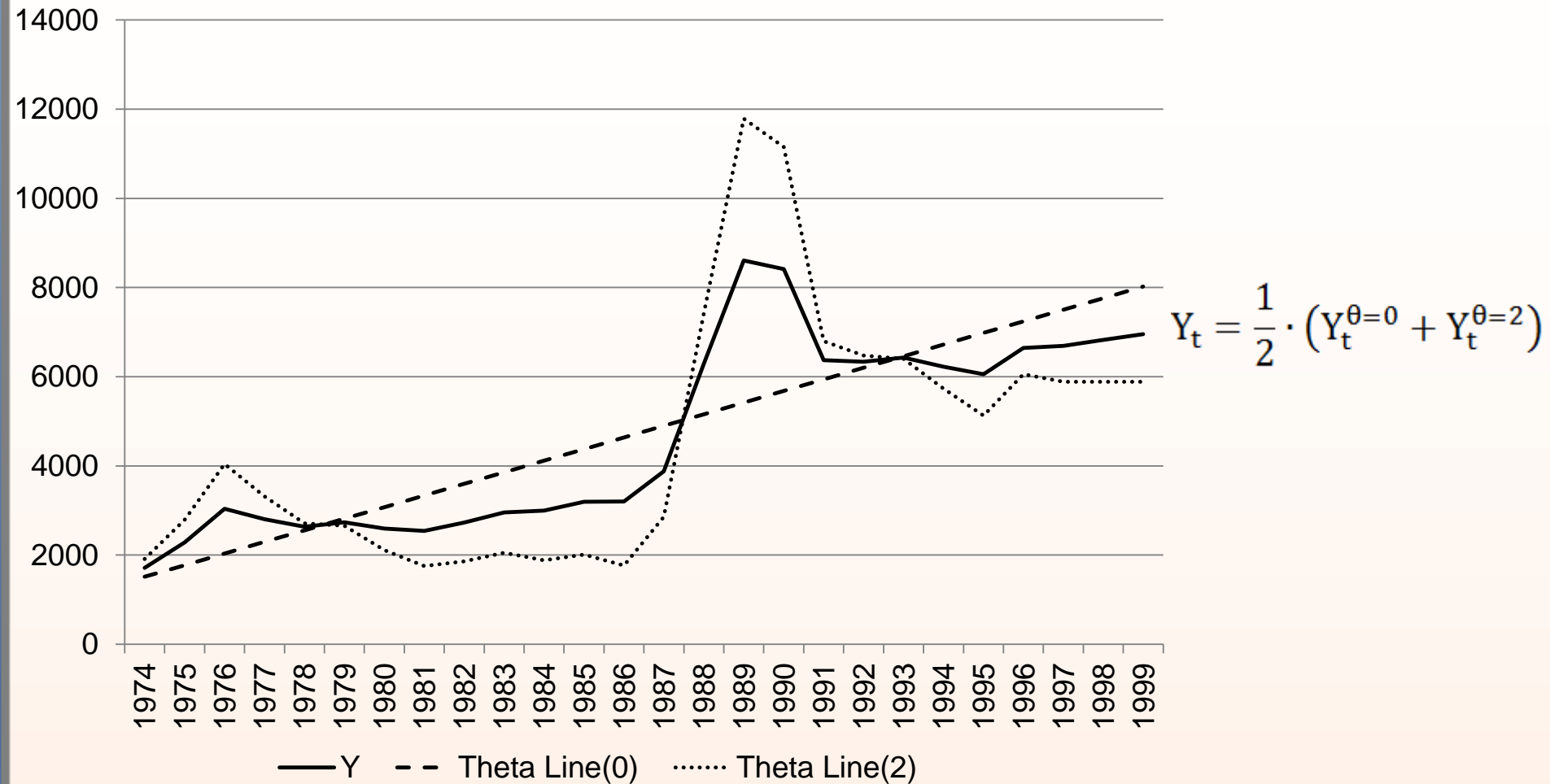
# Γραμμές $\Theta$



# Γραμμές Θ



# Το Κλασικό Μοντέλο $\Theta$



# Το κλασσικό Μοντέλο $\Theta$ στην πράξη

- **Βήμα 0. Τεστ Εποχιακότητας**

Η κάθε χρονοσειρά ελέγχεται για εποχιακή συμπεριφορά με κριτήριο την τιμή του συντελεστή αυτοσυσχέτισης με καθυστέρηση ένα έτος (π.χ. για μηνιαία δεδομένα 12) συγκρινόμενη με την τιμή 1.645 (τιμή της t-κατανομής για πιθανότητα 0.1)

- **Βήμα 1. Αποεποχικοποίηση**

Η χρονοσειρά αποεποχικοποιείται με την κλασσική μέθοδο αποσύνθεσης

- **Βήμα 2. Αποσύνθεση**

Κάθε χρονοσειρά αποσυντίθεται σε δύο γραμμές  $\Theta$ , για  $\theta=0$  και  $\theta=2$ .

- **Βήμα 3. Πρόβλεψη**

Η γραμμή  $\theta=0$  προεκτείνεται με απλή γραμμική παλινδρόμηση (LRL) ενώ η γραμμή  $\theta=2$  με απλή εκθετική εξομάλυνση (SES)

- **Βήμα 4. Συνδυασμός**

Οι προηγούμενες προβλέψεις συνδυάζονται με ίσα βάρη

- **Βήμα 5. Εποχικοποίηση**

Οι τελικές προβλέψεις εποχικοποιούνται



# Μοντέλο $\Theta$

Hyndman & Billah (2003):

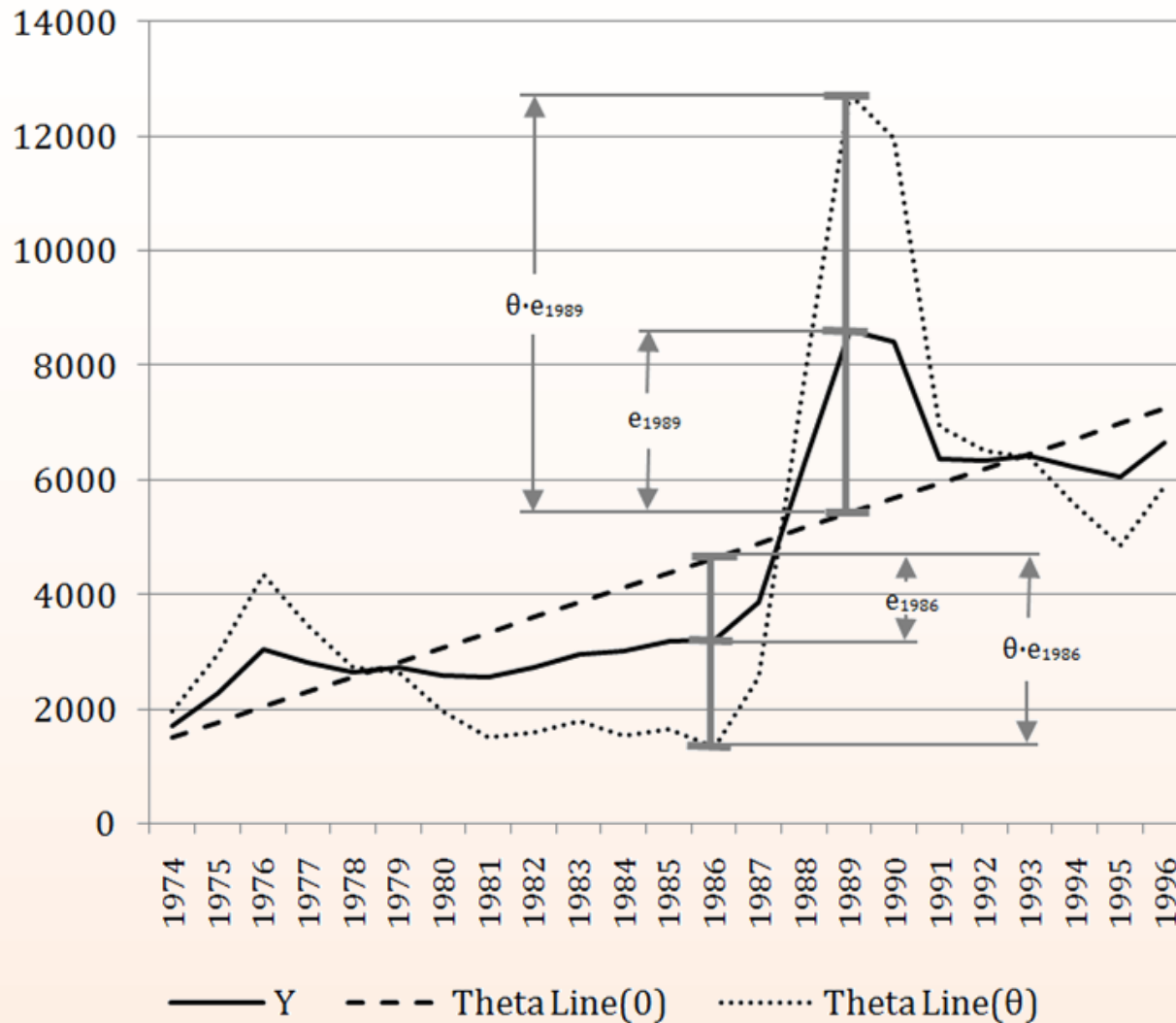
*“SES with drift?”*

# Υπολογίζοντας τις γραμμές Theta

- Για το κλασσικό μοντέλο Theta (παράμετροι 0 και 2) οι γραμμές Theta υπολογίζονται ως εξής:
  - $\text{Theta Line}(0) = \text{LRL}$
  - $\text{Theta Line}(2) = 2 \times \text{Data} - \text{LRL}$
- Γενικότερα ισχύει\*:
  - $\text{Theta Line}(\theta) = \theta \times \text{Data} + (1-\theta) \times \text{LRL}$
  - Ισοδύναμα:  $\text{Theta Line}(\theta) = \text{LRL} + \theta \times e_{\text{LRL}}$

\* Konstantinos Nikolopoulos, Vassilios Assimakopoulos, Nikolaos Bougioukos and Fotios Petropoulos (2008) “Advances in the Theta model”, *Working Paper*

# Υπολογίζοντας τις γραμμές Theta



# Παράδειγμα Theta (1 από 5)

Period	Data	TxC	SxRx100	I=SxR=S	TxCxR
1	4109			101,79	4036,68
2	3874			97,48	3974,13
3	3842	3955,00	97,14	98,54	3899,03
4	3946	3964,25	99,54	102,19	3861,43
5	4207	3984,75	105,58	101,79	4132,95
6	3850	4047,50	95,12	97,48	3949,51
7	4030	4085,00	98,65	98,54	4089,82
8	4260	4108,38	103,69	102,19	4168,71
9	4193	4145,50	101,15	101,79	4119,20
10	4051	4180,63	96,90	97,48	4155,70
11	4126	4222,63	97,71	98,54	4187,24
12	4445	4275,00	103,98	102,19	4349,74
13	4344	4364,13	99,54	101,79	4267,54
14	4319	4436,13	97,36	97,48	4430,63
15	4571	4496,88	101,65	98,54	4638,85
16	4576	4578,13	99,95	102,19	4477,93
17	4699	4620,25	101,70	101,79	4616,30
18	4614	4645,75	99,32	97,48	4733,25
19	4613			98,54	4681,47
20	4738			102,19	4636,46

					min	max	average (w/o min & max)	SI
	105,58	101,15	99,54	101,70	99,54	105,58	101,43	101,79
	95,12	96,90	97,36	99,32	95,12	99,32	97,13	97,48
97,14	98,65	97,71	101,65		97,14	101,65	98,18	98,54
99,54	103,69	103,98	99,95		99,54	103,98	101,82	102,19

sum	398,559
Σ.Κ.	0,9964

# Παράδειγμα Theta (2 από 5)

X	Y	Numerator			Denominator
Period	Data	X-Mean(X)=A	Y-Mean(Y)=B	A*B	(X-Mean(X))^2
1	4036,68	-9,5	-233,67	2219,87	90,25
2	3974,13	-8,5	-296,22	2517,88	72,25
3	3899,03	-7,5	-371,32	2784,92	56,25
4	3861,43	-6,5	-408,92	2657,95	42,25
5	4132,95	-5,5	-137,40	755,68	30,25
6	3949,51	-4,5	-320,84	1443,79	20,25
7	4089,82	-3,5	-180,53	631,86	12,25
8	4168,71	-2,5	-101,64	254,11	6,25
9	4119,20	-1,5	-151,15	226,72	2,25
10	4155,70	-0,5	-114,65	57,32	0,25
11	4187,24	0,5	-83,11	-41,55	0,25
12	4349,74	1,5	79,39	119,09	2,25
13	4267,54	2,5	-2,81	-7,02	6,25
14	4430,63	3,5	160,28	560,98	12,25
15	4638,85	4,5	368,50	1658,24	20,25
16	4477,93	5,5	207,58	1141,71	30,25
17	4616,30	6,5	345,95	2248,64	42,25
18	4733,25	7,5	462,90	3471,78	56,25
19	4681,47	8,5	411,12	3494,53	72,25
20	4636,46	9,5	366,11	3478,06	90,25
21					
22					
23					

b=slope 444.623  
a=constant 3801.8

ThetaLine(0)
LRL
3846.42
3891.05
3935.67
3980.29
4024.92
4069.54
4114.16
4158.79
4203.41
4248.03
4292.66
4337.28
4381.90
4426.53
4471.15
4515.77
4560.40
4605.02
4649.64
4694.27
4738.89
4783.51
4828.14

Average 10,5 4270,329

Sum 29674.52 665

# Παράδειγμα Theta (3 από 5)

Period	Data	ThetaLine(0)	ThetaLine(2)
1	4036,68	3846.42	4226.94
2	3974,13	3891.05	4057.21
3	3899,03	3935.67	3862.39
4	3861,43	3980.29	3742.57
5	4132,95	4024.92	4240.98
6	3949,51	4069.54	3829.48
7	4089,82	4114.16	4065.48
8	4168,71	4158.79	4178.63
9	4119,20	4203.41	4034.99
10	4155,70	4248.03	4063.37
11	4187,24	4292.66	4081.82
12	4349,74	4337.28	4362.20
13	4267,54	4381.90	4153.18
14	4430,63	4426.53	4434.73
15	4638,85	4471.15	4806.55
16	4477,93	4515.77	4440.09
17	4616,30	4560.40	4672.20
18	4733,25	4605.02	4861.48
19	4681,47	4649.64	4713.30
20	4636,46	4694.27	4578.65
21		4738.89	
22		4783.51	
23		4828.14	

SES on ThetaLine(2) with $\alpha=0.5$
4226.94
4131.81
4052.97
3976.00
3918.71
4025.83
3987.67
4038.75
4103.73
4111.46
4133.58
4160.41
4255.08
4261.31
4345.97
4492.41
4485.17
4550.73
4641.99
4661.73
4649.10
4649.10
4649.10

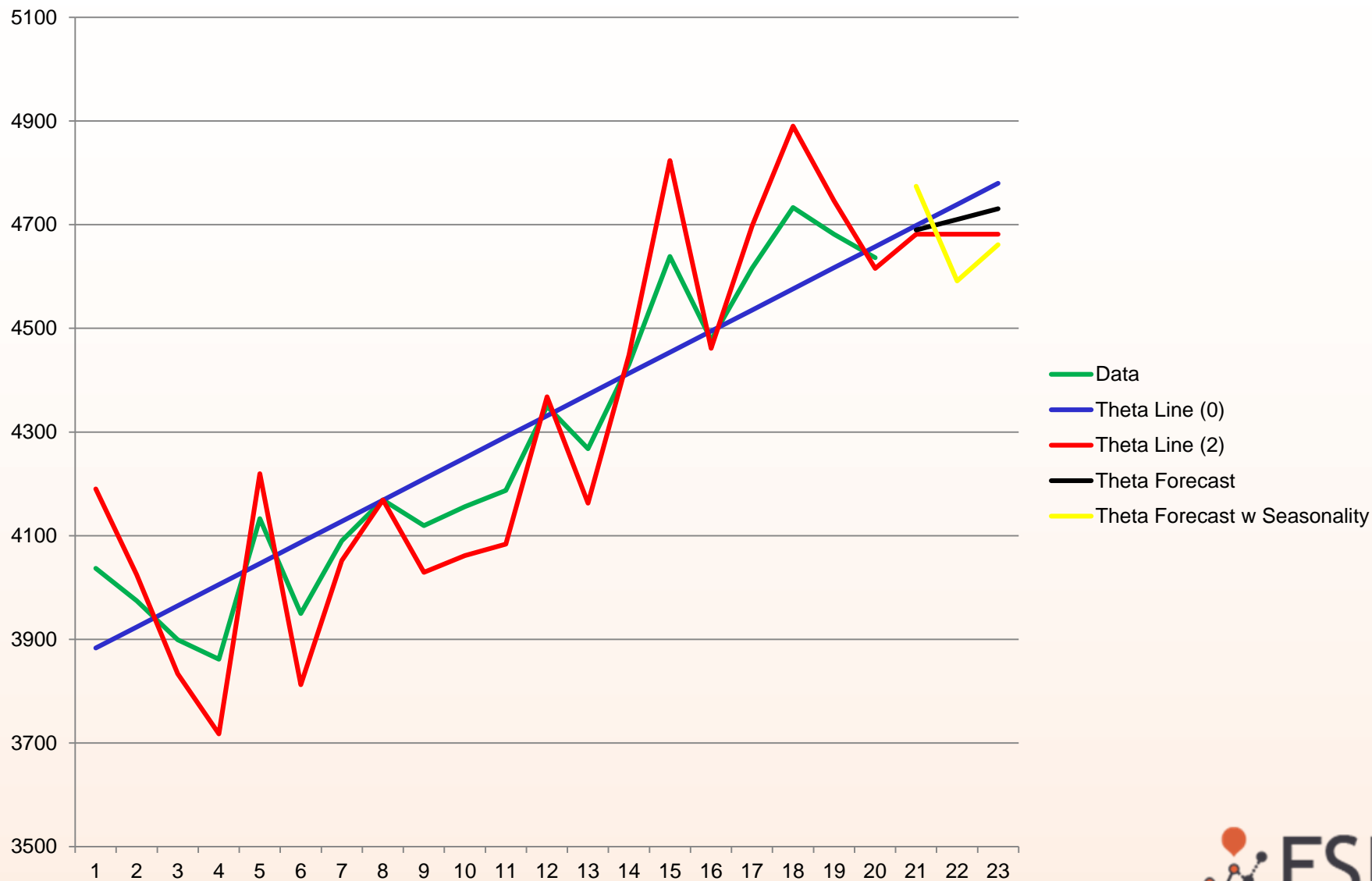
# Παράδειγμα Theta (4 από 5)

Period	Data	ThetaLine(0)	ThetaLine(2)
1	4036,679	3846.42	4226.937
2	3974,128	3891.05	4057.213
3	3899,027	3935.67	3862.39
4	3861,435	3980.29	3742.567
5	4132,955	4024.92	4240.983
6	3949,508	4069.54	3829.48
7	4089,817	4114.16	4065.477
8	4168,706	4158.79	4178.633
9	4119,201	4203.41	4034.99
10	4155,703	4248.03	4063.367
11	4187,242	4292.66	4081.823
12	4349,741	4337.28	4362.2
13	4267,543	4381.90	4153.177
14	4430,63	4426.53	4434.733
15	4638,848	4471.15	4806.55
16	4477,934	4515.77	4440.087
17	4616,295	4560.40	4672.203
18	4733,255	4605.02	4861.48
19	4681,471	4649.64	4713.297
20	4636,462	4694.27	4578.653
21		4738.89	4649.10
22		4783.51	4649.10
23		4828.14	4649.10

Theta Forecast	SI	Theta Forecast w Seasonality
----------------	----	------------------------------

4693.99	101,79	4778.02
4716.30	97,48	4597.45
4738.62	98,54	4669.43

# Παράδειγμα Theta (5 από 5)





# Διαγωνισμοί Πρόβλεψης

*Στόχοι των διαγωνισμών πρόβλεψης:*

- Δημιουργία ερεθίσματων στους ερευνητές για την υλοποίηση νέων μεθόδων πρόβλεψης
- Σύγκριση και ταξινόμηση των μεθόδων πρόβλεψης με κριτήριο την ελαχιστοποίηση του σφάλματος
- Έλεγχος της εγκυρότητας των αποτελεσμάτων προηγούμενων διαγωνισμών πρόβλεψης

# Διαγωνισμοί Πρόβλεψης

*Στόχοι των διαγωνισμών πρόβλεψης:*

- Δημιουργία ερεθίσματων στους ερευνητές για την υλοποίηση νέων μεθόδων πρόβλεψης
- Σύγκριση και ταξινόμηση των μεθόδων πρόβλεψης με κριτήριο την ελαχιστοποίηση του σφάλματος
- Έλεγχος της εγκυρότητας των αποτελεσμάτων προηγούμενων διαγωνισμών πρόβλεψης

# Διαγωνισμός M

## 1982

- Στατιστικά πολύπλοκες ή εξεζητημένες μέθοδοι δεν προσδίδουν απαραίτητα και πιο ακριβείς προβλέψεις από τις πιο απλές.
- Η σχετική κατάταξη της απόδοσης των διαφόρων μεθόδων ποικίλει ανάλογα με το κριτήριο ακρίβειας που χρησιμοποιείται.
- Η ακρίβεια μίας μεθόδου που είναι συνδυασμός απλών μεθόδων συνήθως είναι καλύτερη της ακρίβειας των μεθόδων αν αυτές εφαρμοσθούν ξεχωριστά.
- Η ακρίβεια των διαφόρων μεθόδων εξαρτάται από την έκταση του ορίζοντα πρόβλεψης που εφαρμόζεται.

# Διαγωνισμός Μ3 (2 από 2)

## Συνοπτικά Αποτελέσματα

Μέθοδος Πρόβλεψης	SMAPE των 1428 μηνιαίων χρονοσειρών	SMAPE στο σύνολο των χρονοσειρών (3003)
THETA	13.85	13.01
ForecastPro	13.86	13.19
ForcX	14.45	13.49
COMB S-H-D	14.48	13.52
DAMPEN	14.59	13.63
THETA <sub>sm</sub>	14.66	13.88
RBF	14.77	13.75
B-J automatic	14.81	14.01
AutomatANN	14.93	14.11
SMARTFCS	15.03	14.13

# Διαγωνισμός M4 (1 από 2)

## 2010+

	Datasets	Number of Timeseries	Category
Main Datasets	Business / Industry	1500	Static
	Finance	2000	Multiple Release
	Economics	2000	Static
	Inventory	1500	Multiple Release
	Internet & Telecommunications	1000	Static
	Demographic	1000	Static
	Climate	1000	Static
High Profile Datasets	Economics & Finance	17	Multiple Release (& Live)
	Fortune	35	Live
	GDP Pyramid I	18	Multiple Release (& Live)
	GDP Pyramid II	10	Multiple Release (& Live)
	Industry Pyramid	38	Static

# Διαγωνισμός M4 (2 από 2)

- Το τρίτο συστατικό: Στον διαγωνισμό M4 θα γίνει αξιολόγηση, εκτός των *benchmark* μεθόδων πρόβλεψης και των διαγωνιζομένων μεθόδων πρόβλεψης, των δημοσιευμένων προβλέψεων από ειδικούς (*Bloomberg*, *Fortune* κλπ.)
- Θα εφαρμοσθεί για πρώτη φορά **live forecasting**, όπου θα διαδραματίσουν σημαντικό ρόλο, πλέον των στατιστικών μεθόδων, οι πληροφορίες και τα *inside info* (*soft data*) που θα έχουν στη διάθεσή τους οι συμμετέχοντες.
- Αξιολόγηση στο κέρδος-ζημία, με την εισαγωγή δεικτών που αφορούν μέτρηση κόστους σε αποθέματα αλλά και αποτελέσματα συναλλαγών σε οικονομικές σειρές.
- Χρονοσειρές σε μορφή πυραμίδας με σκοπό την παραγωγή προβλέψεων σε κάθε επίπεδο αυτής και σύγκριση μεταξύ της ακρίβειας των προβλέψεων κάθε επιπέδου.
- Νέες κατηγορίες δεδομένων: *Internet*, Κλίμα, Διακοπτόμενη Ζήτηση.
- Οι διοργανωτές του διαγωνισμού έχουν θέσει ως βασικό όρο αυτού τη δυνατότητα επαλήθευσης των αποτελεσμάτων (**replicability**), οπότε οι διαγωνιζόμενοι θα πρέπει να υποβάλουν, πέραν των προβλέψεων, αναλυτικές πληροφορίες σχετικά με τις μεθόδους και τις αρχικοποιήσεις.

# Πειράματα Βελτιστοποίησης (1 από 2)

- Στόχος η ελαχιστοποίηση του σφάλματος της Μεθόδου Πρόβλεψης  $\Theta$  όταν αυτή εφαρμόζεται στα μηνιαία δεδομένα του Διαγωνισμού M3
- Δείκτες Εποχιακότητας, το πλήθος γραμμών  $\Theta$  που εφαρμόζονται στο μοντέλο και τροποποιήσεις επί των παραμέτρων της εξομάλυνσης
- Το σύνολο των 18 παρατηρήσεων, με το οποίο έγινε και η αξιολόγηση στο διαγωνισμό M3, θεωρείται άγνωστο κατά τη διάρκεια υπολογισμού των μοντέλων πρόβλεψης
- Σε κάθε φάση του πειράματος τροποποιείται μόνο μία παράμετρος, ενώ οι υπόλοιπες κρατώνται σταθερές και σε συμφωνία με το αρχικό μοντέλο της Μεθόδου  $\Theta$

# Πειράματα Βελτιστοποίησης (2 από 2)

- Στη συνέχεια γίνεται μία προσπάθεια ταυτόχρονης εφαρμογής όλων των επιμέρους αλλαγών που οδήγησαν σε βελτίωση του μέσου όρου του σφάλματος, έτσι ώστε να εξετασθεί και να αξιολογηθεί η τελική βελτιστοποίηση
- Επαλήθευση σε ένα dataset 20000+ χρονοσειρών



# Εποχιακότητα

Τροποποίηση της μεθόδου υπολογισμού των Δεικτών Εποχιακότητας

Μέθοδος Υπολογισμού Δεικτών Εποχιακότητας	SMAPE
Κλασσική Μέθοδος Αποσύνθεσης	13.85
James-Stein	13.79
Lemon-Krutchkoff	13.83
<b>Miller-Williams</b>	<b>13.78</b>

Επιμέρους Βελτιστοποίηση:  $(13.85 - 13.78) / 13.85 = 0.5\%$

# Γραμμές $\Theta$ (1 από 2)

Αυτοματοποιημένη μέθοδος υπολογισμού των βαρών των γραμμών  $\Theta(0)$  και  $\Theta(2)$  ούτως ώστε το μοντέλο της πρόβλεψης να προσαρμόζεται βέλτιστα σε ένα σύνολο από 12 παρατηρήσεις που θεωρούνται «κρυφές»

Διάστημα στο οποίο κυμαίνονται τα βάρη των γραμμών $\Theta(0)$ και $\Theta(2)$	SMAPE
50%-50%	13.85
<b>45%-55%</b>	<b>13.65</b>
40%-60%	13.70
35%-65%	13.83
30%-70%	14.00

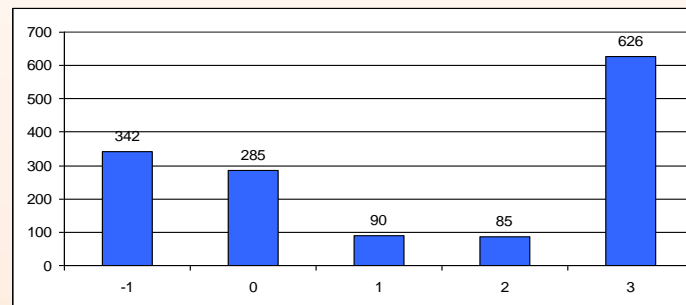
Επιμέρους Βελτιστοποίηση:  $(13.85 - 13.65) / 13.85 = 1.4\%$

# Γραμμές $\Theta$ (2 από 2)

Προστίθενται στο μοντέλο της πρόβλεψης κι άλλες γραμμές  $\Theta$ , πέρα των γραμμών με παραμέτρους  $\theta=0$  και  $\theta=2$ . Συγκεκριμένα ελέγχεται η συνεισφορά των γραμμών με παραμέτρους  $\theta=-1$ ,  $\theta=1$  και  $\theta=3$

Συνδυασμός Γραμμών $\Theta$	SMAPE
$50\% \times \text{Forecast}[L(0)] + 50\% \times \text{Forecast}[L(2)]$	13.85
$33\% \times \text{Forecast}[L(0)] + 33\% \times \text{Forecast}[L(2)] + 33\% \times \text{Forecast}[L(\theta_x)]$	14.34
$45\% \times \text{Forecast}[L(0)] + 45\% \times \text{Forecast}[L(2)] + 10\% \times \text{Forecast}[L(\theta_x)]$	13.71
$47.5\% \times \text{Forecast}[L(0)] + 47.5\% \times \text{Forecast}[L(2)] + 5\% \times \text{Forecast}[L(\theta_x)]$	13.70
$50\% \times \text{Forecast}[L(0)] + 30\% \times \text{Forecast}[L(2)] + 20\% \times \text{Forecast}[L(\theta_x)]$	13.74
<b><math>50\% \times \text{Forecast}[L(0)] + 40\% \times \text{Forecast}[L(2)] + 10\% \times \text{Forecast}[L(\theta_x)]</math></b>	<b>13.68</b>

Επιμέρους Βελτιστοποίηση:  
 $(13.85 - 13.67) / 13.85 = 1.3\%$



# Παράμετρος Εξομάλυνσης (1 από 2)

Αλλαγή των ορίων του διαστήματος επιλογής του συντελεστή εξομάλυνσης

Όρια διαστήματος Επιλογής του συντελεστή εξομάλυνσης	SMAPE
[0, 1]	13.85
[0.1, 1]	13.84
[0.2, 1]	13,89
[0, 0.9]	13.84
[0, 0.8]	13.87
<b>[0.1, 0.9]</b>	<b>13,82</b>

Επιμέρους Βελτιστοποίηση:  $(13.85 - 13.82) / 13.85 = 0.2\%$

# Εφαρμογή Τελικού Μοντέλου

*Εφαρμόστηκαν όλες οι προηγούμενες τροποποιήσεις ταυτόχρονα. Το μοντέλο που προκύπτει καλείται Βελτιστοποιημένο Μοντέλο Θ, και συγκρίνεται με τις καλύτερες μεθόδους πρόβλεψης που συμμετείχαν στο Διαγωνισμό M3*

Μέθοδος Πρόβλεψης	SMAPE	Κατάταξη στο Διαγωνισμό M3
Κλασσικό Μοντέλο Θ	13.85	1
ForecastPro	13.86	2
ForcX	14.45	3
<b>Βελτιστοποιημένο Μοντέλο Θ</b>	<b>13.57</b>	

Συνολική Βελτιστοποίηση:  $(13.85 - 13.57) / 13.85 = 2.024\%$

# Επαλήθευση (1 από 4)

- Over 20000 Timeseries where gathered:
  - o Forecasting Competitions
    - ✓ M-Competition
    - ✓ M2-Competition
    - ✓ M3-Competition
    - ✓ T-Competition
    - ✓ NN3-Competition
  - o FRED - Federal Reserve Bank of St.Liouis
  - o Hyndman, R.J. (n.d.) Time Series Data Library
  - o Collections from Books

# Επαλήθευση (2 από 4)

Holding out Value for the different types of timeseries

Yearly	6
Quarterly	8
Monthly	12
Weekly	12
Daily	14
Other	10

# Επαλήθευση (3 από 4)

Method	sMAPE (all timeseries)
SES	11.03%
Holt	14.79%
Damped	11.02%
Linear Trend	23.45%
Theta Classic	<b>9.64%</b>
Theta Optimized	9.77%



# Επαλήθευση (4 από 4)

Method	Yearly	Quarterly	Monthly	Weekly	Daily	Other
SES	25.40%	21.90%	6.44%	<b>10.83%</b>	19.68%	7.23%
Holt	36.76%	23.20%	8.86%	<b>10.78%</b>	<b>19.42%</b>	<b>5.25%</b>
Damped	28.76%	<b>14.11%</b>	6.78%	11.40%	20.63%	6.52%
Linear Trend	39.21%	74.97%	12.56%	43.42%	36.32%	11.19%
Theta Classic	<b>22.20%</b>	<b>13.93%</b>	<b>6.30%</b>	10.87%	19.76%	<b>5.91%</b>
Theta Optimized	<b>22.84%</b>	15.55%	<b>6.12%</b>	11.24%	<b>19.53%</b>	6.00%

# M4 Competition

100000 series και 18 έγκυρες υποβολές.

## Σημαντικά Ευρήματα

- 12 από τις 17 πιο ακριβείς μεθόδους ήταν «συνδυασμός στατιστικών» μεθόδων.
- Η πιο ακριβής μέθοδος (ακριβείς προβλέψεις και ακριβή διαστήματα προβλέψεων) ήταν μία «υβριδική» μέθοδος που χρησιμοποίησε στατιστικά και machine learning χαρακτηριστικά (Uber Technologies).
- Η δεύτερη πιο ακριβής μέθοδος ήταν συνδυασμός 7 στατιστικών μεθόδων και 1 machine learning, χρησιμοποιώντας βάση που υπολογίζονταν από Machine learning αλγόριθμο με ελαχιστοποίηση των σφαλμάτων στο hold-out tests.
- Η πιο ακριβής και η δεύτερη πιο ακριβής μέθοδος μείωσαν εξαιρετική επιτυχία στο σωστό υπολογισμό των 95% Prediction Intervals.
- Οι έξι αυτούσιες Machine Learning μέθοδοι στον M4 διαγωνισμό, σημείωσαν χειρότερη ακρίβεια από την Comb benchmark μέθοδο και μία μόνο ήταν πιο ακριβής από την Naïve 2.

Makridakis, S., Spiliotis, E. & Assimakopoulos, V. (2018). The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34(4), 802-808.