

Χρηματοδοτικό πρόγραμμα  
«ΦΥΣΙΚΟ ΠΕΡΙΒΑΛΛΟΝ & ΚΑΙΝΟΤΟΜΕΣ ΔΡΑΣΕΙΣ 2022»

Άξονας Προτεραιότητας 3  
«ΕΡΕΥΝΑ ΚΑΙ ΕΦΑΡΜΟΓΗ»

Τίτλος Έργου

Στρατηγική δράση για τον Μετριασμό και την Προσαρμογή της Κλιματικής  
Αλλαγής στα Λιμάνια



**Παραδοτέο 3.1:** Εκτίμηση Χρονοσειρών Κλιματικών δεδομένων με Τεχνικές Βαθιάς  
Μηχανικής Μάθησης

**Υποενότητα εργασίας:** Εκτίμηση Χρονοσειρών με Τεχνικές Βαθιάς Μηχανικής Μάθησης

## Περιεχόμενα

Περιγραφή Παραδοτέου .....	3
Σύνθεση Ομάδας Έργου Παραδοτέου .....	3
<b>1. Εισαγωγή</b> .....	<b>4</b>
2.1. Επισκόπηση Δεδομένων .....	5
2.2. Ανάλυση Δεδομένων .....	7
2.2.1. Ντετερμινιστικά Μοντέλα .....	7
2.2.2. Μοντέλα Μηχανικής Μάθησης .....	10
<b>3. Μοντέλα Πρόβλεψης Ταχύτητας Ανέμου στο Λιμάνι της Χαλκίδας</b> .....	<b>12</b>
3.1. Ανάλυση Βραχυχρόνιας Περιόδου .....	12
3.1.1. Πολική Αναπαράσταση .....	16
3.1.2. Χωρική Συσχέτιση Ανέμου .....	16
3.1.3. Αυτοσυσχέτιση και Μερική Αυτοσυσχέτιση .....	17
3.1.4. Ντετερμινιστικά Μοντέλα .....	22
3.1.5. Μοντέλα Μηχανικής Μάθησης .....	23
3.1.6. Σύγκριση Ευστοχίας μεταξύ Μοντέλων .....	23
3.2. Ανάλυση Μακράς Περιόδου .....	25

## Περιγραφή Παραδοτέου

Στόχος των αλγορίθμων Τεχνητής Νοημοσύνης/Μηχανικής Μάθησης (AI/ML) είναι να παρέχουν ειδοποιήσεις, προβλέψεις ή προτεινόμενες ενέργειες κάτω από άγνωστες περιβαλλοντικές καταστάσεις. Προτείνεται η χρήση ενός συνδυαστικού αλγοριθμικού σχήματος, βασισμένο σε τεχνικές ML. Στόχος του προτεινόμενου αλγορίθμου είναι η εκτίμηση των αγνώστων χρονοσειρών κλιματικών δεδομένων στο πιλοτικό λιμάνι. Οι προβλέψεις θα στηρίζονται τόσο στα δεδομένα που έχουν συλλεχθεί από γειτονικά σημεία αναφοράς, όσο και στις νεοαποκτηθείσες μετρήσεις στο λιμάνι. Συγκεκριμένα, προτείνεται η εκπαίδευση νευρωνικών δικτύων, εξειδικευμένων για την εκτίμηση των χρονοσειρών στα σημεία αναφοράς. Στη συνέχεια, ο προτεινόμενος αλγόριθμος θα αξιοποιήσει τις χωροχρονικές και περιβαλλοντικές συσχετίσεις μεταξύ των τοποθεσιών πιλοτικού λιμένα και σημείων αναφοράς (λιμάνια μελέτης). Για τον σκοπό αυτό, εφαρμόζεται η τεχνική Συνδυαστικής Μηχανικής Μάθησης (Ensemble Machine Learning-EML), σύμφωνα με την οποία η πρόβλεψη των αγνώστων δεδομένων προκύπτει ως σταθμικός μέσος των γνωστών δεδομένων. Για την αποτελεσματική εκτίμηση των βαθμών συνεισφοράς (βάρος) καθεμιάς ανεξάρτητης πρόβλεψης στην τελική πρόβλεψη, υιοθετείται μια τεχνική αξιολόγησης της ομοιότητας (similarity scoring) μεταξύ περιβαλλοντικών συνθηκών στο πιλοτικό λιμάνι και των σημείων αναφοράς. Τούτο επιτυγχάνεται με την αξιοποίηση των δεδομένων που συλλέγονται από τον μετρητικό εξοπλισμό στο λιμάνι ενδιαφέροντος, συσχετίζοντας τις ιστορικές χρονοσειρές με τις νεοαποκτηθείσες. Τα βάρη του σταθμικού μέσου ενημερώνονται περιοδικά όσο οι μετρήσεις στο λιμάνι μελέτης συνεχίζονται, οδηγώντας έτσι σε σταδιακή σύγκλιση του αλγοριθμικού σχήματος.

## Σύνθεση Ομάδας Έργου Παραδοτέου

Βασίλης Νικολακάκης, Επιστ. Συνεργάτης

Αναστάσιος Γιαννόπουλος, Επιστ. Συνεργάτης

Παναγιώτης Τρακάδας, Αναπλ.Καθηγητής

## 1. Εισαγωγή

Η έγκαιρη εισαγωγή μέτρων προσαρμογής για την αντιμετώπιση των επιπτώσεων της Κλιματικής Αλλαγής στους λιμένες αποτελεί βασική προτεραιότητα και μπορούν να μετριαστούν εάν στηριχθούν στην πρόγνωση του κλίματος. Η πρόκληση αυτή απαιτεί την καταγραφή κλιματικών χρονοσειρών υψηλής χωρικής ανάλυσης για το παρόν και το μέλλον, οι οποίες γενικά απουσιάζουν. Αντικείμενο αυτής της μελέτης είναι η ανάπτυξη και ανακατασκευή αξιόπιστων χρονοσειρών ταχύτητας ανέμου υψηλής χωρικής ανάλυσης σε μια τοποθεσία-στόχο από δεδομένα χαμηλής χωρικής ανάλυσης.

Για τον σκοπό αυτό, προτείνεται η εκπαίδευση νευρωνικών δικτύων, εξειδικευμένων για την εκτίμηση χρονοσειρών σε καθορισμένο σημείο, το οποίο για τις ανάγκες του έργου αντιστοιχεί στον πιλοτικό λιμένα μελέτης. Στόχος του προτεινόμενου αλγορίθμου είναι η ανακατασκευή δεδομένων (ταχύτητας) ανέμου άγνωστων χρονοσειρών χρησιμοποιώντας δεδομένα από κοντινά γεωγραφικά σημεία. Για την ανακατασκευή της χρονοσειράς ανέμου εφαρμόζεται η τεχνική Ensemble Machine Learning (EML), σύμφωνα με την οποία η πρόβλεψη άγνωστων δεδομένων προκύπτει ως σταθμισμένος μέσος όρος των γνωστών δεδομένων. Στο πλαίσιο αυτό, ο προτεινόμενος αλγόριθμος αξιοποιεί τους χωροχρονικούς συσχετισμούς υπάρχοντων ιστορικών ανεμολογικών δεδομένων από συγκεκριμένες θέσεις και συγκεκριμένου σημείου αναφοράς (Pilot Port).

Στο Παραδοτέο αυτό, εφαρμόζεται αρχικά μελέτη για την διερεύνηση των καταγεγραμμένων χρονοσειρών, με στόχο την εξαγωγή βασικών συμπερασμάτων σχετικά με τα βασικά μοτίβα και τις στατιστικές ιδιότητες των δεδομένων. Εν συνεχεία, εφαρμόζονται και αξιολογούνται διάφορα μοντέλα (τόσο απλοϊκά όσο και μοντέλα τεχνητής νοημοσύνης) για την πρόβλεψη της ταχύτητας του ανέμου στο Λιμάνι της Χαλκίδας, συμπεριλαμβανομένων: Μοντέλο Απλού Μέσου όρου, Μοντέλο Σταθμικού Μέσου όρου, Μοντέλο Πολλαπλής Γραμμικής Παλινδρόμησης, Μοντέλο Παλινδρόμησης Δέντρου Απόφασης, Μοντέλο Παλινδρόμησης Διανυσμάτων Υποστήριξης, Μοντέλο Παλινδρόμησης Τυχαίου Δάσους, Μοντέλο Παλινδρόμησης Ενίσχυσης Κλίσης και Αυτοπαλινδρομικό Μοντέλο.

## 2. Μεθοδολογικό πλαίσιο

Στο παρόν παραδοτέο περιγράφεται η διαδικασία για την εκτίμηση χρονοσειρών ταχύτητας ανέμου σε μια τοποθεσία-στόχο (Target Location) χρησιμοποιώντας ιστορικές χρονοσειρές ταχύτητας ανέμου από τέσσερις διαφορετικές τοποθεσίες (μεγαλύτερης χωρικής κλίμακας) με Τεχνικές Βαθιάς Μηχανικής Μάθησης. Η περιοχή-στόχος αναφέρεται ως Location of Interest (LoI) και στο πλαίσιο του έργου αντιστοιχεί στο Πιλοτικό Λιμάνι της Χαλκίδας (Pilot Port). Οι τέσσερις τοποθεσίες για τις οποίες υπάρχουν ιστορικές καταγραφές ανεμολογικών σειρών αναφέρονται ως τοποθεσίες αναφοράς ή Reference Points (RPs).

### 2.1. Επισκόπηση Δεδομένων

Ως ιστορικά δεδομένα αξιοποιούνται τα ανοιχτά δεδομένα αναδρομικής ανάλυσης ECMWF-ERA5 για τα πλησιέστερα σημεία στον πιλοτικό λιμένα της Χαλκίδας (σημεία αναφοράς). Το σύνολο των δεδομένων καταγραφής περιλαμβάνει ωριαίες τιμές ταχύτητας ανέμου (σε m/s) στις κατευθύνσεις  $x$  ( $u_{10}$  μεταβλητή) και  $y$  ( $v_{10}$  μεταβλητή) για τέσσερις περιοχές αναφοράς και διάρκεια 54 ετών (1970 έως 2023). Κάθε σύνολο δεδομένων περιλαμβάνει την πληροφορία του Πίνακα 1.

Πίνακας 1. Ανάλυση περιεχόμενης πληροφορίας των δεδομένων

Στήλη δεδομένων	Περιγραφή
Χρόνος	Η χρονική ετικέτα κατά την οποία λήφθηκε το δείγμα (π.χ. 1/1/1970 17:00)
Γεωγραφικό Μήκος	Το γεωγραφικό μήκος του σημείου καταγραφής του ανέμου (σε ΓΣΑ84)
Γεωγραφικός Πλάτος	Το γεωγραφικό πλάτος του σημείου καταγραφής του ανέμου (σε ΓΣΑ84)
Ταχύτητα $u_{10}$	Η ταχύτητα του ανέμου στην $x$ -κατεύθυνση
Ταχύτητα $v_{10}$	Η ταχύτητα του ανέμου στην $y$ -κατεύθυνση

Για την εκτίμηση της επιτυχίας της μεθόδου, η ανακατασκευασμένη χρονοσειρά αξιολογείται στη συνέχεια με βάση τα παρατηρησιακά δεδομένα ανέμου που συγκεντρώνονται από τον σταθμό που έχει εγκατασταθεί, στο πλαίσιο υλοποίησης του έργου και για το σκοπό αυτό, στο πιλοτικό λιμάνι της Χαλκίδας. Μάλιστα, η προβλεπτική αξιολόγηση των απλοϊκών μοντέλων και μοντέλων μηχανικής μάθησης επαναλήφθηκε για δύο χρονικές περιόδους με στόχο να αξιολογηθεί η απόδοση των αλγορίθμων πρόβλεψης τόσο για βραχυχρόνιες όσο και για μακρές χρονικές περιόδους. Συγκεκριμένα, κατά την ανάλυση της βραχυχρόνιας περιόδου επιχειρείται η πρόβλεψη των ανεμολογικών σειρών από 20 Δεκεμβρίου 2023 έως 31 Δεκεμβρίου 2023, ενώ η ανάλυση της

μακράς περιόδου αφορά την πρόβλεψη του ανέμου από 26 Φεβρουαρίου 2024 έως 29 Οκτώβρη 2024.

Στην *Εικόνα 1* παρουσιάζονται οι θέσεις των τεσσάρων σημείων περιφερειακά του λιμένα της Χαλκίδας, καθώς και το LoI που αντιστοιχεί στο Λιμάνι της Χαλκίδας.



**Εικόνα 1.** Σημεία αναφοράς και τοποθεσία ενδιαφέροντος.

## 2.2. Ανάλυση Δεδομένων

Για την ανάπτυξη μοντέλων πρόβλεψης, πραγματοποιήσαμε μια ολοκληρωμένη διερευνητική ανάλυση δεδομένων για να κατανοήσουμε τις κατανομές, τους συσχετισμούς και τις χρονικές εξαρτήσεις των δεδομένων ταχύτητας ανέμου από τις τέσσερις τοποθεσίες αναφοράς. Οπτικοποιήσαμε τα δεδομένα χρησιμοποιώντας ιστογράμματα, χάρτες θερμότητας (heatmaps) και πολικά γραφήματα, ενώ πραγματοποιήσαμε αναλύσεις αυτοσυσχέτισης και μερικής αυτοσυσχέτισης για τον εντοπισμό μοτίβων και εξαρτήσεων. Με βάση τις γνώσεις που αποκτήθηκαν από την διερευνητική ανάλυση, αναπτύχθηκαν διάφορα μοντέλα πρόβλεψης, συμπεριλαμβανομένων των μοντέλων απλού μέσου όρου (SAM), των μοντέλων σταθμικού μέσου όρου (WAM) και μοντέλων μηχανικής μάθησης όπως η γραμμική παλινδρόμηση (LR), η παλινδρόμηση δέντρου αποφάσεων (DTR), η τεχνική SVR, η παλινδρόμηση τυχαίου δάσους (RFR), η παλινδρόμηση ενίσχυσης κλίσης (GBR) και ένα μοντέλο αυτοπαλινδρόμησης (AR).

Επιπλέον, υπολογίστηκαν οι συντελεστές συσχέτισης μεταξύ της ταχύτητας ανέμου της θέσης στόχου και κάθε τοποθεσίας εποχής για να προσδιοριστούν τα βάρη για τα μοντέλα σταθμικού μέσου όρου. Εν συνεχεία, εκπαιδεύτηκαν πολλαπλά μοντέλα μηχανικής μάθησης χρησιμοποιώντας τα δεδομένα ταχύτητας ανέμου από τις τέσσερις τοποθεσίες εποχής ως χαρακτηριστικά εισόδου (features) και την ταχύτητα ανέμου της θέσης στόχου ως μεταβλητή-στόχο (target). Για το μοντέλο αυτοπαλινδρόμησης χρησιμοποιήθηκαν οι δύο τελευταίες τιμές ταχύτητας ανέμου από κάθε θέση εποχής για να προβλεφθεί η επόμενη τιμή στη θέση-στόχο. Η απόδοση όλων των μοντέλων αξιολογήθηκε υπολογίζοντας το μέσο τετραγωνικό σφάλμα (Mean Squared Error, MSE) μεταξύ προβλέψεων και πραγματικών τιμών, ενώ σχεδιάστηκαν οι πραγματικές και οι προβλεπόμενες (από τα μοντέλα) σειρές ταχύτητας ανέμου με στόχο να συγκριθούν οπτικά ως προς την ευστοχία της παρεχόμενης πρόβλεψης. Αυτή η συστηματική προσέγγιση εξασφάλισε την αποτελεσματική αξιοποίηση των ιστορικών δεδομένων για την παραγωγή προβλέψεων ακριβείας για την ταχύτητα ανέμου στην τοποθεσία-στόχου.

### 2.2.1. Ντετερμινιστικά Μοντέλα

Αρχικά, κατασκευάστηκαν ντετερμινιστικά μοντέλα για την πρόβλεψη της ταχύτητας ανέμου στην τοποθεσία στόχο. Τα μοντέλα αυτά βασίστηκαν και τεχνικές απλού και σταθμικού μέσου όρου με σκοπό να βρεθεί η ευστοχία στην πρόβλεψη του ανέμου χρησιμοποιώντας απλά μοντέλα που βασίζονται σε απλούς κανόνες και στατιστικές μετρικές (όπως ο μέσος όρος). Κύριο πλεονέκτημα των μεθόδων αυτών είναι η απλότητά τους, αφού η δημιουργία και η χρησιμοποίησή τους δείχνει

χαμηλή πολυπλοκότητα και γρήγορη κλήση για προβλέψεις. Επιπλέον, η χρήση τέτοιων μοντέλων για λόγους σύγκρισης (baselines) με άλλα πολυπλότερα μοντέλα (όπως αυτά της μηχανικής μάθησης) μπορεί να ποσοτικοποιήσει τα οφέλη των τελευταίων έναντι των απλών προβλεπτικών τεχνικών.

Τα ντετερμινιστικά μοντέλα που υλοποιήθηκαν παρουσιάζονται παρακάτω:

- 1) Μοντέλου Απλού Μέσου Όρου με βάση το τελευταίο έτος (SAM-last):** Το μοντέλο αυτό εκτιμά τις τιμές του ανέμου στη θέση-στόχο με βάση τον μέσο όρο των τιμών στις τοποθεσίες εκπομπής κατά το τελευταίο έτος (2023). Φορμαλιστικά, η εξίσωση που εκφράζει τις προβλέψεις αυτού του μοντέλου είναι:

$$y_{SAM-last}(t) = \frac{1}{4} (X_{1,last}(t) + X_{2,last}(t) + X_{3,last}(t) + X_{4,last}(t))$$

όπου  $y_{SAM-last}(t)$  είναι η προβλεπόμενη τιμή της ταχύτητας ανέμου από το μοντέλο SAM-last τη χρονική στιγμή  $t$  και  $X_{i,last}(t)$  είναι η τιμή της ταχύτητας ανέμου στην τοποθεσία  $i$  ( $i = 1,2,3,4$ ) τη χρονική στιγμή  $t$  του έτους 2023.

- 2) Μοντέλου Απλού Μέσου Όρου με βάση τα έτη σημαντικής συσχέτισης (SAM-sig):** Το μοντέλο αυτό εκτιμά τις τιμές του ανέμου στη θέση-στόχο με βάση τον μέσο όρο των τιμών στις τοποθεσίες εκπομπής κατά τα έτη σημαντικής συσχέτισης. Για το σκοπό αυτό, αρχικά υπολογίζεται ο συντελεστής συσχέτισης Pearson για την ταχύτητα του ανέμου κατά την περίοδο ενδιαφέροντος μεταξύ περιοχής-στόχου και τοποθεσιών αναφοράς για κάθε έτος. Το αποτέλεσμα είναι ένας πίνακας  $4 \times 54$ , ο οποίος περιλαμβάνει τις τιμές συσχέτισης για κάθε μια από τις τέσσερις τοποθεσίες αναφοράς και έτος (1970-2023). Στην συνέχεια, εντοπίζονται τα έτη για τα οποία η τιμή συσχέτισης είναι μεγαλύτερη του 0.15 ή μικρότερη του -0.15 για κάθε σημείο αναφοράς. Τελικά, η εξίσωση που εκφράζει τις προβλέψεις αυτού του μοντέλου είναι:

$$y_{SAM-sig}(t) = \frac{1}{4} (\bar{X}_{1,sig}(t) + \bar{X}_{2,sig}(t) + \bar{X}_{3,sig}(t) + \bar{X}_{4,sig}(t))$$

όπου  $y_{SAM-sig}(t)$  είναι η προβλεπόμενη τιμή της ταχύτητας ανέμου από το μοντέλο SAM-sig τη χρονική στιγμή  $t$  και  $\bar{X}_{i,sig}(t)$  είναι η τιμή της ταχύτητας ανέμου στην τοποθεσία  $i$  ( $i = 1,2,3,4$ ), η οποία έχει προκύψει ως η μέση τιμή των ετών σημαντικής συσχέτισης στην ίδια

χρονική στιγμή. Επομένως, για κάθε τοποθεσία αναφοράς, πρώτα λαμβάνεται ο μέσος όρος μεταξύ των ετών σημαντικής συσχέτισης και, ύστερα, λαμβάνεται ο μέσος όρος μεταξύ των τοποθεσιών αναφοράς.

- 3) Μοντέλου Σταθμικού Μέσου Όρου με βάση όλα τα έτη (WAM-all):** Το μοντέλο αυτό εκτιμά τις τιμές του ανέμου στη θέση-στόχο με βάση τον σταθμισμένο μέσο όρο των τιμών στις τοποθεσίες εκπομπής. Η στάθμιση της συνεισφοράς καθεμιάς τοποθεσίας αναφοράς στον μέσο όρο στηρίζεται στον υπολογισμό βαρών (weights). Τα βάρη υπολογίζονται με βάση την τιμή συσχέτισης της ταχύτητας ανέμου μεταξύ τοποθεσίας-στόχου και τοποθεσία αναφοράς. Η ιδέα είναι ότι οι τοποθεσίες που δείχνουν μεγαλύτερη συσχέτιση με την τοποθεσία-στόχο ως προς την ταχύτητα ανέμου θα έχουν μεγαλύτερα βάρη στον μέσο όρο. Τα βάρη υπολογίζονται με βάση τη μέση συσχέτιση στόχου και αναφοράς σε όλα τα έτη. Έτσι, η προβλέψεις αυτού του μοντέλου δίνονται από την σχέση:

$$y_{WAM-all}(t) = w_1 X_{1,last}(t) + w_2 X_{2,last}(t) + w_3 X_{3,last}(t) + w_4 X_{4,last}(t)$$

όπου  $y_{WAM-all}(t)$  είναι η προβλεπόμενη τιμή της ταχύτητας ανέμου από το μοντέλο WAM-all τη χρονική στιγμή  $t$  και  $X_{i,last}(t)$  είναι η τιμή της ταχύτητας ανέμου στην τοποθεσία  $i$  ( $i = 1,2,3,4$ ), όπως περιγράφηκε για το μοντέλο SAM-last. Οι τιμές  $w_i$  είναι τα βάρη για κάθε τοποθεσία και υπολογίζονται ως μέσος συντελεστής συσχέτισης μεταξύ στόχου και αναφοράς σε όλα τα έτη (1970-2023).

- 4) Μοντέλου Σταθμικού Μέσου Όρου με βάση τα έτη σημαντικής συσχέτισης (WAM-sig):** Το μοντέλο αυτό εκτιμά τις τιμές του ανέμου στη θέση-στόχο με βάση τον σταθμισμένο μέσο όρο των τιμών στις τοποθεσίες εκπομπής. Σε αντίθεση με το μοντέλο WAM-all, εδώ τα βάρη υπολογίζονται με βάση τη μέση συσχέτιση στόχου και αναφοράς μόνο στα έτη που δείχνουν σημαντική συσχέτιση (μεγαλύτερη απόλυτη τιμή από 0.15). Έτσι, η προβλέψεις αυτού του μοντέλου δίνονται από την σχέση:

$$y_{WAM-sig}(t) = w'_1 X_{1,last}(t) + w'_2 X_{2,last}(t) + w'_3 X_{3,last}(t) + w'_4 X_{4,last}(t)$$

όπου  $y_{WAM-sig}(t)$  είναι η προβλεπόμενη τιμή της ταχύτητας ανέμου από το μοντέλο WAM-sig τη χρονική στιγμή  $t$  και  $X_{i,last}(t)$  είναι η τιμή της ταχύτητας ανέμου στην τοποθεσία  $i$  ( $i =$

1,2,3,4), όπως περιγράφηκε για το μοντέλο SAM-last. Οι τιμές  $w'_i$  είναι τα βάρη για κάθε τοποθεσία και υπολογίζονται ως μέσος συντελεστής συσχέτισης μεταξύ στόχου και αναφοράς στα έτη σημαντικής συσχέτισης.

### 2.2.2. Μοντέλα Μηχανικής Μάθησης

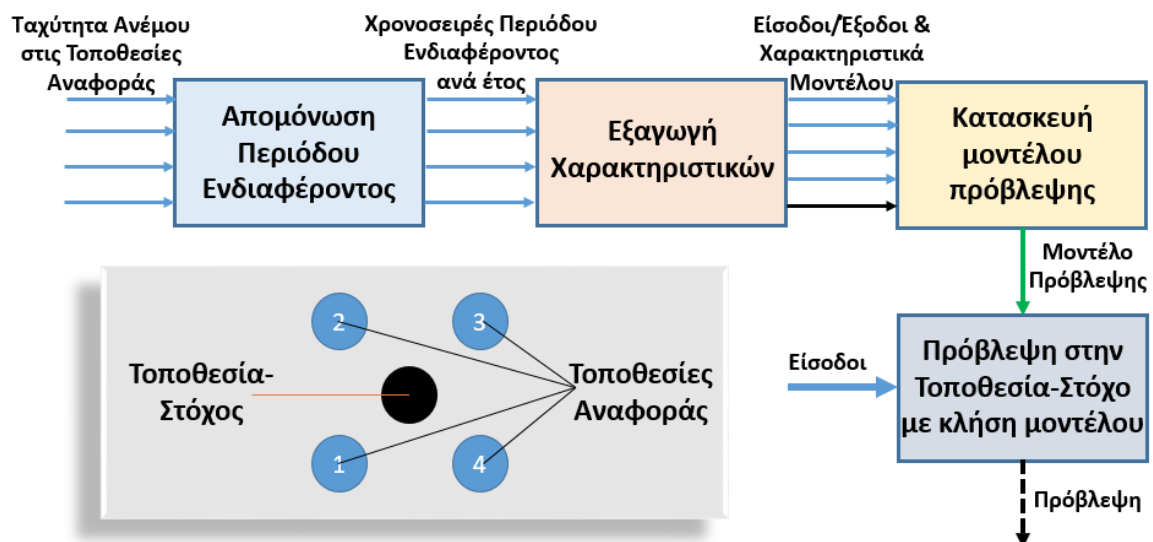
Με στόχο την ακριβέστερη πρόβλεψη, εξετάστηκε μια σειρά από ευρέως χρησιμοποιούμενα μοντέλα μηχανική μάθησης ώστε να παρέχεται και η «μάθηση» μη γραμμικών σχέσεων μεταξύ εισόδων (Σημεία Αναφοράς) και εξόδου (Λιμάνι Χαλκίδας). Για την εκπαίδευση των μοντέλων μηχανικής μάθησης, χρησιμοποιήθηκαν ως εισοδοί οι τιμές της ταχύτητας ανέμου στα 4 σημεία αναφοράς, ενώ ως (επιθυμητή) έξοδος η ταχύτητα ανέμου στο Λιμάνι της Χαλκίδας. Εκπαιδεύτηκαν και εξετάστηκαν τα παρακάτω μοντέλα μηχανικής μάθησης:

- 1) **Πολλαπλή γραμμική παλινδρόμηση (MLR):** Η πολλαπλή γραμμική παλινδρόμηση χρησιμοποιεί έναν γραμμικό συνδυασμό των τιμών ταχύτητας ανέμου και από τις τέσσερις τοποθεσίες εποχής για να προβλέψει την ταχύτητα του ανέμου στη θέση στόχου. Αυτό το μοντέλο υποθέτει μια γραμμική σχέση μεταξύ των προγνωστικών παραγόντων και της μεταβλητής-στόχου, επιτρέποντας την εύκολη ερμηνεία της επίδρασης κάθε θέσης στην πρόβλεψη.
- 2) **Παλινδρόμηση δέντρου αποφάσεων (DTR):** Η παλινδρόμηση δέντρου αποφάσεων χρησιμοποιεί ένα μοντέλο αποφάσεων που μοιάζει με δέντρο για να προβλέψει την ταχύτητα του ανέμου. Χωρίζει αναδρομικά τα δεδομένα σε υποσύνολα με βάση τις τιμές χαρακτηριστικών, δημιουργώντας ένα μοντέλο που μπορεί να συλλάβει μη γραμμικές σχέσεις. Αυτό το μοντέλο είναι ιδιαίτερα χρήσιμο για την κατανόηση της ιεραρχικής δομής της λήψης αποφάσεων στη διαδικασία πρόβλεψης.
- 3) **Παλινδρόμηση διανυσμάτων υποστήριξης (SVR):** Η τεχνική SVR εφαρμόζει τις αρχές των Support Vector Machines σε προβλήματα παλινδρόμησης, στοχεύοντας στην εύρεση της βέλτιστης γραμμής εντός ενός συγκεκριμένου ορίου. Είναι αποτελεσματικό στο χειρισμό δεδομένων υψηλών διαστάσεων και μπορεί να μοντελοποιήσει πολύπλοκες, μη γραμμικές σχέσεις χρησιμοποιώντας διαφορετικές λειτουργίες πυρήνα.
- 4) **Παλινδρόμηση Τυχαίου Δάσους (RFR):** Η τεχνική αυτή δημιουργεί ένα σύνολο δέντρων αποφάσεων, καθένα από τα οποία εκπαιδεύεται σε ένα τυχαίο υποσύνολο των δεδομένων και υπολογίζει κατά μέσο όρο τις προβλέψεις τους. Αυτή η προσέγγιση μειώνει την υπερβολική προσαρμογή και βελτιώνει την ικανότητα γενίκευσης του μοντέλου. Αξιοποιεί

τη δύναμη πολλαπλών δέντρων για να ενισχύσει την ακρίβεια και την ευρωστία της πρόβλεψης.

- 5) **Παλινδρόμηση ενίσχυσης κλίσης (GBR):** Η παλινδρόμηση ενίσχυσης κλίσης δημιουργεί ένα σύνολο ασθενών μοντέλων πρόβλεψης, συνήθως δέντρων αποφάσεων, προσθέτοντας διαδοχικά μοντέλα που διορθώνουν τα σφάλματα των προηγούμενων. Αυτή η μέθοδος επικεντρώνεται στην ελαχιστοποίηση της συνάρτησης απώλειας, οδηγώντας σε εξαιρετικά ακριβείς και αποτελεσματικές προβλέψεις μέσω επαναληπτικών βελτιώσεων.

Σε γενικές γραμμές, η μεθοδολογία επεξεργασίας και κατασκευής μοντέλων πρόβλεψης ανεμολογικών σειρών ακολουθεί τη φιλοσοφία του ακόλουθου διαγράμματος (Εικόνα 2):



Εικόνα 2. Γενικό διάγραμμα επεξεργασίας και κατασκευής μοντέλων πρόβλεψης ανεμολογικών σειρών.

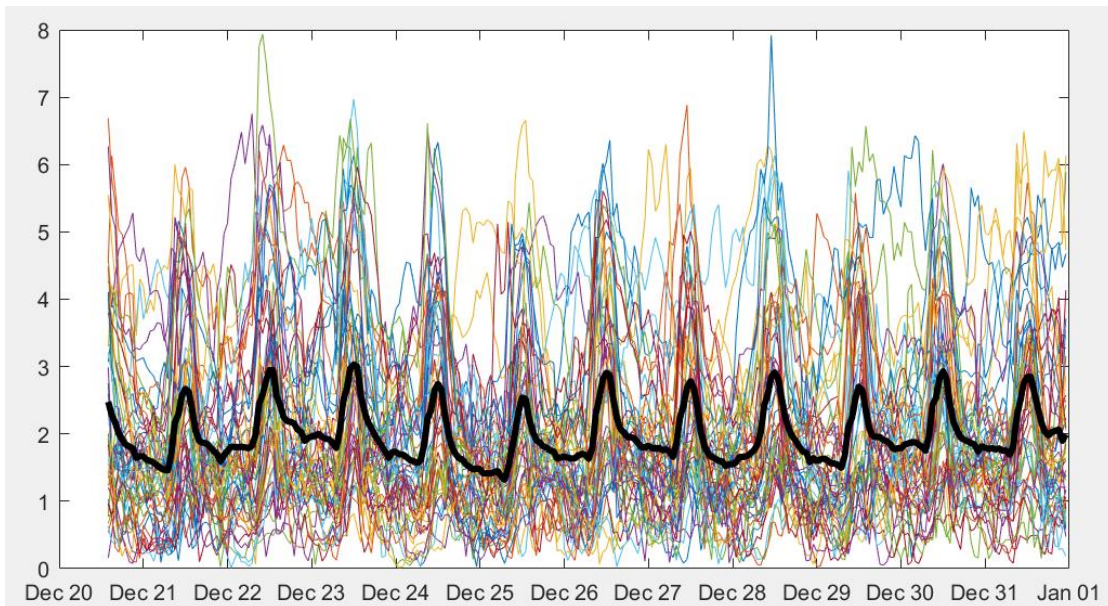
### 3. Μοντέλα Πρόβλεψης Ταχύτητας Ανέμου στο Λιμάνι της Χαλκίδας

Όπως αναφέρθηκε, η προβλεπτική αξιολόγηση των απλοϊκών μοντέλων και μοντέλων μηχανικής μάθησης επαναλήφθηκε για δύο χρονικές περιόδους με στόχο να αξιολογηθεί η απόδοση των αλγορίθμων πρόβλεψης τόσο για βραχυχρόνιες όσο και για μακρές χρονικές περιόδους.

#### 3.1. Ανάλυση Βραχυχρόνιας Περιόδου

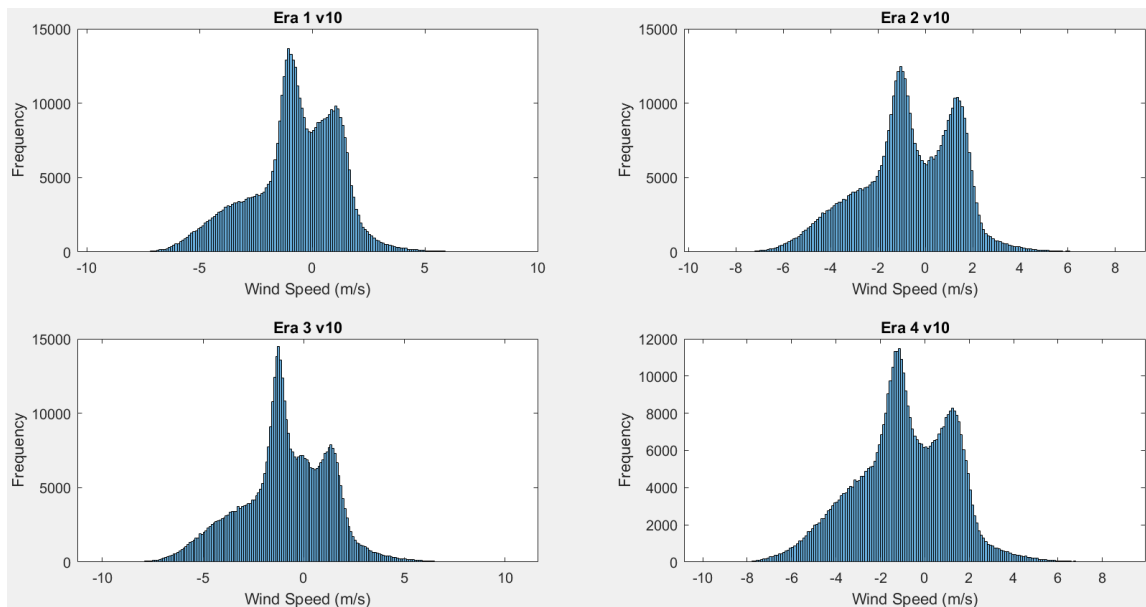
Με την ολοκλήρωση της οργάνωσης των δεδομένων σε πίνακες ανά τοποθεσία αναφοράς, πραγματοποιήθηκε απομόνωση των δεδομένων ενδιαφέροντος. Τα δεδομένα ενδιαφέροντος αναφέρονται στα δείγματα του έτους (ώρα, μέρα, μήνας) για τα οποία υπάρχουν καταγραφές στο Λιμάνι της Χαλκίδας. Στην ακόλουθη ανάλυση εστιάζουμε στην *χρονική περίοδο ενδιαφέροντος* από τις 14:00 ώρα της 20<sup>ης</sup> Δεκεμβρίου έως τις 23:00 ώρα της 31<sup>ης</sup> Δεκεμβρίου.

Δεδομένου ότι τα δείγματα που περιλαμβάνει η χρονική περίοδος ενδιαφέροντος είναι 274, υλοποιήθηκε κώδικας για την δημιουργία ενός πίνακα INTEREST με διαστάσεις 4 x 274 x 54, οποίο απομόνωσε τα δείγματα ενδιαφέροντος σε κάθε τοποθεσία αναφορά και για κάθε έτος. Οι γραμμές του πίνακα αντιστοιχούν στα σημεία αναφοράς, οι στήλες στα δείγματα της περιόδου ενδιαφέροντος και οι σελίδες (τρίτη διάσταση του πίνακα) στα έτη (1970 έως 2023). Με την κατασκευή του πίνακα INTEREST, μπορεί κανείς να σχεδιάσει τις καμπύλες της ταχύτητας ανέμου σε μια τοποθεσία αναφοράς για την περίοδο ενδιαφέροντος. Ενδεικτικά, στην *Εικόνα 3* παρουσιάζεται η ταχύτητα ανέμου στην τοποθεσία αναφοράς 1 για κάθε έτος από το 1970 έως το 2023, ενώ με παχιά μαύρη γραμμή φαίνεται ο μέσος όρος. Παρατηρείται κορύφωση της ταχύτητας ανέμου περίπου στις 12:00 σε καθημερινή βάση.



**Εικόνα 3.** Ταχύτητα ανέμου κατά την περίοδο ενδιαφέροντος στην τοποθεσία αναφοράς 1. Οι πολλαπλές καμπύλες αντιστοιχούν στα έτη 1970 – 2023, ενώ η μαύρη καμπύλη στον μέσο όρο μεταξύ των ετών.

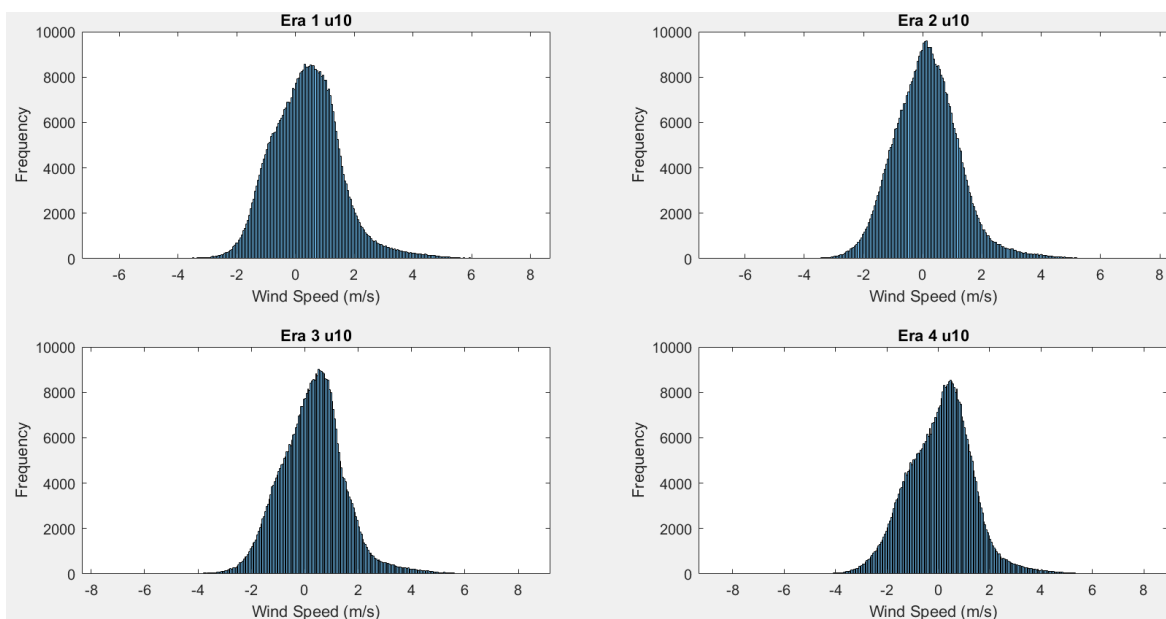
Στη συνέχεια, με σκοπό την επισκόπηση των βασικών ιδιοτήτων των καταγραφόμενων δεδομένων, πραγματοποιήθηκε οπτικοποίηση των κατανομών για τις συνιστώσες του ανέμου, όπως φαίνεται στην **Εικόνα 4**.



**Εικόνα 4.** Κατανομή μεταβλητής  $v_{10}$  στα RPs.

Κάθε ιστόγραμμα αντιπροσωπεύει την κατανομή των τιμών σε ολόκληρη την περίοδο δεδομένων από την 1η Ιανουαρίου 1970 έως τις 31 Δεκεμβρίου 2023 (Εικόνα 5). Οι βασικές παρατηρήσεις που απορρέουν από τις κατανομές είναι οι εξής:

- 1) **Διτροπικές Κατανομές:** Και οι τέσσερις τοποθεσίες εμφανίζουν διτροπικές (bimodal) κατανομές με κορυφές στα  $-2$  m/s και  $2$  m/s. Αυτό δείχνει ότι οι ταχύτητες ανέμου στην κατεύθυνση  $y$  σε αυτές τις τοποθεσίες συχνά κυμαίνονται γύρω από αυτές τις τιμές.
- 2) **Εύρος ταχυτήτων ανέμου:** Το εύρος ταχυτήτων ανέμου και στις τέσσερις εποχές είναι παρόμοιο, με τιμές που κυμαίνονται περίπου από  $-8$  m/s έως  $8$  m/s. Αυτό υποδηλώνει ότι οι συνθήκες ανέμου όσον αφορά την ταχύτητα είναι σχετικά συνεπείς και στις τέσσερις τοποθεσίες.
- 3) **Υψηλότερες συχνότητες:** Οι υψηλότερες συχνότητες για ταχύτητες ανέμου είναι περίπου  $-2$  m/s και  $2$  m/s σε όλες τις τοποθεσίες, υποδηλώνοντας ένα κοινό μοτίβο στη συμπεριφορά του ανέμου κατά την κατεύθυνση  $y$ .



Εικόνα 5. Κατανομή μεταβλητής  $u_{10}$  στα RPs.

Ως προς την κατεύθυνση  $x$  του ανέμου, παρατηρούνται τα εξής:

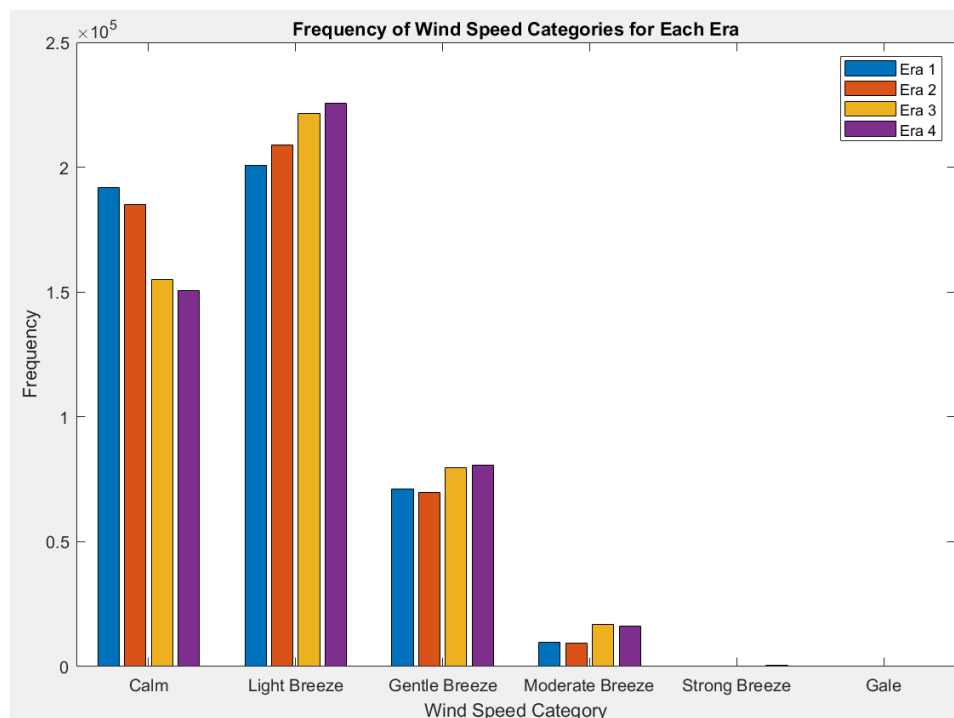
- 1) **Μονότροπες κατανομές:** Και οι τέσσερις τοποθεσίες εμφανίζουν μονότροπικές (unimodal) κατανομές με κέντρο το  $0$ . Αυτό δείχνει ότι οι ταχύτητες ανέμου στην κατεύθυνση  $x$  σε αυτές τις τοποθεσίες συχνά κυμαίνονται στο  $0$ .

- 2) **Εύρος ταχυτήτων ανέμου:** Το εύρος ταχυτήτων ανέμου και στις τέσσερις περιοχές είναι παρόμοιο, με τιμές που κυμαίνονται περίπου από -4 έως 4 m/s. Αυτό υποδηλώνει ότι οι συνθήκες ανέμου όσον αφορά την ταχύτητα είναι σχετικά συνεπείς σε αυτές τις τοποθεσίες.
- 3) **Υψηλότερες συχνότητες:** Οι υψηλότερες συχνότητες για ταχύτητες ανέμου είναι περίπου 0 m/s σε όλες τις περιοχές, υποδηλώνοντας ένα κοινό μοτίβο στη συμπεριφορά του ανέμου κατά την κατεύθυνση  $\chi$ .

Στην συνέχεια, η κατηγοριοποίηση των τιμών του ανέμου σύμφωνα με τον Πίνακα 2, υποδεικνύει ότι οι κατηγορίες ανέμου έχουν παρόμοια συχνότητα εμφάνισης σε όλες τις περιοχές.

**Πίνακας 2.** Κατηγοριοποίηση εντάσεων ανέμου (m/s)

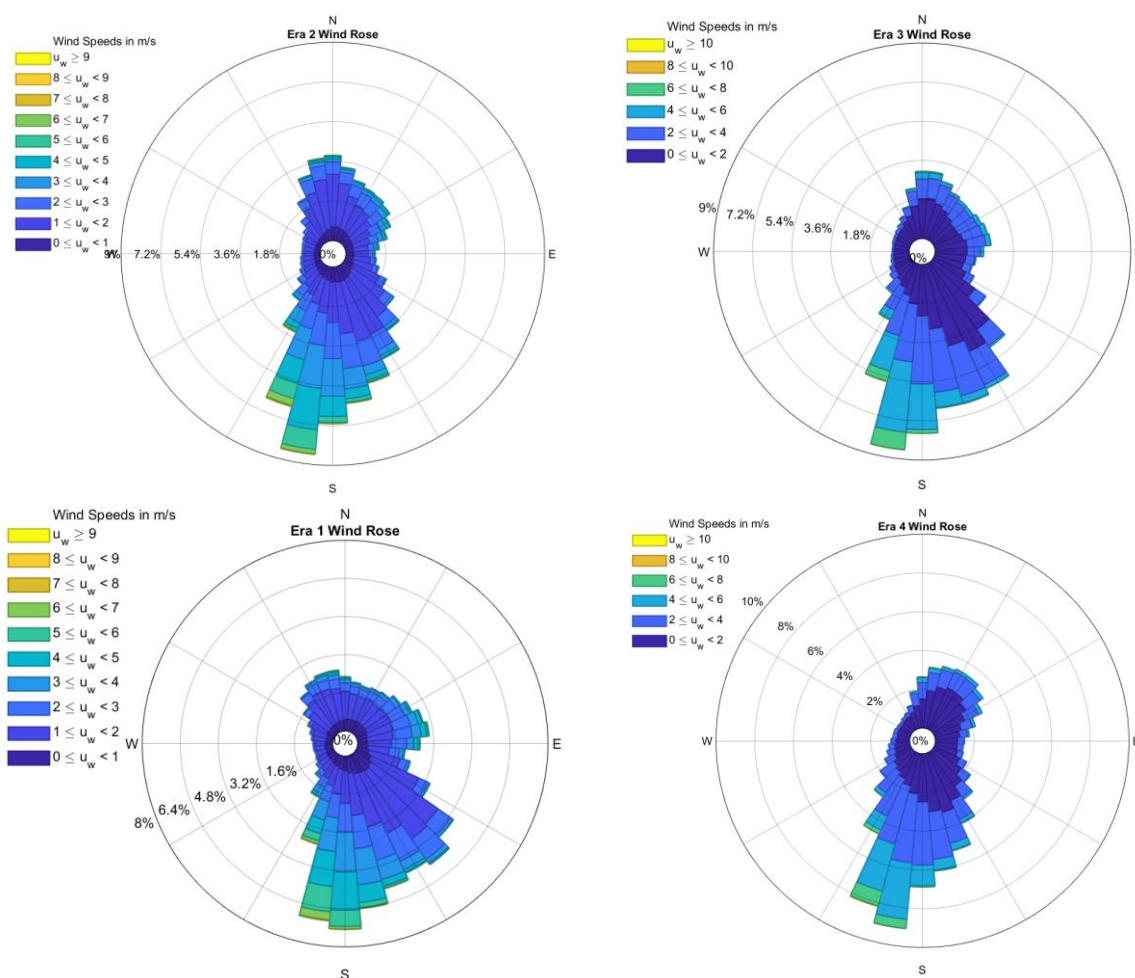
Κατηγορία ανέμου	Εύρος τιμών (m/s)
Calm	0 – 1.5
Light Breeze	1.5 – 3.3
Gentle Breeze	3.3 – 5.5
Moderate Breeze	5.5 – 8
Strong Breeze	8 – 10.8
Gale	>10.8



**Εικόνα 6.** Συχνότητα εμφάνισης κατηγοριών ανέμου.

### 3.1.1. Πολική Αναπαράσταση

Κατά την πολική αναπαράσταση του ανέμου, μετατρέπουμε τις καρτεσιανές μεταβλητές  $u$ ,  $v$  σε πολική αναπαράσταση (πλάτος και γωνία). Στην *Εικόνα 7* παρουσιάζονται οι κατανομές του ανέμου ανά συχνότητα εμφάνισης και διεύθυνση (Wind Rose Plots) για τα 4 σημεία αναφοράς.

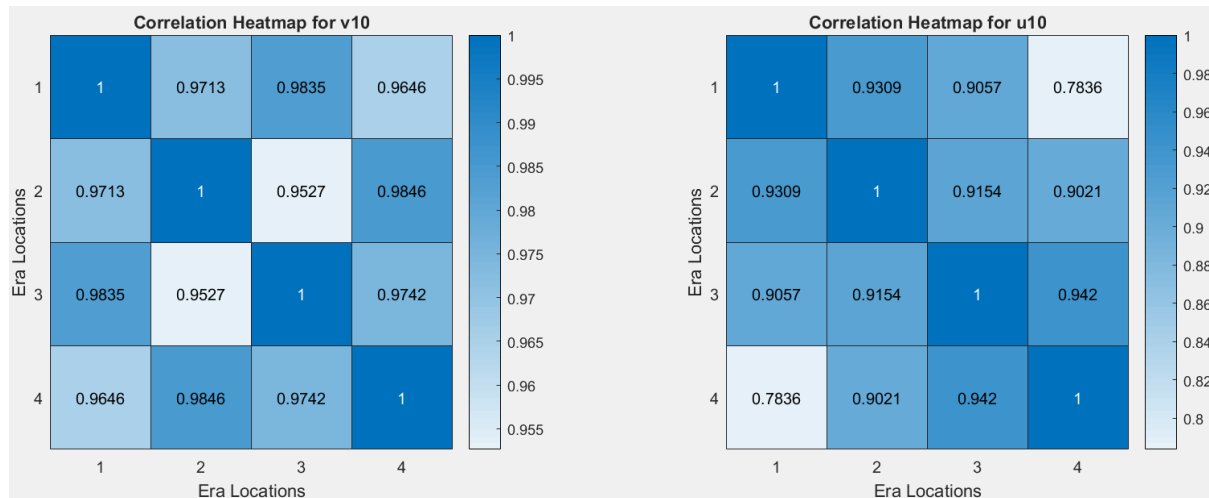


*Εικόνα 7. Πολική αναπαράσταση ανέμου στα τέσσερα σημεία αναφοράς.*

Παρατηρούμε ότι η πλειονότητα των τιμών της διεύθυνσης του ανέμου καταγράφηκε στις 200 μοίρες (νοτιάς), ενώ πολύ λίγα δείγματα καταγράφηκαν στις γωνίες 90 και 170 μοιρών.

### 3.1.2. Χωρική Συσχέτιση Ανέμου

Στη συνέχεια, υπολογίστηκαν οι συντελεστές συσχέτισης Pearson, με στόχο να ποσοτικοποιηθεί η ομοιότητα μεταξύ των ανεμολογικών σειρών στις διαφορετικές τοποθεσίες, όπως φαίνεται στην *Εικόνα 8*.

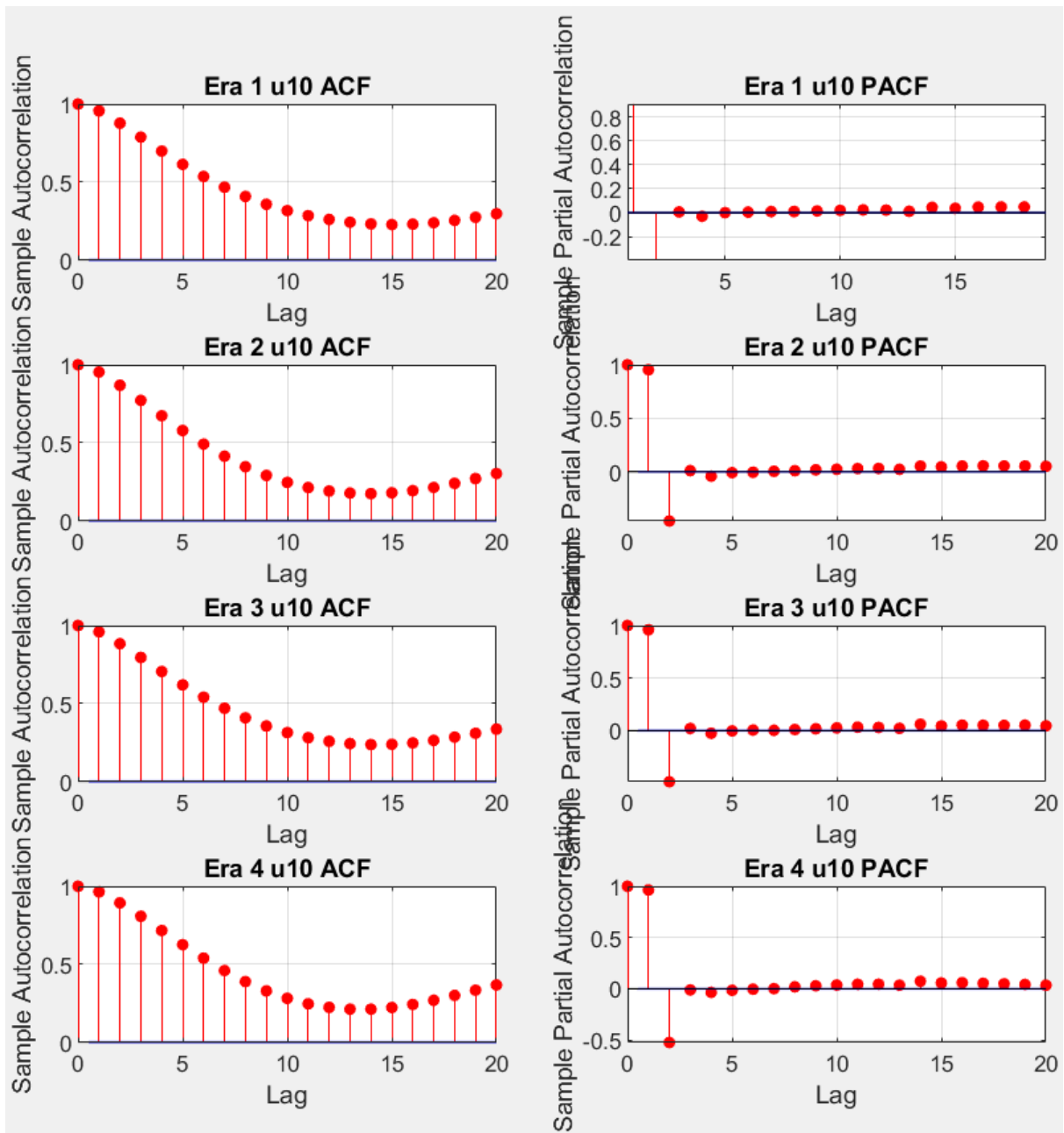


**Εικόνα 8.** Συσχέτιση ανεμολογικών δεδομένων μεταξύ των τεσσάρων περιοχών

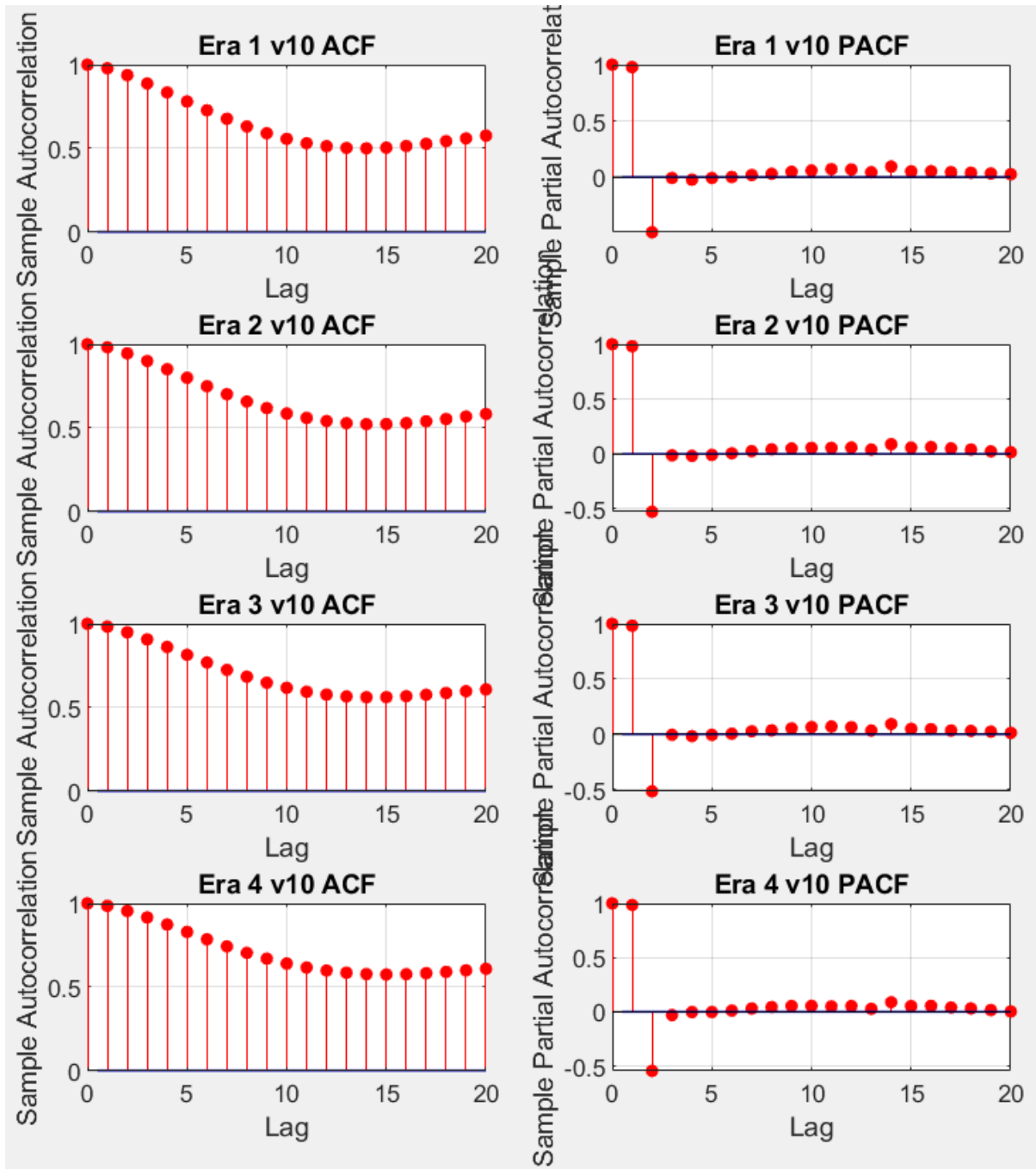
Οι συσχετίσεις μεταξύ των διαφορετικών περιοχών είναι πολύ υψηλές, όλες πάνω από 0.95, υποδηλώνοντας ισχυρές γραμμικές σχέσεις. Η υψηλότερη συσχέτιση (0.9846) είναι μεταξύ της Περιοχής 2 και 4, υποδηλώνοντας πολύ παρόμοια μοτίβα ταχύτητας ανέμου στην κατεύθυνση  $y$  για αυτές τις τοποθεσίες. Δεδομένης της υψηλής συσχέτισης, αποδεικνύεται ότι είναι χρήσιμο να αξιοποιηθούν οι τέσσερις περιοχές για την πρόβλεψη μιας ενδιάμεσης περιοχής. Η ενδιάμεση περιοχή αναμένεται να έχει και αυτή υψηλή συσχέτιση με τις γειτονικές της, αν και άλλοι παράγοντες (υψόμετρο, χρόνος) μπορεί να μειώσουν τη συσχέτιση μεταξύ των τεσσάρων περιοχών και μιας ενδιάμεσης. Για το σκοπό αυτό, στην παρούσα αναφορά θα ελεγχθούν και Μοντέλα Μέσου Όρου, τα οποία υποθέτουν υψηλή γραμμική συσχέτιση του ανέμου στην περιοχή ενδιαφέροντος με τις καταγραφές στα σημεία αναφοράς.

### 3.1.3. Αυτοσυσχέτιση και Μερική Αυτοσυσχέτιση

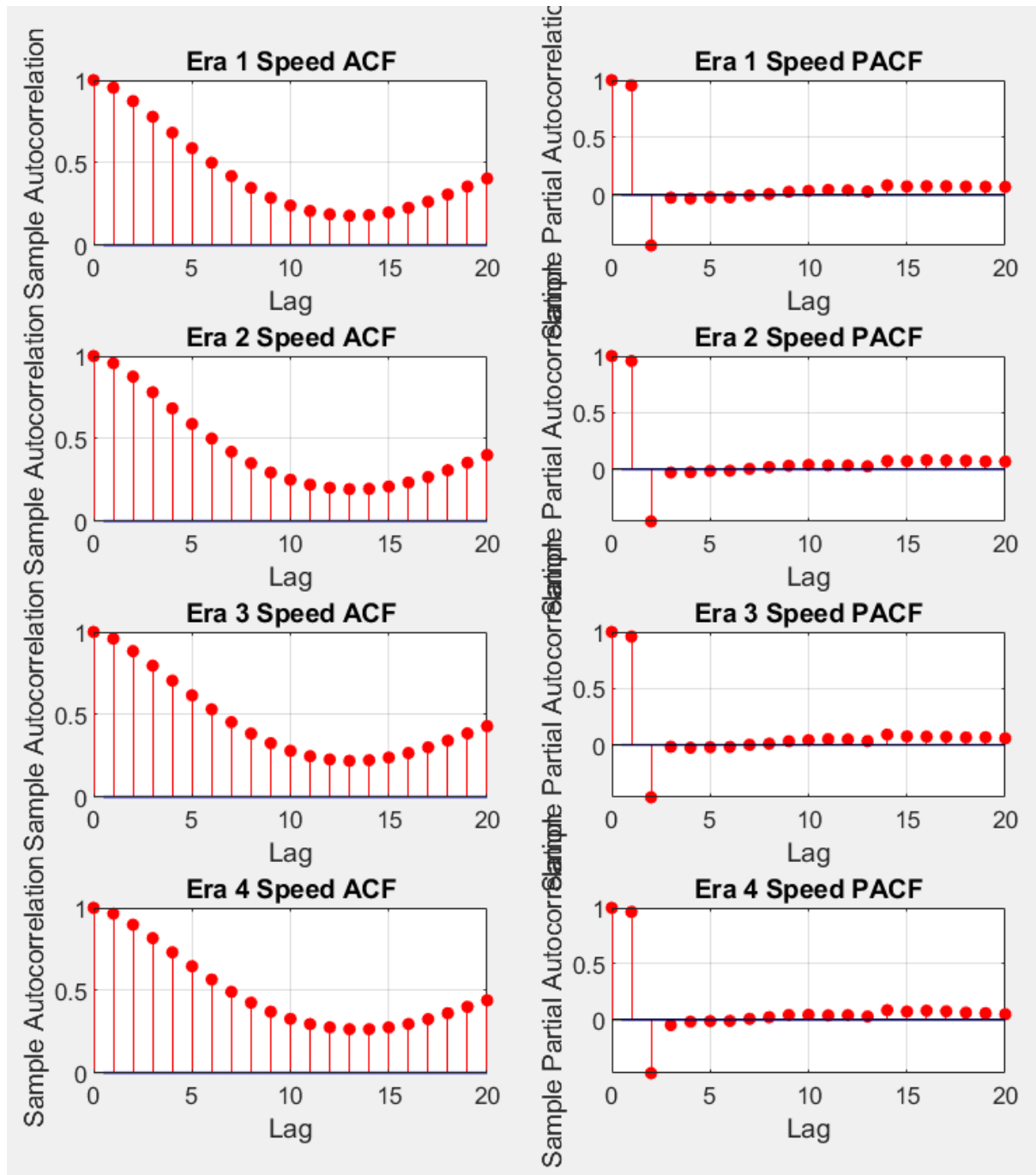
Με σκοπό την εύρεση χρονικών εξαρτήσεων μιας δεδομένης τιμής του ανέμου με παρελθοντικές τιμές της χρονοσειράς, υπολογίστηκε η αυτοσυσχέτιση (ACF) και η μερική αυτοσυσχέτιση (PACF) των τεσσάρων χρονοσειρών. Οι Εικόνες 9 – 12 δείχνουν την αυτοσυσχέτιση (αριστερή στήλη) και την μερική αυτοσυσχέτιση (δεξιά στήλη) των μεταβλητών  $u$ ,  $v$ , speed και direction για κάθε περιοχή (γραμμές).



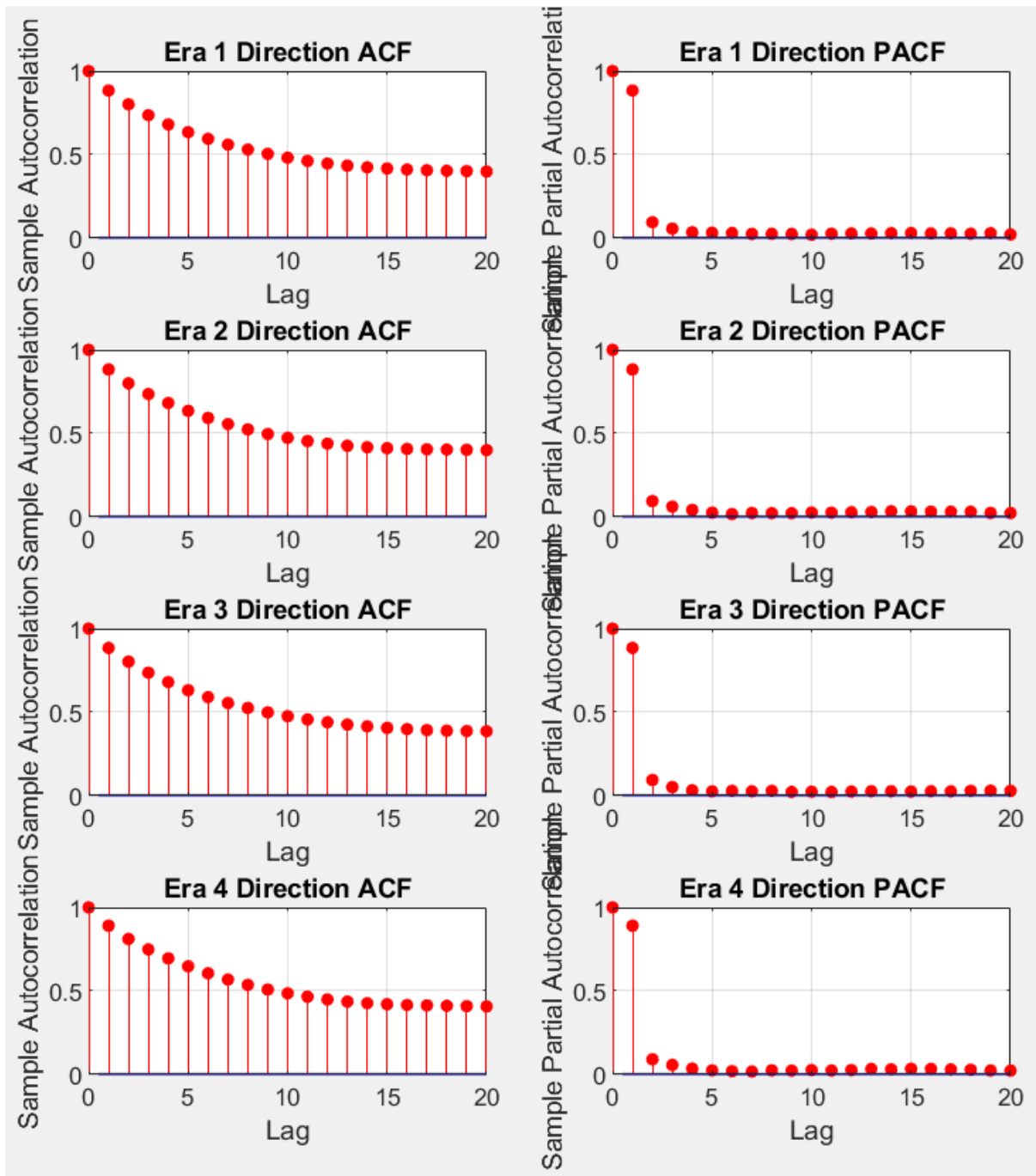
Εικόνα 9. Αυτοσυσχέτιση και Μερική Αυτοσυσχέτιση για την μεταβλητή  $u$ .



Εικόνα 10. Αυτοσυσχέτιση και Μερική Αυτοσυσχέτιση για την μεταβλητή  $v$ .



**Εικόνα 11.** Αυτοσυσχέτιση και Μερική Αυτοσυσχέτιση για την μεταβλητή speed (πλάτος ανέμου).



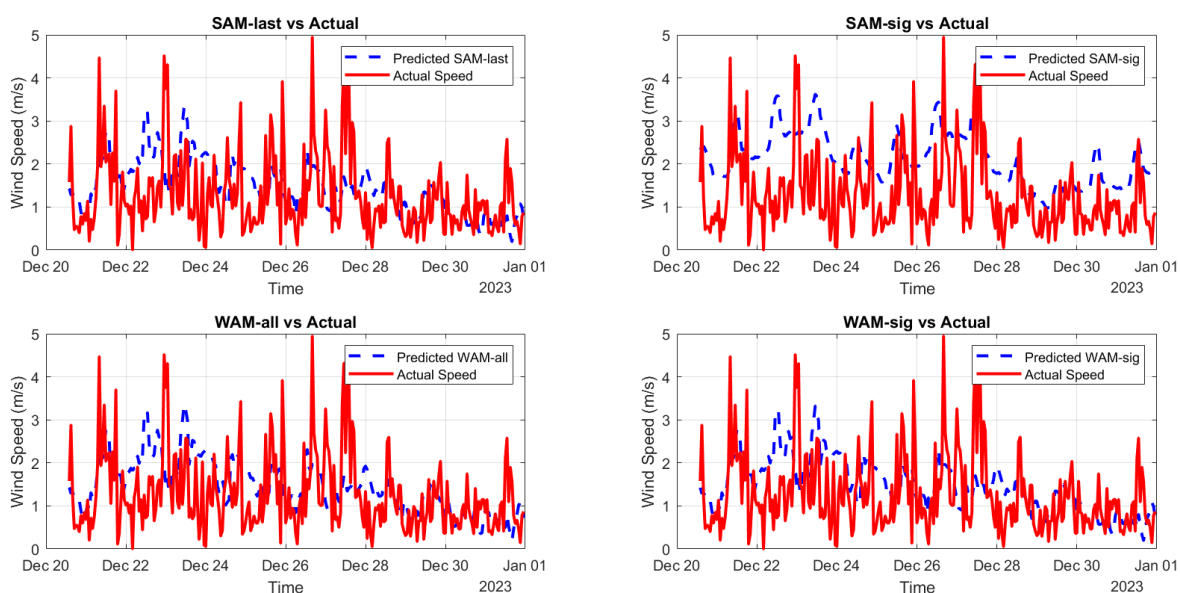
Εικόνα 12. Αυτοσυσχέτιση και Μερική Αυτοσυσχέτιση για την μεταβλητή direction (διεύθυνση ανέμου).

- 1) **Ισχυρή Αυτοσυσχέτιση:** Όλες οι περιοχές παρουσιάζουν ισχυρή αυτοσυσχέτιση στην ταχύτητα του ανέμου, όπως υποδεικνύεται από σημαντικές αυτοσυσχετίσεις μέχρι και την τιμή της υστέρησης (lag) 10.
- 2) **Συνέπειες μοντελοποίησης:** Τα διαγράμματα PACF υποδηλώνουν ότι ένα αυτοπαλινδρομικό μοντέλο (autoregressive model) τάξης 1 ή 2 ( $AR(1)$  ή  $AR(2)$ ) μπορεί να είναι κατάλληλο για τη μοντελοποίηση των δεδομένων ταχύτητας ανέμου σε κάθε εποχή.

Συνολικά, αυτά τα ευρήματα υποδεικνύουν ότι η ταχύτητα του ανέμου σε μια δεδομένη χρονική στιγμή επηρεάζεται σε μεγάλο βαθμό από τις τιμές του πρόσφατου παρελθόντος και τα αυτοπαλινδρομικά μοντέλα θα μπορούσαν να συλλάβουν αποτελεσματικά αυτή τη χρονική εξάρτηση.

### 3.1.4. Ντετερμινιστικά Μοντέλα

Στη συνέχεια εφαρμογής των τεσσάρων ντετερμινιστικών μοντέλων, σχεδιάσαμε τις καμπύλες πρόβλεψης και πραγματικών τιμών για κάθε μοντέλο, όπως φαίνεται στην *Εικόνα 13*.



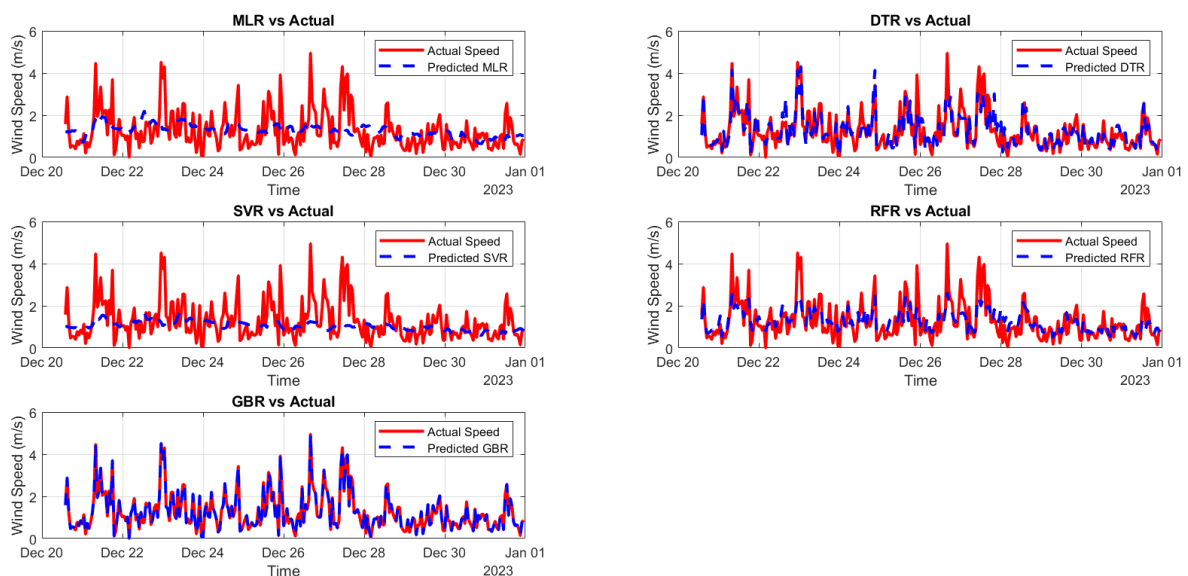
**Εικόνα 13.** Πραγματικές και προβλεπόμενες τιμές ταχύτητας ανέμου στην τοποθεσία-στόχο με χρήση ντετερμινιστικών μοντέλων.

Παρατηρείται μη βέλτιση ευστοχία των ντετερμινιστικών μοντέλων, με τις καμπύλες πρόβλεψης να μην ταιριάζουν με τις πραγματικές καμπύλες. Ωστόσο, και τα τέσσερα ντετερμινιστικά μοντέλα παρέχουν μια γρήγορη και μικρής πολυπλοκότητας εκτίμηση της ταχύτητας του ανέμου στο Λιμάνι της Χαλκίδας, παρουσιάζοντας σχετική ευστοχία στην πρόβλεψη της τάσης (trend) ή της κλίσης των

ανεμολογικών σειρών, με χαμηλή όμως ακρίβεια πρόβλεψης της πραγματικής τιμής της ταχύτητας ανέμου. Η μη βέλτιστη απόδοση των μοντέλων αυτών δικαιολογείται απόλυτα, δεδομένου ότι η κατασκευή τους στηρίζεται σε παραδοχές μέσου όρου και γραμμικής σχέσης μεταξύ των ανεμολογικών σειρών στο Λίμανι και αυτών στα σημεία αναφοράς. Τέτοιες παραδοχές παρέχουν μια καλή εκτίμηση, αλλά σε ορισμένες περιπτώσεις –όπως στην περίπτωση μας– αποτυγχάνουν να εκτιμήσουν μη γραμμικές και πολύπλοκες σχέσεις μεταξύ εισόδων-εξόδου και, τελικά, να προβλέψουν με ακρίβεια την μεταβλητή-στόχο.

### 3.1.5. Μοντέλα Μηχανικής Μάθησης

Εκπαιδευόντας και τα 5 μοντέλα, σχεδιάσαμε τις καμπύλες πρόβλεψης, όπως φαίνεται στην *Εικόνα 14*. Παρατηρείται ότι όλα τα μοντέλα μηχανικής μάθησης πετυχαίνουν μεγάλη ευστοχία στην πρόβλεψη της ταχύτητας ανέμου, με τα μοντέλα DTR και GBR να δείχνουν υψηλότερης ακρίβειας προβλέψεις.

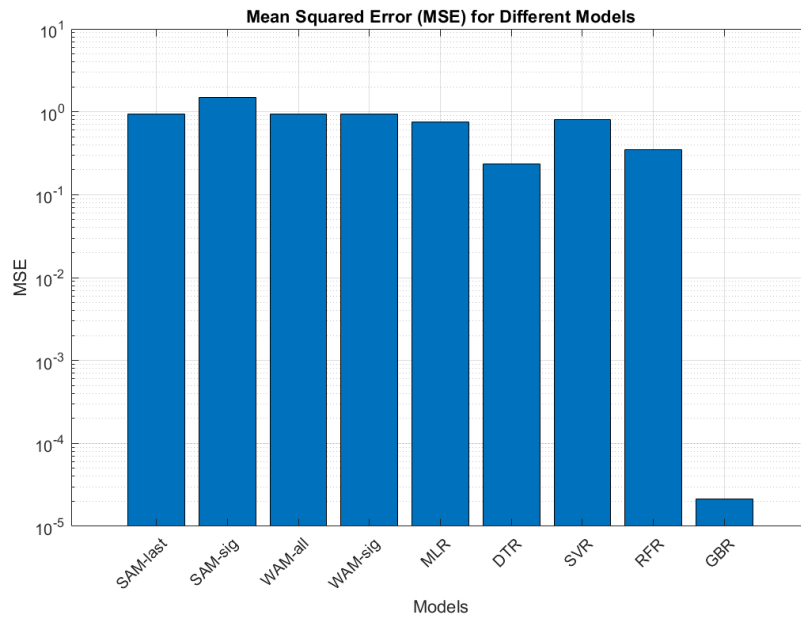


**Εικόνα 14.** Πραγματικές και προβλεπόμενες τιμές ταχύτητας ανέμου στην τοποθεσία-στόχο με χρήση μοντέλων μηχανικής μάθησης.

### 3.1.6. Σύγκριση Ευστοχίας μεταξύ Μοντέλων

Στην ενότητα αυτή παρουσιάζεται η συγκριτική αποτίμηση ως προς της πρόβλεψη που πετυχαίνουν όλα τα μοντέλα που χρησιμοποιήθηκαν στην παρούσα τεχνική έκθεση. Η μετρική αξιολόγησης για

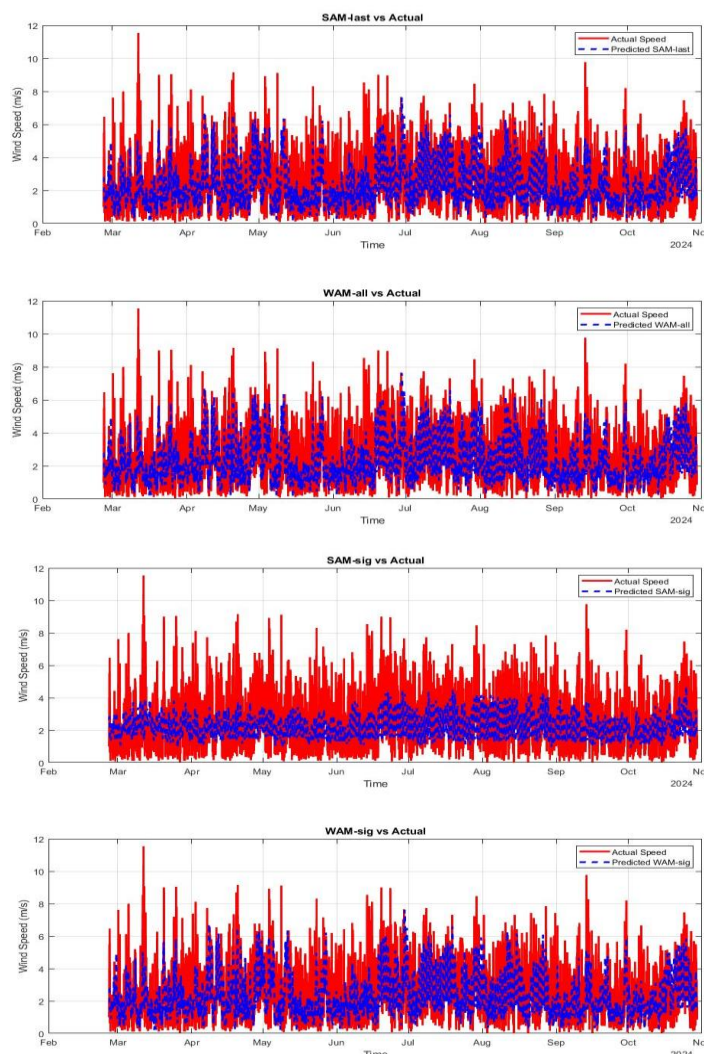
κάθε μοντέλο είναι το μέσο τετραγωνικό σφάλμα μεταξύ προβλέψεων και πραγματικών τιμών. Η Εικόνα 15 δείχνει το μέσο τετραγωνικό σφάλμα (MSE) καθεμιάς μεθόδου.



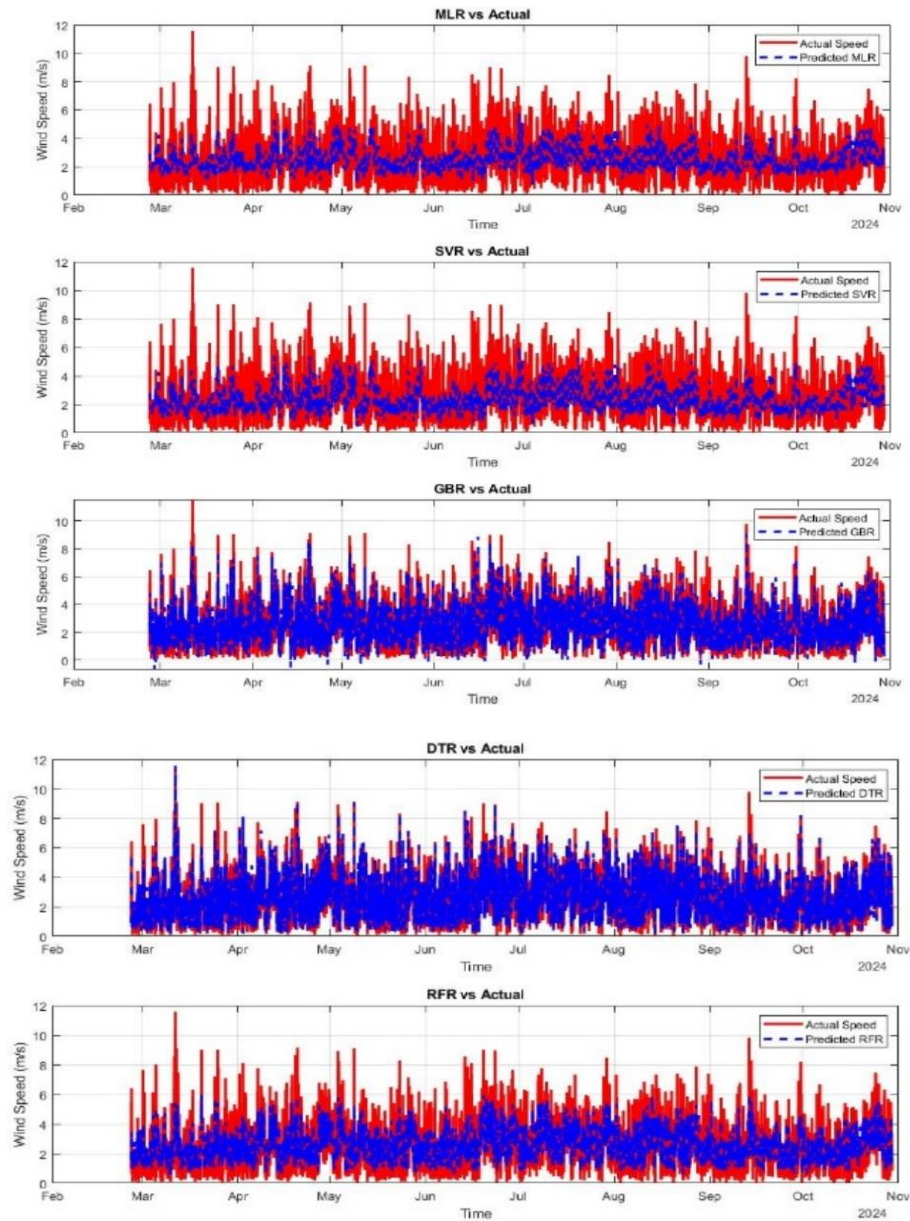
**Εικόνα 15.** Μέσο Τετραγωνικό Σφάλμα διαφορετικών μοντέλων πρόβλεψης ταχύτητας ανέμου στο Λιμάνι της Χαλκίδας.

### 3.2. Ανάλυση Μακράς Περιόδου

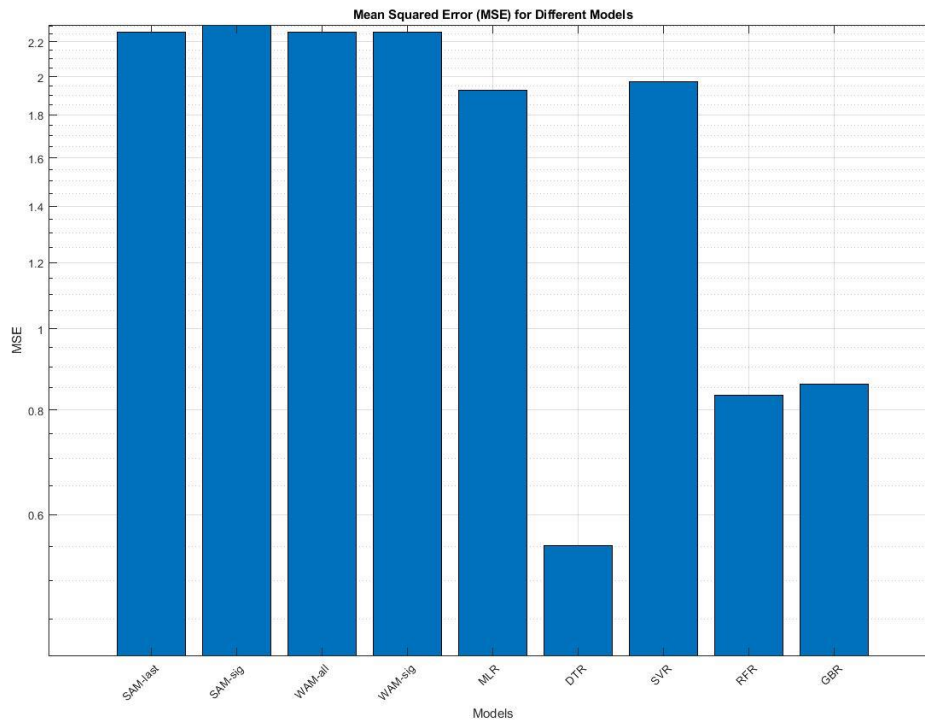
Ακολουθώντας ακριβώς την ίδια μεθοδολογία με αυτή των Ενοτήτων 3.1, αυτή η Ενότητα εστιάζει στην αξιολόγηση των προβλεπτικών μοντέλων στην (μακρά) περίοδο ενδιαφέροντος 26 Φεβρουαρίου 2024 έως 29 Οκτωβρίου. Ο λόγος που επεκτείνεται το εύρος πρόβλεψης από μερικές μέρες (όπως στην Ενότητα 3.2) σε μήνες (συγκεκριμένα, 8 μήνες και 3 μέρες) είναι να αποκαλυφθεί ο βαθμός στον οποίο τα προτεινόμενα μοντέλα προσφέρουν ακριβείς προβλέψεις (και) σε μεγαλύτερες περιόδους. Για το σκοπό αυτό, αναπτύσσονται ακριβώς οι ίδιες τεχνικές με την μόνη διαφοροποίηση να είναι το εύρος της πρόβλεψης. Παρακάτω, παρουσιάζονται τα αποτελέσματα των μοντέλων μέσου όρου, καθώς των μοντέλων μηχανικής μάθησης (Εικόνες 16, 17), ενώ συνοψίζονται και τα μέσα τετραγωνικά σφάλματα όλων των μεθόδων για συγκριτικούς σκοπούς (Εικόνα 18).



**Εικόνα 16.** Πραγματικές και προβλεπόμενες τιμές ταχύτητας ανέμου κατά την περίοδο 26/02/2024 έως 29/10/2024 στην τοποθεσία-στόχο με χρήση ντετερμινιστικών μοντέλων.

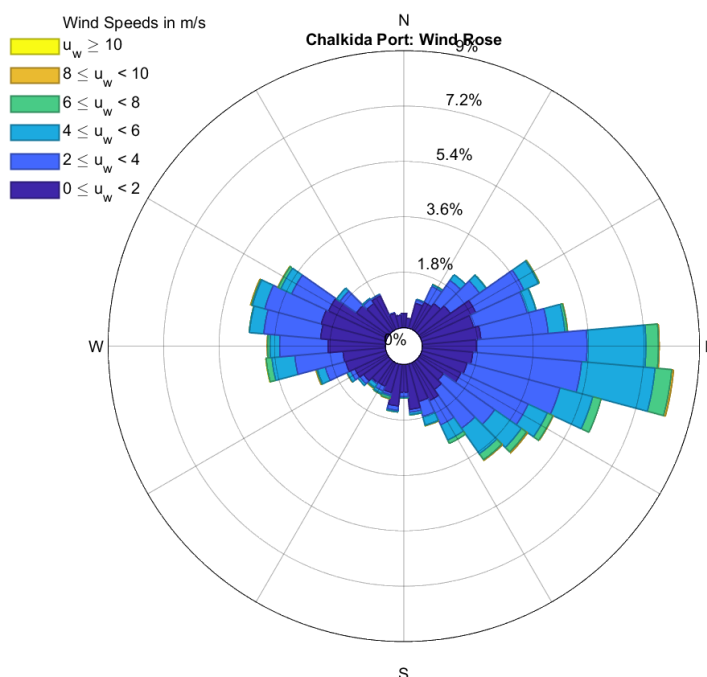


**Εικόνα 17.** Πραγματικές και προβλεπόμενες τιμές ταχύτητας ανέμου κατά την περίοδο 26.02.2024 έως 29.10.2024 στην τοποθεσία – στόχο με χρήση μοντέλων μηχανικής μάθησης.



**Εικόνα 18.** Μέσο Τετραγωνικό Σφάλμα διαφορετικών μοντέλων πρόβλεψης ταχύτητας ανέμου κατά τη μακρά περίοδο στο Λιμάνι Χαλκίδας.

Είναι σημαντικό να τονιστεί ότι, κατά την εφαρμογή της προτεινόμενης μεθοδολογίας, η απόδοση της προβλεπτικής ικανότητας των μοντέλων μπορεί να μεταβληθεί ανάλογα με την περίοδο ενδιαφέροντος που λαμβάνεται υπόψιν. Πράγματι, όπως φάνηκε στην Ενότητα 3.1, για βραχυχρόνιες προβλέψεις του ανέμου, το μοντέλο GBR έδειξε την μεγαλύτερη ακρίβεια προβλέψεων (ισοδύναμα, το μικρότερο MSE), γεγονός που το καθιστά κατάλληλο για πρόβλεψη του ανέμου όταν τα δεδομένα ενδιαφέροντος είναι μικρής διάρκειας. Αντίθετα, η εφαρμογή στη μακρά περίοδο της παρούσας ενότητας υποδεικνύει ότι το ακριβέστερο μοντέλο είναι το DTR, με μέσο τετραγωνικό σφάλμα μικρότερο του 0.6. Με βάση τα ανωτέρω, είναι σημαντικό να τονιστεί ότι η προτεινόμενη μέθοδος είναι γενικού σκοπού και μπορεί να εφαρμοστεί για οποιαδήποτε περίοδο πρόβλεψης των ανεμολογικών δεδομένων, υπό τον αστερίσκο ότι το δύναται να μεταβάλλεται το ακριβέστερο μοντέλο. Επιπλέον, δεν είναι δυνατό να γνωρίζει κανείς ποιο μοντέλο θα είναι ακριβέστερο εκ των προτέρων, δεδομένου ότι η αναπτυχθείσα μεθοδολογία εξαρτάται αμιγώς από τα εκάστοτε δεδομένα (data-driven method).



Εικόνα 19. Πολική αναπαράσταση ανέμου στο Λιμάνι Χαλκίδας.

Ως προς την υψηλότερη ικανότητα πρόβλεψης του ανέμου στο λιμάνι της Χαλκίδας που επιδεικνύουν τα μοντέλα μηχανικής μάθησης (έναντι των μοντέλων μέσου όρου), είναι σημαντικό να σημειωθεί η διαφορά προσέγγισης που ακολουθούν. Όπως φαίνεται στην *Εικόνα 19*, η πολική αναπαράσταση της κατανομής του ανέμου στο λιμάνι ενδιαφέροντος είναι εμφανώς διαφορετική από αυτή που φαίνεται στα τέσσερα σημεία αναφοράς (βλ. *Εικόνα 5*). Το γεγονός αυτό αποδεικνύει ότι γραμμικά μοντέλα, όπως αυτά του μέσου όρου ή αυτό της γραμμικής παλινδρόμησης, δεν είναι ικανά να ανακατασκευάσουν με ευστοχία τα δεδομένα στην περιοχή ενδιαφέροντος, δεδομένου ότι η σχέση που συνδέει τα μοτίβα του ανέμου στα σημεία αναφοράς με αυτά στο Λιμάνι της Χαλκίδας είναι μη γραμμική. Αυτό μπορεί να οφείλεται σε πολύπλοκα μορφολογικά ή καιρικά φαινόμενα που διαφοροποιούν τον άνεμο σε διαφορετικές περιοχές, όπως π.χ. οι διαφορές υψομέτρου, η θερμοκρασία, το κλίμα ή/και η ύπαρξη θάλασσας πλησίον του Λιμανιού. Με βάση τα ανωτέρω, η υψηλότερη ακρίβεια ανακατασκευής του ανέμου παρέχεται από μη γραμμικά μοντέλα πρόβλεψης, όπως το DTR, γεγονός που επαληθεύει ότι η πρόβλεψη του ανέμου σε συγκεκριμένες περιοχές απαιτεί μη γραμμικούς συνδυασμούς γνωστών δεδομένων από περιοχές αναφοράς. Σε τέτοιες περιπτώσεις, τα μοντέλα μηχανικής μάθησης είναι ικανότερα να αποτυπώσουν τα μοτίβα του ανέμου, όπως αποδεικνύεται και στα αποτελέσματα της παρούσας ανάλυσης. Τέλος, περισσότερα δεδομένα και αναλύσεις από την ίδια περιοχή ή/και από άλλες περιοχές που επιχειρείται



ανακατασκευή ανέμου είναι δυνατό να ισχυροποιήσουν περαιτέρω τα ευρήματα του παρόντος ή να φανερώσουν με υψηλότερη ακρίβεια τα μοτίβα του ανέμου.