



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**  
**ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ**  
**ΤΟΜΕΑΣ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ ΚΑΙ ΠΕΡΙΒΑΛΛΟΝΤΟΣ**

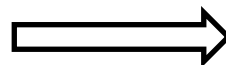
**Βραχυπρόθεσμη τοπική μετεωρολογική  
πρόγνωση με αναζήτηση ανάλογων καταστάσεων**

Γεώργιος Θεοδωρόπουλος  
Επιβλέπων καθηγητής Δημήτριος Κουτσογιάννης

Αθήνα 29 Οκτωβρίου 2012

# Εισαγωγή

Δημιουργία ενός μοντέλου

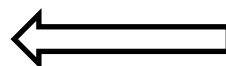


Στόχος

Βραχυπρόθεσμη τοπική  
μετεωρολογική πρόγνωση



Αναζήτηση Ανάλογων  
Καταστάσεων στην Ιστορική  
Χρονοσειρά



Δυο καταστάσεις που έχουν  
παρόμοιο παρελθόν θα  
εξελιχτούν με παρόμοιο τρόπο

- Το μοντέλο υλοποιήθηκε με την δημιουργία ενός αλγορίθμου
- Μετεωρολογικός σταθμός του Εθνικού Μετσόβιου Πολυτεχνείου (βρίσκεται στους πρόποδες του Υμηττού)

Μετεωρολογικές Μεταβλητές

- 1) Θερμοκρασία
- 2) Υγρασία
- 3) Ταχύτητα Ανέμου
- 4) Βροχόπτωση

Χρονικά βήματα:

- 1) Δέκα λεπτών
- 2) Μίας Ώρας
- 3) Μίας ημέρας

Χρονικός ορίζοντας πρόβλεψης τα δέκα βήματα

# Χρονοσειρές

- Βασική απαίτηση η ύπαρξη αξιόπιστων επεξεργασμένων ιστορικών χρονοσειρών
- Περίοδος λειτουργίας του σταθμού από το 1993
- Δυο περίοδοι λειτουργίας:
  - 1<sup>η</sup> Ένα όργανο μέτρησης για κάθε μετεωρολογική μεταβλητή
  - 2<sup>η</sup> Ύπαρξη και εφεδρικού οργάνου μέτρησης
- Προβλήματα πρωτογενών χρονοσειρών:
  - Μη σταθερό χρονικό βήμα ( $\approx 10$  min)
  - Πιθανώς λανθασμένη τιμή μέτρησης
  - Μη ενιαία χρονοσειρά για όλη την εξεταζόμενη περίοδο
  - Έλλειψη χρονοσειρών για διαφορετικό χρονικό βήμα



# Επεξεργασία Πρωτογενών Χρονοσειρών

- **1<sup>ο</sup> βήμα** Μετατροπή χρονοσειρών από χρονικά ακανόνιστες σε αυστηρά κανονικοποιημένες  
Ακανόνιστο χρονικό βήμα  $\approx 10$  min  $\longrightarrow$  Αυστηρό χρονικό βήμα 10 min
- **2<sup>ο</sup> βήμα** Έλεγχος χρονικής συνέπειας  
Διαφορά δύο διαδοχικών μετρήσεων < Τιμή ελέγχου
- **3<sup>ο</sup> βήμα** Έλεγχος εύρους
- Τιμή ελέγχου και τα όρια ελέγχου του εύρους
  - Από την περιοχή που βρίσκεται ο σταθμός
  - Την μετεωρολογική μεταβλητή
- Επιλέχτηκε να μην χρησιμοποιηθούν οι μετρήσεις της περιόδου 1993 με 1996 λόγω πολλών λανθασμένων τιμών

# Ενοποίηση Δευτερογενών Χρονοσειρών

- Στόχος η δημιουργία τεσσάρων χρονοσειρών, μία για κάθε μετεωρολογική μεταβλητή
- Η κάθε χρονοσειρα προήρθε από ενοποιήσει των δευτερογενών χρονοσειρών
- Για την ενοποίηση έγινε έλεγχος με χρήση παλινδρόμηση μεταξύ της κοινής περιόδου λειτουργίας των οργάνων
- Θεωρήθηκαν ως πιο σωστές οι πιο πρόσφατες χρονοσειρές

Γραμμική παλινδρόμηση χρησιμοποιήθηκε:

- Θερμοκρασία
- Υγρασία

$$Y = aX + b \quad \alpha = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2} \quad b = \bar{y} - a\bar{x}$$

Παλινδρόμηση με εξίσωση δύναμης χρησιμοποιήθηκε :

- Ταχύτητα Ανέμου
- Βροχόπτωση

$$Y = bX^a \quad \xrightarrow{\text{Τροποποίηση}} \quad \ln(Y) = \ln(b) + a\ln(X)$$
$$\alpha = \frac{\sum(\ln(x) - \overline{\ln(x)})(\ln(y) - \overline{\ln(y)})}{\sum(\ln(x) - \overline{\ln(x)})^2} \quad b = e^{\overline{\ln(y)} - a\overline{\ln(x)}}$$

# Συνάθροιση χρονοσειρών

- Δημιουργία ωριαίας και ημερήσιας χρονοσειράς
- Έλεγχος κατανομής σφάλματος για τον καθορισμό μέγιστου αριθμού ελλειπουσών τιμών
- Δεν έγινε έλεγχος στην μετεωρολογική μεταβλητή της βροχόπτωσης

## Ωριαία συνάθροιση:

- Έλεγχος για 1 έως 5 ελλείπουσες τιμές σε μια ώρα (τυχαία επιλογή)
- Επιλέχτηκε ως μέγιστος αριθμός ελλειπουσών τιμών οι τέσσερις

## Ημερήσια Συνάθροιση:

### Ελέγχθηκαν τρεις τρόποι:

- Απευθείας συνάθροιση από ωριαία σε ημερήσια
- Χρήση τριών μετρήσεων με συντελεστές βαρύτητας (Θερμοκρασία, Υγρασία)
- **Ενδιάμεση εξάωρη συνάθροιση**

# Μεθοδολογία Αναλόγων (1)

- Βασική αρχή της μεθόδου: Η δυναμική ενός συστήματος αποτυπώνεται σε μια χρονοσειρά  $\longrightarrow$  αρκεί να βρεθούν ανάλογες παρόμοιες καταστάσεις.
- Χρήση της μέσης τιμής ώστε να εξαχθεί η τιμή πρόβλεψης

Βασικά ζητήματα που εξεταστήκαν:

1. Καθορισμός παρούσας κατάστασης
2. Καθορισμός της αποδέκτης χρονικής απόστασης
3. Πλήθος αναλόγων τιμών

## Καθορισμός Διανύσματος Παρούσας Κατάσταση (ΔΠΚ):

- Γραμμική ΔΠΚ =  $(\chi_n, \chi_{n-1}, \dots, \chi_{n-(m-1)})$   $m=5, 8, 12$
- Λογαριθμική ΔΠΚ =  $(\chi_{n+1-2^0}, \chi_{n+1-2^1}, \dots, \chi_{n+1-2^m})$   $m=4, 5, 6$

## Απόρριψη Μετρήσεων

- Με βάση την τιμή μέτρησης  $|\chi_\lambda - \chi_n| \leq \varepsilon + \max\{\Delta ΠΚ\} - \min\{\Delta ΠΚ\}$
- Με βάση τη χρονική απόσταση

# Μεθοδολογία Αναλόγων (2)

## Εύρεση ανάλογων καταστάσεων:

- Απόσταση  
(Ευκλείδεια Νόρμα)

$$A_k = \sqrt{(x_v - y_k)^2 + (x_{v-1} - y_{k-1})^2 + \dots + (x_{v-(m-1)} - y_{k-(m-1)})^2}$$

- Μέγιστη Διαφορά (Μέγιστη Νόρμα)

$$A_k = \max_{1 \leq i \leq m} |x_{v-(i-1)} - y_{k-(i-1)}|$$

- Χρήση μέσου όρου τιμών για την εξαγωγή πρόβλεψης με πλήθος 10, 14, 20

## Πολυμεταβλητό μοντέλο

- Εξαγωγή πρόβλεψης εξετάζοντας και τις τέσσερις μετεωρολογικές μεταβλητές παράλληλα
- Για την δημιουργία ‘οδηγός’ ήταν η θερμοκρασία
- Δημιουργήθηκαν τέσσερις λίστες  $A_k$  μία για κάθε μεταβλητή
- Κανονικοποίηση των αποστάσεων

$$A_{ki}^{sum} = A_{ki}^T + A_{ki}^H + A_{ki}^{WS} + A_{ki}^P$$

## Δημιουργία Περιβάλλουσας

- Για χρονικό βήμα μίας ώρας και μίας ημέρας δημιουργήθηκε περιβάλλουσα από την μέγιστη και ελάχιστη τιμή



# Αλγόριθμος Μοντέλου

- Η δημιουργία του αλγορίθμου έγινε με χρήση της γλώσσας προγραμματισμού Python
- Βασική συνθήκη για τη λειτουργία του αλγορίθμου ήταν η ύπαρξη σταθερού χρονικού βήματος
- Ο αλγόριθμος δημιουργούσε δυο χρονοσειρές, μια ιστορική και μία μελλοντική
- Γινόταν έλεγχος για να μην υπάρχουν κενές τιμές κατά την δημιουργία του ΔΠΚ και στις πραγματικές μελλοντικές μετρήσεις
- Στα χρονικά βήματα του ενός δεκαλέπτου και της μίας ώρας δημιουργήθηκαν 5000 διαφορετικά ζευγάρια ιστορικών – μελλοντικών χρονοσειρών, ενώ στο ημερήσιο χρονικό βήμα 1000

Αλγόριθμος για την ενημέρωση του κοινού:

- Χρήση της βιβλιοθήκης Pthelma
- Γινόταν κανονικοποίηση στις πρωτογενείς χρονοσειρές
- Εφεδρικό σενάριο που χρειαζόταν λιγότερα στοιχεία για τη δημιουργία του ΔΠΚ

# Σενάρια και Μέθοδοι

- Χρονικό βήμα ενός Δεκαλέπτου
  - 3 διαφορετικές τιμές στο πλήθος των στοιχείων του ΔΠΚ
  - 2 χρονικές αποστάσεις
  - 3 διαφορετικές τιμές στο πλήθος των αναλόγων καταστάσεων
  - Συνολικά 18 σενάρια για κάθε μέθοδο
- Χρονικά βήματα μίας ώρας και μίας ημέρας
  - 3 διαφορετικές τιμές στα στοιχεία του ΔΠΚ
  - 3 χρονικές αποστάσεις
  - 3 διαφορετικές τιμές στο πλήθος των αναλόγων καταστάσεων
  - Συνολικά 27 σενάρια για κάθε μέθοδο
- Μέθοδοι:
  - 2 μοντέλα Απλό και Πολυμεταβλητό
  - 3 μέθοδοι συνδυασμού δημιουργίας ΔΠΚ και αναζήτησης των αναλόγων καταστάσεων
  - Συνολικά 6 διαφορετικές μέθοδοι για κάθε μετεωρολογική μεταβλητή
- Συνολικά εξετάστηκαν 108 διαφορετικοί συνδυασμοί για το χρονικό βήμα δέκα λεπτών και 162 διαφορετικοί συνδυασμοί τόσο για το χρονικό βήμα μίας ώρας όσο και για μίας ημέρας

# Επεξεργασία Αποτελεσμάτων

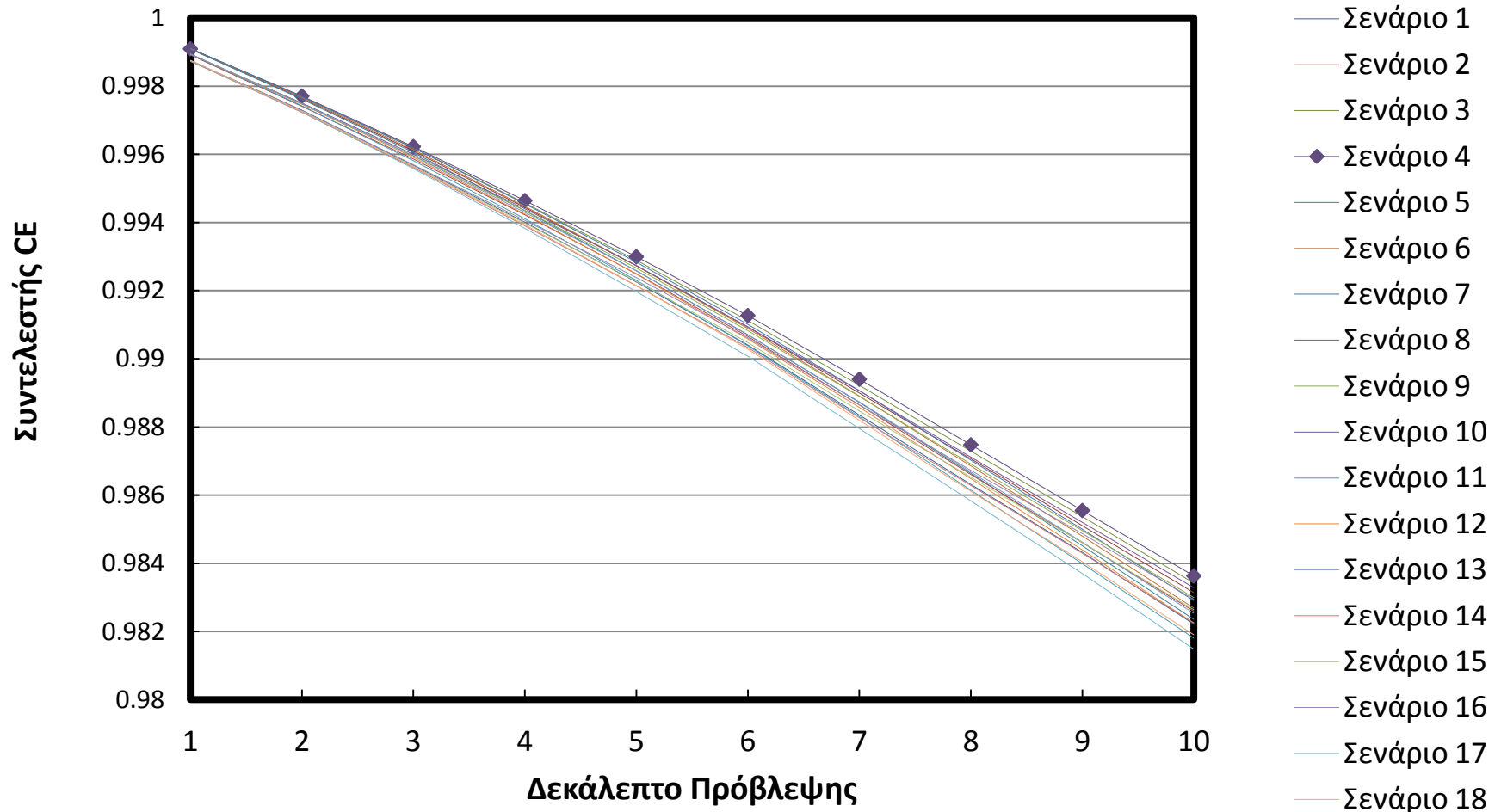
Χρησιμοποιήθηκαν τρεις στατιστικοί συντελεστές:

- Συντελεστής Αποδοτικότητας CE (Nash and Sutcliffe, 1970)
- RMSE
- BIAS
- Στόχος η δημιουργία απλού μοντέλου που θα χρησιμοποιεί λίγα στοιχεία για την μετεωρολογική πρόγνωση

Σύγκριση του μοντέλου με δυο απλά μοντέλα

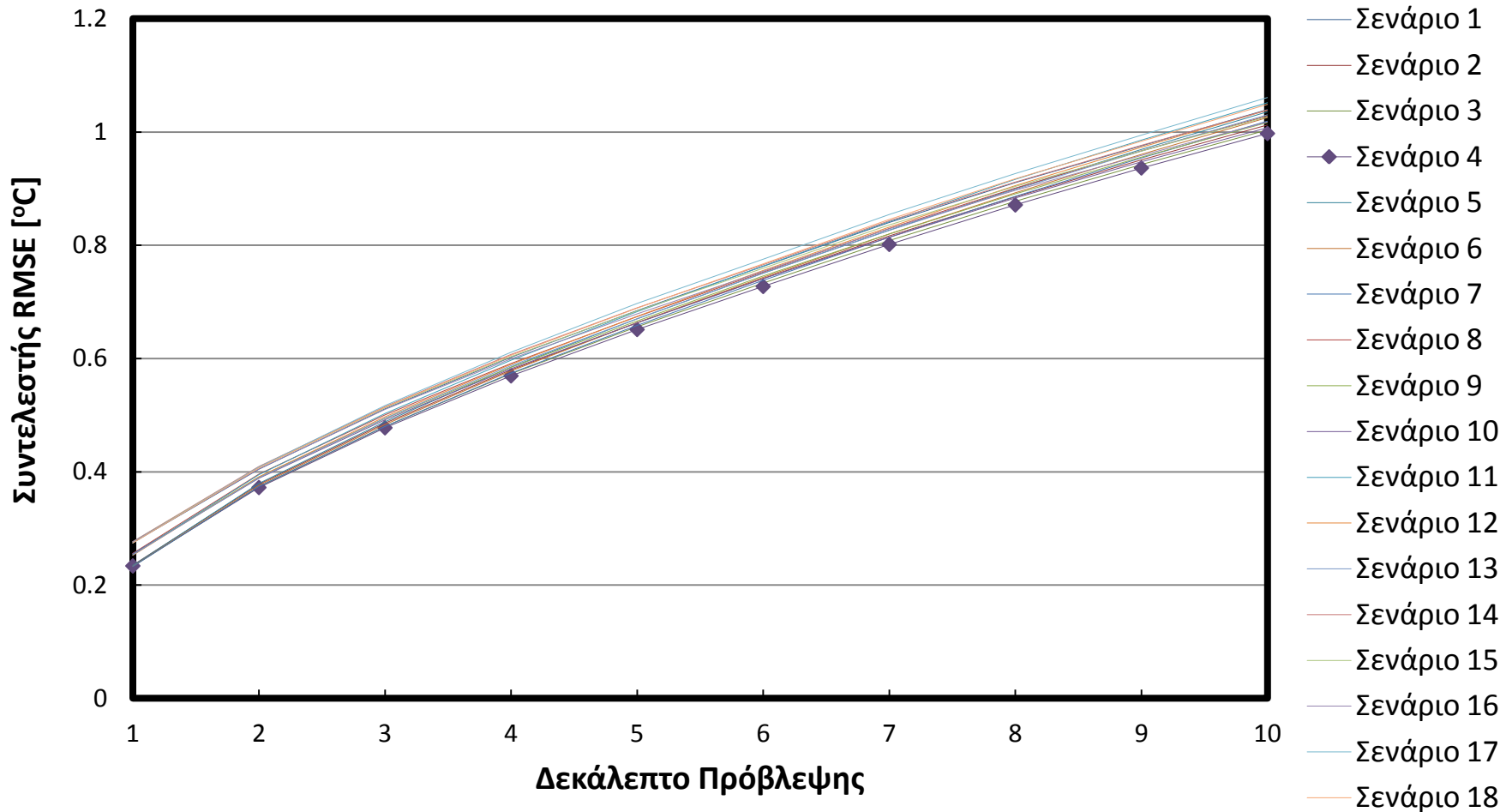
- Benchmark 1
- Benchmark 2

# Επιλογή Σεναρίου (1)



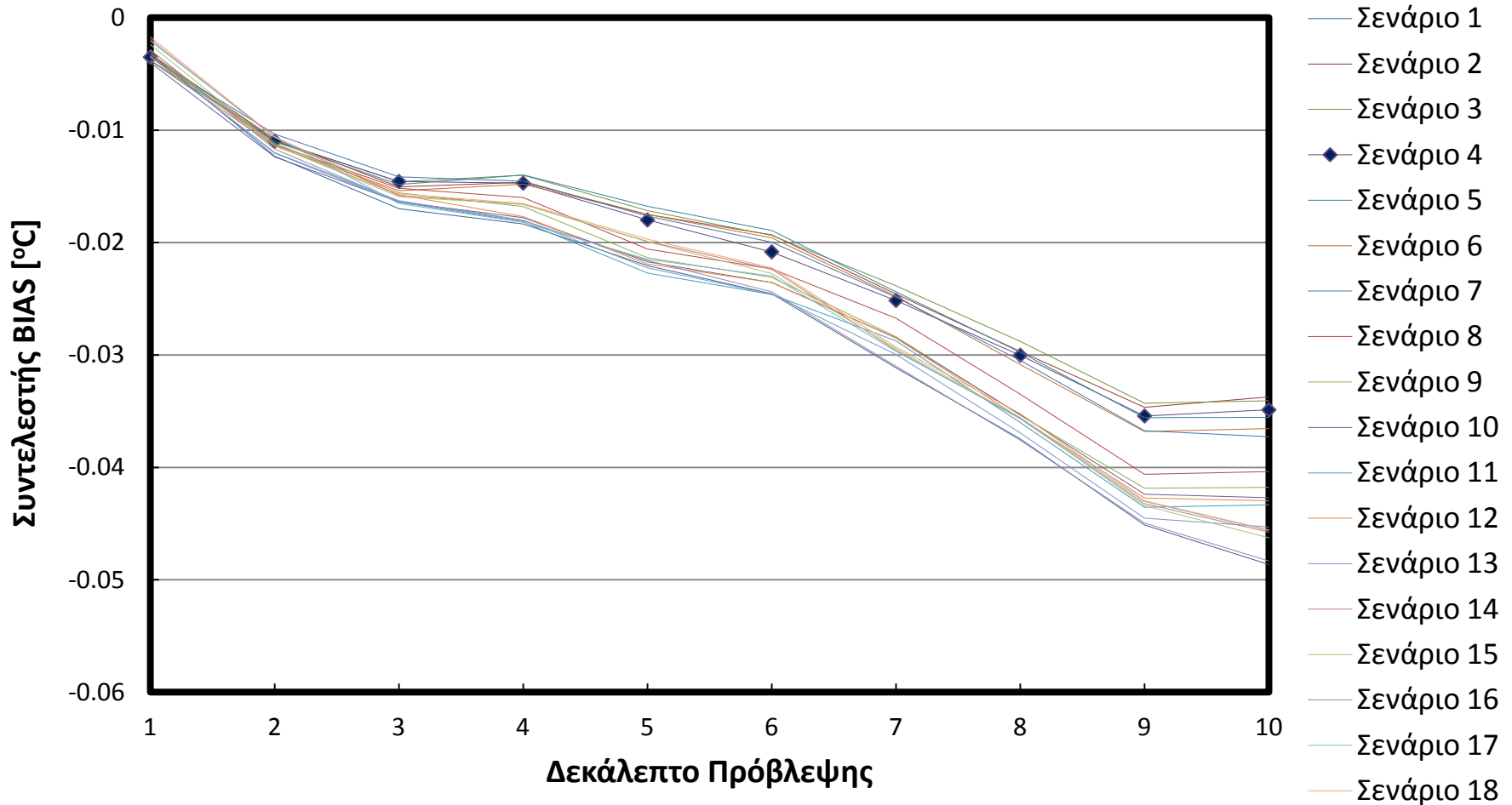
Διάγραμμα συντελεστή CE για τα σενάρια της μεθόδου με μία μεταβλητή και λογαριθμικό ΔΠΚ για την μετεωρολογική μεταβλητή της θερμοκρασίας και για χρονικό βήμα 10 λεπτών

# Επιλογή Σεναρίου (2)



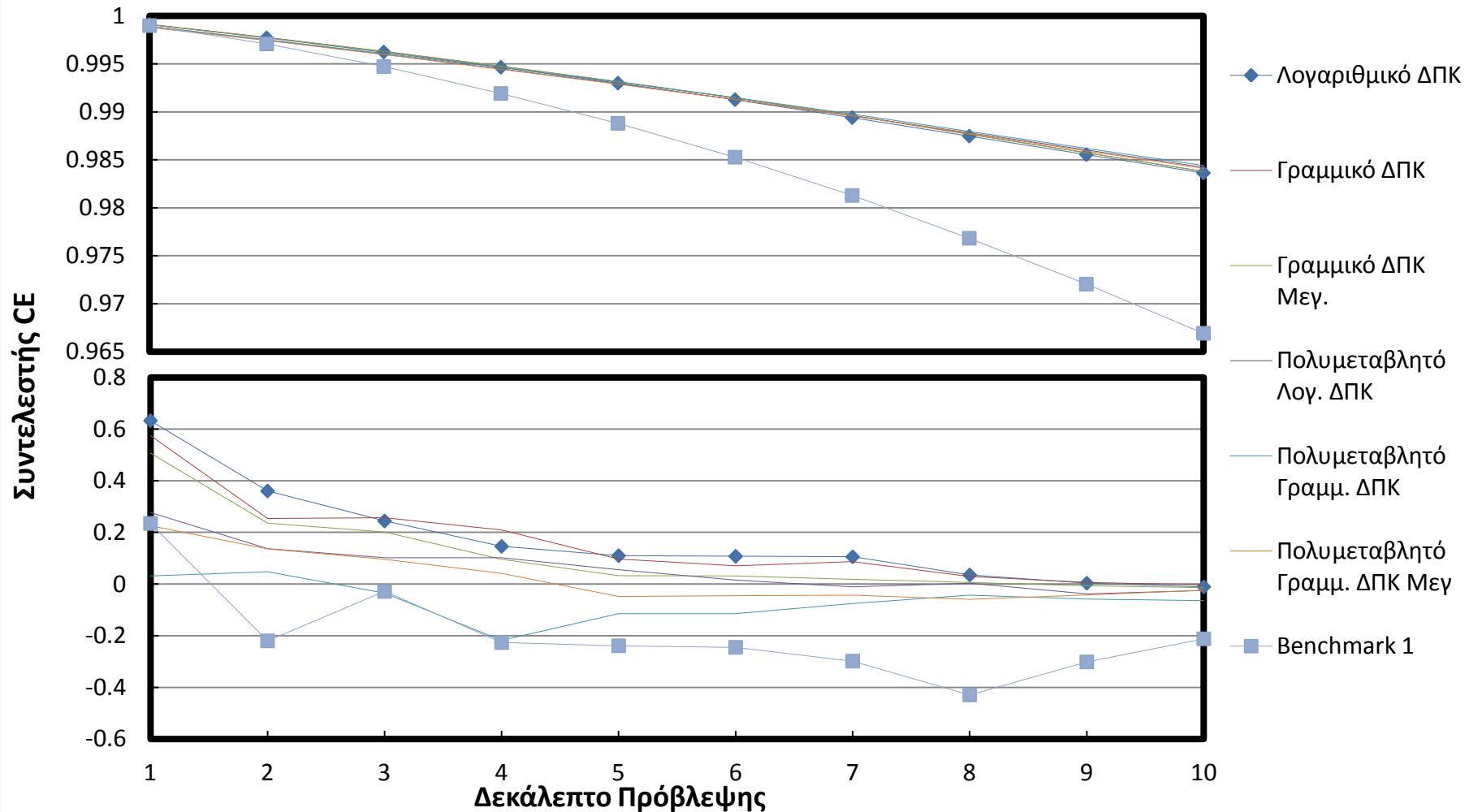
Διάγραμμα συντελεστή RMSE για τα σενάρια της μεθόδου με μία μεταβλητή και λογαριθμικό ΔΠΚ για την μετεωρολογική μεταβλητή της θερμοκρασίας και για χρονικό βήμα 10 λεπτών

# Επιλογή Σεναρίου (3)



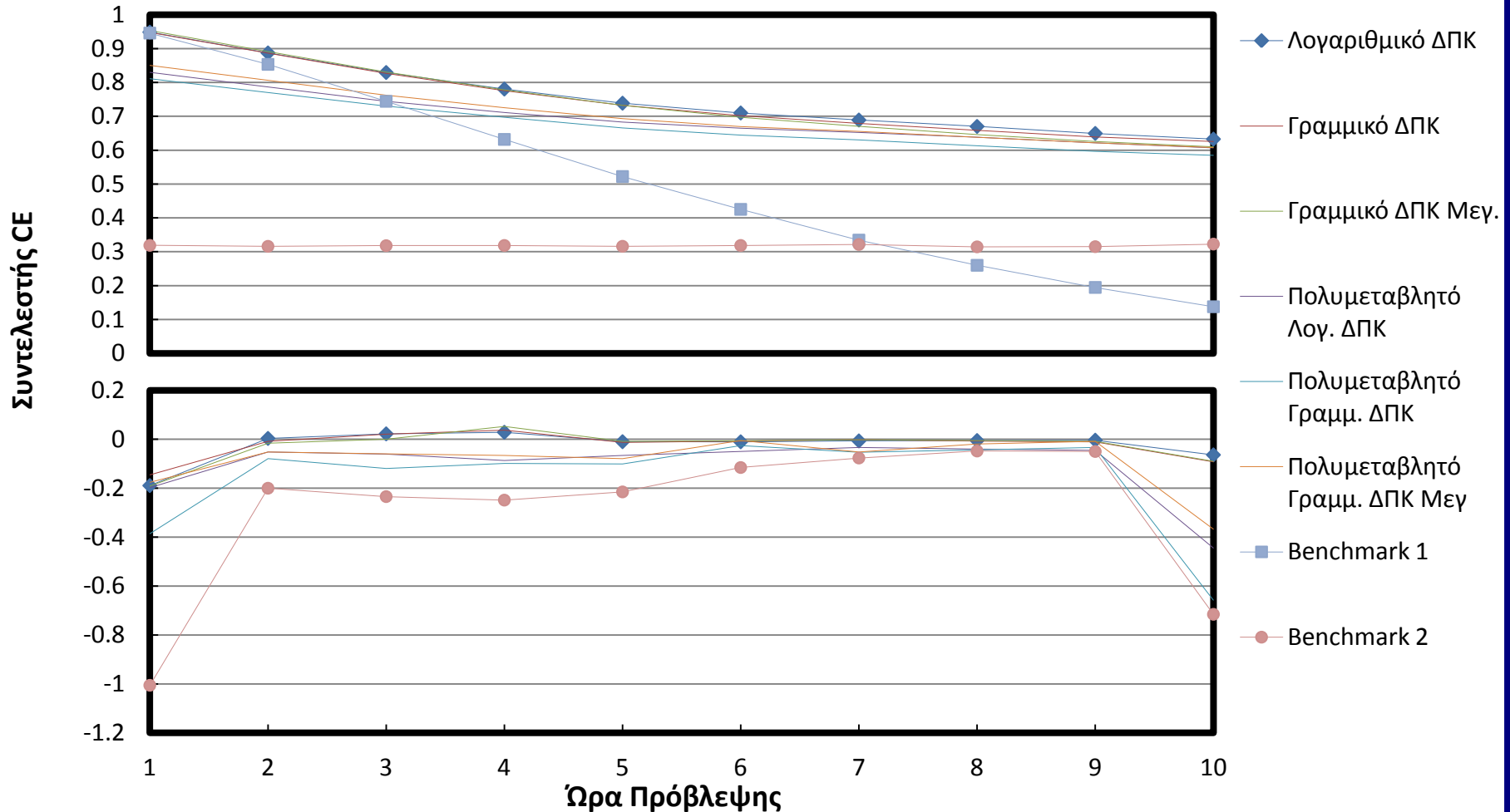
Διάγραμμα συντελεστή BIAS για τα σενάρια της μεθόδου με μία μεταβλητή και λογαριθμικό ΔΠΚ για την μετεωρολογική μεταβλητή της θερμοκρασίας και για χρονικό βήμα 10 λεπτών

# Επιλογή Μεθόδου (Χρονικό Βήμα Δέκα Λεπτών)



Διάγραμμα συντελεστή CE των διαφορετικών μεθόδων για τις μετεωρολογικές μεταβλητές της θερμοκρασίας (πάνω), βροχόπτωσης (κάτω) και για χρονικό βήμα δέκα λεπτών

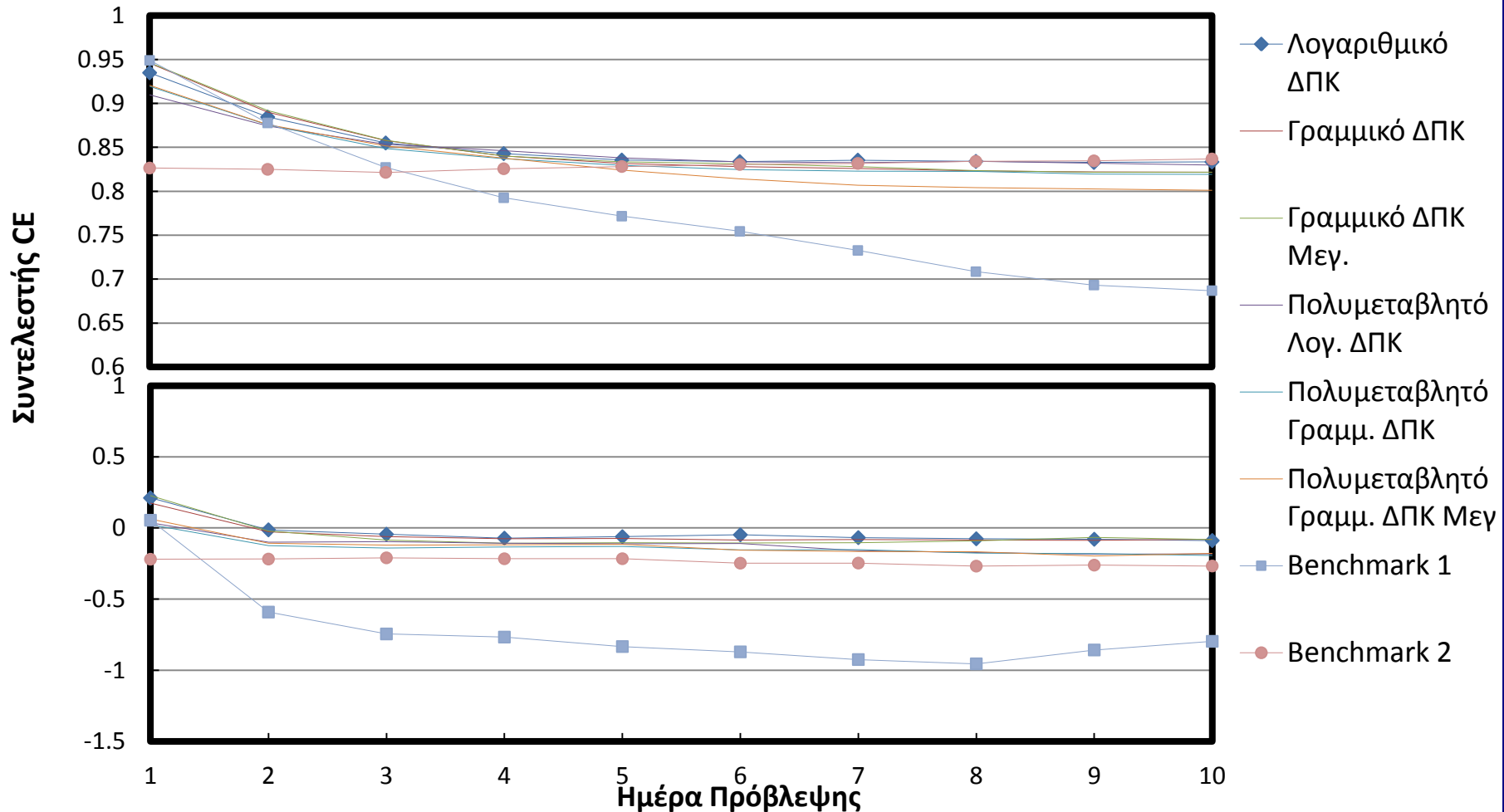
# Επιλογή Μεθόδου (Χρονικό Βήμα Μίας Ώρας)



Διάγραμμα συντελεστή CE των διαφορετικών μεθόδων για τις μετεωρολογικές μεταβλητές της υγρασίας (πάνω), βροχόπτωσης (κάτω) και για χρονικό βήμα μίας ώρας



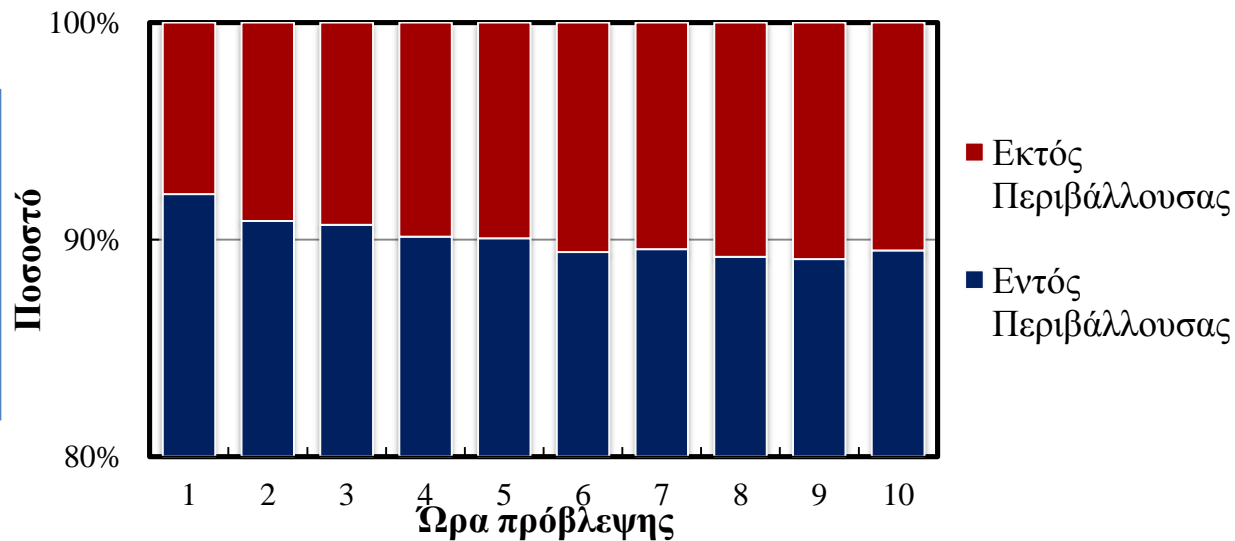
# Επιλογή Μεθόδου (Χρονικό Βήμα Μίας Ημέρας)



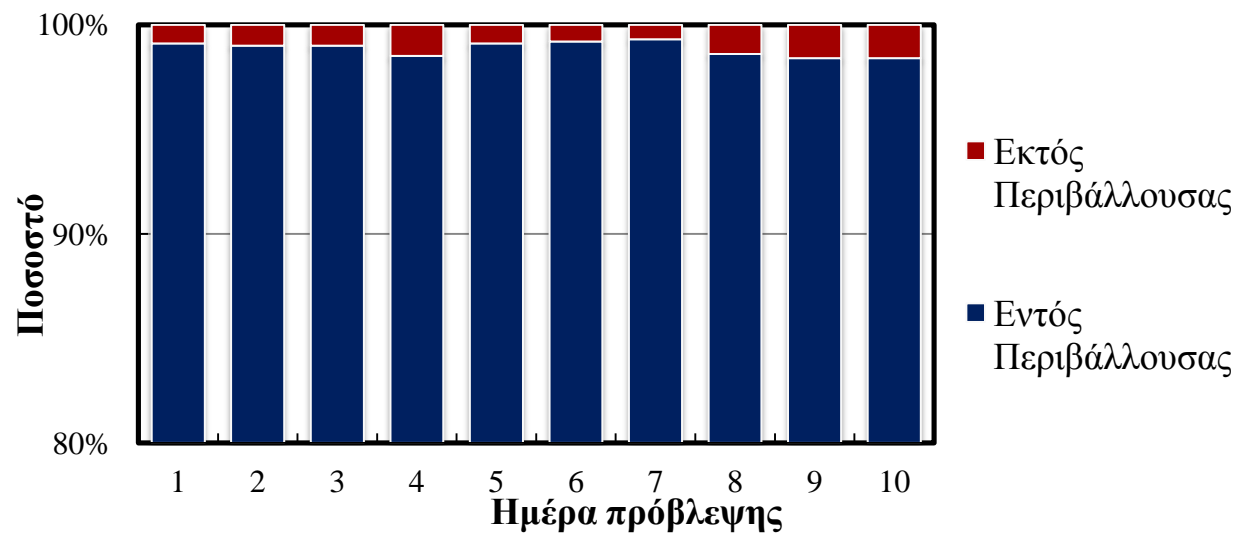
Διάγραμμα συντελεστή CE των διαφορετικών μεθόδων για τις μετεωρολογικές μεταβλητές της θερμοκρασίας (πάνω), ταχύτητας ανέμου (κάτω) και για χρονικό βήμα μίας ημέρας

# Περιβάλλουσα

Ραβδογράμματα με τα ποσοστά των εντός και εκτός τιμών της περιβάλλουσας, για την μετεωρολογική μεταβλητή της θερμοκρασίας και για χρονικό βήμα μίας ώρας



Ραβδογράμματα με τα ποσοστά των εντός και εκτός τιμών της περιβάλλουσας, για την μετεωρολογική μεταβλητή της βροχόπτωσης και για χρονικό βήμα μίας ημέρας



# Συμπεράσματα

- Βασικότερο πλεονέκτημα της μεθόδου είναι η απλότητα καθώς και η χρήση απλών μαθηματικών για την εξαγωγή πρόγνωσης
- Το μοντέλο ανταποκρίνεται πολύ καλά στις περισσότερες περιπτώσεις
- Το μοντέλο αδυνατεί να έχει ικανοποιητικά αποτελέσματα για την μετεωρολογική μεταβλητή της βροχόπτωσης καθώς και για την ταχύτητα ανέμου για το χρονικό βήμα της μιας ημέρας
- Το πολυμεταβλητό είναι υποδεέστερο του απλού και σε κάποιες περιπτώσεις χειρότερο από τα μοντέλα Benchmark.
- Στα τρία ζητήματα που εξετάστηκαν
  - ❑ Η λογαριθμική μέθοδος ΔΠΚ εμφανίζεται να είναι η καλύτερη
  - ❑ Ελάχιστη χρονική απόσταση 120 min (για τα χρονικά βήματα του δεκάλεπτου και της ώρας)
  - ❑ Χρήση του μέγιστου αριθμού αναλόγων τιμών
- Η περιβάλλουσα εμφανίζει πολύ ικανοποιητικά αποτελέσματα για τις μετεωρολογικές μεταβλητές της θερμοκρασίας, υγρασίας και ταχύτητας ανέμου

Σας ευχαριστώ για την προσοχή σας !