



University of
East London

Διπλωματική Εργασία

Τοπική βραχυπρόθεσμη πρόγνωση καιρού με machine learning και deep learning

Κυριάκος Κάρτας (AM: u2977385)

Επιβλέπων: Αντώνιος Αρχοντής

Αθήνα, Μάιος 2026

Δήλωση Συγγραφέα Διπλωματικής Εργασίας

Ο κάτωθι υπογεγραμμένος Κυριάκος Κάρτας με αριθμό μητρώου u2977385 φοιτητής του **UEL** δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών

Κυριάκος Κάρτας

Ευχαριστίες – Αφιερώσεις

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες σε όλους όσοι με στήριξαν κατά τη διάρκεια εκπόνησης της παρούσας διπλωματικής εργασίας.

Αρχικά, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κ. Αρχοντή, για την καθοδήγηση, την υποστήριξη και τις πολύτιμες παρατηρήσεις του σε όλη τη διάρκεια της εργασίας. Η συμβολή του υπήρξε ιδιαίτερα σημαντική τόσο για την επιστημονική διαμόρφωση της εργασίας όσο και για την ολοκλήρωσή της.

Ιδιαίτερες ευχαριστίες οφείλω στη σύντροφό μου, Δέσποινα, για την αμέριστη στήριξη, την υπομονή και την κατανόησή της καθ' όλη τη διάρκεια αυτής της προσπάθειας. Η παρουσία της, η ενθάρρυνση και η πίστη της σε εμένα αποτέλεσαν σημαντική πηγή δύναμης.

Τέλος, θα ήθελα να ευχαριστήσω βαθιά τους γονείς μου, για τη συνεχή αγάπη, στήριξη και εμπιστοσύνη τους σε κάθε στάδιο της ακαδημαϊκής και προσωπικής μου πορείας. Η συμπαράστασή τους υπήρξε καθοριστική και τους αφιερώνω με ευγνωμοσύνη ένα σημαντικό μέρος αυτής της προσπάθειας.

26/05/2026

Κυριάκος Κάρτας

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει τη δυνατότητα τοπικής βραχυπρόθεσμης πρόβλεψης της θερμοκρασίας αέρα με χρήση τεχνικών μηχανικής μάθησης και βαθιάς μάθησης, αξιοποιώντας αποκλειστικά ιστορικές παρατηρήσεις από έναν μεμονωμένο μετεωρολογικό σταθμό Weathercloud. Το πρόβλημα διαμορφώνεται ως αυστηρά τοπική πρόβλεψη από παρατηρήσεις ενός μόνο σταθμού, χωρίς χρήση δεδομένων αριθμητικής πρόγνωσης καιρού, δορυφορικών δεδομένων, δεδομένων επανανάλυσης ή παρατηρήσεων από πολλαπλούς σταθμούς.

Για την υλοποίηση της μελέτης αναπτύχθηκε επαναλήψιμο πειραματικό λογισμικό σε Python. Τα πρωτογενή δεδομένα από αρχεία Weathercloud CSV μετατράπηκαν σε ενιαίο σχήμα δεδομένων και προετοιμάστηκαν μέσω του MetDataPy, το οποίο χρησιμοποιήθηκε για κανονικοποίηση χρονικών σημάνσεων, ποιοτικό έλεγχο, διαχείριση χρονικών κενών, παραγωγή παράγωγων μετεωρολογικών χαρακτηριστικών, χαρακτηριστικών χρονικής υστέρησης, κυλιόμενων στατιστικών και στόχων πρόβλεψης. Το κύριο πείραμα επικεντρώθηκε στην πρόβλεψη της θερμοκρασίας αέρα για έξι χρονικούς ορίζοντες: 10 λεπτά, 1 ώρα, 3 ώρες, 6 ώρες, 12 ώρες και 24 ώρες.

Τα δεδομένα διαχωρίστηκαν χρονολογικά σε σύνολα εκπαίδευσης, επικύρωσης και δοκιμής, ενώ η κλιμάκωση των χαρακτηριστικών εφαρμόστηκε με βάση μόνο το σύνολο εκπαίδευσης, ώστε να αποφευχθεί η διαρροή πληροφορίας από μεταγενέστερες χρονικές περιόδους. Αξιολογήθηκαν 11 μοντέλα από τρεις οικογένειες: απλά μοντέλα αναφοράς, παραδοσιακά μοντέλα μηχανικής μάθησης και Deep Learning μοντέλα. Η αξιολόγηση πραγματοποιήθηκε με τις μετρικές MAE, RMSE, MAPE και persistence skill score.

Τα αποτελέσματα δείχνουν ότι η πρόβλεψη θερμοκρασίας από έναν μόνο τοπικό μετεωρολογικό σταθμό είναι εφικτή κυρίως στους ορίζοντες από 1 έως 12 ώρες. Στον ορίζοντα των 10 λεπτών, το μοντέλο persistence παρέμεινε ιδιαίτερα ισχυρό, με οριακή αριθμητική υπεροχή του Ridge. Στους ορίζοντες της 1 και των 3 ωρών επικράτησε το Random Forest, ενώ στους ορίζοντες των 6, 12 και 24 ωρών το Gradient Boosting πέτυχε την καλύτερη απόδοση. Στο συγκεκριμένο dataset και υπό το συγκεκριμένο training budget, τα μοντέλα Deep Learning δεν ξεπέρασαν τα ισχυρότερα tree-based Machine Learning μοντέλα.

Συνολικά, η εργασία δείχνει ότι οι ιστορικές τοπικές παρατηρήσεις περιέχουν αξιοποιήσιμη πληροφορία για βραχυπρόθεσμη πρόβλεψη θερμοκρασίας. Παράλληλα, η μείωση της βελτίωσης έναντι του μοντέλου persistence στους μεγαλύτερους χρονικούς ορίζοντες αναδεικνύει τα πρακτικά όρια μιας καθαρά τοπικής προσέγγισης χωρίς εξωτερικές προγνωστικές πηγές.

Λέξεις-Κλειδιά: βραχυπρόθεσμη πρόγνωση καιρού, μηχανική μάθηση, βαθιά μάθηση, μετεωρολογικές χρονοσειρές, τοπική πρόβλεψη καιρού, μετεωρολογικός σταθμός, πρόβλεψη θερμοκρασίας

Abstract

This dissertation investigates local short-term air temperature forecasting using machine learning and deep learning techniques, based exclusively on historical observations from a single Weathercloud meteorological station. The forecasting problem is formulated as a strictly observation-only and single-station task, without the use of Numerical Weather Prediction outputs, satellite data, reanalysis datasets, or observations from multiple stations.

A reproducible experimental pipeline was developed in Python to support the full forecasting workflow. Raw Weathercloud CSV data were converted into a unified canonical schema and prepared through MetDataPy, which was used for timestamp normalization, quality-control flagging, gap handling, derived meteorological features, lag features, rolling features, and forecast horizon creation. The main experiment focused on air temperature prediction across six forecast horizons: 10 minutes, 1 hour, 3 hours, 6 hours, 12 hours, and 24 hours.

The dataset was split chronologically into training, validation, and test sets, and feature scaling was fitted only on the training set to prevent data leakage. A total of 11 models were evaluated across three model families: baseline models, traditional Machine Learning models, and Deep Learning models. Model performance was assessed using MAE, RMSE, MAPE, and persistence skill score.

The results indicate that useful temperature forecasting from a single local station is feasible mainly for horizons between 1 and 12 hours. At the 10-minute horizon, the persistence baseline remained highly competitive, with Ridge showing only marginal numerical improvement. Random Forest achieved the best performance at the 1-hour and 3-hour horizons, while Gradient Boosting achieved the best results at the 6-hour, 12-hour, and 24-hour horizons. The evaluated Deep Learning models did not outperform the strongest tree-based Machine Learning models in this experimental setting.

Overall, the study shows that local historical observations contain useful predictive information for short-term temperature forecasting. However, the decreasing improvement over the persistence baseline at longer horizons highlights the practical limits of a purely local observation-only approach when no external forecasting sources are used.

Keywords: short-term weather forecasting, machine learning, deep learning, meteorological time series, single-station forecasting, observation-only forecasting, temperature prediction

Πίνακας Περιεχομένων

Δήλωση Συγγραφέα Διπλωματικής Εργασίας	ii
Ευχαριστίες – Αφιερώσεις	iii
Περίληψη	iv
Abstract.....	v
Πίνακας Περιεχομένων.....	vi
Ευρετήριο Πινάκων	ix
Ευρετήριο Σχημάτων.....	x
Κεφάλαιο 1. Εισαγωγή	1
1.1 Πλαίσιο, σκοπός και στόχοι της εργασίας	1
1.2 Μεθοδολογία	2
1.3 Περιορισμοί	2
1.4 Ορισμοί.....	3
1.5 Οργάνωση Κεφαλαίων.....	4
Κεφάλαιο 2. Βιβλιογραφική έρευνα – Σχετικές προσπάθειες	5
2.1 Εισαγωγή	5
2.2 Μεθοδολογία Βιβλιογραφικής Αναζήτησης.....	5
2.2.1 Πηγές Αναζήτησης.....	5
2.2.2 Λέξεις-Κλειδιά Αναζήτησης	6
2.2.3 Κριτήρια Επιλογής Άρθρων	6
2.2.4 Διαδικασία Φιλτραρίσματος.....	6
2.2.5 Μεθοδολογία Ανάλυσης Άρθρων	7
2.3 Κατηγοριοποίηση των Ερευνητικών Προσεγγίσεων Πρόγνωσης Καιρού	7
2.3.1 Κατηγοριοποίηση με βάση τον Χρονικό Ορίζοντα Πρόγνωσης	7
2.3.2 Κατηγοριοποίηση με βάση τη Γεωγραφική Κλίμακα Πρόγνωσης	8
2.3.3 Κατηγοριοποίηση με βάση τη Μεθοδολογία Πρόγνωσης	9

2.4	Κατηγοριοποίηση των Χρησιμοποιούμενων Μοντέλων Πρόγνωσης	11
2.4.1	Στατιστικά Μοντέλα	11
2.4.2	Μοντέλα Μηχανικής Μάθησης	11
2.4.3	Μοντέλα Βαθιάς Μάθησης	12
2.5	Κατηγοριοποίηση των Δεδομένων και Τρόπων Συλλογής	13
2.5.1	Δεδομένα Μετεωρολογικών Σταθμών	14
2.5.2	Δεδομένα Numerical Weather Prediction (NWP).....	14
2.5.3	Δεδομένα Δορυφόρων	14
2.5.4	Συνδυασμός Πολλαπλών Πηγών Δεδομένων	15
2.6	Σύνοψη Βιβλιογραφικής Ανασκόπησης και Κύρια Ευρήματα	16
2.6.1	Ευρήματα ως προς τον Χρονικό Ορίζοντα Πρόγνωσης	16
2.6.2	Ευρήματα ως προς τη Γεωγραφική Κλίμακα Πρόγνωσης	16
2.6.3	Ευρήματα ως προς τα Χρησιμοποιούμενα Μοντέλα	16
2.6.4	Ευρήματα ως προς τα Δεδομένα	17
2.6.5	Συνολική Αποτύπωση των Ευρημάτων.....	17
2.7	Ερευνητικό Κενό (Research Gap).....	17
2.8	Ερευνητικές Υποθέσεις.....	18
Κεφάλαιο 3. Μεθοδολογία		20
3.1	Εισαγωγή στη Μεθοδολογία	20
3.2	Συνολική Αρχιτεκτονική Συστήματος.....	22
3.3	Συλλογή Δεδομένων	25
3.4	Προεπεξεργασία Δεδομένων	25
3.5	Δημιουργία Χαρακτηριστικών (Feature Engineering)	27
3.6	Δημιουργία Πειραματικών Συνόλων Δεδομένων	30
3.7	Μοντέλα Πρόβλεψης.....	31
3.7.1	Baseline Models	31
3.7.2	Μοντέλα Machine Learning.....	32
3.7.3	Μοντέλα Deep Learning.....	32
3.8	Εκπαίδευση Μοντέλων	33

3.9	Αξιολόγηση Μοντέλων	34
3.9.1	Μετρικές Αξιολόγησης.....	34
3.9.2	Συγκριτική Αξιολόγηση Μοντέλων	35
Κεφάλαιο 4. Αποτελέσματα		36
4.1	Παρουσίαση του Λογισμικού.....	36
4.2	Σχόλια και Ανατροφοδότηση.....	39
4.3	Πειραματική Διάταξη και Δεδομένα Αξιολόγησης	40
4.4	Συγκεντρωτικά Αποτελέσματα Πρόβλεψης	44
4.5	Αναλυτική Αξιολόγηση ανά Οικογένεια Μοντέλων.....	46
4.5.1	Baseline Models	47
4.5.2	Μοντέλα Machine Learning.....	48
4.5.3	Μοντέλα Deep Learning.....	50
4.5.4	Συγκριτική Ερμηνεία	51
4.6	Οπτικοποίηση και Διαγνωστική Ανάλυση Αποτελεσμάτων	52
4.7	Παραγόμενα Artifacts και Πρακτικά Αποτελέσματα Χρήσης του Λογισμικού	58
Κεφάλαιο 5. Συζήτηση - Συμπεράσματα - Μελλοντικές Επεκτάσεις.....		61
5.1	Ανακεφαλαίωση.....	61
5.2	Συζήτηση και Συμπεράσματα.....	62
5.3	Αξιοποίηση και Πρακτικές Προεκτάσεις της Έρευνας.....	64
5.4	Περιορισμοί και Μελλοντικές Επεκτάσεις	65
5.5	Τελικό Συμπέρασμα.....	68
Βιβλιογραφία		70
Παράρτημα Α – Αναπαραγωγή Πειραματικής Εκτέλεσης		72

Ευρετήριο Πινάκων

Πίνακας 4.1 - Σύνοψη τεχνικής αξιολόγησης του λογισμικού ως προς τις μεθοδολογικές απαιτήσεις της εργασίας.	39
Πίνακας 4.2 - Σύνοψη αρχικού και προεπεξεργασμένου dataset.	41
Πίνακας 4.3 - Χρονικοί ορίζοντες πρόβλεψης που αξιολογήθηκαν στο πείραμα αναφοράς. 41	
Πίνακας 4.4 - Οικογένειες μοντέλων και επιμέρους μοντέλα που αξιολογήθηκαν στο πείραμα αναφοράς, με σύντομη περιγραφή κάθε μοντέλου και τον ρόλο που έπαιξε στη συνολική σύγκριση.	43
Πίνακας 4.5 - Καλύτερο μοντέλο ανά χρονικό ορίζοντα με MAE, RMSE και persistence skill score.	44
Πίνακας 4.6 - Απόδοση baseline μοντέλων ανά ορίζοντα με MAE, RMSE και skill score.	48
Πίνακας 4.7 - Απόδοση μοντέλων Machine Learning ανά χρονικό ορίζοντα.	50
Πίνακας 4.8 - Απόδοση μοντέλων Deep Learning ανά ορίζοντα με MAE, RMSE, epochs trained και best validation loss.	51
Πίνακας 4.9 - Ενδεικτικά χρονικά όρια splits ανά χρονικό ορίζοντα στο πείραμα αναφοράς. 59	
Πίνακας 4.10 - Σύνοψη βασικών παραγόμενων αρχείων στο πείραμα αναφοράς.	59
Πίνακας 5.1 - Κύρια ευρήματα της εργασίας.	63
Πίνακας 5.2 - Πρακτικές και ερευνητικές προεκτάσεις της εργασίας.	65
Πίνακας 5.3 - Περιορισμοί της εργασίας και πιθανές μελλοντικές επεκτάσεις.	67

Ευρετήριο Σχημάτων

Σχήμα 2.1 - Κατανομή των μελετών ανά χρονικό ορίζοντα πρόγνωσης.	8
Σχήμα 2.2 - Κατανομή των μελετών ανά γεωγραφική κλίμακα.	9
Σχήμα 2.3 - Κατανομή των μελετών ανά οικογένεια μοντέλου.	10
Σχήμα 2.4 - Κατανομή των αρχιτεκτονικών βαθιάς μάθησης στη βιβλιογραφία.	13
Σχήμα 2.5 - Κατανομή των μελετών ανά πηγή δεδομένων.	15
Σχήμα 3.1 - Αρχιτεκτονική του Προτεινόμενου Συστήματος Πρόβλεψης Καιρού.	21
Σχήμα 3.2 - Πειραματικό Workflow της Προτεινόμενης Μεθοδολογίας.	24
Σχήμα 4.1 - Συνολική ροή του λογισμικού από raw Weathercloud CSV files έως metrics, plots και run snapshot.	37
Σχήμα 4.2 - Σύγκριση MAE του καλύτερου μοντέλου και του persistence baseline ανά χρονικό ορίζοντα.	45
Σχήμα 4.3 - Μεταβολή σφάλματος ανά ορίζοντα.	46
Σχήμα 4.4 - Σύγκριση καλύτερου μοντέλου ανά οικογένεια.	47
Σχήμα 4.5 - Διαγνωστική σύγκριση MAE όλων των μοντέλων ανά χρονικό ορίζοντα.	52
Σχήμα 4.6 - Σύγκριση RMSE όλων των μοντέλων ανά χρονικό ορίζοντα.	53
Σχήμα 4.7 - Μεταβολή MAE ανά χρονικό ορίζοντα για κάθε μοντέλο.	54
Σχήμα 4.8 - Heatmap persistence skill score ανά μοντέλο και χρονικό ορίζοντα.	54
Σχήμα 4.9 - Σύγκριση καλύτερου baseline, καλύτερου Machine Learning και καλύτερου Deep Learning μοντέλου ανά χρονικό ορίζοντα.	55
Σχήμα 4.10 - Scatter plot πραγματικών και προβλεπόμενων τιμών για το Gradient Boosting στον ορίζοντα των 12 ωρών.	56
Σχήμα 4.11 - Time-series plot πραγματικής και προβλεπόμενης θερμοκρασίας για το Gradient Boosting στον ορίζοντα των 12 ωρών.	57
Σχήμα 4.12 - Residual diagnostics για το Gradient Boosting στον ορίζοντα των 12 ωρών. .	57

Κεφάλαιο 1. Εισαγωγή

1.1 Πλαίσιο, σκοπός και στόχοι της εργασίας

Η πρόγνωση καιρού αποτελεί ένα από τα σημαντικότερα πεδία εφαρμογής της ανάλυσης χρονοσειρών, της μηχανικής μάθησης και της βαθιάς μάθησης. Η ακριβής εκτίμηση μετεωρολογικών μεταβλητών, ακόμη και για μικρούς χρονικούς ορίζοντες, μπορεί να υποστηρίξει πρακτικές αποφάσεις σε τομείς όπως η γεωργία, η ενέργεια, οι μεταφορές, η καθημερινή δραστηριότητα και η τοπική διαχείριση κινδύνων. Παραδοσιακά, η πρόγνωση καιρού βασίζεται σε αριθμητικά μοντέλα πρόγνωσης καιρού, τα οποία αξιοποιούν φυσικές εξισώσεις και δεδομένα μεγάλης χωρικής κλίμακας. Ωστόσο, η ανάπτυξη τεχνικών μηχανικής μάθησης έχει δημιουργήσει νέες δυνατότητες για την αξιοποίηση ιστορικών παρατηρήσεων με σκοπό την εκτίμηση μελλοντικών τιμών.

Η παρούσα εργασία εντάσσεται στο πεδίο της τοπικής βραχυπρόθεσμης πρόγνωσης καιρού με χρήση δεδομένων από μετεωρολογικό σταθμό. Ειδικότερα, εξετάζεται κατά πόσο είναι εφικτή η πρόβλεψη της θερμοκρασίας αέρα σε τοπική κλίμακα, χρησιμοποιώντας αποκλειστικά ιστορικές παρατηρήσεις από έναν μεμονωμένο σταθμό Weathercloud. Η εργασία δεν χρησιμοποιεί αριθμητικά μοντέλα πρόγνωσης καιρού, δορυφορικά δεδομένα, reanalysis datasets ή παρατηρήσεις από πολλαπλούς σταθμούς. Με αυτόν τον τρόπο, το πρόβλημα παραμένει αυστηρά observation-only και single-station.

Σκοπός της εργασίας είναι η ανάπτυξη και αξιολόγηση ενός επαναλήψιμου πειραματικού πλαισίου για βραχυπρόθεσμη πρόβλεψη θερμοκρασίας, βασισμένου σε τεχνικές Machine Learning και Deep Learning. Η εργασία εξετάζει αν οι ιστορικές τοπικές παρατηρήσεις ενός σταθμού περιέχουν επαρκή πληροφορία για χρήσιμη πρόβλεψη σε χρονικούς ορίζοντες από 10 λεπτά έως 24 ώρες.

Οι βασικοί στόχοι της εργασίας είναι οι εξής:

- η μελέτη της σχετικής βιβλιογραφίας για πρόγνωση καιρού με χρήση μηχανικής μάθησης και βαθιάς μάθησης,
- η ανάπτυξη pipeline για εισαγωγή, προεπεξεργασία και μετατροπή μετεωρολογικών χρονοσειρών σε supervised learning datasets,
- η αξιοποίηση του MetDataPy ως επίσημου επιπέδου προετοιμασίας μετεωρολογικών δεδομένων,
- η εκπαίδευση και σύγκριση baseline, Machine Learning και Deep Learning μοντέλων,
- η αξιολόγηση των μοντέλων σε πολλαπλούς χρονικούς ορίζοντες,
- η διερεύνηση των ορίων μιας πρόγνωσης που βασίζεται αποκλειστικά σε τοπικές παρατηρήσεις,
- η παραγωγή τεκμηριωμένων artifacts, πινάκων μετρικών και διαγραμμάτων για την υποστήριξη της ανάλυσης.

Η εργασία οργανώνεται γύρω από ένα πρακτικό λογισμικό pipeline, το οποίο εκτελεί όλα τα βασικά στάδια του πειράματος: εισαγωγή δεδομένων, προεπεξεργασία, feature engineering, δημιουργία supervised datasets, χρονολογικό διαχωρισμό δεδομένων, εκπαίδευση μοντέλων, αξιολόγηση και παραγωγή αποτελεσμάτων.

1.2 Μεθοδολογία

Η μεθοδολογία της εργασίας περιλαμβάνει δύο βασικές κατευθύνσεις: βιβλιογραφική έρευνα και πειραματική υλοποίηση. Η βιβλιογραφική έρευνα χρησιμοποιήθηκε για την κατανόηση των σύγχρονων προσεγγίσεων πρόγνωσης καιρού με τεχνικές μηχανικής μάθησης και βαθιάς μάθησης. Μέσα από αυτήν εντοπίστηκαν βασικές κατηγορίες μοντέλων, διαφορετικές πηγές δεδομένων, χρονικοί ορίζοντες πρόγνωσης και κενά στη σχετική βιβλιογραφία.

Η πειραματική μεθοδολογία βασίστηκε στην ανάπτυξη λογισμικού pipeline σε Python. Τα πρωτογενή δεδομένα προέρχονται από εξαγωγές CSV της πλατφόρμας Weathercloud και αφορούν έναν τοπικό μετεωρολογικό σταθμό. Τα δεδομένα εισάγονται και μετατρέπονται σε ενιαίο canonical schema, με κοινές μονάδες μέτρησης και χρονική σήμανση σε UTC. Η προετοιμασία των μετεωρολογικών χρονοσειρών γίνεται μέσω του MetDataPy, το οποίο χρησιμοποιείται για timestamp normalization, quality-control flagging, gap handling, derived meteorological features, lag features, rolling features και forecast horizon creation.

Στη συνέχεια, τα δεδομένα μετατρέπονται σε supervised learning μορφή. Για κάθε χρονικό ορίζοντα πρόβλεψης δημιουργείται ξεχωριστή στήλη στόχου, η οποία αντιστοιχεί στη μελλοντική τιμή της θερμοκρασίας. Οι χρονικοί ορίζοντες που εξετάζονται είναι 10 λεπτά, 1 ώρα, 3 ώρες, 6 ώρες, 12 ώρες και 24 ώρες.

Τα δεδομένα διαχωρίζονται χρονολογικά σε training, validation και test set. Δεν εφαρμόζεται τυχαία ανάμειξη, καθώς το πρόβλημα αφορά χρονοσειρές και η χρήση μελλοντικών πληροφοριών θα οδηγούσε σε data leakage. Η κλιμάκωση των χαρακτηριστικών εφαρμόζεται μετά τον διαχωρισμό, με τους scalers να προσαρμόζονται αποκλειστικά στο training set.

Για την αξιολόγηση εξετάζονται τρεις οικογένειες μοντέλων:

- baseline models: persistence, moving_average, climatology,
- παραδοσιακά μοντέλα Machine Learning: linear_regression, ridge, random_forest, gradient_boosting και support vector regression (SVR),
- μοντέλα Deep Learning: LSTM, GRU, TCN.

Η απόδοση των μοντέλων αξιολογείται στο test set με τις μετρικές MAE, RMSE, MAPE και persistence skill score. Η χρήση πολλαπλών μετρικών επιτρέπει πληρέστερη αξιολόγηση, καθώς το MAE παρέχει άμεση ερμηνεία του μέσου σφάλματος, το RMSE δίνει μεγαλύτερη βαρύτητα στα μεγάλα σφάλματα, ενώ το persistence skill score δείχνει αν ένα μοντέλο βελτιώνει ουσιαστικά την απλή persistence πρόβλεψη.

1.3 Περιορισμοί

Η παρούσα εργασία έχει συγκεκριμένους περιορισμούς, οι οποίοι καθορίζουν το πλαίσιο μέσα στο οποίο πρέπει να ερμηνευθούν τα αποτελέσματα.

Πρώτον, η εργασία βασίζεται σε δεδομένα από έναν μόνο μετεωρολογικό σταθμό. Η επιλογή αυτή είναι σκόπιμη, καθώς στόχος είναι η διερεύνηση ενός single-station forecasting setup. Ωστόσο, περιορίζει τη δυνατότητα γενίκευσης των συμπερασμάτων σε άλλες γεωγραφικές περιοχές ή σε διαφορετικά μικροκλίματα.

Δεύτερον, το κύριο πείραμα επικεντρώνεται στην πρόβλεψη της θερμοκρασίας αέρα. Άλλες μεταβλητές, όπως υγρασία, βροχόπτωση, πίεση ή ταχύτητα ανέμου, χρησιμοποιούνται ως είσοδοι, αλλά δεν αποτελούν κύριους στόχους πρόβλεψης στην παρούσα εκτέλεση.

Τρίτον, η εργασία δεν χρησιμοποιεί εξωτερικές προγνωστικές πηγές, όπως Numerical Weather Prediction outputs, δορυφορικά δεδομένα ή δεδομένα από γειτονικούς σταθμούς. Αυτό καθιστά το πρόβλημα πιο καθαρό μεθοδολογικά, αλλά περιορίζει την προβλεπτική ικανότητα σε μεγαλύτερους χρονικούς ορίζοντες.

Τέταρτον, η αξιολόγηση πραγματοποιείται offline, με ιστορικό train/validation/test διαχωρισμό. Δεν εξετάζεται πλήρες real-time deployment, συνεχής ενημέρωση μοντέλων ή παραγωγική λειτουργία σε πραγματικό περιβάλλον.

Τέλος, η εργασία αξιολογεί συγκεκριμένες οικογένειες μοντέλων και συγκεκριμένη παραμετροποίηση. Περαιτέρω βελτιστοποίηση υπερπαραμέτρων, περισσότερα random seeds ή πρόσθετες αρχιτεκτονικές θα μπορούσαν να δώσουν διαφορετικά αποτελέσματα.

1.4 Ορισμοί

Για την καλύτερη κατανόηση της εργασίας, παρατίθενται οι βασικοί όροι που χρησιμοποιούνται.

Μετεωρολογική χρονοσειρά: Ακολουθία μετεωρολογικών παρατηρήσεων ταξινομημένων χρονικά, όπως θερμοκρασία, υγρασία, πίεση ή ταχύτητα ανέμου.

Βραχυπρόθεσμη πρόγνωση: Πρόβλεψη μετεωρολογικής μεταβλητής για μικρό χρονικό ορίζοντα, από λίγα λεπτά έως μερικές ώρες ή μία ημέρα.

Observation-only πρόγνωση: Πρόγνωση που βασίζεται αποκλειστικά σε ιστορικές παρατηρήσεις και δεν χρησιμοποιεί εξωτερικά forecast products ή NWP δεδομένα.

Single-station πρόγνωση: Πρόγνωση που χρησιμοποιεί δεδομένα από έναν μόνο μετεωρολογικό σταθμό.

Forecast horizon: Το χρονικό διάστημα μεταξύ της στιγμής πρόβλεψης και της μελλοντικής τιμής που προβλέπεται.

Lag feature: Χαρακτηριστικό που αντιστοιχεί σε παρελθοντική τιμή μιας μεταβλητής, π.χ. η θερμοκρασία μία ώρα πριν.

Rolling feature: Χαρακτηριστικό που υπολογίζεται από κυλιόμενο χρονικό παράθυρο, όπως μέσος όρος, τυπική απόκλιση, ελάχιστη ή μέγιστη τιμή προηγούμενων παρατηρήσεων.

Baseline model: Απλό μοντέλο αναφοράς που χρησιμοποιείται για να αξιολογηθεί αν ένα πιο σύνθετο μοντέλο προσφέρει ουσιαστική βελτίωση.

Persistence model: Baseline μοντέλο που προβλέπει ότι η μελλοντική τιμή θα είναι ίση με την τελευταία διαθέσιμη παρατήρηση.

Machine Learning: Οικογένεια μεθόδων που μαθαίνουν σχέσεις από δεδομένα με σκοπό την πρόβλεψη ή ταξινόμηση.

Deep Learning: Υποκατηγορία της μηχανικής μάθησης που χρησιμοποιεί νευρωνικά δίκτυα πολλών επιπέδων για την εκμάθηση σύνθετων προτύπων.

Data leakage: Κατάσταση κατά την οποία πληροφορία από το μέλλον ή από το test set επηρεάζει άμεσα ή έμμεσα την εκπαίδευση του μοντέλου.

Chronological split: Διαχωρισμός δεδομένων με βάση τη χρονική σειρά, ώστε τα παλαιότερα δεδομένα να χρησιμοποιούνται για εκπαίδευση και τα νεότερα για αξιολόγηση.

MAE: Mean Absolute Error, δηλαδή μέσο απόλυτο σφάλμα πρόβλεψης.

RMSE: Root Mean Squared Error, δηλαδή τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος.

MAPE: Mean Absolute Percentage Error, δηλαδή μέσο ποσοστιαίο απόλυτο σφάλμα.

Persistence skill score: Μετρική που δείχνει τη βελτίωση ενός μοντέλου σε σχέση με το persistence baseline.

1.5 Οργάνωση Κεφαλαίων

Η εργασία οργανώνεται σε πέντε βασικά κεφάλαια.

Στο Κεφάλαιο 1 παρουσιάζεται το γενικό πλαίσιο της εργασίας, ο σκοπός, οι στόχοι, η μεθοδολογική προσέγγιση, οι περιορισμοί και οι βασικοί όροι που χρησιμοποιούνται.

Στο Κεφάλαιο 2 παρουσιάζεται η βιβλιογραφική ανασκόπηση. Εξετάζονται σχετικές ερευνητικές προσπάθειες στην πρόγνωση καιρού με τεχνικές Machine Learning και Deep Learning, οι κατηγορίες μοντέλων που χρησιμοποιούνται, οι πηγές δεδομένων και το ερευνητικό κενό που καλύπτει η παρούσα εργασία.

Στο Κεφάλαιο 3 περιγράφεται αναλυτικά η μεθοδολογία. Παρουσιάζονται η αρχιτεκτονική του συστήματος, η συλλογή και προεπεξεργασία δεδομένων, το feature engineering, η δημιουργία supervised datasets, τα μοντέλα πρόβλεψης, η διαδικασία εκπαίδευσης και οι μετρικές αξιολόγησης.

Στο Κεφάλαιο 4 παρουσιάζονται το λογισμικό που αναπτύχθηκε, η πειραματική διάταξη, τα αποτελέσματα των μοντέλων, οι συγκριτικοί πίνακες, τα διαγράμματα και τα παραγόμενα artifacts.

Στο Κεφάλαιο 5 γίνεται συζήτηση των αποτελεσμάτων, ανακεφαλαίωση των βασικών ευρημάτων, παρουσίαση των συμπερασμάτων, πρακτικές προεκτάσεις της εργασίας, περιορισμοί και προτάσεις για μελλοντική έρευνα.

Κεφάλαιο 2. Βιβλιογραφική έρευνα – Σχετικές προσπάθειες

2.1 Εισαγωγή

Η πρόγνωση καιρού αποτελεί ένα από τα σημαντικότερα προβλήματα πρόγνωσης χρονοσειρών, με εφαρμογές σε τομείς όπως η γεωργία, η ενέργεια, οι μεταφορές και η πολιτική προστασία. Παραδοσιακά, η πρόγνωση καιρού βασίζεται σε αριθμητικά μοντέλα πρόγνωσης (Numerical Weather Prediction - NWP), τα οποία χρησιμοποιούν φυσικές εξισώσεις για την προσομοίωση της ατμόσφαιρας.

Τα τελευταία χρόνια, η αυξημένη διαθεσιμότητα μετεωρολογικών δεδομένων και η ανάπτυξη τεχνικών μηχανικής μάθησης έχουν οδηγήσει στη χρήση data-driven προσεγγίσεων για την πρόγνωση καιρού (Zhang et al., 2025; Fang et al., 2021). Οι προσεγγίσεις αυτές αξιοποιούν ιστορικά δεδομένα με σκοπό την πρόβλεψη μελλοντικών τιμών μετεωρολογικών μεταβλητών, όπως η θερμοκρασία, η υγρασία, η ταχύτητα ανέμου και η βροχόπτωση.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η πρόγνωση καιρού σε τοπική κλίμακα, χρησιμοποιώντας δεδομένα από μεμονωμένους μετεωρολογικούς σταθμούς (Hewage et al., 2020; Dos Santos et al., 2025). Η προσέγγιση αυτή επιτρέπει την ανάπτυξη εξειδικευμένων μοντέλων πρόγνωσης για συγκεκριμένες γεωγραφικές περιοχές, αξιοποιώντας αποκλειστικά ιστορικές μετρήσεις.

Στο κεφάλαιο αυτό παρουσιάζεται η βιβλιογραφική ανασκόπηση των σχετικών εργασιών, η κατηγοριοποίηση των προσεγγίσεων πρόγνωσης καιρού, καθώς και η ανάλυση του state of the art. Επιπλέον, παρουσιάζονται τα βασικά ευρήματα της βιβλιογραφίας και προσδιορίζεται το ερευνητικό κενό στο οποίο στοχεύει η παρούσα εργασία.

2.2 Μεθοδολογία Βιβλιογραφικής Αναζήτησης

Η βιβλιογραφική ανασκόπηση πραγματοποιήθηκε με στόχο την αποτύπωση του state of the art στη βραχυπρόθεσμη πρόγνωση καιρού με χρήση τεχνικών μηχανικής μάθησης και βαθιάς μάθησης. Η διαδικασία ακολούθησε συστηματική προσέγγιση αναζήτησης, επιλογής, φιλτραρίσματος και κατηγοριοποίησης της σχετικής βιβλιογραφίας. Το τελικό σώμα της ανασκόπησης περιλαμβάνει τόσο εμπειρικές μελέτες όσο και ανασκοπικές ή θεμελιώδεις εργασίες, ώστε να αποτυπώνεται τόσο το άμεσο ερευνητικό υπόβαθρο όσο και το ευρύτερο state of the art του πεδίου.

2.2.1 Πηγές Αναζήτησης

Η αναζήτηση των άρθρων πραγματοποιήθηκε στις ακόλουθες επιστημονικές βάσεις δεδομένων:

- Google Scholar
- IEEE Xplore
- ScienceDirect
- SpringerLink
- arXiv

Οι βάσεις αυτές επιλέχθηκαν λόγω της ευρείας κάλυψης επιστημονικών δημοσιεύσεων στους τομείς της μηχανικής μάθησης, της τεχνητής νοημοσύνης και της μετεωρολογίας.

2.2.2 Λέξεις-Κλειδιά Αναζήτησης

Η αναζήτηση πραγματοποιήθηκε χρησιμοποιώντας συνδυασμούς λέξεων-κλειδίων που σχετίζονται με το αντικείμενο της εργασίας. Οι βασικές λέξεις-κλειδιά που χρησιμοποιήθηκαν ήταν:

- weather forecasting machine learning
- deep learning weather prediction
- local weather forecasting
- station-based weather prediction
- time series weather forecasting
- LSTM weather forecasting
- transformer weather forecasting
- meteorological time series forecasting

Οι λέξεις-κλειδιά χρησιμοποιήθηκαν είτε μεμονωμένα είτε σε συνδυασμούς, ώστε να εντοπιστούν όσο το δυνατόν πιο σχετικές εργασίες.

2.2.3 Κριτήρια Επιλογής Άρθρων

Από τα αποτελέσματα της αρχικής αναζήτησης, τα άρθρα επιλέχθηκαν με βάση τα ακόλουθα κριτήρια:

- Συνάφεια με το αντικείμενο της πρόγνωσης καιρού
- Χρήση τεχνικών μηχανικής μάθησης ή βαθιάς μάθησης
- Χρήση πραγματικών μετεωρολογικών δεδομένων
- Δημοσίευση σε επιστημονικά περιοδικά ή συνέδρια
- Χρονική επικαιρότητα της εργασίας

Επιπλέον, δόθηκε ιδιαίτερη έμφαση σε μελέτες που εξετάζουν πρόγνωση καιρού σε τοπική κλίμακα, καθώς αυτές σχετίζονται άμεσα με το αντικείμενο της παρούσας εργασίας.

2.2.4 Διαδικασία Φιλτραρίσματος

Η αρχική αναζήτηση απέδωσε συνολικά περίπου 60 επιστημονικές εργασίες. Στη συνέχεια πραγματοποιήθηκε φιλτράρισμα με βάση τον τίτλο, την περίληψη, τη διαθεσιμότητα πλήρους κειμένου και τη συνάφεια με το αντικείμενο της μελέτης. Αποκλείστηκαν εργασίες χωρίς σαφή πειραματική αξιολόγηση, εργασίες που εστίαζαν αποκλειστικά σε κλιματική ανάλυση χωρίς σαφές forecasting task, καθώς και εργασίες που δεν παρείχαν επαρκή πληροφορία για τα δεδομένα, τα μοντέλα ή τις μετρικές αξιολόγησης.

Μετά τη διαδικασία αυτή, επιλέχθηκαν 20 εργασίες που θεωρήθηκαν οι πλέον σχετικές με το αντικείμενο της εργασίας. Οι εργασίες αυτές κατηγοριοποιήθηκαν με βάση:

- τον χρονικό ορίζοντα πρόγνωσης
- τη γεωγραφική κλίμακα πρόγνωσης
- τα χρησιμοποιούμενα μοντέλα
- τον τρόπο συλλογής δεδομένων

Η κατηγοριοποίηση αυτή επέτρεψε τη συστηματική ανάλυση της βιβλιογραφίας και την εξαγωγή συμπερασμάτων σχετικά με το state of the art.

2.2.5 Μεθοδολογία Ανάλυσης Άρθρων

Για κάθε επιλεγμένη εργασία καταγράφηκαν τα ακόλουθα χαρακτηριστικά:

- Συγγραφείς και έτος δημοσίευσης
- Στόχος της μελέτης
- Χρησιμοποιούμενα δεδομένα
- Μέθοδος πρόγνωσης
- Μετρικές αξιολόγησης
- Βασικά αποτελέσματα

Η καταγραφή αυτή επέτρεψε τη συγκριτική ανάλυση των μελετών και τη δημιουργία πινάκων και διαγραμμάτων που παρουσιάζονται στις επόμενες ενότητες. Για τις ανάγκες της κατηγοριοποίησης, κάθε άρθρο ταξινομήθηκε με βάση τον κύριο ερευνητικό του προσανατολισμό, παρότι ορισμένες μελέτες θα μπορούσαν να ενταχθούν σε περισσότερες από μία κατηγορίες.

2.3 Κατηγοριοποίηση των Ερευνητικών Προσεγγίσεων Πρόγνωσης Καιρού

Με βάση τη βιβλιογραφική ανασκόπηση, οι ερευνητικές προσεγγίσεις πρόγνωσης καιρού μπορούν να κατηγοριοποιηθούν σύμφωνα με τρία βασικά κριτήρια:

- Τον χρονικό ορίζοντα πρόγνωσης
- Τη γεωγραφική κλίμακα πρόγνωσης
- Τη μεθοδολογία πρόγνωσης

Η κατηγοριοποίηση αυτή επιτρέπει τη συστηματική ανάλυση των μελετών και τη σύγκριση των διαφορετικών προσεγγίσεων που χρησιμοποιούνται στη βιβλιογραφία.

2.3.1 Κατηγοριοποίηση με βάση τον Χρονικό Ορίζοντα Πρόγνωσης

Οι μελέτες που εξετάστηκαν κατηγοριοποιούνται σε τρεις βασικές κατηγορίες, ανάλογα με τον χρονικό ορίζοντα πρόγνωσης:

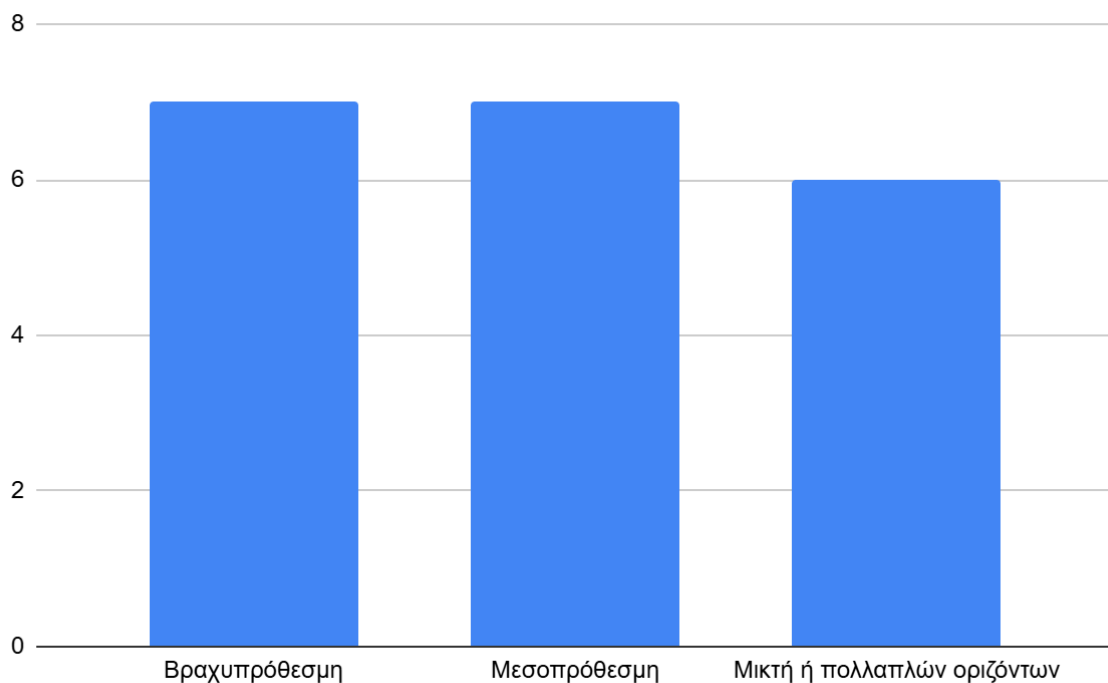
- Βραχυπρόθεσμη πρόγνωση (Short-term forecasting)
- Μεσοπρόθεσμη πρόγνωση (Medium-term forecasting)
- Μακροπρόθεσμη πρόγνωση (Long-term forecasting)

Η βραχυπρόθεσμη πρόγνωση αφορά χρονικούς ορίζοντες από λίγα λεπτά έως μερικές ώρες ή ημέρες και χρησιμοποιείται κυρίως σε εφαρμογές που απαιτούν άμεση πληροφόρηση, όπως η πρόγνωση βροχόπτωσης ή ανέμου (Shi et al., 2015; Hewage et al., 2020). Οι περισσότερες εργασίες που χρησιμοποιούν τεχνικές μηχανικής μάθησης επικεντρώνονται στη βραχυπρόθεσμη πρόγνωση, καθώς τα μοντέλα αυτά παρουσιάζουν καλύτερη απόδοση σε μικρούς χρονικούς ορίζοντες (Rahman et al., 2024; Abdulla et al., 2022; Li & Qian, 2024).

Η μεσοπρόθεσμη πρόγνωση αφορά χρονικούς ορίζοντες από λίγες ημέρες έως περίπου μία εβδομάδα (Bi et al., 2023; Lam et al., 2023; Nguyen et al., 2024). Στην κατηγορία αυτή συναντώνται κυρίως υβριδικές προσεγγίσεις που συνδυάζουν αριθμητικά μοντέλα πρόγνωσης και τεχνικές μηχανικής μάθησης.

Η μακροπρόθεσμη πρόγνωση αφορά χρονικούς ορίζοντες μεγαλύτερους της μίας εβδομάδας και χρησιμοποιείται κυρίως σε κλιματικές μελέτες. Οι εργασίες αυτής της κατηγορίας βασίζονται κυρίως σε μεγάλα datasets και αριθμητικά μοντέλα πρόγνωσης.

Στο επιλεγμένο σώμα των 20 άρθρων, η μεγαλύτερη συγκέντρωση μελετών εντοπίζεται στη βραχυπρόθεσμη πρόγνωση καιρού, γεγονός που επιβεβαιώνει το αυξημένο ενδιαφέρον της ερευνητικής κοινότητας για την ανάπτυξη μοντέλων μικρού χρονικού ορίζοντα.



Σχήμα 2.1 - Κατανομή των μελετών ανά χρονικό ορίζοντα πρόγνωσης.

Η ποσοτική αποτύπωση των επιλεγμένων μελετών ως προς τον χρονικό ορίζοντα παρουσιάζεται στο Σχήμα 2.1. Παρατηρείται ισχυρή παρουσία βραχυπρόθεσμων εφαρμογών, αλλά και σημαντική εκπροσώπηση μεσοπρόθεσμης πρόγνωσης, κυρίως μέσω πρόσφατων global AI μοντέλων. Οι ανασκοπήσεις και οι εργασίες χωρίς ενιαίο lead time πρέπει να ληφθούν υπόψη ερμηνευτικά και όχι ως άμεσα συγκρίσιμα πειραματικά setups.

2.3.2 Κατηγοριοποίηση με βάση τη Γεωγραφική Κλίμακα Πρόγνωσης

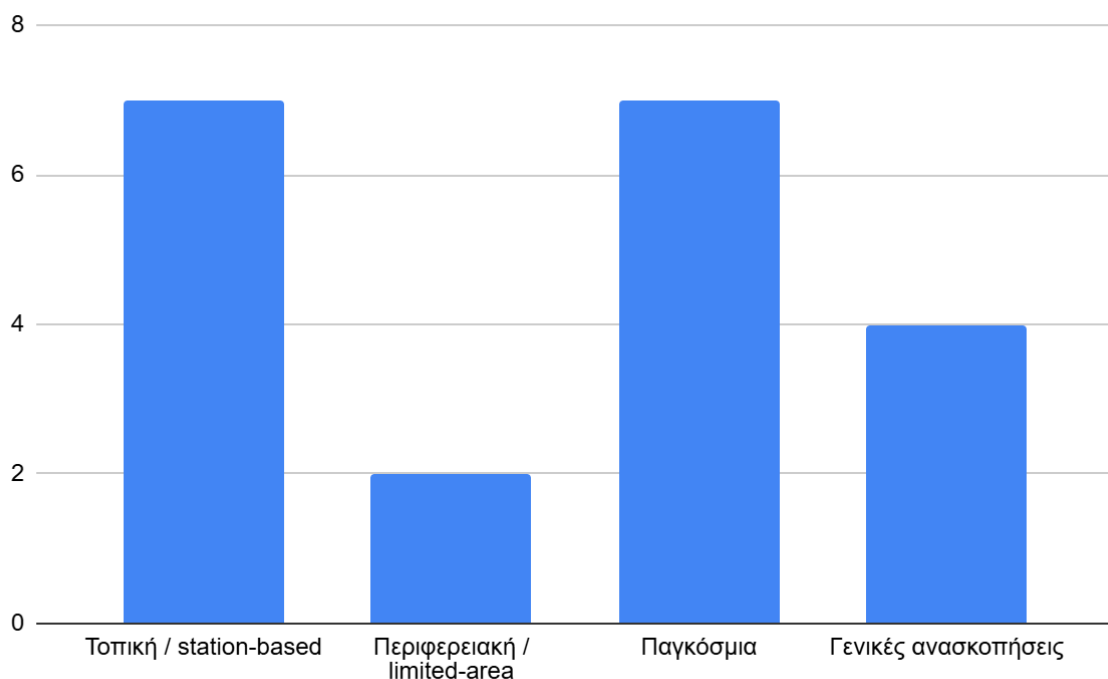
Οι μελέτες κατηγοριοποιήθηκαν επίσης με βάση τη γεωγραφική κλίμακα πρόγνωσης σε:

- Παγκόσμια πρόγνωση (Global forecasting)
- Περιφερειακή πρόγνωση (Regional forecasting)
- Τοπική πρόγνωση (Local forecasting)

Η παγκόσμια πρόγνωση αφορά μοντέλα που καλύπτουν ολόκληρο τον πλανήτη και βασίζονται κυρίως σε μεγάλα datasets και αριθμητικά μοντέλα πρόγνωσης ή reanalysis δεδομένα (Pathak et al., 2022; Bi et al., 2023; Lam et al., 2023). Τα τελευταία χρόνια, η κατηγορία αυτή έχει παρουσιάσει σημαντική ανάπτυξη με τη χρήση deep learning μοντέλων (Lam et al., 2023; Bi et al., 2023; Kochkov et al., 2024; Price et al., 2025).

Η περιφερειακή πρόγνωση αφορά συγκεκριμένες γεωγραφικές περιοχές και χρησιμοποιείται συχνά για τη βελτίωση της ακρίβειας των προβλέψεων σε συγκεκριμένες περιοχές (Xu et al., 2025).

Η τοπική πρόγνωση αφορά δεδομένα από συγκεκριμένους μετεωρολογικούς σταθμούς ή πολύ περιορισμένες γεωγραφικές περιοχές. Στο σώμα των επιλεγμένων άρθρων, οι μελέτες αυτής της κατηγορίας είναι λιγότερες σε σχέση με τις global και regional προσεγγίσεις, γεγονός που αναδεικνύει το ερευνητικό ενδιαφέρον της παρούσας εργασίας.



Σχήμα 2.2 - Κατανομή των μελετών ανά γεωγραφική κλίμακα.

Το Σχήμα 2.2 συνοψίζει τη γεωγραφική κλίμακα των μελετών. Η βιβλιογραφία εμφανίζει σαφή πώλωση μεταξύ παγκόσμιων πλεγματοποιημένων προσεγγίσεων και τοπικών ή station-based εφαρμογών, ενώ οι αυστηρά regional ή limited-area υλοποιήσεις παραμένουν λιγότερες. Η παρατήρηση αυτή ενισχύει τη σκοπιμότητα της παρούσας εργασίας, η οποία στοχεύει ρητά σε πρόγνωση τοπικής κλίμακας.

2.3.3 Κατηγοριοποίηση με βάση τη Μεθοδολογία Πρόγνωσης

Οι μελέτες κατηγοριοποιήθηκαν επίσης με βάση τη μεθοδολογία πρόγνωσης σε τρεις βασικές κατηγορίες:

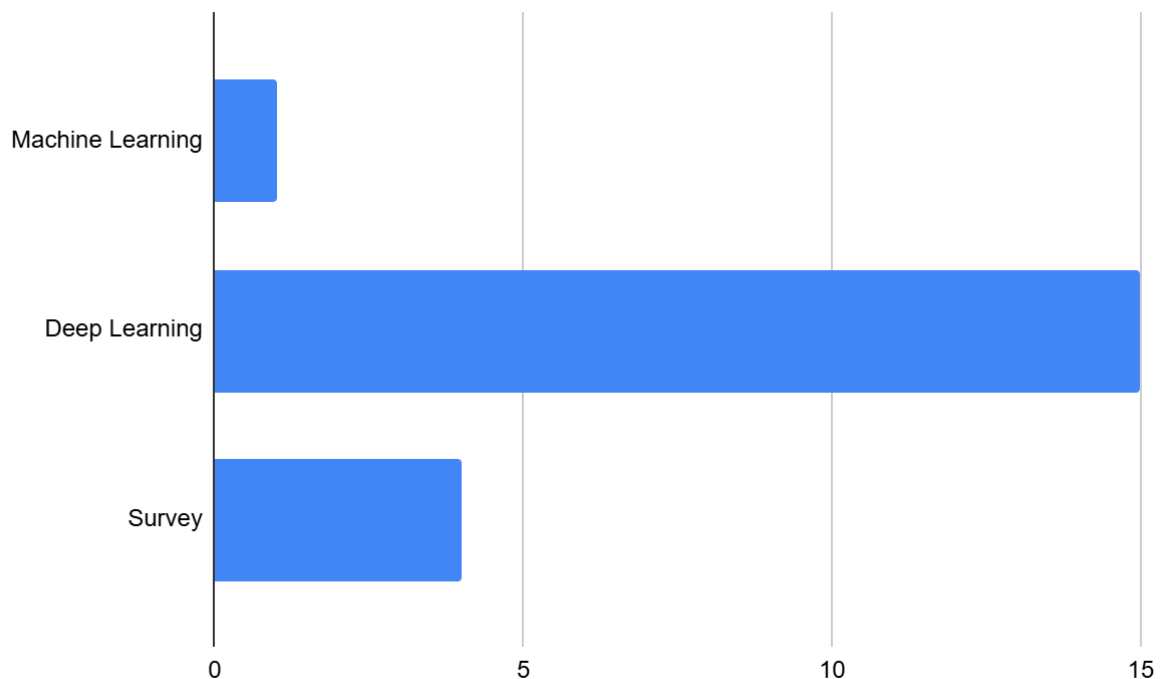
- Στατιστικές μέθοδοι

- Μέθοδοι μηχανικής μάθησης
- Μέθοδοι βαθιάς μάθησης

Οι στατιστικές μέθοδοι περιλαμβάνουν παραδοσιακές τεχνικές πρόγνωσης χρονοσειρών, όπως ARIMA και γραμμική παλινδρόμηση (Zhang et al., 2025; Lim & Zohren, 2021). Οι μέθοδοι αυτές χρησιμοποιούνται κυρίως ως baseline μοντέλα (Zhang et al., 2025; Rahman et al., 2024).

Οι μέθοδοι μηχανικής μάθησης περιλαμβάνουν αλγορίθμους όπως Random Forest, Support Vector Regression και Gradient Boosting (Rahman et al., 2024; Zhang et al., 2025). Οι μέθοδοι αυτές μπορούν να μοντελοποιήσουν μη γραμμικές σχέσεις μεταξύ των μεταβλητών (Rahman et al., 2024; Zhang et al., 2025).

Οι μέθοδοι βαθιάς μάθησης περιλαμβάνουν μοντέλα όπως LSTM, GRU, CNN και Transformer. Οι μέθοδοι αυτές παρουσιάζουν υψηλή απόδοση σε προβλήματα πρόγνωσης χρονοσειρών και χρησιμοποιούνται όλο και περισσότερο στη μετεωρολογική πρόγνωση (Lim & Zohren, 2021; Lu et al., 2025; Fang et al., 2021).



Σχήμα 2.3 - Κατανομή των μελετών ανά οικογένεια μοντέλου.

Στο Σχήμα 2.3 αποτυπώνεται η μετατόπιση της βιβλιογραφίας προς Deep Learning μοντέλα. Η συγκέντρωση μελετών στην κατηγορία DL είναι συμβατή με τη σύγχρονη τάση για data-driven πρόγνωση τόσο σε global όσο και σε local κλίμακα. Παράλληλα, η παρουσία baselines και κλασικών ML προσεγγίσεων παραμένει μεθοδολογικά κρίσιμη για τη δίκαιη συγκριτική αξιολόγηση.

Από την παραπάνω κατηγοριοποίηση προκύπτει ότι η σύγχρονη έρευνα επικεντρώνεται κυρίως στη χρήση τεχνικών βαθιάς μάθησης για τη βραχυπρόθεσμη πρόγνωση καιρού. Παράλληλα, στο επιλεγμένο corpus παρατηρείται μικρότερη παρουσία μελετών που εξετάζουν τοπική πρόγνωση με χρήση δεδομένων από μεμονωμένους μετεωρολογικούς σταθμούς. Το

εύρημα αυτό αποτελεί βασική ένδειξη ενός πιο ειδικού ερευνητικού κενού, το οποίο εξετάζεται στις επόμενες ενότητες.

2.4 Κατηγοριοποίηση των Χρησιμοποιούμενων Μοντέλων Πρόγνωσης

Στη βιβλιογραφία που εξετάστηκε παρατηρείται μεγάλη ποικιλία μοντέλων πρόγνωσης καιρού, τα οποία μπορούν να κατηγοριοποιηθούν σε τρεις βασικές ομάδες. Ενώ η ενότητα 2.3 παρουσίασε τη γενική μεθοδολογική κατηγοριοποίηση των μελετών, η παρούσα ενότητα εστιάζει στα επιμέρους μοντέλα που χρησιμοποιούνται συχνότερα στη βιβλιογραφία:

- Στατιστικά μοντέλα
- Μοντέλα μηχανικής μάθησης
- Μοντέλα βαθιάς μάθησης

Η κατηγοριοποίηση αυτή επιτρέπει τη σύγκριση των διαφορετικών προσεγγίσεων και την ανάλυση της εξέλιξης των μοντέλων πρόγνωσης καιρού.

2.4.1 Στατιστικά Μοντέλα

Τα στατιστικά μοντέλα αποτελούν τις παραδοσιακές μεθόδους πρόγνωσης χρονοσειρών και χρησιμοποιούνται συχνά ως baseline μοντέλα για τη σύγκριση με πιο σύγχρονες προσεγγίσεις. Τα πιο συχνά χρησιμοποιούμενα στατιστικά μοντέλα που εντοπίστηκαν στη βιβλιογραφία είναι:

- ARIMA
- SARIMA
- Linear Regression
- Persistence models

Τα μοντέλα αυτά χαρακτηρίζονται από σχετικά χαμηλή υπολογιστική πολυπλοκότητα και εύκολη ερμηνεία των αποτελεσμάτων. Ωστόσο, παρουσιάζουν περιορισμένη ικανότητα μοντελοποίησης μη γραμμικών σχέσεων, γεγονός που περιορίζει την απόδοσή τους σε πολύπλοκα μετεωρολογικά δεδομένα.

Παρόλα αυτά, αρκετές μελέτες χρησιμοποιούν τα στατιστικά μοντέλα ως σημείο αναφοράς για την αξιολόγηση πιο σύνθετων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης. Η χρήση ισχυρών baselines, όπως τα persistence models και συναφείς στατιστικές προσεγγίσεις, είναι μεθοδολογικά κρίσιμη για τη δίκαιη αποτίμηση της προστιθέμενης αξίας των πιο σύνθετων μοντέλων.

2.4.2 Μοντέλα Μηχανικής Μάθησης

Τα μοντέλα μηχανικής μάθησης χρησιμοποιούνται ευρέως στη σύγχρονη βιβλιογραφία για την πρόγνωση καιρού. Τα μοντέλα αυτά μπορούν να διαχειριστούν μη γραμμικές σχέσεις μεταξύ των μεταβλητών και να αξιοποιήσουν μεγάλο όγκο δεδομένων.

Τα πιο συχνά χρησιμοποιούμενα μοντέλα μηχανικής μάθησης που εντοπίστηκαν είναι:

- Random Forest

- Support Vector Regression (SVR)
- Gradient Boosting
- Decision Trees
- k-Nearest Neighbors (k-NN)

Τα μοντέλα αυτά χρησιμοποιούνται κυρίως για την πρόγνωση μεταβλητών όπως:

- θερμοκρασία
- υγρασία
- ταχύτητα ανέμου
- βροχόπτωση

Σύμφωνα με τη βιβλιογραφία, τα μοντέλα μηχανικής μάθησης παρουσιάζουν καλύτερη απόδοση από τα παραδοσιακά στατιστικά μοντέλα, ιδιαίτερα όταν τα δεδομένα παρουσιάζουν μη γραμμικές σχέσεις. Παρότι η σύγχρονη βιβλιογραφία δίνει μεγαλύτερη έμφαση σε αρχιτεκτονικές βαθιάς μάθησης, οι προσεγγίσεις μηχανικής μάθησης εξακολουθούν να αποτελούν ουσιαστικό σημείο σύγκρισης, ιδιαίτερα σε σενάρια με περιορισμένο όγκο δεδομένων ή αυξημένη ανάγκη ερμηνευσιμότητας.

2.4.3 Μοντέλα Βαθιάς Μάθησης

Τα μοντέλα βαθιάς μάθησης (Deep Learning μοντέλα) αποτελούν τη νεότερη τάση στην πρόγνωση καιρού και έχουν παρουσιάσει σημαντικές βελτιώσεις στην ακρίβεια των προβλέψεων. Τα μοντέλα αυτά είναι ιδιαίτερα κατάλληλα για την ανάλυση χρονοσειρών και την εξαγωγή πολύπλοκων προτύπων από δεδομένα.

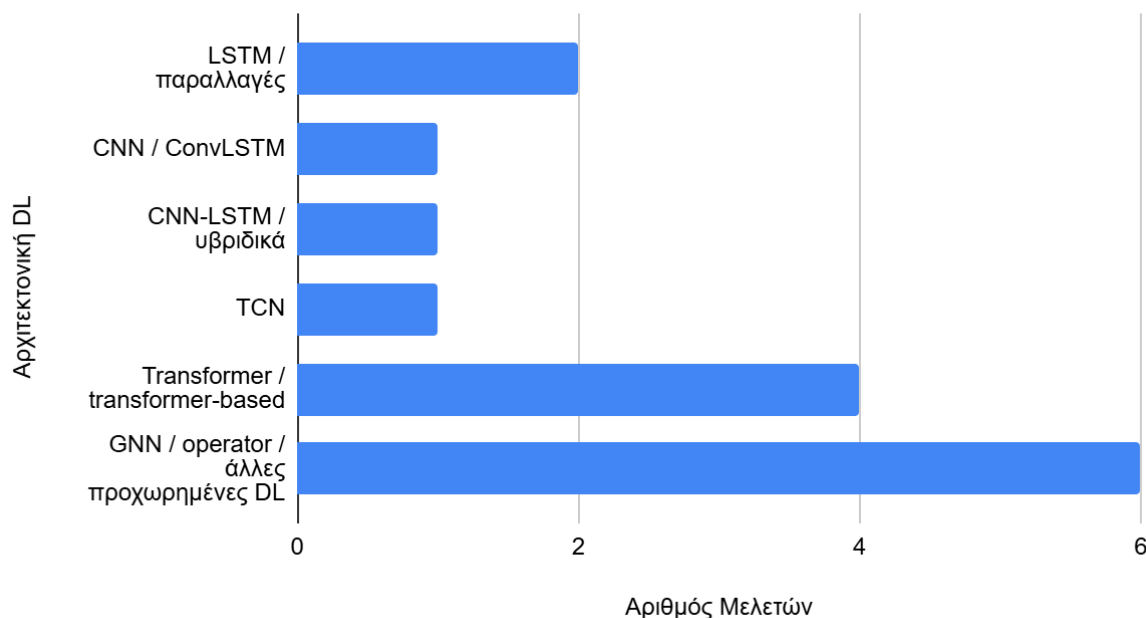
Τα πιο συχνά χρησιμοποιούμενα Deep Learning μοντέλα που εντοπίστηκαν στη βιβλιογραφία είναι:

- Long Short-Term Memory (LSTM)
- Gated Recurrent Units (GRU)
- Recurrent Neural Networks (RNN)
- Convolutional Neural Networks (CNN)
- CNN-LSTM
- Temporal Convolutional Networks (TCN)
- Transformer

Τα μοντέλα αυτά χρησιμοποιούνται κυρίως σε προβλήματα βραχυπρόθεσμης πρόγνωσης καιρού και παρουσιάζουν υψηλή απόδοση σε σχέση με τα παραδοσιακά μοντέλα.

Ειδικότερα, τα μοντέλα LSTM εμφανίζονται συχνότερα στη βιβλιογραφία, λόγω της ικανότητάς τους να μοντελοποιούν χρονικές εξαρτήσεις σε δεδομένα χρονοσειρών (Abdulla et al., 2022; Lim & Zohren, 2021). Τα τελευταία χρόνια, παρατηρείται επίσης αυξημένη χρήση transformer-based μοντέλων, τα οποία παρουσιάζουν ιδιαίτερα ενθαρρυντικά αποτελέσματα (Alerskans et al., 2022; Dos Santos et al., 2025; Nguyen et al., 2024).

Αριθμός Μελετών έναντι Αρχιτεκτονική DL



Σχήμα 2.4 - Κατανομή των αρχιτεκτονικών βαθιάς μάθησης στη βιβλιογραφία.

Στο Σχήμα 2.4 αποτυπώνεται η ποικιλία των αρχιτεκτονικών DL που χρησιμοποιούνται στη σύγχρονη βιβλιογραφία. Οι recurrent, convolutional, hybrid και transformer-based προσεγγίσεις συνυπάρχουν, ενώ τα τελευταία έτη ενισχύεται ιδιαίτερα η παρουσία transformer-derived μοντέλων. Η χαρτογράφηση αυτή καθοδηγεί τη λογική επιλογής των υποψήφιων μοντέλων για το πειραματικό μέρος της εργασίας.

Από την ανάλυση των μοντέλων προκύπτει ότι η σύγχρονη βιβλιογραφία μετατοπίζεται σταδιακά από τα παραδοσιακά στατιστικά μοντέλα προς τις τεχνικές βαθιάς μάθησης. Παράλληλα, παρατηρείται ότι τα μοντέλα μηχανικής μάθησης εξακολουθούν να χρησιμοποιούνται ευρέως, ιδιαίτερα σε εφαρμογές με περιορισμένο όγκο δεδομένων.

Η παρατήρηση αυτή είναι σημαντική για τον σχεδιασμό της παρούσας εργασίας, καθώς υποδεικνύει την ανάγκη σύγκρισης διαφορετικών κατηγοριών μοντέλων πρόγνωσης.

2.5 Κατηγοριοποίηση των Δεδομένων και Τρόπων Συλλογής

Η απόδοση των μοντέλων πρόγνωσης καιρού εξαρτάται σε μεγάλο βαθμό από τα διαθέσιμα δεδομένα και τον τρόπο συλλογής τους. Για τον λόγο αυτό, οι μελέτες που εξετάστηκαν κατηγοριοποιήθηκαν με βάση την πηγή και τον τρόπο συλλογής των δεδομένων.

Από τη βιβλιογραφική ανασκόπηση προέκυψαν τέσσερις βασικές κατηγορίες δεδομένων:

- Δεδομένα μετεωρολογικών σταθμών
- Δεδομένα Numerical Weather Prediction (NWP)
- Δεδομένα δορυφόρων
- Συνδυασμός πολλαπλών πηγών δεδομένων

Η κατηγοριοποίηση αυτή επιτρέπει τη σύγκριση των διαφορετικών προσεγγίσεων και τη διερεύνηση της επίδρασης των δεδομένων στην ακρίβεια της πρόγνωσης.

2.5.1 Δεδομένα Μετεωρολογικών Σταθμών

Τα δεδομένα μετεωρολογικών σταθμών αποτελούν μία από τις πιο συχνά χρησιμοποιούμενες πηγές δεδομένων στη βιβλιογραφία, ιδιαίτερα σε τοπικά forecasting setups (Hewage et al., 2020; Abdulla et al., 2022; Dos Santos et al., 2025). Οι μετρήσεις αυτές προέρχονται από σταθερούς σταθμούς και περιλαμβάνουν μεταβλητές όπως:

- θερμοκρασία
- υγρασία
- ατμοσφαιρική πίεση
- ταχύτητα ανέμου
- διεύθυνση ανέμου
- βροχόπτωση

Τα δεδομένα αυτά χρησιμοποιούνται κυρίως σε μελέτες τοπικής πρόγνωσης καιρού, καθώς επιτρέπουν την ανάπτυξη μοντέλων για συγκεκριμένες γεωγραφικές περιοχές (Hewage et al., 2020; Li & Qian, 2024; Sun et al., 2026). Ωστόσο, τα δεδομένα από έναν μόνο σταθμό ενδέχεται να παρουσιάζουν περιορισμένη χωρική πληροφορία, γεγονός που μπορεί να επηρεάσει την ακρίβεια των προβλέψεων.

2.5.2 Δεδομένα Numerical Weather Prediction (NWP)

Τα δεδομένα Numerical Weather Prediction (NWP) προέρχονται από αριθμητικά μοντέλα πρόγνωσης καιρού και χρησιμοποιούνται ευρέως στη βιβλιογραφία. Συναφείς αλλά διακριτές κατηγορίες αποτελούν τα reanalysis ή analysis datasets, τα οποία παρέχουν πλεγματοποιημένες ανακατασκευές της ατμοσφαιρικής κατάστασης και χρησιμοποιούνται συχνά ως εκπαιδευτικά ή συγκριτικά δεδομένα σε σύγχρονα AI weather models. Για τον λόγο αυτό, στην παρούσα ανασκόπηση ο όρος NWP χρησιμοποιείται ευρύτερα, ώστε να καλύπτει τόσο forecast model outputs όσο και reanalysis-based εισόδους όταν αυτές αποτελούν την κύρια πληροφοριακή βάση της μελέτης.

Οι μελέτες που χρησιμοποιούν NWP ή reanalysis δεδομένα παρουσιάζουν συνήθως υψηλή ακρίβεια, καθώς αξιοποιούν μεγάλο όγκο δεδομένων και φυσικές παραμέτρους (Pathak et al., 2022; Bi et al., 2023; Lam et al., 2023; Kochkov et al., 2024). Ωστόσο, η χρήση αυτών των δεδομένων απαιτεί σημαντική υπολογιστική ισχύ και πρόσβαση σε εξειδικευμένες πηγές δεδομένων.

2.5.3 Δεδομένα Δορυφόρων

Τα δορυφορικά δεδομένα χρησιμοποιούνται κυρίως σε μελέτες μεγάλης κλίμακας και περιλαμβάνουν πληροφορίες όπως:

- νεφοκάλυψη
- θερμοκρασία επιφάνειας
- υγρασία ατμόσφαιρας

Τα δεδομένα αυτά προσφέρουν μεγάλη χωρική κάλυψη, αλλά συχνά απαιτούν επεξεργασία πριν χρησιμοποιηθούν σε μοντέλα πρόγνωσης.

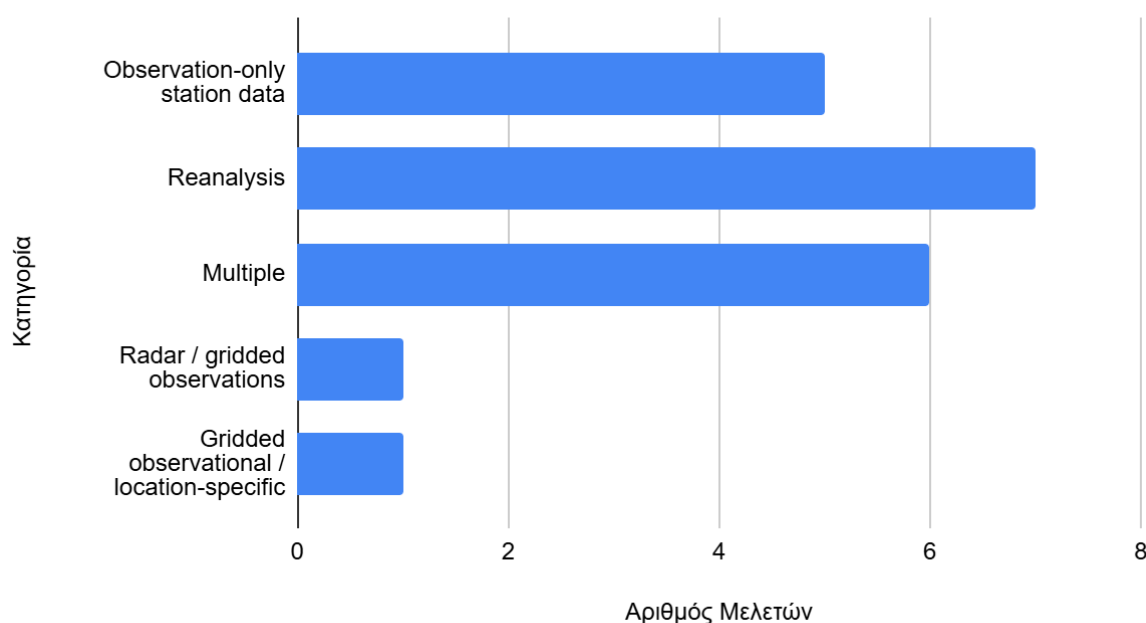
2.5.4 Συνδυασμός Πολλαπλών Πηγών Δεδομένων

Πολλές σύγχρονες μελέτες χρησιμοποιούν συνδυασμό διαφορετικών πηγών δεδομένων, όπως (Allen et al., 2025; Alerskans et al., 2022):

- μετεωρολογικούς σταθμούς
- δορυφορικά δεδομένα
- αριθμητικά μοντέλα πρόγνωσης

Η προσέγγιση αυτή επιτρέπει τη βελτίωση της ακρίβειας των μοντέλων, καθώς αξιοποιείται μεγαλύτερος όγκος πληροφοριών. Ωστόσο, η διαχείριση πολλαπλών πηγών δεδομένων αυξάνει την πολυπλοκότητα των μοντέλων.

Αριθμός Μελετών έναντι Κατηγορία



Σχήμα 2.5 - Κατανομή των μελετών ανά πηγή δεδομένων.

Το Σχήμα 2.5 παρουσιάζει τις κύριες πηγές δεδομένων που αξιοποιούνται στη βιβλιογραφία. Η έρευνα υψηλής επίδρασης στη global πρόγνωση βασίζεται κυρίως σε reanalysis, αναλύσεις ή NWP-derived πλεγματοποιημένα δεδομένα, ενώ τα αυστηρά observation-only station-based setups είναι σαφώς λιγότερα. Η διαπίστωση αυτή ενισχύει το ερευνητικό κενό που καλύπτει η παρούσα εργασία.

Από την ανάλυση των πηγών δεδομένων προκύπτει ότι, στο επιλεγμένο σύνολο άρθρων, πολλές μελέτες χρησιμοποιούν συνδυασμό πολλαπλών πηγών δεδομένων ή δεδομένα από αριθμητικά μοντέλα πρόγνωσης. Αντίθετα, μικρότερος αριθμός μελετών χρησιμοποιεί αποκλειστικά δεδομένα από μεμονωμένους μετεωρολογικούς σταθμούς.

Το εύρημα αυτό αναδεικνύει τη σημασία της παρούσας εργασίας, η οποία επικεντρώνεται στην ανάπτυξη μοντέλων πρόγνωσης χρησιμοποιώντας δεδομένα από έναν τοπικό μετεωρολογικό σταθμό.

2.6 Σύνοψη Βιβλιογραφικής Ανασκόπησης και Κύρια Ευρήματα

Η ανάλυση των επιλεγμένων μελετών επιτρέπει την εξαγωγή σημαντικών συμπερασμάτων σχετικά με τις σύγχρονες τάσεις στην πρόγνωση καιρού με χρήση τεχνικών μηχανικής μάθησης και βαθιάς μάθησης (Zhang et al., 2025; Fang et al., 2021; Lu et al., 2025). Τα ευρήματα αυτά προκύπτουν από την κατηγοριοποίηση των μελετών με βάση τον χρονικό ορίζοντα πρόγνωσης, τη γεωγραφική κλίμακα, τα χρησιμοποιούμενα μοντέλα και τις πηγές δεδομένων.

2.6.1 Ευρήματα ως προς τον Χρονικό Ορίζοντα Πρόγνωσης

Στο επιλεγμένο σώμα των άρθρων, η μεγαλύτερη συκέντρωση μελετών αφορά τη βραχυπρόθεσμη πρόγνωση καιρού. Οι μελέτες αυτές εξετάζουν χρονικούς ορίζοντες από λίγα λεπτά έως μερικές ώρες ή ημέρες. Η αυξημένη ερευνητική δραστηριότητα στη βραχυπρόθεσμη πρόγνωση σχετίζεται με την ανάγκη για άμεση και ακριβή πληροφόρηση σε εφαρμογές όπως η διαχείριση ενέργειας και οι μεταφορές.

Αντίθετα, μικρότερος αριθμός μελετών επικεντρώνεται στη μεσοπρόθεσμη και μακροπρόθεσμη πρόγνωση. Οι μελέτες αυτές βασίζονται κυρίως σε μεγάλα datasets και σε συνδυασμό αριθμητικών μοντέλων πρόγνωσης και τεχνικών μηχανικής μάθησης.

2.6.2 Ευρήματα ως προς τη Γεωγραφική Κλίμακα Πρόγνωσης

Η ανάλυση των μελετών έδειξε ότι μεγάλο μέρος της βιβλιογραφίας του corpus αφορά παγκόσμια ή περιφερειακή πρόγνωση καιρού. Οι μελέτες αυτές χρησιμοποιούν συνήθως μεγάλα datasets, πλεγματοποιημένα ατμοσφαιρικά δεδομένα και σύνθετα Deep Learning μοντέλα.

Αντίθετα, περιορισμένος αριθμός μελετών εξετάζει την πρόγνωση καιρού σε τοπική κλίμακα. Οι μελέτες αυτές χρησιμοποιούν δεδομένα από μεμονωμένους μετεωρολογικούς σταθμούς και επικεντρώνονται σε συγκεκριμένες γεωγραφικές περιοχές.

Το εύρημα αυτό υποδηλώνει ότι η τοπική πρόγνωση καιρού αποτελεί ερευνητικό πεδίο με περιθώρια περαιτέρω διερεύνησης.

2.6.3 Ευρήματα ως προς τα Χρησιμοποιούμενα Μοντέλα

Η βιβλιογραφική ανασκόπηση δείχνει ότι τα Deep Learning μοντέλα χρησιμοποιούνται όλο και περισσότερο στη σύγχρονη πρόγνωση καιρού (Zhang et al., 2025; Fang et al., 2021; Lu et al., 2025). Τα μοντέλα αυτά εμφανίζουν υψηλή ακρίβεια και δυνατότητα μοντελοποίησης πολύπλοκων χρονικών εξαρτήσεων, ιδιαίτερα όταν είναι διαθέσιμος μεγάλος όγκος δεδομένων ή σύνθετες χωροχρονικές αναπαραστάσεις.

Παράλληλα, τα μοντέλα μηχανικής μάθησης εξακολουθούν να χρησιμοποιούνται ευρέως, ιδιαίτερα σε περιπτώσεις όπου τα διαθέσιμα δεδομένα είναι περιορισμένα (Rahman et al., 2024; Zhang et al., 2025). Τα στατιστικά μοντέλα χρησιμοποιούνται κυρίως ως baseline μοντέλα για σύγκριση.

Η αυξανόμενη χρήση deep learning μοντέλων υποδεικνύει τη μετατόπιση της έρευνας προς πιο σύνθετες και αποδοτικές προσεγγίσεις.

2.6.4 Ευρήματα ως προς τα Δεδομένα

Η ανάλυση των πηγών δεδομένων δείχνει ότι πολλές μελέτες χρησιμοποιούν συνδυασμό δεδομένων από πολλαπλές πηγές, όπως μετεωρολογικούς σταθμούς, δορυφορικά δεδομένα και αριθμητικά μοντέλα πρόγνωσης (Allen et al., 2025; Alerskans et al., 2022; Xu et al., 2025). Το στοιχείο αυτό αυξάνει συνήθως την πληροφοριακή πληρότητα των εισόδων, αλλά μεταβάλλει και τη φύση του προβλήματος σε σχέση με ένα καθαρά station-based, observation-only setup.

Αντίθετα, λιγότερες μελέτες χρησιμοποιούν αποκλειστικά δεδομένα από μεμονωμένους μετεωρολογικούς σταθμούς. Η προσέγγιση αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς επιτρέπει την ανάπτυξη μοντέλων πρόγνωσης για συγκεκριμένες γεωγραφικές περιοχές.

2.6.5 Συνολική Αποτύπωση των Ευρημάτων

Συνοψίζοντας, από τη βιβλιογραφική ανασκόπηση των 20 επιλεγμένων άρθρων προκύπτουν τα εξής:

- Οι περισσότερες μελέτες επικεντρώνονται στη βραχυπρόθεσμη πρόγνωση καιρού
- Η πλειονότητα των μελετών χρησιμοποιεί Deep Learning μοντέλα
- Πολλές μελέτες χρησιμοποιούν δεδομένα από πολλαπλές πηγές
- Περιορισμένος αριθμός μελετών εξετάζει τοπική πρόγνωση καιρού
- Λίγες μελέτες χρησιμοποιούν δεδομένα από έναν μόνο μετεωρολογικό σταθμό

Τα παραπάνω ευρήματα αποτελούν τη βάση για τον προσδιορισμό του ερευνητικού κενού, το οποίο παρουσιάζεται στην επόμενη ενότητα.

2.7 Ερευνητικό Κενό (Research Gap)

Η βιβλιογραφική ανασκόπηση ανέδειξε σημαντικές εξελίξεις στη χρήση τεχνικών μηχανικής μάθησης και βαθιάς μάθησης για την πρόγνωση καιρού (Zhang et al., 2025; Fang et al., 2021; Lu et al., 2025). Παρά τη σημαντική πρόοδο, η ανάλυση του επιλεγμένου σώματος μελετών δείχνει ότι η σχετική βιβλιογραφία παραμένει αποσπασματική ως προς ορισμένους συγκεκριμένους συνδυασμούς χαρακτηριστικών, οι οποίοι δεν έχουν διερευνηθεί επαρκώς στο ίδιο πειραματικό πλαίσιο.

Ένα από τα βασικά ευρήματα είναι ότι η πλειονότητα των μελετών επικεντρώνεται σε πρόγνωση μεγάλης ή περιφερειακής κλίμακας (Pathak et al., 2022; Bi et al., 2023; Lam et al., 2023; Xu et al., 2025). Οι μελέτες αυτές χρησιμοποιούν μεγάλα datasets και σύνθετα Deep Learning μοντέλα. Αντίθετα, περιορισμένος αριθμός μελετών εξετάζει την πρόγνωση καιρού σε τοπική κλίμακα, χρησιμοποιώντας δεδομένα από μεμονωμένους μετεωρολογικούς σταθμούς.

Επιπλέον, πολλές μελέτες βασίζονται σε συνδυασμό δεδομένων από πολλαπλές πηγές, όπως αριθμητικά μοντέλα πρόγνωσης, δορυφορικά δεδομένα και δεδομένα μετεωρολογικών σταθμών (Allen et al., 2025; Alerskans et al., 2022; Xu et al., 2025). Η προσέγγιση αυτή αυξάνει την ακρίβεια των προβλέψεων, αλλά παράλληλα αυξάνει την πολυπλοκότητα των μοντέλων και την ανάγκη για μεγάλες υπολογιστικές υποδομές.

Αντίθετα, η πρόγνωση καιρού με χρήση δεδομένων από έναν μόνο μετεωρολογικό σταθμό αποτελεί λιγότερο μελετημένη προσέγγιση, παρά την ύπαρξη ορισμένων σχετικών μελετών (Hewage et al., 2020; Dos Santos et al., 2025; Sun et al., 2026). Η διερεύνηση αυτής της

προσέγγισης παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς επιτρέπει την ανάπτυξη πιο απλών και υπολογιστικά ελαφρύτερων μοντέλων πρόγνωσης, χωρίς εξάρτηση από εξωτερικές εισόδους NWP.

Επιπλέον, από τη βιβλιογραφία παρατηρείται ότι λίγες μελέτες εξετάζουν ταυτόχρονα τη σύγκριση διαφορετικών κατηγοριών μοντέλων πρόγνωσης, όπως στατιστικά μοντέλα, μοντέλα μηχανικής μάθησης και Deep Learning μοντέλα, χρησιμοποιώντας το ίδιο dataset (Rahman et al., 2024; Dos Santos et al., 2025).

Τέλος, παρατηρείται περιορισμένος αριθμός μελετών που εξετάζουν πολλαπλούς χρονικούς ορίζοντες πρόγνωσης στο ίδιο πειραματικό πλαίσιο (Lim & Zohren, 2021; Sun et al., 2026; Dos Santos et al., 2025). Η προσέγγιση αυτή θα μπορούσε να προσφέρει μια πιο ολοκληρωμένη αξιολόγηση των μοντέλων πρόγνωσης.

Με βάση τα παραπάνω, το ερευνητικό κενό της παρούσας εργασίας δεν αφορά την πλήρη απουσία σχετικής βιβλιογραφίας, αλλά τον περιορισμένο αριθμό μελετών που συνδυάζουν, στο ίδιο ερευνητικό πλαίσιο, τα εξής χαρακτηριστικά:

- Πρόγνωση καιρού σε τοπική κλίμακα
- Χρήση δεδομένων από έναν μετεωρολογικό σταθμό
- Σύγκριση διαφορετικών κατηγοριών μοντέλων
- Εξέταση πολλαπλών χρονικών οριζόντων πρόγνωσης

Η κάλυψη των παραπάνω σημείων αποτελεί τη βασική συνεισφορά της παρούσας εργασίας και οδηγεί στον καθορισμό των στόχων της έρευνας, οι οποίοι παρουσιάζονται στην επόμενη ενότητα.

2.8 Ερευνητικές Υποθέσεις

Με βάση τη βιβλιογραφική ανασκόπηση και το ερευνητικό κενό που εντοπίστηκε στις προηγούμενες ενότητες, διατυπώνονται οι ακόλουθες ερευνητικές υποθέσεις. Οι υποθέσεις αυτές καθοδηγούν τον σχεδιασμό της πειραματικής διαδικασίας και εστιάζουν στην αξιολόγηση ενός τοπικού, observation-only και single-station πλαισίου βραχυπρόθεσμης πρόγνωσης καιρού.

Η πρώτη ερευνητική υπόθεση είναι ότι θα μελετηθεί η ικανότητα ενός observation-only, single-station μοντέλου να παρέχει χρήσιμη βραχυπρόθεσμη πρόγνωση καιρού για συγκεκριμένες μετεωρολογικές μεταβλητές, με χρήση αποκλειστικά station-based χρονοσειρών.

Η δεύτερη ερευνητική υπόθεση είναι ότι θα αξιολογηθούν και θα συγκριθούν Deep Learning μοντέλα με παραδοσιακά στατιστικά μοντέλα, τα οποία θα περιλαμβάνουν σύνθετες χρονικές εξαρτήσεις και μη γραμμικές σχέσεις μεταξύ των μεταβλητών.

Η τρίτη ερευνητική υπόθεση είναι ότι θα μελετηθεί η απόδοση όλων των μοντέλων καθώς θα αυξάνεται ο χρονικός ορίζοντας πρόγνωσης, δημιουργώντας αποτελέσματα ανά χρονικό ορίζοντα.

Η τέταρτη ερευνητική υπόθεση είναι ότι θα μελετηθούν τα όρια ενός καθαρά τοπικού, observation-based forecasting setup και η συγκριτική αποτίμηση baseline, machine learning και deep learning μοντέλων στο ίδιο dataset.

Οι παραπάνω ερευνητικές υποθέσεις καθορίζουν τη μεθοδολογική προσέγγιση της εργασίας, την επιλογή των μοντέλων, τη διαμόρφωση των πειραματικών σεναρίων και τη συγκριτική αξιολόγηση των αποτελεσμάτων που παρουσιάζονται στα επόμενα κεφάλαια.

Κεφάλαιο 3. Μεθοδολογία

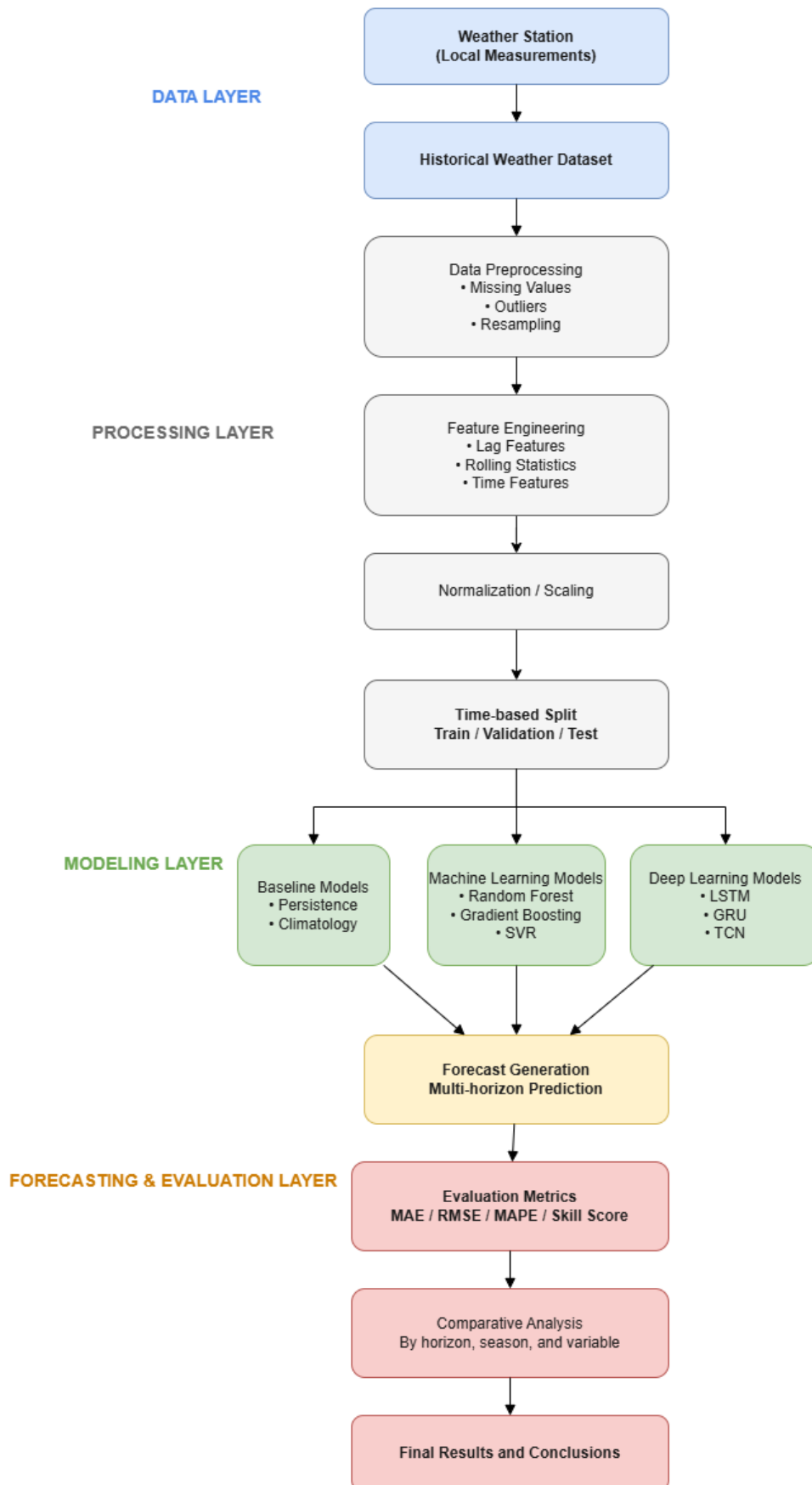
3.1 Εισαγωγή στη Μεθοδολογία

Η παρούσα εργασία ακολουθεί μια συστηματική πειραματική μεθοδολογία για την ανάπτυξη και αξιολόγηση μοντέλων βραχυπρόθεσμης πρόγνωσης καιρού σε τοπική κλίμακα, με χρήση τεχνικών Machine Learning και Deep Learning. Η μεθοδολογία σχεδιάστηκε με στόχο τη δημιουργία ενός επαναλήψιμου και επεκτάσιμου πειραματικού πλαισίου, το οποίο επιτρέπει τη συγκριτική αξιολόγηση διαφορετικών αλγορίθμων πρόβλεψης σε πολλαπλούς χρονικούς ορίζοντες, χρησιμοποιώντας αποκλειστικά ιστορικά δεδομένα από έναν μεμονωμένο μετεωρολογικό σταθμό.

Η διαδικασία περιλαμβάνει τη συλλογή και προεπεξεργασία μετεωρολογικών δεδομένων, τη δημιουργία χαρακτηριστικών (feature engineering), τη διαμόρφωση πειραματικών συνόλων δεδομένων, την εκπαίδευση μοντέλων πρόβλεψης και την αξιολόγηση των αποτελεσμάτων με χρήση κατάλληλων μετρικών σφάλματος. Επιπλέον, εφαρμόζεται συγκριτική ανάλυση μεταξύ διαφορετικών οικογενειών μοντέλων, συμπεριλαμβανομένων baseline μοντέλων, παραδοσιακών αλγορίθμων Machine Learning και σύγχρονων αρχιτεκτονικών Deep Learning.

Η παρούσα μεθοδολογία έχει σχεδιαστεί ώστε να υποστηρίζει τη μελέτη των δυνατοτήτων και των περιορισμών ενός πλαισίου πρόγνωσης που βασίζεται αποκλειστικά σε παρατηρήσεις από έναν μεμονωμένο μετεωρολογικό σταθμό, καθώς και τη διερεύνηση της επίδρασης του χρονικού ορίζοντα πρόγνωσης στην απόδοση διαφορετικών κατηγοριών μοντέλων.

Η συνολική αρχιτεκτονική του συστήματος παρουσιάζεται στο Σχήμα 3.1, ενώ η πειραματική διαδικασία και το workflow της μεθοδολογίας παρουσιάζονται στο Σχήμα 3.2.



Σχήμα 3.1 - Αρχιτεκτονική του Προτεινόμενου Συστήματος Πρόβλεψης Καιρού.

3.2 Συνολική Αρχιτεκτονική Συστήματος

Η αρχιτεκτονική του προτεινόμενου συστήματος πρόγνωσης καιρού αποτελείται από διακριτά στάδια επεξεργασίας δεδομένων και εκπαίδευσης μοντέλων. Το σύστημα σχεδιάστηκε με modular προσέγγιση, ώστε να επιτρέπεται η σαφής οργάνωση της πειραματικής διαδικασίας και η εύκολη προσαρμογή επιμέρους σταδίων χωρίς να αλλοιώνεται η συνοχή του συνολικού πλαισίου.

Η διαδικασία ξεκινά με τη συλλογή ιστορικών μετεωρολογικών δεδομένων από έναν μεμονωμένο τοπικό μετεωρολογικό σταθμό. Τα δεδομένα αυτά περιλαμβάνουν βασικές μετεωρολογικές μεταβλητές, όπως η θερμοκρασία, η σχετική υγρασία, η ατμοσφαιρική πίεση, η ταχύτητα ανέμου, η διεύθυνση ανέμου και η βροχόπτωση, καθώς και πρόσθετες μεταβλητές όπου αυτές είναι διαθέσιμες. Η επιλογή αυτής της αρχιτεκτονικής ανταποκρίνεται στον στόχο της εργασίας για ανάπτυξη και αξιολόγηση μοντέλων πρόγνωσης σε τοπική κλίμακα, με χρήση αποκλειστικά δεδομένων παρατήρησης.

Στη συνέχεια, πραγματοποιείται προεπεξεργασία των δεδομένων, η οποία περιλαμβάνει καθαρισμό, έλεγχο ποιότητας, διαχείριση ελλিপών τιμών και κανονικοποίηση, με στόχο τη δημιουργία ενός αξιόπιστου και συνεπούς συνόλου εισόδου. Μετά την προεπεξεργασία, εφαρμόζεται διαδικασία δημιουργίας χαρακτηριστικών (feature engineering), μέσω της οποίας παράγονται χρονικά, στατιστικά και παράγωγα μετεωρολογικά χαρακτηριστικά που ενισχύουν την πληροφοριακή αξία των δεδομένων.

Τα επεξεργασμένα δεδομένα χρησιμοποιούνται στη συνέχεια για τη δημιουργία των πειραματικών συνόλων εκπαίδευσης, επικύρωσης και δοκιμής, με διατήρηση της χρονικής ακολουθίας των παρατηρήσεων, ώστε να αποφεύγεται η διαρροή πληροφορίας. Πάνω σε αυτό το ενιαίο πειραματικό υπόβαθρο εφαρμόζονται διαφορετικές κατηγορίες μοντέλων πρόγνωσης, συμπεριλαμβανομένων baseline μοντέλων, παραδοσιακών αλγορίθμων Machine Learning και μοντέλων Deep Learning.

Στο τελικό στάδιο, η απόδοση των μοντέλων αξιολογείται με χρήση κατάλληλων μετρικών σφάλματος και πραγματοποιείται συγκριτική ανάλυση ως προς την ικανότητά τους να υποστηρίξουν βραχυπρόθεσμη πρόγνωση καιρού σε πολλαπλούς χρονικούς ορίζοντες. Με τον τρόπο αυτό, η συνολική αρχιτεκτονική του συστήματος λειτουργεί ως ένα ενιαίο πειραματικό πλαίσιο για τη διερεύνηση των δυνατοτήτων και των περιορισμών ενός πλαισίου πρόγνωσης που βασίζεται σε έναν μόνο σταθμό και αποκλειστικά σε δεδομένα παρατήρησης.

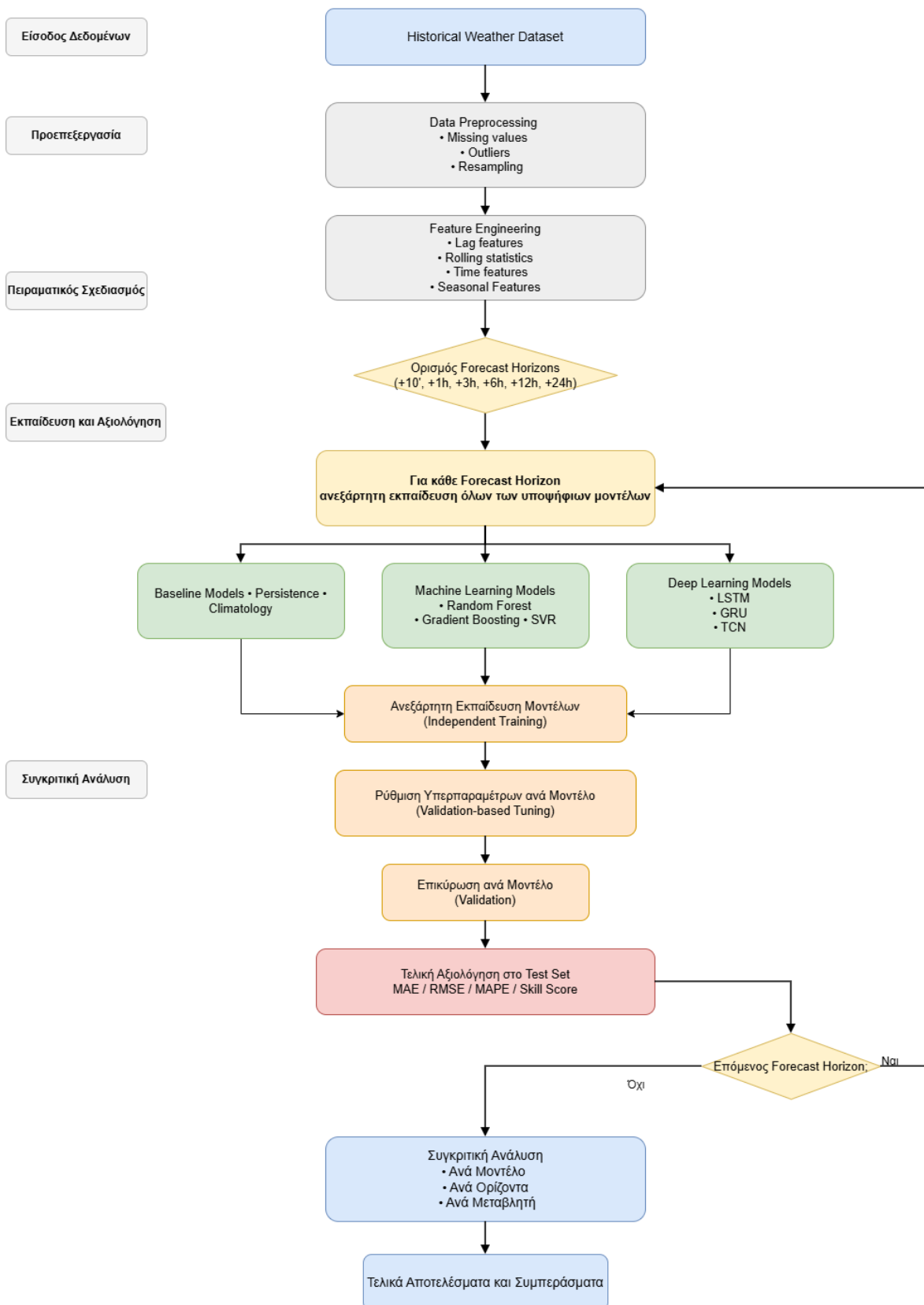
Πειραματική Διαδικασία

Η πειραματική διαδικασία της εργασίας ακολουθεί μια επαναληπτική προσέγγιση, η οποία επιτρέπει τη συστηματική σύγκριση διαφορετικών μοντέλων πρόγνωσης και διαφορετικών χρονικών οριζόντων πρόβλεψης. Η διαδικασία ξεκινά με τον καθορισμό των χρονικών οριζόντων πρόβλεψης και συνεχίζεται με την ανεξάρτητη εκπαίδευση και αξιολόγηση όλων των υποψήφιων οικογενειών μοντέλων για κάθε χρονικό ορίζοντα.

Για κάθε χρονικό ορίζοντα πρόβλεψης, εφαρμόζονται διαφορετικές κατηγορίες μοντέλων, συμπεριλαμβανομένων baseline μοντέλων, παραδοσιακών αλγορίθμων Machine Learning και μοντέλων Deep Learning. Κάθε μοντέλο εκπαιδεύεται και αξιολογείται στο ίδιο πειραματικό πλαίσιο, χρησιμοποιώντας σύνολα δεδομένων εκπαίδευσης, επικύρωσης και δοκιμής που διαμορφώνονται με ενιαία μεθοδολογική λογική.

Η διαδικασία εφαρμόζεται για κάθε χρονικό ορίζοντα πρόβλεψης και για όλα τα υποψήφια μοντέλα, επιτρέποντας τη συγκριτική αξιολόγηση των αποτελεσμάτων και τη διερεύνηση της επίδρασης του forecasting horizon στην απόδοση των διαφορετικών κατηγοριών μοντέλων. Το συνολικό workflow παρουσιάζεται στο Σχήμα 3.2.

Διάγραμμα Πειραματικής Ροής (Workflow / Experimental Pipeline)



Σχήμα 3.2 - Πειραματικό Workflow της Προτεινόμενης Μεθοδολογίας.

3.3 Συλλογή Δεδομένων

Η ποιότητα και η αξιοπιστία των δεδομένων αποτελούν κρίσιμο παράγοντα για την ανάπτυξη μοντέλων πρόγνωσης καιρού υψηλής ακρίβειας. Στην παρούσα εργασία χρησιμοποιούνται ιστορικά μετεωρολογικά δεδομένα από έναν μεμονωμένο τοπικό μετεωρολογικό σταθμό, με στόχο την ανάπτυξη και αξιολόγηση μοντέλων βραχυπρόθεσμης πρόγνωσης καιρού σε τοπική κλίμακα.

Τα δεδομένα περιλαμβάνουν βασικές μετεωρολογικές μεταβλητές, οι οποίες χρησιμοποιούνται ευρέως στη βιβλιογραφία για την πρόβλεψη καιρικών συνθηκών. Συγκεκριμένα, το σύνολο δεδομένων περιλαμβάνει τη θερμοκρασία αέρα (Temperature), τη σχετική υγρασία (Humidity), την ατμοσφαιρική πίεση (Pressure), την ταχύτητα ανέμου (Wind Speed), τη διεύθυνση ανέμου (Wind Direction), τη βροχόπτωση (Precipitation), καθώς και ενδεχόμενες πρόσθετες μεταβλητές, εφόσον αυτές είναι διαθέσιμες και κριθούν κατάλληλες για το πειραματικό πλαίσιο.

Τα δεδομένα συλλέγονται σε τακτά χρονικά διαστήματα, σχηματίζοντας χρονοσειρές υψηλής χρονικής ανάλυσης. Η χρονικά εξαρτώμενη μορφή τους επιτρέπει την εφαρμογή τόσο παραδοσιακών τεχνικών Machine Learning, όσο και σύγχρονων μεθόδων Deep Learning. Παράλληλα, η χρονική περίοδος κάλυψης των δεδομένων είναι επαρκής, ώστε να περιλαμβάνει διαφορετικές εποχικές μεταβολές και ποικιλία καιρικών συνθηκών, στοιχείο που ενισχύει τη χρονική κάλυψη του πειραματικού δείγματος των μοντέλων πρόβλεψης.

Για την υποστήριξη της εισαγωγής, ενοποίησης και προετοιμασίας των μετεωρολογικών δεδομένων αξιοποιείται το MetDataPy, ένα εργαλείο που αναπτύχθηκε στο πλαίσιο της παρούσας εργασίας. Το MetDataPy δεν αποτελεί μοντέλο πρόγνωσης, αλλά βιβλιοθήκη προεπεξεργασίας και προετοιμασίας μετεωρολογικών χρονοσειρών για downstream εφαρμογές machine learning και deep learning (Kartas, 2026).

Στο πλαίσιο της παρούσας μεθοδολογίας, το MetDataPy χρησιμοποιείται για τη μετατροπή των πρωτογενών δεδομένων του μετεωρολογικού σταθμού σε ενιαία και επαναχρησιμοποιήσιμη μορφή. Η λειτουργικότητά του περιλαμβάνει mapping-driven ingestion ετερογενών αρχείων, ενοποίηση σχήματος δεδομένων, κανονικοποίηση μονάδων μέτρησης, διαχείριση χρονικών δεικτών, εισαγωγή χρονικών κενών, flagging προβληματικών παρατηρήσεων, δημιουργία παράγωγων μετεωρολογικών μεταβλητών και παραγωγή ML-ready datasets με lags, forecast horizons και χρονικά έγκυρο διαχωρισμό δεδομένων.

Η αξιοποίηση του MetDataPy ενισχύει την επαναληψιμότητα της πειραματικής διαδικασίας, καθώς οι αποφάσεις προεπεξεργασίας γίνονται πιο σαφείς, τεκμηριωμένες και επαναχρησιμοποιήσιμες. Με τον τρόπο αυτό, το εργαλείο λειτουργεί ως υποδομή προετοιμασίας δεδομένων, ενώ η εκπαίδευση και αξιολόγηση των μοντέλων πρόγνωσης πραγματοποιείται σε επόμενα στάδια της μεθοδολογίας.

3.4 Προεπεξεργασία Δεδομένων

Η προεπεξεργασία των δεδομένων αποτελεί κρίσιμο στάδιο στην ανάπτυξη μοντέλων πρόγνωσης καιρού, καθώς οι μετεωρολογικές χρονοσειρές συχνά περιλαμβάνουν χρονικά κενά, ελλιπείς τιμές, ασυνέπειες μορφοποίησης, θόρυβο και μη ρεαλιστικές παρατηρήσεις. Για τον λόγο αυτό, πριν από τη δημιουργία χαρακτηριστικών και την εκπαίδευση των μοντέλων εφαρμόζεται οργανωμένο πρωτόκολλο προετοιμασίας, με στόχο τη δημιουργία ενός συνεπούς και αξιόπιστου συνόλου δεδομένων.

Τα πρωτογενή δεδομένα εισάγονται από αρχεία CSV που εξάγονται από την πλατφόρμα Weathercloud. Κάθε γραμμή αντιστοιχεί σε μία χρονική παρατήρηση του μετεωρολογικού σταθμού, ενώ οι στήλες αντιστοιχούν στις διαθέσιμες μετεωρολογικές μεταβλητές. Στο διαθέσιμο δείγμα, η χρονική σήμανση δίνεται στη στήλη Date (Europe/Athens), οι παρατηρήσεις έχουν χρονική ανάλυση 10 λεπτών, και οι μονάδες μέτρησης περιλαμβάνονται στα ονόματα των πεδίων.

Μετά την εισαγωγή, τα δεδομένα χαρτογραφούνται μέσω του MetDataPy σε ενιαίο canonical schema, με κοινή ονοματολογία, συνεπείς μετρικές μονάδες και χρονικό δείκτη σε UTC. Η ενιαία δομή περιλαμβάνει τις βασικές εξωτερικές μετεωρολογικές μεταβλητές που παρέχονται από τον σταθμό, όπως θερμοκρασία, σχετική υγρασία, ατμοσφαιρική πίεση, μέση ταχύτητα ανέμου, ριπή ανέμου, διεύθυνση ανέμου, βροχόπτωση, ένταση βροχής και ηλιακή ακτινοβολία.

Σε επίπεδο υλοποίησης, η βασική canonical δομή του dataset περιλαμβάνει κατ' ελάχιστον τα πεδία `ts_utc: datetime`, `temp_c: float`, `rh_pct: float`, `pres_hpa: float`, `wspd_ms: float`, `gust_ms: float`, `wdir_deg: float`, `rain_mm: float`, `rain_rate_mmh: float` και `solar_wm2: float`. Η δομή αυτή επιτρέπει την ομοιόμορφη επεξεργασία των μεταβλητών και τη μετατροπή των χρονοσειρών σε μορφή κατάλληλη για αλγορίθμους machine learning και deep learning.

Στη συνέχεια εφαρμόζεται ποιοτικός έλεγχος των παρατηρήσεων. Ο έλεγχος αυτός περιλαμβάνει τον εντοπισμό τιμών που παραβιάζουν φυσικά ή μετεωρολογικά αποδεκτά όρια, όπως αρνητική σχετική υγρασία, σχετική υγρασία μεγαλύτερη από 100%, αρνητική ταχύτητα ανέμου, αρνητική βροχόπτωση ή μη ρεαλιστικές τιμές ατμοσφαιρικής πίεσης. Επιπλέον, εξετάζονται απότομες μεταβολές που μπορεί να οφείλονται σε σφάλματα αισθητήρων ή μετάδοσης δεδομένων, καθώς και περιπτώσεις όπου μία μεταβλητή παραμένει σταθερή για ασυνήθιστα μεγάλο χρονικό διάστημα.

Ο ποιοτικός έλεγχος υλοποιείται μέσω του MetDataPy, το οποίο εφαρμόζει μη καταστροφική λογική επισήμανσης (non-destructive flagging). Με την προσέγγιση αυτή, ύποπτες ή μη ρεαλιστικές παρατηρήσεις δεν αφαιρούνται αυτόματα κατά τον αρχικό έλεγχο, αλλά επισημαίνονται με αντίστοιχα quality-control flags. Με τον τρόπο αυτό, η πολιτική διατήρησης ή αποκλεισμού παρατηρήσεων από το τελικό supervised dataset παραμένει τεκμηριωμένη και αναπαραγώγιμη.

Η διαχείριση των ελλিপών τιμών πραγματοποιείται ανάλογα με τη διάρκεια του κενού και τη φύση της μεταβλητής. Μικρά κενά σε συνεχείς μεταβλητές, όπως η θερμοκρασία, η υγρασία και η ατμοσφαιρική πίεση, μπορούν να αντιμετωπιστούν με χρονικά συνεπή συμπλήρωση ή παρεμβολή. Αντίθετα, για μεταβλητές γεγονότων ή συσσώρευσης, όπως η βροχόπτωση, η συμπλήρωση γίνεται πιο συντηρητικά, ώστε να μην δημιουργούνται τεχνητά καιρικά φαινόμενα. Μεγαλύτερα χρονικά κενά δεν συμπληρώνονται αυθαίρετα και μπορούν να οδηγήσουν σε αποκλεισμό των αντίστοιχων εγγραφών από το supervised dataset, όταν επηρεάζουν τα απαιτούμενα lags ή τον αντίστοιχο forecast horizon.

Κατά τη χρονική ευθυγράμμιση της χρονοσειράς, το MetDataPy μπορεί να εισάγει τις αναμενόμενες χρονικές εγγραφές που απουσιάζουν και να τις επισημαίνει με βοηθητικά πεδία, όπως `gap: boolean`. Αντίστοιχα, κατά τον ποιοτικό έλεγχο μπορούν να παραχθούν συνοπτικά πεδία, όπως `qc_any: boolean`, τα οποία δηλώνουν αν μία παρατήρηση έχει αποτύχει σε κάποιον έλεγχο ποιότητας.

Στο πειραματικό pipeline, τα αριθμητικά quality-control και gap-related πεδία μπορούν να συμπεριληφθούν ως βοηθητικά χαρακτηριστικά εισόδου, χωρίς να αποτελούν κύριες

μετεωρολογικές μεταβλητές ή στόχους πρόβλεψης. Με τον τρόπο αυτό, τα μοντέλα μπορούν να αξιοποιήσουν πληροφορία για την πληρότητα και την αξιοπιστία των παρατηρήσεων.

Μετά τον καθαρισμό, τον ποιοτικό έλεγχο και τη διαχείριση των χρονικών κενών, εφαρμόζεται κλιμάκωση των αριθμητικών μεταβλητών. Στην παρούσα εργασία χρησιμοποιείται κυρίως Min-Max scaling στο διάστημα $[0, 1]$, ώστε οι μεταβλητές με διαφορετικές μονάδες μέτρησης να μετατραπούν σε συγκρίσιμη αριθμητική κλίμακα. Για μία μεταβλητή x , η μετασχηματισμένη τιμή x' υπολογίζεται ως:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

όπου x_{min} και x_{max} υπολογίζονται αποκλειστικά από το training set. Η ίδια μετασχηματιστική συνάρτηση εφαρμόζεται στη συνέχεια στα validation και test sets, ώστε να αποφευχθεί η διαρροή πληροφορίας από μελλοντικά δεδομένα. Η πρακτική αυτή είναι απαραίτητη σε προβλήματα πρόγνωσης χρονοσειρών, καθώς οποιαδήποτε χρήση στατιστικών από το validation ή test set κατά την προεπεξεργασία θα μπορούσε να οδηγήσει σε μεροληπτική αξιολόγηση των μοντέλων.

Η συνολική διαδικασία προεπεξεργασίας υποστηρίζεται από το MetDataPy, το οποίο λειτουργεί ως υποδομή εισαγωγής, ενοποίησης, χρονικής ευθυγράμμισης, ποιοτικού ελέγχου και προετοιμασίας των μετεωρολογικών χρονοσειρών. Το εργαλείο δεν αποτελεί μοντέλο πρόγνωσης, αλλά χρησιμοποιείται για τη δημιουργία τεκμηριωμένων και επαναλήψιμων datasets που μπορούν να αξιοποιηθούν στα επόμενα στάδια της πειραματικής διαδικασίας.

Η διαδικασία προεπεξεργασίας αντιστοιχεί στο στάδιο Data Preprocessing της αρχιτεκτονικής του συστήματος, όπως παρουσιάζεται στο Σχήμα 3.1, καθώς και στο αρχικό στάδιο του πειραματικού workflow του Σχήματος 3.2.

3.5 Δημιουργία Χαρακτηριστικών (Feature Engineering)

Η διαδικασία δημιουργίας χαρακτηριστικών (Feature Engineering) αποτελεί ένα από τα σημαντικότερα στάδια στην ανάπτυξη μοντέλων πρόβλεψης καιρού, καθώς επιτρέπει την εξαγωγή επιπλέον πληροφορίας από τις διαθέσιμες μετεωρολογικές μεταβλητές και συμβάλλει στη βελτίωση της απόδοσης των αλγορίθμων μηχανικής μάθησης.

Στο πλαίσιο της παρούσας εργασίας, εφαρμόζεται διαδικασία δημιουργίας χαρακτηριστικών κατάλληλων για forecasting εφαρμογές, με στόχο τον εμπλουτισμό της αναπαράστασης των μετεωρολογικών χρονοσειρών. Για την υποστήριξη της διαδικασίας αυτής αξιοποιείται το εργαλείο MetDataPy.

Αρχικά, δημιουργούνται χρονικά χαρακτηριστικά (temporal features), τα οποία αποτυπώνουν τη χρονική θέση κάθε παρατήρησης μέσα σε ημερήσιους και εποχικούς κύκλους. Η επιλογή τους βασίζεται στο γεγονός ότι βασικές μετεωρολογικές μεταβλητές, όπως η θερμοκρασία, η υγρασία και η ηλιακή ακτινοβολία, εμφανίζουν έντονη ημερήσια και ετήσια περιοδικότητα. Για τον λόγο αυτό, από τον χρονικό δείκτη των δεδομένων εξάγονται η ώρα της ημέρας και η ημέρα του έτους. Η ώρα της ημέρας χρησιμοποιείται για την αποτύπωση της ημερήσιας περιοδικότητας, ενώ η ημέρα του έτους χρησιμοποιείται για την αποτύπωση της ετήσιας/εποχικής μεταβολής των μετεωρολογικών συνθηκών.

Για χρονικές μεταβλητές με κυκλική συμπεριφορά δεν χρησιμοποιείται αποκλειστικά η απλή ακέραια αναπαράσταση, καθώς αυτή μπορεί να εισάγει τεχνητές ασυνέχειες στα όρια του κύκλου. Για παράδειγμα, η ώρα 23:00 μιας ημέρας και η ώρα 00:00 της επόμενης ημέρας είναι διαδοχικές χρονικά, αλλά αν η ώρα αναπαρασταθεί ως απλός ακέραιος αριθμός, οι τιμές 23 και 0 εμφανίζονται αριθμητικά απομακρυσμένες. Για την αντιμετώπιση αυτού του προβλήματος, εφαρμόζεται κυκλική κωδικοποίηση με χρήση ημιτόνου και συνημιτόνου.

Συγκεκριμένα, για την ώρα της ημέρας δημιουργούνται τα χαρακτηριστικά:

$$\begin{aligned} hour_{sin} &= \sin\left(\frac{2\pi \cdot hour}{24}\right) \\ hour_{cos} &= \cos\left(\frac{2\pi \cdot hour}{24}\right) \end{aligned}$$

Αντίστοιχα, για την ετήσια/εποχική περιοδικότητα δημιουργούνται χαρακτηριστικά με βάση την ημέρα του έτους:

$$\begin{aligned} day_of_year_{sin} &= \sin\left(\frac{2\pi \cdot day_of_year}{365}\right) \\ day_of_year_{cos} &= \cos\left(\frac{2\pi \cdot day_of_year}{365}\right) \end{aligned}$$

Η δημιουργία χρονικών χαρακτηριστικών υποστηρίζεται από το MetDataPy μέσω της λειτουργικότητας `calendar features`. Στο πειραματικό pipeline, τα χρονικά χαρακτηριστικά διαμορφώνονται ώστε να περιλαμβάνουν κυκλική κωδικοποίηση της ώρας της ημέρας και της ημέρας του έτους. Με αυτόν τον τρόπο, τα μοντέλα μπορούν να αξιοποιήσουν την ημερήσια και εποχική περιοδικότητα των μετεωρολογικών δεδομένων χωρίς να αντιμετωπίζουν τις κυκλικές χρονικές μεταβλητές ως γραμμικές αριθμητικές τιμές.

Στη συνέχεια, δημιουργούνται χαρακτηριστικά χρονικής υστέρησης (`lag features`), τα οποία επιτρέπουν στα μοντέλα να αξιοποιούν την ιστορική πληροφορία των μετεωρολογικών μεταβλητών. Δεδομένου ότι τα πρωτογενή δεδομένα έχουν χρονική ανάλυση 10 λεπτών, τα `lag features` ορίζονται σε αριθμό χρονικών βημάτων της χρονοσειράς. Στην παρούσα εργασία χρησιμοποιούνται τα χρονικά βήματα $L = \{1, 3, 6, 12, 24, 72, 144\}$, τα οποία αντιστοιχούν σε ιστορικές τιμές 10 λεπτών, 30 λεπτών, 1 ώρας, 2 ωρών, 4 ωρών, 12 ωρών και 24 ωρών πριν από τη χρονική στιγμή πρόβλεψης.

Για κάθε επιλεγμένο `lag` $l \in L$, δημιουργείται αντίστοιχο χαρακτηριστικό για κάθε βασική μετεωρολογική μεταβλητή. Για παράδειγμα, το χαρακτηριστικό `temp_c_lag6` αντιστοιχεί στην τιμή της θερμοκρασίας μία ώρα πριν, ενώ το `rh_pct_lag144` αντιστοιχεί στην τιμή της σχετικής υγρασίας 24 ώρες πριν. Η χρήση πολλαπλών χρονικών υστερήσεων επιτρέπει στα μοντέλα να αξιοποιούν τόσο πρόσφατες μεταβολές όσο και ημερήσια επαναλαμβανόμενα μοτίβα.

Επιπλέον, δημιουργούνται χαρακτηριστικά κυλίσμενων στατιστικών (`rolling features`) για τις βασικές αριθμητικές μετεωρολογικές μεταβλητές του `canonical dataset`. Συγκεκριμένα, τα `rolling features` εφαρμόζονται στα πεδία `temp_c`, `rh_pct`, `pres_hpa`, `wspd_ms`, `gust_ms`, `rain_mm`, `rain_rate_mmh` και `solar_wm2`. Για κάθε μεταβλητή υπολογίζονται κυλίσμενος μέσος όρος, κυλίσμενη τυπική απόκλιση, ελάχιστη και μέγιστη τιμή.

Τα χρονικά παράθυρα που χρησιμοποιούνται είναι 1 ώρα, 3 ώρες, 6 ώρες και 24 ώρες. Δεδομένου ότι η χρονική ανάλυση των δεδομένων είναι 10 λεπτά, τα παράθυρα αυτά αντιστοιχούν αντίστοιχα σε 6, 18, 36 και 144 χρονικά βήματα. Με αυτόν τον τρόπο, τα μοντέλα μπορούν να αξιοποιήσουν τόσο βραχυπρόθεσμες μεταβολές όσο και μεγαλύτερης διάρκειας τάσεις των μετεωρολογικών μεταβλητών.

Η διεύθυνση ανέμου δεν χρησιμοποιείται απευθείας ως `wdir_deg` για τον υπολογισμό απλών κυλιόμενων στατιστικών, επειδή αποτελεί κυκλική μεταβλητή. Πριν από τη χρήση της, μετασχηματίζεται σε δύο κυκλικά χαρακτηριστικά, `wdir_sin` και `wdir_cos`, με βάση τις σχέσεις:

$$\begin{aligned}wdir_{sin} &= \sin\left(\frac{2\pi \cdot wdir_deg}{360}\right) \\wdir_{cos} &= \cos\left(\frac{2\pi \cdot wdir_deg}{360}\right)\end{aligned}$$

Τα κυλιόμενα στατιστικά που αφορούν τη διεύθυνση ανέμου υπολογίζονται πάνω στα `wdir_sin` και `wdir_cos`, ώστε να αποφεύγεται η λανθασμένη γραμμική μεταχείριση γωνιακών τιμών, όπως 350° και 10°.

Παράλληλα, δημιουργούνται παράγωγα μετεωρολογικά χαρακτηριστικά (derived meteorological features), τα οποία προκύπτουν από συνδυασμό βασικών μετρήσεων του σταθμού. Στην παρούσα εργασία χρησιμοποιούνται τα χαρακτηριστικά dew point, vapour pressure deficit (VPD), heat index και wind chill, τα οποία παράγονται μέσω του `MetDataPy` ως μέρος της διαδικασίας εμπλουτισμού χαρακτηριστικών.

Το dew point υπολογίζεται από τη θερμοκρασία αέρα και τη σχετική υγρασία, δηλαδή από τα πεδία `temp_c` και `rh_pct`. Το χαρακτηριστικό αυτό εκφράζει τη θερμοκρασία στην οποία ο αέρας θα έφτανε σε κορεσμό υδρατμών, εφόσον η πίεση παρέμενε σταθερή. Η χρήση του είναι σχετική με τη μοντελοποίηση της υγρασίας και των συνθηκών συμπύκνωσης.

Το vapour pressure deficit (VPD) υπολογίζεται επίσης από τη θερμοκρασία αέρα και τη σχετική υγρασία. Εκφράζει τη διαφορά μεταξύ της κορεσμένης και της πραγματικής πίεσης υδρατμών και χρησιμοποιείται ως δείκτης ξηρότητας της ατμόσφαιρας. Η σχέση του με τη θερμοκρασία και την υγρασία το καθιστά χρήσιμο συμπληρωματικό χαρακτηριστικό για την περιγραφή τοπικών μετεωρολογικών συνθηκών.

Το heat index υπολογίζεται από τη θερμοκρασία αέρα και τη σχετική υγρασία και εκφράζει την αντιληπτή θερμοκρασία υπό συνθήκες θερμικής επιβάρυνσης. Το χαρακτηριστικό αυτό εισάγει συνδυαστική πληροφορία για τη θερμοκρασία και την υγρασία, η οποία δεν αποτυπώνεται πλήρως όταν οι δύο μεταβλητές χρησιμοποιούνται ανεξάρτητα.

Το wind chill υπολογίζεται από τη θερμοκρασία αέρα και την ταχύτητα ανέμου, δηλαδή από τα πεδία `temp_c` και `wspd_ms`. Εκφράζει την αντιληπτή θερμοκρασία υπό την επίδραση του ανέμου και μπορεί να προσφέρει πρόσθετη πληροφορία σε περιόδους χαμηλότερων θερμοκρασιών και αυξημένης ανεμολογικής δραστηριότητας.

Τα παραπάνω παράγωγα χαρακτηριστικά δεν αποτελούν ανεξάρτητες νέες παρατηρήσεις, αλλά μετασχηματισμούς βασικών μετεωρολογικών μεταβλητών. Η χρήση τους αποσκοπεί στον εμπλουτισμό της αναπαράστασης εισόδου των μοντέλων με φυσικά ερμηνεύσιμες σχέσεις μεταξύ θερμοκρασίας, υγρασίας και ανέμου. Οι υπολογισμοί πραγματοποιούνται από το

MetDataPy, το οποίο παρέχει τα συγκεκριμένα derived variables ακολουθώντας καθιερωμένες διατυπώσεις της βιβλιογραφίας (Kartas, 2026).

Η διαδικασία δημιουργίας χαρακτηριστικών αντιστοιχεί στο στάδιο Feature Engineering της αρχιτεκτονικής του συστήματος (Σχήμα 3.1) και στο αρχικό μέρος του πειραματικού workflow (Σχήμα 3.2), όπου τα προεπεξεργασμένα δεδομένα μετατρέπονται σε σύνολο χαρακτηριστικών κατάλληλο για την εκπαίδευση μοντέλων Machine Learning και Deep Learning.

3.6 Δημιουργία Πειραματικών Συνόλων Δεδομένων

Μετά τη διαδικασία προεπεξεργασίας και δημιουργίας χαρακτηριστικών, ακολουθεί η δημιουργία των πειραματικών συνόλων δεδομένων, τα οποία χρησιμοποιούνται για την εκπαίδευση και αξιολόγηση των μοντέλων πρόβλεψης. Η σωστή διαμόρφωση των συνόλων δεδομένων είναι ιδιαίτερα σημαντική στις εφαρμογές πρόβλεψης χρονοσειρών, καθώς απαιτείται η διατήρηση της χρονικής ακολουθίας των δεδομένων.

Αρχικά, τα δεδομένα μετατρέπονται σε μορφή κατάλληλη για επιβλεπόμενη μάθηση (supervised learning). Η διαδικασία αυτή περιλαμβάνει τη μετατροπή των χρονοσειρών σε σύνολα εισόδων και εξόδων, όπου οι προηγούμενες τιμές των μεταβλητών χρησιμοποιούνται για την πρόβλεψη μελλοντικών τιμών. Η προσέγγιση αυτή είναι ιδιαίτερα σημαντική για την εφαρμογή αλγορίθμων Machine Learning και Deep Learning σε δεδομένα χρονοσειρών.

Στη συνέχεια, ορίζονται διαφορετικοί χρονικοί ορίζοντες πρόβλεψης (forecast horizons), προκειμένου να αξιολογηθεί η απόδοση των μοντέλων σε διαφορετικά χρονικά διαστήματα. Οι χρονικοί ορίζοντες που εξετάζονται περιλαμβάνουν βραχυπρόθεσμη πρόβλεψη (π.χ. 1 ώρα), μεσοπρόθεσμη βραχυχρόνια πρόβλεψη (π.χ. 3 ώρες) και εκτεταμένη βραχυπρόθεσμη πρόβλεψη (π.χ. 6 ή 12 ώρες). Η χρήση πολλαπλών χρονικών οριζόντων επιτρέπει την αξιολόγηση της απόδοσης των μοντέλων σε διαφορετικά επίπεδα δυσκολίας.

Έπειτα, τα δεδομένα διαχωρίζονται σε τρία βασικά σύνολα: σύνολο εκπαίδευσης (training set), σύνολο επικύρωσης (validation set) και σύνολο δοκιμής (test set). Ο διαχωρισμός πραγματοποιείται με χρονολογική σειρά (time-based split) και όχι με τυχαία δειγματοληψία ή τυχαίοποίηση των παρατηρήσεων. Η επιλογή αυτή είναι απαραίτητη, επειδή τα δεδομένα αποτελούν χρονοσειρά και το ζητούμενο είναι η πρόβλεψη μελλοντικών τιμών από προηγούμενες παρατηρήσεις.

Συγκεκριμένα, τα παλαιότερα δεδομένα χρησιμοποιούνται για την εκπαίδευση των μοντέλων, το επόμενο χρονικό τμήμα χρησιμοποιείται για την επικύρωση και ρύθμιση υπερπαραμέτρων, ενώ το τελευταίο χρονικό τμήμα διατηρείται ως ανεξάρτητο σύνολο δοκιμής. Με τον τρόπο αυτό, η αξιολόγηση προσομοιώνει ρεαλιστικό σενάριο πρόγνωσης, στο οποίο το μοντέλο εκπαιδεύεται σε ιστορικά δεδομένα και αξιολογείται σε μεταγενέστερες παρατηρήσεις.

Η κατανομή των δεδομένων πραγματοποιείται με αναλογία 70% για το training set, 15% για το validation set και 15% για το test set, διατηρώντας πάντα τη χρονική σειρά των εγγραφών. Η τυχαία ανάμειξη των δεδομένων αποφεύγεται, καθώς θα μπορούσε να οδηγήσει σε data leakage, επιτρέποντας σε πληροφορία από μελλοντικές παρατηρήσεις να επηρεάσει έμμεσα την εκπαίδευση ή τη ρύθμιση των μοντέλων.

Η διαδικασία δημιουργίας πειραματικών συνόλων δεδομένων αντιστοιχεί στο στάδιο Dataset Preparation της αρχιτεκτονικής του συστήματος (Σχήμα 3.1), καθώς και στο αρχικό μέρος του

πειραματικού workflow (Σχήμα 3.2), όπου τα επεξεργασμένα δεδομένα μετατρέπονται σε μορφή κατάλληλη για εκπαίδευση και αξιολόγηση των μοντέλων πρόβλεψης.

3.7 Μοντέλα Πρόβλεψης

Στην παρούσα εργασία εξετάζονται διαφορετικές κατηγορίες μοντέλων πρόβλεψης με στόχο τη συγκριτική αξιολόγηση της απόδοσης τεχνικών Machine Learning και Deep Learning στη βραχυπρόθεσμη πρόγνωση καιρού. Η επιλογή των μοντέλων βασίζεται τόσο στη σχετική βιβλιογραφία όσο και στην καταλληλότητά τους για δεδομένα χρονοσειρών, καθώς και στην ανάγκη σύγκρισης απλών και πιο σύνθετων προσεγγίσεων στο ίδιο πειραματικό πλαίσιο.

Για τη συστηματική αξιολόγηση των μοντέλων, τα πειράματα οργανώνονται σε τρεις βασικές κατηγορίες: Baseline Models, Machine Learning Models και Deep Learning Models. Η προσέγγιση αυτή επιτρέπει τη σύγκριση της απόδοσης απλών μεθόδων με πιο σύνθετες τεχνικές, παρέχοντας μια ολοκληρωμένη εικόνα της αποτελεσματικότητας των διαφορετικών μοντέλων πρόβλεψης. Η συνολική διαδικασία επιλογής και αξιολόγησης των μοντέλων παρουσιάζεται στο πειραματικό workflow του Σχήματος 3.2, όπου για κάθε χρονικό ορίζοντα πρόβλεψης εκπαιδεύονται διαφορετικές οικογένειες μοντέλων και πραγματοποιείται συγκριτική αξιολόγηση.

3.7.1 Baseline Models

Η χρήση baseline models αποτελεί καθιερωμένη πρακτική σε προβλήματα πρόγνωσης χρονοσειρών, καθώς παρέχει ένα απλό σημείο αναφοράς για την αξιολόγηση πιο σύνθετων μοντέλων. Σε μελέτες πρόγνωσης καιρού και μετεωρολογικών χρονοσειρών, απλές προσεγγίσεις όπως το Persistence Model και στατιστικές μέθοδοι αναφοράς χρησιμοποιούνται συχνά ώστε να εξεταστεί αν τα μοντέλα machine learning και deep learning προσφέρουν ουσιαστική βελτίωση έναντι απλών κανόνων πρόβλεψης (Hewage et al., 2020; Rahman et al., 2024; Zhang et al., 2025). Για τον λόγο αυτό, στην παρούσα εργασία τα baseline μοντέλα χρησιμοποιούνται ως ελάχιστο σημείο σύγκρισης για την αποτίμηση της προστιθέμενης αξίας των πιο σύνθετων προσεγγίσεων.

Στην παρούσα εργασία εξετάζονται τα ακόλουθα baseline μοντέλα:

Persistence Model

Το μοντέλο Persistence αποτελεί μία από τις απλούστερες μεθόδους πρόβλεψης χρονοσειρών. Σύμφωνα με την προσέγγιση αυτή, η μελλοντική τιμή μιας μεταβλητής θεωρείται ίση με την τελευταία διαθέσιμη παρατήρηση. Παρά την απλότητά του, το μοντέλο Persistence αποτελεί σημαντικό σημείο αναφοράς, ιδιαίτερα σε βραχυπρόθεσμες προβλέψεις.

Moving Average

Το μοντέλο Moving Average βασίζεται στον υπολογισμό του μέσου όρου προηγούμενων τιμών της χρονοσειράς. Η μέθοδος αυτή επιτρέπει την εξομάλυνση των δεδομένων και την παραγωγή πιο σταθερών προβλέψεων.

Climatology

Το μοντέλο Climatology χρησιμοποιείται ως στατικό μετεωρολογικό baseline και προβλέπει τη μελλοντική τιμή βάσει ιστορικού μέσου όρου που υπολογίζεται αποκλειστικά από το training set.

3.7.2 Μοντέλα Machine Learning

Στην κατηγορία των μοντέλων Machine Learning εξετάζονται παραδοσιακοί αλγόριθμοι παλινδρόμησης και πρόβλεψης χρονοσειρών, οι οποίοι αξιοποιούν τα χαρακτηριστικά που παράγονται κατά το στάδιο του feature engineering. Τα μοντέλα αυτά επιλέχθηκαν λόγω της ευρείας χρήσης τους στη βιβλιογραφία, αλλά και επειδή αποτελούν ισχυρά σημεία σύγκρισης σε προβλήματα πρόγνωσης που βασίζονται σε δομημένες εισόδους και μη γραμμικές σχέσεις μεταξύ μεταβλητών.

Τα μοντέλα Machine Learning που εξετάζονται περιλαμβάνουν:

Linear Regression

Η γραμμική παλινδρόμηση αποτελεί μία από τις βασικότερες μεθόδους πρόβλεψης και χρησιμοποιείται ως σημείο αναφοράς για πιο σύνθετα μοντέλα.

Ridge Regression

Το Ridge Regression αποτελεί regularized εκδοχή της γραμμικής παλινδρόμησης με L_2 ποινή και χρησιμοποιείται ως πιο σταθερή γραμμική εναλλακτική έναντι της απλής Linear Regression.

Random Forest

Το Random Forest αποτελεί ensemble μέθοδο, η οποία βασίζεται στη δημιουργία πολλαπλών δέντρων απόφασης. Η μέθοδος αυτή είναι ιδιαίτερα αποτελεσματική στη διαχείριση μη γραμμικών σχέσεων μεταξύ των μεταβλητών.

Gradient Boosting

Το Gradient Boosting αποτελεί ισχυρή τεχνική ensemble learning, η οποία χρησιμοποιείται ευρέως σε προβλήματα πρόβλεψης χρονοσειρών και σε εφαρμογές παλινδρόμησης υψηλής απόδοσης.

Support Vector Regression (SVR)

Το Support Vector Regression αποτελεί επέκταση των Support Vector Machines για προβλήματα παλινδρόμησης και χρησιμοποιείται συχνά σε εφαρμογές πρόβλεψης λόγω της ικανότητάς του να διαχειρίζεται σύνθετες σχέσεις στα δεδομένα.

3.7.3 Μοντέλα Deep Learning

Για την πρόβλεψη μετεωρολογικών χρονοσειρών εξετάζονται επίσης μοντέλα Deep Learning, τα οποία είναι ιδιαίτερα κατάλληλα για την ανάλυση χρονικά εξαρτώμενων δεδομένων. Η επιλογή των μοντέλων αυτών βασίζεται στην ικανότητά τους να μαθαίνουν σύνθετες χρονικές εξαρτήσεις και μη γραμμικά πρότυπα απευθείας από τις ακολουθίες των δεδομένων.

Τα μοντέλα Deep Learning που εξετάζονται περιλαμβάνουν:

Long Short-Term Memory (LSTM)

Τα LSTM αποτελούν ειδική κατηγορία Recurrent Neural Networks και χρησιμοποιούνται ευρέως για την πρόβλεψη χρονοσειρών. Η αρχιτεκτονική αυτή επιτρέπει την εκμάθηση μακροχρόνιων εξαρτήσεων μεταξύ των δεδομένων.

Gated Recurrent Unit (GRU)

Τα GRU αποτελούν απλοποιημένη εκδοχή των LSTM και παρουσιάζουν χαμηλότερη υπολογιστική πολυπλοκότητα, διατηρώντας παράλληλα υψηλή απόδοση.

Temporal Convolutional Networks (TCN)

Τα Temporal Convolutional Networks αποτελούν σύγχρονη αρχιτεκτονική Deep Learning για δεδομένα χρονοσειρών και χρησιμοποιούνται ολοένα και περισσότερο στη βιβλιογραφία λόγω της ικανότητάς τους να μοντελοποιούν χρονικές εξαρτήσεις με αποδοτικό τρόπο.

Η διαδικασία εκπαίδευσης και αξιολόγησης των μοντέλων πραγματοποιείται για κάθε χρονικό ορίζοντα πρόβλεψης, όπως παρουσιάζεται στο Σχήμα 3.2, επιτρέποντας τη συγκριτική αξιολόγηση των διαφορετικών προσεγγίσεων στο ίδιο πειραματικό πλαίσιο.

3.8 Εκπαίδευση Μοντέλων

Η εκπαίδευση των μοντέλων αποτελεί ένα από τα βασικά στάδια της προτεινόμενης μεθοδολογίας, καθώς μέσω αυτής αξιολογείται η ικανότητα διαφορετικών προσεγγίσεων πρόβλεψης να μοντελοποιούν τη δυναμική των μετεωρολογικών χρονοσειρών. Στην παρούσα εργασία, η διαδικασία εκπαίδευσης οργανώνεται με τρόπο που επιτρέπει τη συγκριτική αξιολόγηση διαφορετικών μοντέλων πρόβλεψης για βραχυπρόθεσμη τοπική πρόγνωση καιρού υπό ενιαίο πειραματικό πλαίσιο.

Η εκπαίδευση των μοντέλων πραγματοποιείται χρησιμοποιώντας τα σύνολα δεδομένων που δημιουργήθηκαν στο προηγούμενο στάδιο της μεθοδολογίας, δηλαδή το σύνολο εκπαίδευσης (training set), το σύνολο επικύρωσης (validation set) και το σύνολο δοκιμής (test set). Η χρήση διακριτών συνόλων δεδομένων είναι απαραίτητη για την ορθή ανάπτυξη και αξιολόγηση των μοντέλων, καθώς επιτρέπει τη ρύθμιση των παραμέτρων και των υπερπαραμέτρων χωρίς να επηρεάζεται η αντικειμενικότητα της τελικής αξιολόγησης.

Η διαδικασία εκπαίδευσης εφαρμόζεται ξεχωριστά για κάθε μοντέλο και για κάθε χρονικό ορίζοντα πρόβλεψης. Με τον τρόπο αυτό, όλα τα μοντέλα αξιολογούνται υπό τις ίδιες συνθήκες, γεγονός που διασφαλίζει τη συγκρισιμότητα των αποτελεσμάτων. Επιπλέον, για κάθε χρονικό ορίζοντα πρόβλεψης χρησιμοποιούνται τα ίδια σύνολα εισόδου και ο ίδιος διαχωρισμός σε training, validation και test δεδομένα, ώστε οι διαφορές στην απόδοση να αποδίδονται κυρίως στα χαρακτηριστικά της ίδιας της μοντελοποιητικής προσέγγισης.

Για τα μοντέλα Machine Learning, η εκπαίδευση βασίζεται στα χαρακτηριστικά που έχουν εξαχθεί κατά το στάδιο του feature engineering. Η επιλογή των υπερπαραμέτρων πραγματοποιείται με χρήση του συνόλου επικύρωσης, με στόχο τη βελτιστοποίηση της απόδοσης κάθε μοντέλου χωρίς αξιοποίηση του συνόλου δοκιμής. Η προσέγγιση αυτή είναι ιδιαίτερα σημαντική, καθώς συμβάλλει στη βελτίωση της ικανότητας γενίκευσης των μοντέλων σε μη παρατηρημένα δεδομένα.

Στην περίπτωση των μοντέλων Deep Learning, η εκπαίδευση πραγματοποιείται μέσω επαναληπτικής διαδικασίας βελτιστοποίησης των παραμέτρων του δικτύου. Κατά την εκπαίδευση εφαρμόζονται πρακτικές παρακολούθησης της απόδοσης στο σύνολο επικύρωσης, με στόχο τη σταθερότητα της μάθησης και τον περιορισμό του κινδύνου υπερεκπαίδευσης. Η προσέγγιση αυτή είναι ιδιαίτερα σημαντική σε προβλήματα πρόβλεψης χρονοσειρών, όπου η αυξημένη πολυπλοκότητα των μοντέλων μπορεί να οδηγήσει σε μειωμένη ικανότητα γενίκευσης.

Η εκπαίδευση των baseline μοντέλων εντάσσεται στο ίδιο πειραματικό πλαίσιο, ώστε να είναι δυνατή η άμεση σύγκριση της απόδοσής τους με πιο σύνθετες προσεγγίσεις. Με τον τρόπο αυτό διασφαλίζεται ότι όλα τα μοντέλα συγκρίνονται με κοινή μεθοδολογική λογική, κοινό διαχωρισμό δεδομένων και ενιαία διαδικασία προεπεξεργασίας και κανονικοποίησης. Η

διαδικασία εκπαίδευσης αντιστοιχεί στο στάδιο Model Training του συνολικού workflow της μεθοδολογίας και αποτελεί κρίσιμο βήμα για τη διερεύνηση των ερευνητικών υποθέσεων της εργασίας. Μέσα από την ενιαία αυτή πειραματική προσέγγιση, καθίσταται δυνατή η αξιόπιστη σύγκριση baseline, machine learning και deep learning μοντέλων ως προς την ικανότητά τους να υποστηρίξουν βραχυπρόθεσμη πρόγνωση καιρού σε τοπική κλίμακα.

3.9 Αξιολόγηση Μοντέλων

Η αξιολόγηση των μοντέλων αποτελεί κρίσιμο στάδιο της προτεινόμενης μεθοδολογίας, καθώς επιτρέπει τη συγκριτική ανάλυση της απόδοσης των διαφορετικών τεχνικών πρόβλεψης. Στην παρούσα εργασία, η αξιολόγηση των μοντέλων πραγματοποιείται χρησιμοποιώντας το σύνολο δοκιμής (test set), το οποίο δεν χρησιμοποιείται κατά τη διαδικασία εκπαίδευσης ή ρύθμισης των υπερπαραμέτρων.

Η χρήση ανεξάρτητου συνόλου δοκιμής επιτρέπει την αντικειμενική αξιολόγηση της απόδοσης των μοντέλων και διασφαλίζει ότι τα αποτελέσματα αντικατοπτρίζουν την πραγματική ικανότητα γενίκευσης των μοντέλων σε νέα δεδομένα. Για την αξιολόγηση των μοντέλων πρόβλεψης χρησιμοποιούνται κατάλληλες μετρικές σφάλματος, οι οποίες αξιοποιούνται ευρέως στη βιβλιογραφία για την αξιολόγηση μοντέλων πρόβλεψης χρονοσειρών.

3.9.1 Μετρικές Αξιολόγησης

Στην παρούσα εργασία χρησιμοποιούνται οι ακόλουθες μετρικές αξιολόγησης:

Mean Absolute Error (MAE)

Η μετρική Mean Absolute Error (MAE) υπολογίζει τη μέση απόλυτη διαφορά μεταξύ των πραγματικών και των προβλεπόμενων τιμών. Η μετρική αυτή είναι ιδιαίτερα χρήσιμη, καθώς παρέχει μια εύκολα κατανοητή εκτίμηση του σφάλματος πρόβλεψης.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

όπου y_i είναι η πραγματική τιμή και $\hat{y}_i$ η προβλεπόμενη τιμή.

Root Mean Squared Error (RMSE)

Η μετρική Root Mean Squared Error (RMSE) χρησιμοποιείται ευρέως στην αξιολόγηση μοντέλων πρόβλεψης και δίνει μεγαλύτερη βαρύτητα στα μεγαλύτερα σφάλματα.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Η μετρική αυτή είναι ιδιαίτερα χρήσιμη σε εφαρμογές πρόβλεψης καιρού, όπου μεγάλα σφάλματα μπορεί να έχουν σημαντικό αντίκτυπο.

Mean Absolute Percentage Error (MAPE)

Η μετρική Mean Absolute Percentage Error (MAPE) εκφράζει το σφάλμα πρόβλεψης ως

ποσοστό, επιτρέποντας τη σύγκριση της σχετικής ακρίβειας των μοντέλων. Ωστόσο, η χρήση της είναι καταλληλότερη σε περιπτώσεις όπου οι πραγματικές τιμές δεν είναι μηδενικές ή πολύ κοντά στο μηδέν.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Η μετρική αυτή χρησιμοποιείται συμπληρωματικά προς τις MAE και RMSE, ώστε να παρέχεται και ποσοστιαία αποτίμηση του σφάλματος όπου αυτό είναι μεθοδολογικά κατάλληλο.

Persistence Skill Score

Το persistence skill score χρησιμοποιείται για να εκτιμηθεί η σχετική βελτίωση ενός μοντέλου έναντι του persistence baseline. Στην παρούσα εργασία υπολογίζεται με βάση το τετραγωνικό σφάλμα ως:

$$Skill = 1 - \frac{RMSE_{model}^2}{RMSE_{persistence}^2}$$

Τιμή μεγαλύτερη του μηδενός δηλώνει βελτίωση έναντι του persistence baseline, ενώ αρνητική τιμή δηλώνει χειρότερη απόδοση από το baseline.

3.9.2 Συγκριτική Αξιολόγηση Μοντέλων

Η αξιολόγηση των μοντέλων πραγματοποιείται για κάθε χρονικό ορίζοντα πρόβλεψης και για κάθε μοντέλο ξεχωριστά. Στη συνέχεια, πραγματοποιείται συγκριτική ανάλυση των αποτελεσμάτων, με στόχο τον εντοπισμό του μοντέλου με την καλύτερη απόδοση υπό το ίδιο πειραματικό πλαίσιο.

Η συγκριτική αξιολόγηση επιτρέπει:

- τη σύγκριση μεταξύ baseline, Machine Learning και Deep Learning μοντέλων,
- την αξιολόγηση της επίδρασης του χρονικού ορίζοντα πρόβλεψης στην απόδοση των μοντέλων,
- την επιλογή του καταλληλότερου μοντέλου πρόβλεψης για το υπό εξέταση πρόβλημα.

Η διαδικασία αξιολόγησης αντιστοιχεί στο στάδιο Model Evaluation του πειραματικού workflow, όπως παρουσιάζεται στο Σχήμα 3.2. Τα αποτελέσματα της συγκριτικής ανάλυσης παρουσιάζονται στο Κεφάλαιο 4.

Κεφάλαιο 4. Αποτελέσματα

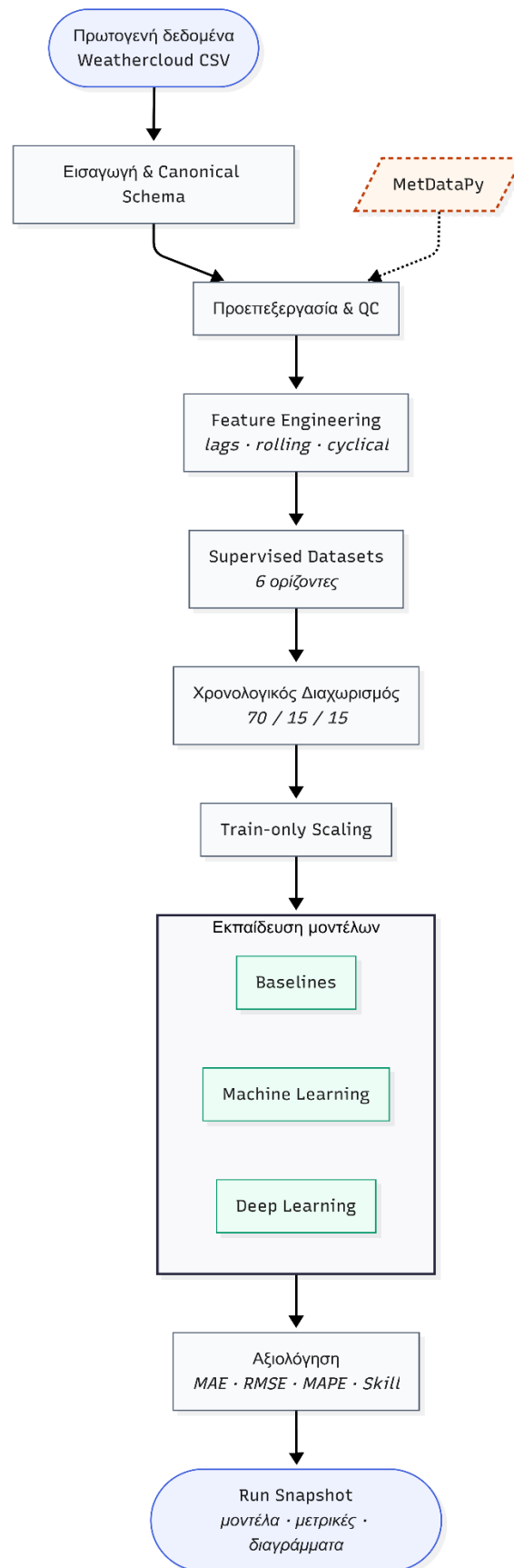
4.1 Παρουσίαση του Λογισμικού

Στο πλαίσιο της παρούσας εργασίας αναπτύχθηκε ένα ολοκληρωμένο λογισμικό πειραματικής πρόγνωσης καιρού, το οποίο υποστηρίζει τη διαδικασία από την εισαγωγή πρωτογενών δεδομένων μετεωρολογικού σταθμού έως την εκπαίδευση, αξιολόγηση και αρχειοθέτηση μοντέλων πρόβλεψης. Το λογισμικό σχεδιάστηκε ειδικά για το ερευνητικό αντικείμενο της εργασίας, δηλαδή τη βραχυπρόθεσμη τοπική πρόγνωση με χρήση αποκλειστικά ιστορικών παρατηρήσεων από έναν μεμονωμένο μετεωρολογικό σταθμό.

Η βασική λειτουργία του συστήματος είναι η αυτοματοποίηση της πειραματικής διαδικασίας σε τοπικό ή server-side περιβάλλον εκτέλεσης. Ο χρήστης τοποθετεί τα αρχεία εξαγωγής Weathercloud στον κατάλληλο φάκελο εισόδου και στη συνέχεια εκτελεί το pipeline μέσω γραμμής εντολών. Το σύστημα αναλαμβάνει την εισαγωγή των δεδομένων, την προεπεξεργασία, τη δημιουργία χαρακτηριστικών, τη δημιουργία supervised datasets για διαφορετικούς χρονικούς ορίζοντες, την εκπαίδευση μοντέλων, την αξιολόγηση και την παραγωγή αναφορών και διαγραμμάτων.

Ο πλήρης πηγαίος κώδικας του πειραματικού pipeline, μαζί με τα αρχεία ρύθμισης και την τεχνική τεκμηρίωση, είναι διαθέσιμος σε δημόσιο GitHub repository: <https://github.com/kkartas/local-weather-forecasting-ml>. Το repository περιλαμβάνει τον Python κώδικα του forecasting pipeline, configuration files, οδηγίες εγκατάστασης και εκτέλεσης, documentation για τα επιμέρους στάδια του pipeline, καθώς και test suite για κρίσιμες λειτουργίες της πειραματικής διαδικασίας.

Η τεκμηρίωση του repository καλύπτει τη διαδικασία από τα raw Weathercloud CSV αρχεία έως την παραγωγή supervised datasets, trained models, fitted scalers, predictions, metrics, plots και frozen run snapshots. Με αυτόν τον τρόπο, το λογισμικό δεν λειτουργεί μόνο ως βοηθητικός κώδικας εκτέλεσης, αλλά ως αναπαραγωγίμη ερευνητική υποδομή για την τελική πειραματική αξιολόγηση της εργασίας.



Σχήμα 4.1 - Συνολική ροή του λογισμικού από raw Weathercloud CSV files έως metrics, plots και run snapshot.

Η αρχιτεκτονική του λογισμικού ακολουθεί διακριτά στάδια. Αρχικά, το στάδιο ingest διαβάζει τα πρωτογενή αρχεία Weathercloud, τα χαρτογραφεί σε ενιαίο canonical schema και μετατρέπει τις χρονικές σημάνσεις σε UTC. Στη συνέχεια, το στάδιο preprocess εφαρμόζει έλεγχο ποιότητας, χειρισμό χρονικών κενών και δημιουργία μετεωρολογικών χαρακτηριστικών. Το στάδιο train δημιουργεί τα supervised datasets, εκτελεί τον χρονολογικό διαχωρισμό σε training, validation και test set, εφαρμόζει κλιμάκωση χωρίς διαρροή πληροφορίας και εκπαιδεύει τα διαθέσιμα baseline, Machine Learning και Deep Learning μοντέλα. Τέλος, το στάδιο evaluate παράγει πίνακες μετρικών, αναφορές και γραφήματα αξιολόγησης.

Κεντρικό στοιχείο του λογισμικού είναι η αξιοποίηση του MetDataPy ως επίσημου επιπέδου προετοιμασίας μετεωρολογικών χρονοσειρών. Το MetDataPy αναλαμβάνει τις επαναχρησιμοποιήσιμες λειτουργίες προετοιμασίας δεδομένων, όπως canonical schema mapping, κανονικοποίηση μονάδων, timestamp normalization, quality-control flagging, gap handling, derived meteorological features, lag features, rolling features και horizon target creation. Το παρόν repository αναλαμβάνει την πειραματική εντοχική διαμόρφωση, δηλαδή configuration, training, evaluation, reports, plots και artifact management.

Για τη χρήση του λογισμικού, ο χρήστης εγκαθιστά τις απαραίτητες εξαρτήσεις, τοποθετεί τα raw Weathercloud CSV αρχεία στον φάκελο data/raw/ και εκτελεί την πλήρη πειραματική διαδικασία με την εντολή:

```
python -m weather_forecasting_pipeline run-all --config configs/default.yaml
```

Εναλλακτικά, τα στάδια μπορούν να εκτελεστούν ανεξάρτητα:

```
python -m weather_forecasting_pipeline ingest --config configs/default.yaml
```

```
python -m weather_forecasting_pipeline preprocess --config configs/default.yaml
```

```
python -m weather_forecasting_pipeline train --config configs/default.yaml
```

```
python -m weather_forecasting_pipeline evaluate --config configs/default.yaml
```

Τα αποτελέσματα της εκτέλεσης αποθηκεύονται σε οργανωμένη δομή φακέλων. Τα ενδιάμεσα datasets αποθηκεύονται στους φακέλους data/interim/ και data/processed/, ενώ τα μοντέλα, οι scalars, οι μετρικές, τα διαγράμματα και οι αναφορές αποθηκεύονται στον φάκελο artifacts/. Επιπλέον, το λογισμικό υποστηρίζει τη δημιουργία πλήρους στιγμιότυπου εκτέλεσης στον φάκελο runs/<run_id>/, ώστε κάθε πείραμα να μπορεί να αρχειοθετηθεί και να αναπαραχθεί.

Σημαντικό επίτευγμα του λογισμικού είναι ότι μετατρέπει μία πολύπλοκη και επιρρεπή σε σφάλματα πειραματική διαδικασία σε επαναλήψιμο workflow. Η χρήση ενιαίου configuration file επιτρέπει τον έλεγχο των forecast horizons, των μοντέλων, των παραμέτρων εκπαίδευσης, των paths και των μετρικών αξιολόγησης. Παράλληλα, η αποθήκευση split metadata, fitted scalars, predictions και metrics επιτρέπει τον έλεγχο της μεθοδολογικής ορθότητας και την αναπαραγωγή των αποτελεσμάτων.

Ένα ακόμη επίτευγμα είναι η ενσωμάτωση κανόνων αποφυγής data leakage. Τα δεδομένα διαχωρίζονται χρονολογικά, χωρίς τυχαία ανάμειξη, ενώ οι scalars προσαρμόζονται αποκλειστικά στο training set και στη συνέχεια εφαρμόζονται στα validation και test sets. Με τον τρόπο αυτό, η αξιολόγηση προσομοιώνει ρεαλιστικό σενάριο πρόβλεψης, όπου το μοντέλο έχει πρόσβαση μόνο σε παρελθοντική πληροφορία.

Το λογισμικό συνοδεύεται από αυτοματοποιημένα tests για κρίσιμα σημεία της πειραματικής διαδικασίας, όπως chronological splits, train-only scaling, metric calculation και MetDataPy integration. Η ύπαρξη test suite, documentation και frozen run snapshots ενισχύει την αξιοπιστία του πειραματικού πλαισίου και καθιστά το λογισμικό κατάλληλο όχι μόνο για την παρούσα εργασία, αλλά και για μελλοντικές επεκτάσεις σε άλλους στόχους πρόβλεψης, διαφορετικούς χρονικούς ορίζοντες ή επιπλέον μετεωρολογικούς σταθμούς.

4.2 Σχόλια και Ανατροφοδότηση

Η αξιολόγηση του λογισμικού δεν περιορίστηκε μόνο στην αριθμητική απόδοση των μοντέλων πρόβλεψης, αλλά περιλάμβανε και τεχνική αποτίμηση της λειτουργικότητας, της επαναληψιμότητας και της καταλληλότητας του συστήματος για το πειραματικό πλαίσιο της εργασίας. Δεδομένου ότι το λογισμικό αναπτύχθηκε κυρίως ως ερευνητικό εργαλείο και όχι ως τελικό εμπορικό προϊόν, η ανατροφοδότηση εστιάζει στην ορθότητα της πειραματικής διαδικασίας, στη σαφήνεια της χρήσης και στην αξιοπιστία των παραγόμενων αποτελεσμάτων.

Σε επίπεδο τεχνικής αξιολόγησης, εξετάστηκε κατά πόσο το λογισμικό υλοποιεί τις βασικές μεθοδολογικές απαιτήσεις της εργασίας. Συγκεκριμένα, ελέγχθηκε ότι η διαδικασία παραμένει observation-only, ότι δεν χρησιμοποιούνται δεδομένα Numerical Weather Prediction ή εξωτερικά προϊόντα πρόγνωσης, ότι τα δεδομένα διαχωρίζονται χρονολογικά, και ότι οι μετασχηματισμοί κλιμάκωσης εφαρμόζονται χωρίς διαρροή πληροφορίας από το validation ή test set. Η αξιολόγηση αυτή είναι ιδιαίτερα σημαντική, καθώς σε προβλήματα πρόγνωσης χρονοσειρών μικρές αποκλίσεις στην προεπεξεργασία μπορούν να οδηγήσουν σε υπεραισιόδοξα αποτελέσματα.

	Μεθοδολογική απαίτηση	Τεκμηρίωση
1	Observation-only είσοδος (καμία NWP / δορυφορική / reanalysis πηγή)	Επιβεβαίωση από το ingest stage· το pipeline δέχεται αποκλειστικά Weathercloud CSV.
2	Single-station setup	Όλα τα δεδομένα προέρχονται από έναν τοπικό σταθμό· καμία γεωγραφική σύνθεση.
3	Χρονολογικός διαχωρισμός 70/15/15 χωρίς ανάμειξη	Καταγεγραμμένα χρονικά όρια ανά ορίζοντα στα split metadata (Πίνακας 4.9).
4	Train-only scaling (αποφυγή leakage)	Οι scalers προσαρμόζονται μόνο στο training set και εφαρμόζονται στα val/test.
5	Ανεξάρτητη εκπαίδευση ανά στόχο και χρονικό ορίζοντα	11 μοντέλα × 6 ορίζοντες = 66 ανεξάρτητα εκπαιδευμένα μοντέλα.
6	Παραγωγή structured artifacts	Predictions, scalers, models, metrics CSV, plots, reports (Πίνακας 4.10).
7	Αναπαραγωγικότητα μέσω frozen run snapshot	Κάθε εκτέλεση παράγει αυτοτελή φάκελο runs/<id>/ με όλα τα artifacts.
8	Αυτοματοποιημένα tests κρίσιμων σημείων (splits, scaling, metrics, integrations)	Test suite που καλύπτει MetDataPy integration, χρονολογικά splits, train-only scaling, υπολογισμό μετρικών και smoke paths μοντέλων.

Πίνακας 4.1 - Σύνοψη τεχνικής αξιολόγησης του λογισμικού ως προς τις μεθοδολογικές απαιτήσεις της εργασίας.

Η χρήση του MetDataPy ως υποχρεωτικού επιπέδου προετοιμασίας δεδομένων αξιολογήθηκε θετικά, καθώς διαχωρίζει με σαφή τρόπο τις ευθύνες του συστήματος. Το MetDataPy αναλαμβάνει τη μετεωρολογική προετοιμασία των χρονοσειρών, ενώ το παρόν λογισμικό επικεντρώνεται στην πειραματική ενορχήστρωση, την εκπαίδευση και την αξιολόγηση. Ο διαχωρισμός αυτός μειώνει την επανάληψη κώδικα, κάνει τη μεθοδολογία πιο καθαρή και επιτρέπει τη μελλοντική βελτίωση της προεπεξεργασίας χωρίς να απαιτείται ριζική αλλαγή στο forecasting pipeline.

Ιδιαίτερα χρήσιμη κρίθηκε και η παραγωγή οργανωμένων artifacts. Κάθε εκτέλεση μπορεί να παράγει ενδιάμεσα datasets, split metadata, fitted scalers, αποθηκευμένα μοντέλα, predictions, μετρικές αξιολόγησης, plots και συνοπτική αναφορά. Με αυτόν τον τρόπο, τα αποτελέσματα δεν παραμένουν μόνο ως έξοδος στην κονσόλα, αλλά αρχειοθετούνται σε μορφή που μπορεί να ελεγχθεί, να συγκριθεί και να χρησιμοποιηθεί στη συγγραφή της εργασίας.

Η ύπαρξη run snapshots αποτελεί επίσης σημαντικό στοιχείο ανατροφοδότησης ως προς την πρακτική χρησιμότητα του λογισμικού. Το snapshot μιας εκτέλεσης συγκεντρώνει σε έναν φάκελο τα configuration files, τα αποτελέσματα, τα predictions, τα plots και τα reports που αντιστοιχούν σε συγκεκριμένο πείραμα. Αυτό επιτρέπει την επιστροφή σε παλαιότερη εκτέλεση χωρίς αμφιβολία για το ποια δεδομένα, ποια μοντέλα και ποιες ρυθμίσεις χρησιμοποιήθηκαν.

Από πλευράς χρήσης, το λογισμικό αξιολογείται ως κατάλληλο για τεχνικό χρήστη που είναι εξοικειωμένος με Python και γραμμή εντολών. Οι βασικές εντολές είναι περιορισμένες και ακολουθούν τη λογική των διακριτών σταδίων του pipeline. Ωστόσο, η χρήση του από μη τεχνικό χρήστη θα απαιτούσε πρόσθετη διεπαφή ή αναλυτικότερο εγχειρίδιο χρήσης, καθώς η τρέχουσα μορφή του συστήματος είναι προσανατολισμένη σε ερευνητική και πειραματική χρήση.

Η αξιολόγηση μέσω αυτοματοποιημένων tests έδειξε ότι το λογισμικό καλύπτει κρίσιμα σημεία της λειτουργικότητας, όπως η ενσωμάτωση με το MetDataPy, η δημιουργία χρονολογικών splits, η αποφυγή data leakage, ο υπολογισμός μετρικών, η εκπαίδευση μοντέλων και η παραγωγή artifacts. Η ύπαρξη αυτών των tests ενισχύει την αξιοπιστία του συστήματος και μειώνει τον κίνδυνο να εισαχθούν σφάλματα σε μελλοντικές αλλαγές του pipeline.

Συνολικά, η ανατροφοδότηση δείχνει ότι το λογισμικό πέτυχε τον βασικό του στόχο: να μετατρέψει μια σύνθετη διαδικασία πρόγνωσης χρονοσειρών σε επαναλήψιμο, τεκμηριωμένο και ελέγξιμο pipeline. Παρά τους περιορισμούς ως προς τη φιλικότητα προς μη τεχνικούς χρήστες, το σύστημα προσφέρει σταθερή βάση για την εκτέλεση των πειραμάτων της εργασίας και για μελλοντική επέκταση σε νέους στόχους, μοντέλα ή μετεωρολογικούς σταθμούς.

4.3 Πειραματική Διάταξη και Δεδομένα Αξιολόγησης

Για την αξιολόγηση του λογισμικού και των μοντέλων πρόβλεψης πραγματοποιήθηκε πλήρης πειραματική εκτέλεση του pipeline με χρήση πραγματικών δεδομένων από τον μετεωρολογικό σταθμό Weathercloud της Νέας Τρίγλιας. Το συγκεκριμένο πείραμα αποτέλεσε τη βασική εκτέλεση αναφοράς για την εξαγωγή των αποτελεσμάτων του Κεφαλαίου 4 και αποθηκεύτηκε ως frozen run snapshot, ώστε τα παραγόμενα αρχεία, οι ρυθμίσεις, τα μοντέλα και οι μετρικές να μπορούν να ελεγχθούν και να αναπαραχθούν.

Το dataset αποτελείται από μετεωρολογικές παρατηρήσεις χρονικής ανάλυσης 10 λεπτών. Τα πρωτογενή δεδομένα προέρχονται από αρχεία CSV της πλατφόρμας Weathercloud και

καλύπτουν χρονική περίοδο από τις 28 Φεβρουαρίου 2023 έως τις 16 Μαΐου 2026. Μετά το στάδιο εισαγωγής, το canonical dataset περιείχε 168.829 εγγραφές και 10 βασικές μετεωρολογικές μεταβλητές. Μετά την προεπεξεργασία, την εισαγωγή χρονικών κενών, τον ποιοτικό έλεγχο και τη δημιουργία πρόσθετων χαρακτηριστικών, το prepared dataset περιείχε 168.847 χρονικές εγγραφές και 234 στήλες.

Στάδιο	Εγγραφές	Στήλες
Canonical (μετά εισαγωγή)	168.829	10
Prepared (μετά προεπεξεργασία)	168.847	234
Supervised (ανά ορίζοντα)	134.604 – 136.238	1.866

Πίνακας 4.2 - Σύνοψη αρχικού και προεπεξεργασμένου dataset.

Ο στόχος πρόβλεψης του πειράματος ήταν η θερμοκρασία αέρα, δηλαδή η μεταβλητή temp_c. Οι υπόλοιπες μετεωρολογικές μεταβλητές χρησιμοποιήθηκαν ως είσοδοι των μοντέλων. Σε αυτές περιλαμβάνονται η σχετική υγρασία, η ατμοσφαιρική πίεση, η μέση ταχύτητα ανέμου, η ριπή ανέμου, η διεύθυνση ανέμου, η βροχόπτωση, η ένταση βροχόπτωσης, η ηλιακή ακτινοβολία και ο δείκτης UV. Επιπλέον, αξιοποιήθηκαν παράγωγα μετεωρολογικά χαρακτηριστικά, όπως dew point, vapour pressure deficit, heat index και wind chill.

Η πρόβλεψη πραγματοποιήθηκε για έξι χρονικούς ορίζοντες. Οι ορίζοντες αυτοί επιλέχθηκαν ώστε να καλύπτουν τόσο πολύ βραχυπρόθεσμη πρόβλεψη όσο και πρόβλεψη εντός της ίδιας ημέρας. Δεδομένου ότι η συχνότητα των δεδομένων είναι 10 λεπτά, οι χρονικοί ορίζοντες εκφράζονται ως αριθμός βημάτων της χρονοσειράς.

Label	Βήματα 10 λεπτών	Πραγματικός ορίζοντας
m10	1	10 λεπτά
h01	6	1 ώρα
h03	18	3 ώρες
h06	36	6 ώρες
h12	72	12 ώρες
h24	144	24 ώρες

Πίνακας 4.3 - Χρονικοί ορίζοντες πρόβλεψης που αξιολογήθηκαν στο πείραμα αναφοράς.

Για κάθε χρονικό ορίζοντα δημιουργήθηκε ξεχωριστό supervised dataset, με αντίστοιχη μελλοντική στήλη στόχο της μορφής temp_c_t+<steps>. Με τον τρόπο αυτό, κάθε μοντέλο εκπαιδεύτηκε ανεξάρτητα για κάθε ορίζοντα πρόβλεψης. Η επιλογή αυτή διασφαλίζει ότι η σύγκριση των μοντέλων πραγματοποιείται με σαφή και ελεγχόμενο τρόπο, καθώς κάθε ορίζοντας αντιμετωπίζεται ως ξεχωριστό πρόβλημα πρόβλεψης.

Η δημιουργία χαρακτηριστικών περιλάμβανε lag features, rolling statistics, κυκλική κωδικοποίηση χρονικών μεταβλητών και κυκλική κωδικοποίηση της διεύθυνσης ανέμου. Τα lag features κάλυπταν προηγούμενες παρατηρήσεις σε χρονικά βήματα 1, 3, 6, 12, 24, 72 και 144, τα οποία αντιστοιχούν σε ιστορική πληροφορία από 10 λεπτά έως 24 ώρες πριν. Τα rolling features υπολογίστηκαν σε παράθυρα 6, 18, 36 και 144 βημάτων, δηλαδή σε παράθυρα 1 ώρας, 3 ωρών, 6 ωρών και 24 ωρών.

Τα δεδομένα διαχωρίστηκαν χρονολογικά σε training, validation και test set με αναλογία 70%, 15% και 15% αντίστοιχα. Δεν εφαρμόστηκε τυχαία ανάμειξη των παρατηρήσεων, καθώς κάτι

τέτοιο θα παραβίαζε τη χρονική δομή του προβλήματος και θα μπορούσε να οδηγήσει σε data leakage. Το training set χρησιμοποιήθηκε για την εκπαίδευση των μοντέλων, το validation set για επιλογές σχετικές με τη διαδικασία εκπαίδευσης και early stopping, ενώ το test set χρησιμοποιήθηκε αποκλειστικά για την τελική αξιολόγηση.

Για κάθε forecast horizon, η κατανομή των supervised δειγμάτων στα training, validation και test sets έγινε με χρονική συνέπεια ως προς τη μελλοντική τιμή-στόχο $t+h$. Επομένως, ένα δείγμα εκπαίδευσης δεν περιλαμβάνει στόχο πρόβλεψης του οποίου το timestamp ανήκει σε μεταγενέστερο validation ή test partition. Η επιλογή αυτή ενισχύει την αποφυγή data leakage μεταξύ των χρονικών συνόλων και καθιστά την τελική αξιολόγηση πιο αντιπροσωπευτική ενός πραγματικού forecasting σεναρίου.

Η κλιμάκωση των αριθμητικών χαρακτηριστικών εφαρμόστηκε μετά τον χρονολογικό διαχωρισμό των δεδομένων. Οι scalers προσαρμόστηκαν αποκλειστικά στο training set και στη συνέχεια εφαρμόστηκαν στα validation και test sets. Με αυτόν τον τρόπο αποφεύχθηκε η χρήση στατιστικών πληροφοριών από μεταγενέστερες χρονικές περιόδους κατά την εκπαίδευση. Η ίδια αρχή εφαρμόστηκε και στην κλιμάκωση του στόχου για τα Deep Learning μοντέλα, όπου οι προβλέψεις επαναμετατράπηκαν στις αρχικές μονάδες μέτρησης πριν τον υπολογισμό των μετρικών.

Στο πείραμα αξιολογήθηκαν τρεις οικογένειες μοντέλων: baseline models, παραδοσιακά μοντέλα Machine Learning και μοντέλα Deep Learning. Στα baseline models περιλαμβάνονται τα persistence, moving_average και climatology. Στα μοντέλα Machine Learning περιλαμβάνονται linear_regression, ridge, random_forest, gradient_boosting και support vector regression (SVR). Στα μοντέλα Deep Learning περιλαμβάνονται LSTM, GRU και TCN. Για κάθε έναν από τους έξι χρονικούς ορίζοντες εκπαιδεύτηκαν και αξιολογήθηκαν όλα τα παραπάνω μοντέλα, οδηγώντας συνολικά σε **66 εκπαιδευμένα και αξιολογημένα μοντέλα** στο πείραμα αναφοράς.

Οικογένεια	Μοντέλο	Σύντομη περιγραφή	Ρόλος στο πείραμα
Baseline	persistence	Προβλέπει ότι η μελλοντική τιμή ισούται με την τελευταία παρατήρηση.	Κύριο σημείο αναφοράς· βάση υπολογισμού του skill score.
Baseline	moving_average	Πρόβλεψη ίση με τον κυλιόμενο μέσο όρο πρόσφατων παρατηρήσεων.	Δευτερεύον baseline για έλεγχο της επίδρασης εξομάλυνσης.
Baseline	climatology	Πρόβλεψη βάσει εποχικού/χρονικού μέσου όρου του training set.	Στατικό μετεωρολογικό σημείο αναφοράς.
Machine Learning	linear_regression	Απλή γραμμική παλινδρόμηση χωρίς regularization.	Αξιολόγηση γραμμικής υπόθεσης· ανέδειξε αστάθεια RMSE σε μεγαλύτερους ορίζοντες.
Machine Learning	ridge	Γραμμική παλινδρόμηση με L2 regularization.	Regularized γραμμική εναλλακτική·

Οικογένεια	Μοντέλο	Σύντομη περιγραφή	Ρόλος στο πείραμα
			επικράτησε στον ορίζοντα 10 min.
Machine Learning	random_forest	Ensemble δέντρων αποφάσεων με bagging.	Επικράτησε στους ορίζοντες 1 h και 3 h.
Machine Learning	gradient_boosting	Ensemble δέντρων αποφάσεων με sequential boosting.	Επικράτησε στους ορίζοντες 6 h, 12 h και 24 h· καλύτερο μοντέλο συνολικά.
Machine Learning	support_vector_regression (SVR)	Παλινδρόμηση με RBF kernel.	Ανταγωνιστικό σε ενδιάμεσους ορίζοντες, αλλά χωρίς επικράτηση και με υψηλό υπολογιστικό κόστος.
Deep Learning	LSTM	Recurrent neural network με long short-term memory cells.	Πιο σταθερό DL μοντέλο· καλύτερη DL επίδοση σε ορίζοντες ≥ 3 h.
Deep Learning	GRU	Recurrent neural network με gated recurrent units.	Ανταγωνιστικό σε ενδιάμεσους ορίζοντες, με αστάθεια στις 24 h.
Deep Learning	TCN	Temporal Convolutional Network με dilated convolutions.	Καλή απόδοση σε μικρούς ορίζοντες (10 min, 1 h), αστάθεια στις 12 h.

Πίνακας 4.4 - Οικογένειες μοντέλων και επιμέρους μοντέλα που αξιολογήθηκαν στο πείραμα αναφοράς, με σύντομη περιγραφή κάθε μοντέλου και τον ρόλο που έπαιξε στη συνολική σύγκριση.

Σύνολο: 11 μοντέλα $\times$ 6 χρονικοί ορίζοντες = **66 εκπαιδευμένα και αξιολογημένα μοντέλα** στο πείραμα αναφοράς.

Η αξιολόγηση πραγματοποιήθηκε με χρήση των μετρικών MAE, RMSE, MAPE και persistence skill score. Οι μετρικές MAE και RMSE εκφράζονται σε βαθμούς Κελσίου και επιτρέπουν την άμεση ερμηνεία του σφάλματος πρόβλεψης. Το MAPE χρησιμοποιήθηκε συμπληρωματικά, με προσοχή λόγω της ιδιαιτερότητας της θερμοκρασίας κοντά σε μηδενικές τιμές. Το persistence skill score χρησιμοποιήθηκε για την εκτίμηση της βελτίωσης κάθε μοντέλου σε σχέση με το persistence baseline, το οποίο αποτελεί ισχυρό σημείο αναφοράς σε προβλήματα βραχυπρόθεσμης πρόγνωσης.

Η πειραματική διάταξη αυτή επιτρέπει την απάντηση των βασικών ερευνητικών ερωτημάτων της εργασίας: αν η πρόβλεψη θερμοκρασίας από έναν μόνο τοπικό σταθμό είναι εφικτή, ποια οικογένεια μοντέλων αποδίδει καλύτερα ανά χρονικό ορίζοντα, και σε ποιο σημείο η διαθέσιμη πληροφορία των τοπικών παρατηρήσεων παύει να προσφέρει ουσιαστική βελτίωση έναντι απλών baseline προσεγγίσεων.

4.4 Συγκεντρωτικά Αποτελέσματα Πρόβλεψης

Μετά την ολοκλήρωση της πειραματικής εκτέλεσης, τα μοντέλα αξιολογήθηκαν στο test set για κάθε χρονικό ορίζοντα πρόβλεψης. Η αξιολόγηση πραγματοποιήθηκε με βάση τις μετρικές MAE, RMSE, MAPE και persistence skill score. Στην παρούσα ενότητα παρουσιάζονται τα συγκεντρωτικά αποτελέσματα, με έμφαση στο καλύτερο μοντέλο ανά ορίζοντα και στη μεταβολή της απόδοσης καθώς αυξάνεται ο χρονικός ορίζοντας πρόβλεψης.

Ο Πίνακας 4.5 παρουσιάζει το μοντέλο με τη χαμηλότερη τιμή MAE για κάθε χρονικό ορίζοντα. Το MAE επιλέγεται ως βασική μετρική παρουσίασης, επειδή εκφράζει το μέσο απόλυτο σφάλμα σε βαθμούς Κελσίου και είναι άμεσα ερμηνεύσιμο στο πλαίσιο πρόβλεψης θερμοκρασίας.

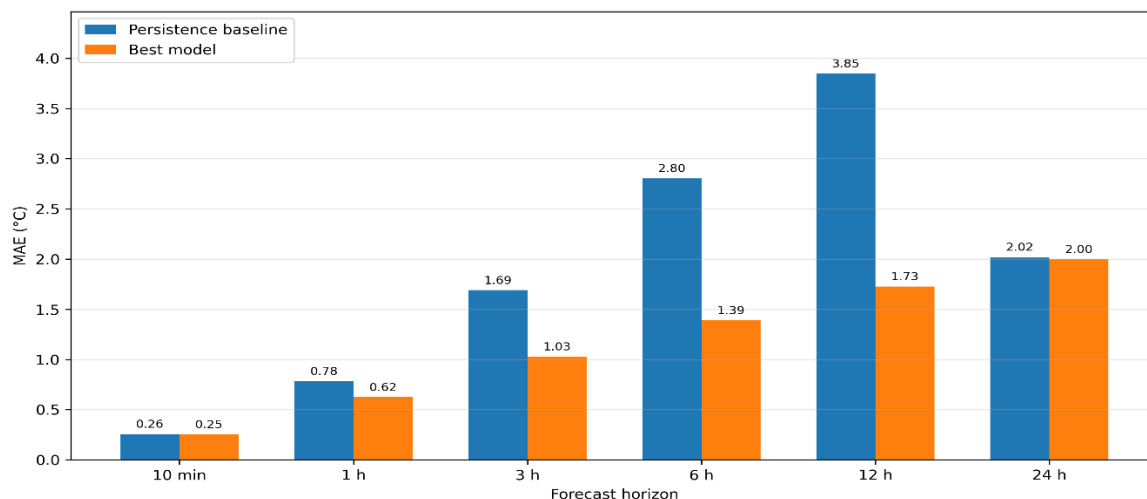
Ορίζοντας	Καλύτερο μοντέλο	MAE (°C)	RMSE (°C)	Skill vs Persistence
10 λεπτά	Ridge	0.254	0.381	0.022
1 ώρα	Random Forest	0.624	0.871	0.374
3 ώρες	Random Forest	1.029	1.377	0.652
6 ώρες	Gradient Boosting	1.389	1.826	0.745
12 ώρες	Gradient Boosting	1.726	2.293	0.752
24 ώρες	Gradient Boosting	1.998	2.634	0.058

Πίνακας 4.5 - Καλύτερο μοντέλο ανά χρονικό ορίζοντα με MAE, RMSE και persistence skill score.

Στον πολύ βραχυπρόθεσμο ορίζοντα των 10 λεπτών, το persistence baseline παραμένει εξαιρετικά ισχυρό, με MAE 0.257 °C και RMSE 0.386 °C. Το ridge εμφανίζει αμελητέα οριακή υπεροχή με MAE 0.254 °C και RMSE 0.381 °C (skill = +0.022). Πρόκειται για στατιστικά σχεδόν ισοδύναμη επίδοση, και η πρακτική ερμηνεία παραμένει ότι σε τόσο σύντομο χρονικό ορίζοντα η τελευταία διαθέσιμη παρατήρηση αποτελεί ήδη πολύ ισχυρή πρόβλεψη και τα πιο σύνθετα μοντέλα έχουν περιορισμένο περιθώριο ουσιαστικής βελτίωσης.

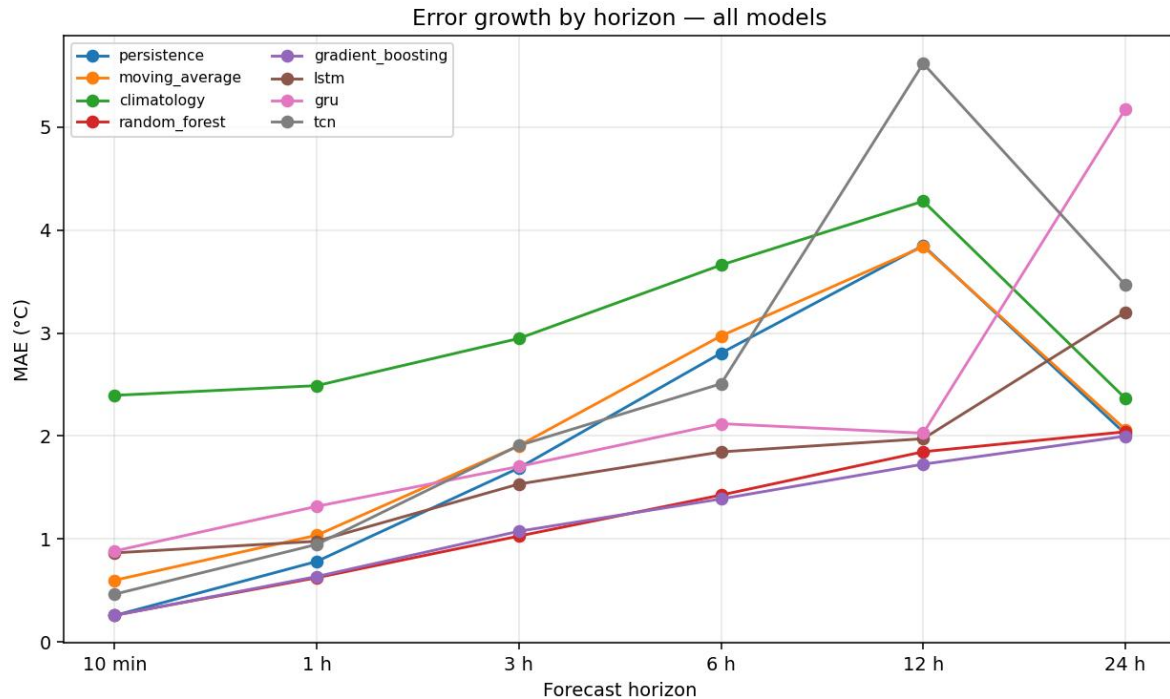
Αντίθετα, στους ορίζοντες της 1 και των 3 ωρών, το καλύτερο μοντέλο είναι το Random Forest. Το μοντέλο αυτό επιτυγχάνει MAE 0.624 °C στη 1 ώρα και 1.029 °C στις 3 ώρες, βελτιώνοντας την απόδοση σε σχέση με το persistence baseline. Η βελτίωση αυτή δείχνει ότι, σε χρονικούς ορίζοντες μεγαλύτερους των λίγων λεπτών, τα engineered χαρακτηριστικά, τα lag features και οι rolling στατιστικές αρχίζουν να προσφέρουν πρόσθετη πληροφορία πέρα από την τελευταία παρατήρηση.

Στους ορίζοντες των 6, 12 και 24 ωρών, το Gradient Boosting εμφανίζεται ως το καλύτερο μοντέλο. Η καλύτερη σχετική απόδοση παρατηρείται στους ορίζοντες των 6 και 12 ωρών, όπου το persistence skill score φτάνει περίπου το 0.75. Αυτό σημαίνει ότι το Gradient Boosting μειώνει σημαντικά το τετραγωνικό σφάλμα σε σχέση με το persistence baseline. Το εύρημα αυτό αποτελεί ένα από τα βασικά αποτελέσματα της εργασίας, καθώς δείχνει ότι ένα μοντέλο Machine Learning μπορεί να αξιοποιήσει την πληροφορία των τοπικών παρατηρήσεων για να παράγει χρήσιμες βραχυπρόθεσμες προβλέψεις θερμοκρασίας χωρίς εξωτερικά δεδομένα πρόγνωσης.



Σχήμα 4.2 - Σύγκριση MAE του καλύτερου μοντέλου και του persistence baseline ανά χρονικό ορίζοντα.

Η μεταβολή του σφάλματος ανά ορίζοντα δείχνει ότι η δυσκολία του προβλήματος αυξάνεται όσο απομακρύνεται ο χρονικός ορίζοντας από την τρέχουσα παρατήρηση. Ωστόσο, η αύξηση αυτή δεν είναι απολύτως γραμμική. Ιδιαίτερο ενδιαφέρον παρουσιάζει η σύγκριση μεταξύ των 12 και 24 ωρών. Παρότι θα αναμενόταν το σφάλμα στις 24 ώρες να είναι μεγαλύτερο από το σφάλμα στις 12 ώρες, παρατηρείται ότι τόσο το persistence όσο και το Gradient Boosting εμφανίζουν χαμηλότερο MAE στις 24 ώρες από ό,τι στις 12 ώρες. Το φαινόμενο αυτό μπορεί να ερμηνευθεί από την έντονη ημερήσια περιοδικότητα της θερμοκρασίας: η θερμοκρασία στο $t+24h$ αντιστοιχεί στην ίδια περίπου ώρα της επόμενης ημέρας, ενώ η θερμοκρασία στο $t+12h$ βρίσκεται συχνά σε αντίθετη φάση του ημερήσιου κύκλου.



Σχήμα 4.3 - Μεταβολή σφάλματος ανά ορίζοντα.

Παρά τη σχετική μείωση του MAE στις 24 ώρες, το persistence skill score δείχνει ότι η ουσιαστική βελτίωση των μοντέλων έναντι του persistence σχεδόν μηδενίζεται στον 24ωρο ορίζοντα. Το Gradient Boosting παραμένει οριακά καλύτερο, με skill score 0.058, όμως η διαφορά είναι πολύ μικρότερη σε σχέση με τους ορίζοντες των 6 και 12 ωρών. Αυτό δείχνει ότι το διαθέσιμο σήμα των τοπικών παρατηρήσεων εξαντλείται σταδιακά και ότι, χωρίς εξωτερικές πηγές πληροφορίας όπως NWP outputs, η πρόβλεψη σε μεγαλύτερους ορίζοντες γίνεται σημαντικά δυσκολότερη.

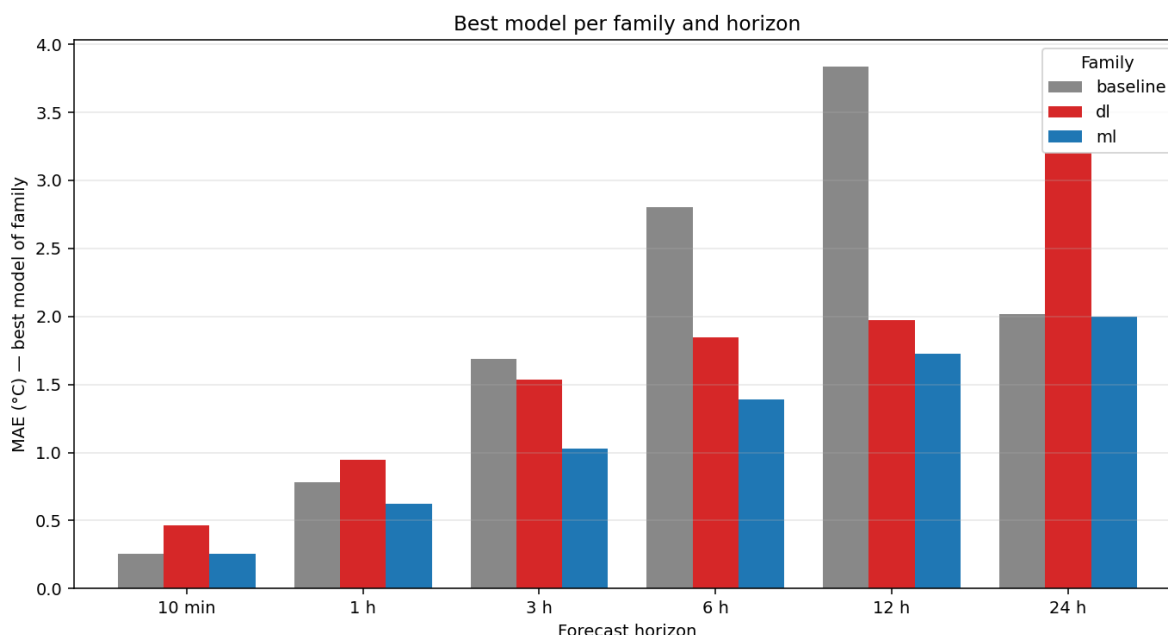
Συνολικά, τα συγκεντρωτικά αποτελέσματα οδηγούν σε τρία βασικά συμπεράσματα. Πρώτον, η απλή persistence προσέγγιση είναι εξαιρετικά ισχυρή σε πολύ μικρούς χρονικούς ορίζοντες και αποτελεί απαραίτητο σημείο αναφοράς. Δεύτερον, τα μοντέλα Machine Learning, και ειδικότερα τα Random Forest και Gradient Boosting, προσφέρουν ουσιαστική βελτίωση στους ενδιάμεσους ορίζοντες πρόβλεψης. Τρίτον, η χρησιμότητα των καθαρά τοπικών παρατηρήσεων μειώνεται αισθητά στον ορίζοντα των 24 ωρών, όπου τα σύνθετα μοντέλα συγκλίνουν σε απόδοση με το persistence baseline.

Τα ευρήματα αυτά υποστηρίζουν την κεντρική ερευνητική υπόθεση της εργασίας: η βραχυπρόθεσμη πρόβλεψη θερμοκρασίας από έναν μόνο τοπικό μετεωρολογικό σταθμό είναι εφικτή και μπορεί να αποδώσει χρήσιμα αποτελέσματα, ιδίως στους ορίζοντες από 1 έως 12 ώρες. Ταυτόχρονα, αναδεικνύουν και το βασικό όριο της observation-only προσέγγισης, καθώς η απουσία εξωτερικών προγνωστικών εισόδων περιορίζει τη δυνατότητα ουσιαστικής βελτίωσης σε μεγαλύτερους χρονικούς ορίζοντες.

4.5 Αναλυτική Αξιολόγηση ανά Οικογένεια Μοντέλων

Η συγκεντρωτική εικόνα των αποτελεσμάτων δείχνει ποιο μοντέλο επικράτησε ανά χρονικό ορίζοντα, όμως για την πληρέστερη ερμηνεία των ευρημάτων είναι χρήσιμη η ανάλυση ανά

οικογένεια μοντέλων. Η σύγκριση αυτή επιτρέπει να φανεί όχι μόνο ποιο μοντέλο είχε τη χαμηλότερη τιμή σφάλματος, αλλά και ποια κατηγορία προσεγγίσεων είναι καταλληλότερη για το συγκεκριμένο πρόβλημα πρόβλεψης θερμοκρασίας από έναν μόνο τοπικό μετεωρολογικό σταθμό.



Σχήμα 4.4 - Σύγκριση καλύτερου μοντέλου ανά οικογένεια.

4.5.1 Baseline Models

Τα baseline μοντέλα χρησιμοποιήθηκαν ως σημεία αναφοράς για την αποτίμηση της προστιθέμενης αξίας των πιο σύνθετων μοντέλων. Στην παρούσα εκτέλεση αξιολογήθηκαν τρία baseline μοντέλα: persistence, moving average και climatology.

Το persistence model εμφάνισε την καλύτερη απόδοση στον ορίζοντα των 10 λεπτών, με MAE 0.257 °C και RMSE 0.386 °C. Το αποτέλεσμα αυτό επιβεβαιώνει ότι σε πολύ μικρούς χρονικούς ορίζοντες η τελευταία διαθέσιμη τιμή της θερμοκρασίας αποτελεί εξαιρετικά ισχυρή πρόβλεψη. Η θερμοκρασία, ως σχετικά αργά μεταβαλλόμενη μετεωρολογική μεταβλητή, δεν αλλάζει συνήθως απότομα μέσα σε διάστημα λίγων λεπτών, επομένως ένα απλό persistence μοντέλο είναι δύσκολο να ξεπεραστεί.

Καθώς ο ορίζοντας πρόβλεψης αυξάνεται, το σφάλμα του persistence αυξάνεται σημαντικά μέχρι τις 12 ώρες. Συγκεκριμένα, το MAE αυξάνεται από 0.257 °C στα 10 λεπτά σε 3.849 °C στις 12 ώρες. Η αύξηση αυτή δείχνει ότι η απλή υπόθεση σταθερότητας της τελευταίας παρατήρησης χάνει σταδιακά την ισχύ της όσο το μοντέλο καλείται να προβλέψει πιο μακρινές χρονικές στιγμές.

Το moving average baseline δεν εμφάνισε ουσιαστική βελτίωση έναντι του persistence στους περισσότερους ορίζοντες. Αν και η χρήση κυλιόμενου μέσου μπορεί να εξομαλύνει τον θόρυβο, στην παρούσα περίπτωση φαίνεται ότι η εξομάλυνση αφαιρεί χρήσιμη βραχυπρόθεσμη πληροφορία από τη χρονοσειρά. Το μοντέλο αυτό ήταν χειρότερο από το persistence στους

ορίζοντες έως και τις 6 ώρες, ενώ στις 12 ώρες είχε οριακά καλύτερο MAE από το persistence, χωρίς όμως να αποτελεί συνολικά ισχυρή προσέγγιση.

Το climatology baseline είχε τη χαμηλότερη συνολική απόδοση μεταξύ των baseline μοντέλων. Η συμπεριφορά αυτή είναι αναμενόμενη, καθώς η κλιματολογική προσέγγιση βασίζεται σε μέσες τιμές και δεν αξιοποιεί την τρέχουσα δυναμική της χρονοσειράς. Παρόλα αυτά, η παρουσία του είναι χρήσιμη, επειδή λειτουργεί ως ένα απλό μετεωρολογικό σημείο αναφοράς και δείχνει ότι τα μοντέλα που αξιοποιούν πρόσφατες παρατηρήσεις υπερτερούν σημαντικά έναντι μιας στατικής εποχικής ή χρονικής μέσης τιμής.

Μοντέλο	Ορίζοντας	MAE (°C)	RMSE (°C)	Skill
persistence	10 min	0.257	0.386	0.000
persistence	1 h	0.782	1.100	0.000
persistence	3 h	1.689	2.333	0.000
persistence	6 h	2.805	3.616	0.000
persistence	12 h	3.849	4.606	0.000
persistence	24 h	2.021	2.714	0.000
moving_average	10 min	0.599	0.843	-3.778
moving_average	1 h	1.036	1.454	-0.747
moving_average	3 h	1.906	2.598	-0.240
moving_average	6 h	2.972	3.782	-0.094
moving_average	12 h	3.839	4.593	+0.006
moving_average	24 h	2.064	2.764	-0.037
climatology	10 min	2.395	3.125	-64.679
climatology	1 h	2.489	3.209	-7.512
climatology	3 h	2.947	3.680	-1.489
climatology	6 h	3.660	4.456	-0.519
climatology	12 h	4.280	5.067	-0.210
climatology	24 h	2.365	3.085	-0.292

Πίνακας 4.6 - Απόδοση baseline μοντέλων ανά ορίζοντα με MAE, RMSE και skill score.

Συνολικά, τα baseline αποτελέσματα δείχνουν ότι το persistence αποτελεί το βασικό και ισχυρότερο σημείο αναφοράς της εργασίας. Η ύπαρξη ισχυρού baseline είναι σημαντική, διότι αποτρέπει την υπερεκτίμηση της απόδοσης πιο σύνθετων μοντέλων. Ένα μοντέλο Machine Learning ή Deep Learning έχει πρακτική αξία μόνο εφόσον καταφέρνει να βελτιώσει ουσιαστικά την απόδοση σε σχέση με το persistence.

4.5.2 Μοντέλα Machine Learning

Τα μοντέλα Machine Learning εμφάνισαν την καλύτερη συνολική απόδοση στο πείραμα. Ιδιαίτερα τα ensemble tree-based μοντέλα, δηλαδή το Random Forest και το Gradient Boosting, αξιοποίησαν αποτελεσματικά τα engineered χαρακτηριστικά που δημιουργήθηκαν από το pipeline.

Το Random Forest ήταν το καλύτερο μοντέλο στους ορίζοντες της 1 και των 3 ωρών. Στη 1 ώρα πέτυχε MAE 0.624 °C και RMSE 0.871 °C, ενώ στις 3 ώρες πέτυχε MAE 1.029 °C και RMSE 1.377 °C. Η απόδοση αυτή δείχνει ότι το Random Forest μπορεί να εκμεταλλευτεί μη γραμμικές σχέσεις μεταξύ των lag features, rolling statistics και μετεωρολογικών μεταβλητών, προσφέροντας σαφή βελτίωση σε σχέση με το persistence.

Το Gradient Boosting ήταν το καλύτερο μοντέλο στους ορίζοντες των 6, 12 και 24 ωρών. Το αποτέλεσμα αυτό δείχνει ότι το Gradient Boosting είναι ιδιαίτερα αποτελεσματικό στην αξιοποίηση των σύνθετων χαρακτηριστικών του dataset, ειδικά σε ορίζοντες όπου το persistence δεν επαρκεί πλέον, αλλά η τοπική ιστορική πληροφορία εξακολουθεί να περιέχει χρήσιμο προβλεπτικό σήμα.

Η linear_regression αξιολογήθηκε και παρουσίασε μικτές επιδόσεις. Αν και οι τιμές MAE ήταν σε ορισμένες περιπτώσεις σχετικά λογικές, οι τιμές RMSE αυξήθηκαν υπερβολικά στους μεγαλύτερους χρονικούς ορίζοντες. Συγκεκριμένα, στους ορίζοντες των 3, 6, 12 και 24 ωρών καταγράφηκαν RMSE 7.692 °C, 13.400 °C, 12.397 °C και 13.863 °C αντίστοιχα. Η μεγάλη απόκλιση μεταξύ MAE και RMSE δείχνει την παρουσία ακραίων προβλέψεων, πιθανότατα λόγω πολυσυγγραμμικότητας στο μεγάλο σύνολο χαρακτηριστικών που περιλαμβάνει εκατοντάδες lag και rolling μεταβλητές. Το εύρημα αυτό κατέστησε αναγκαία την αξιολόγηση και ενός regularized γραμμικού μοντέλου, του ridge, ως πιο σταθερής γραμμικής εναλλακτικής.

Το ridge βελτίωσε σαφώς την επίδοση στον ορίζοντα των 10 λεπτών (MAE 0.254 °C, RMSE 0.381 °C, skill +0.022) και είναι η μοναδική περίπτωση στο πείραμα στην οποία γραμμικό μοντέλο επικρατεί. Ωστόσο, δεν επιλύει το πρόβλημα ακραίων προβλέψεων στους μεγαλύτερους ορίζοντες: το RMSE παραμένει υψηλό (5.420 °C στις 3 ώρες, 10.738 °C στις 6 ώρες, 11.026 °C στις 12 ώρες και 11.562 °C στις 24 ώρες, με αρνητικό skill score). Επομένως, το Ridge τεκμηριώνεται ως κατάλληλο μόνο για τον πολύ βραχυπρόθεσμο ορίζοντα και ως αναλυτική σύγκριση έναντι του linear_regression, όχι ως πρακτική επιλογή για τους μεγαλύτερους ορίζοντες.

Το Support Vector Regression (SVR) αξιολογήθηκε επίσης. Είχε ανταγωνιστική επίδοση σε ορισμένους ενδιάμεσους ορίζοντες (skill +0.454 στις 3 ώρες, +0.692 στις 6 ώρες, +0.711 στις 12 ώρες), αλλά δεν επικράτησε σε κανέναν χρονικό ορίζοντα. Επιπλέον, εμφάνισε αρνητικό persistence skill score στους ορίζοντες 10 λεπτών, 1 ώρας και 24 ωρών, ενώ το υπολογιστικό του κόστος ήταν ιδιαίτερα υψηλό για το πλήρες dataset. Στο παρόν πειραματικό πλαίσιο, το SVR δεν προκύπτει ως πρακτικά κατάλληλη επιλογή, καθώς δεν επικράτησε σε κανέναν ορίζοντα και είχε δυσανάλογο υπολογιστικό κόστος.

Μοντέλο	Ορίζοντας	MAE (°C)	RMSE (°C)	Skill
linear_regression	10 min	0.401	0.715	-2.437
linear_regression	1 h	0.755	1.977	-2.230
linear_regression	3 h	1.270	7.692	-9.871
linear_regression	6 h	1.709	13.400	-12.733
linear_regression	12 h	1.921	12.397	-6.244
linear_regression	24 h	2.366	13.863	-25.091
ridge	10 min	0.254	0.381	+0.022
ridge	1 h	0.663	1.313	-0.424
ridge	3 h	1.169	5.420	-4.398
ridge	6 h	1.590	10.738	-7.820

Μοντέλο	Ορίζοντας	MAE (°C)	RMSE (°C)	Skill
ridge	12 h	1.873	11.026	-4.730
ridge	24 h	2.314	11.562	-17.147
random_forest	10 min	0.260	0.397	-0.062
random_forest	1 h	0.624	0.871	+0.374
random_forest	3 h	1.029	1.377	+0.652
random_forest	6 h	1.427	1.942	+0.712
random_forest	12 h	1.847	2.495	+0.707
random_forest	24 h	2.041	2.701	+0.009
gradient_boosting	10 min	0.258	0.395	-0.048
gradient_boosting	1 h	0.636	0.874	+0.368
gradient_boosting	3 h	1.075	1.413	+0.633
gradient_boosting	6 h	1.389	1.826	+0.745
gradient_boosting	12 h	1.726	2.293	+0.752
gradient_boosting	24 h	1.998	2.634	+0.058
svr	10 min	0.671	1.141	-7.760
svr	1 h	0.977	1.405	-0.632
svr	3 h	1.263	1.724	+0.454
svr	6 h	1.472	2.007	+0.692
svr	12 h	1.858	2.474	+0.711
svr	24 h	2.298	3.090	-0.296

Πίνακας 4.7 - Απόδοση μοντέλων *Machine Learning* ανά χρονικό ορίζοντα.

Συνολικά, τα αποτελέσματα των μοντέλων *Machine Learning* δείχνουν ότι τα ensemble tree-based μοντέλα αποτελούν την πιο κατάλληλη οικογένεια για το συγκεκριμένο πρόβλημα. Το Random Forest αποδίδει καλύτερα στους βραχύτερους από τους μη τετριμμένους ορίζοντες, ενώ το Gradient Boosting εμφανίζεται ως το πιο σταθερό και αποτελεσματικό μοντέλο στους μεσαίους και μεγαλύτερους ορίζοντες.

4.5.3 Μοντέλα Deep Learning

Τα μοντέλα Deep Learning που αξιολογήθηκαν ήταν τα LSTM, GRU και TCN. Τα μοντέλα αυτά επιλέχθηκαν επειδή είναι κατάλληλα για προβλήματα χρονοσειρών και μπορούν θεωρητικά να μάθουν χρονικές εξαρτήσεις απευθείας από ακολουθιακά δεδομένα. Στο συγκεκριμένο πείραμα, όμως, δεν ξεπέρασαν τα καλύτερα tree-based *Machine Learning* μοντέλα.

Το LSTM ήταν το πιο σταθερό από τα τρία Deep Learning μοντέλα. Εμφάνισε την καλύτερη επίδοση μεταξύ των DL μοντέλων στους ορίζοντες των 3, 6, 12 και 24 ωρών. Για παράδειγμα, στις 6 ώρες πέτυχε MAE 1.846 °C, ενώ στις 12 ώρες πέτυχε MAE 1.974 °C. Οι τιμές αυτές είναι καλύτερες από τα απλά baseline μοντέλα σε αρκετούς ορίζοντες, αλλά παραμένουν χειρότερες από τις αντίστοιχες επιδόσεις του Gradient Boosting.

Το GRU παρουσίασε μεγαλύτερη αστάθεια. Αν και σε ορισμένους ορίζοντες είχε αποδεκτή συμπεριφορά, στον ορίζοντα των 24 ωρών εμφάνισε πολύ υψηλό σφάλμα, με MAE 5.175 °C και

RMSE 6.060 °C. Το αποτέλεσμα αυτό δείχνει ότι η συγκεκριμένη εκπαίδευση δεν γενίκευσε ικανοποιητικά στον μεγαλύτερο χρονικό ορίζοντα.

Το TCN είχε σχετικά καλή απόδοση στους μικρότερους ορίζοντες μεταξύ των DL μοντέλων, ειδικά στα 10 λεπτά και στη 1 ώρα. Ωστόσο, παρουσίασε σημαντική αστάθεια σε μεγαλύτερους ορίζοντες. Στις 12 ώρες, για παράδειγμα, το TCN εμφάνισε MAE 5.617 °C και RMSE 6.575 °C, τιμές πολύ χειρότερες από τα αντίστοιχα Machine Learning μοντέλα και από το LSTM. Η συμπεριφορά αυτή δείχνει ότι, στην παρούσα διαμόρφωση, η αρχιτεκτονική TCN δεν ήταν επαρκώς σταθερή για όλους τους χρονικούς ορίζοντες.

Μοντέλο	Ορίζοντας	MAE (°C)	RMSE (°C)	Skill	Epochs trained	Best val. loss
LSTM	10 min	0.865	1.182	-8.405	10	0.0288
LSTM	1 h	0.978	1.304	-0.405	20	0.0448
LSTM	3 h	1.534	1.981	+0.279	20	0.0791
LSTM	6 h	1.846	2.385	+0.565	20	0.1115
LSTM	12 h	1.974	2.561	+0.691	20	0.1470
LSTM	24 h	3.204	4.068	-1.246	9	0.1605
GRU	10 min	0.883	1.280	-10.012	20	0.0340
GRU	1 h	1.317	1.763	-1.569	18	0.0429
GRU	3 h	1.704	2.175	+0.131	20	0.0548
GRU	6 h	2.120	2.709	+0.439	18	0.1252
GRU	12 h	2.027	2.609	+0.679	20	0.1605
GRU	24 h	5.175	6.060	-3.986	6	0.1937
TCN	10 min	0.464	0.596	-1.391	14	0.0046
TCN	1 h	0.947	1.221	-0.231	15	0.0255
TCN	3 h	1.908	2.434	-0.089	8	0.0908
TCN	6 h	2.507	3.298	+0.168	13	0.1140
TCN	12 h	5.617	6.575	-1.037	7	0.1366
TCN	24 h	3.466	4.323	-1.537	14	0.1441

Πίνακας 4.8 - Απόδοση μοντέλων Deep Learning ανά ορίζοντα με MAE, RMSE, epochs trained και best validation loss.

Το βασικό συμπέρασμα από την αξιολόγηση των Deep Learning μοντέλων είναι ότι, παρότι μπορούν να μοντελοποιήσουν χρονικές εξαρτήσεις, δεν ξεπέρασαν τα tree-based Machine Learning μοντέλα στο συγκεκριμένο dataset. Αυτό μπορεί να οφείλεται στο γεγονός ότι τα engineered χαρακτηριστικά, όπως lags, rolling statistics, κυκλικές χρονικές μεταβλητές, QC flags και παράγωγα μετεωρολογικά χαρακτηριστικά, παρέχουν ήδη πολύ ισχυρή αναπαράσταση για τα μοντέλα δέντρων. Επιπλέον, το διαθέσιμο dataset, αν και αρκετά μεγάλο για κλασικά ML μοντέλα, παραμένει περιορισμένο σε σχέση με τις απαιτήσεις πιο σύνθετων Deep Learning αρχιτεκτονικών.

4.5.4 Συγκριτική Ερμηνεία

Η συνολική σύγκριση των αποτελεσμάτων δείχνει ότι τα tree-based Machine Learning μοντέλα αποτέλεσαν την πιο αποτελεσματική κατηγορία για το συγκεκριμένο πρόβλημα. Το Random Forest πέτυχε την καλύτερη επίδοση στους ορίζοντες της 1 και των 3 ωρών, ενώ το Gradient

Boosting επικράτησε στους ορίζοντες των 6, 12 και 24 ωρών. Η καλύτερη σχετική επίδοση παρατηρήθηκε στους ορίζοντες των 6 και 12 ωρών, όπου το Gradient Boosting πέτυχε persistence skill score περίπου 0.75.

Το αποτέλεσμα αυτό είναι ιδιαίτερα σημαντικό, επειδή δείχνει ότι οι τοπικές ιστορικές παρατηρήσεις περιέχουν αξιοποιήσιμη πληροφορία πέρα από την τελευταία διαθέσιμη τιμή. Τα lag features, τα rolling statistics, οι κυκλικές μεταβλητές χρόνου, η κυκλική αναπαράσταση της διεύθυνσης ανέμου, τα QC flags και τα παράγωγα μετεωρολογικά χαρακτηριστικά επιτρέπουν στα tree-based μοντέλα να εκμεταλλευτούν μη γραμμικές σχέσεις στη χρονοσειρά.

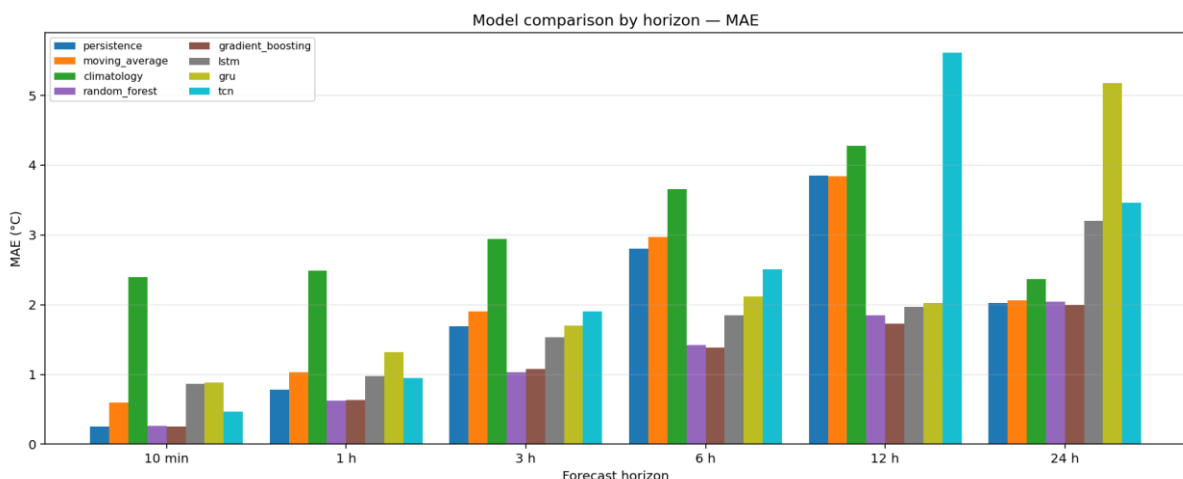
Τα baseline μοντέλα παραμένουν απαραίτητα για την ερμηνεία των αποτελεσμάτων. Το persistence είναι ιδιαίτερα ισχυρό στα 10 λεπτά, όπου η θερμοκρασία μεταβάλλεται συνήθως αργά. Στον ορίζοντα των 24 ωρών, το persistence επωφελείται από την ημερήσια περιοδικότητα της θερμοκρασίας, επειδή η πρόβλεψη αφορά περίπου την ίδια ώρα της επόμενης ημέρας. Για τον λόγο αυτό, η επίδοση στις 24 ώρες πρέπει να ερμηνεύεται όχι μόνο με βάση το MAE, αλλά και με βάση τη βελτίωση έναντι του persistence.

Τα Deep Learning μοντέλα παρείχαν χρήσιμο σημείο σύγκρισης, αλλά δεν αποτέλεσαν την καλύτερη πρακτική επιλογή. Το LSTM ήταν το πιο σταθερό DL μοντέλο, όμως παρέμεινε πίσω από το Gradient Boosting στους σημαντικότερους ορίζοντες. Το GRU και το TCN παρουσίασαν αστάθειες σε ορισμένες περιπτώσεις, γεγονός που δείχνει ότι η αυξημένη πολυπλοκότητα ενός μοντέλου δεν συνεπάγεται απαραίτητα καλύτερη απόδοση στο συγκεκριμένο dataset.

Συνολικά, η συγκριτική αξιολόγηση δείχνει ότι η καταλληλότερη πρακτική επιλογή για το συγκεκριμένο πείραμα είναι τα tree-based Machine Learning μοντέλα, με κύριο εκπρόσωπο το Gradient Boosting για τους ορίζοντες των 6 και 12 ωρών.

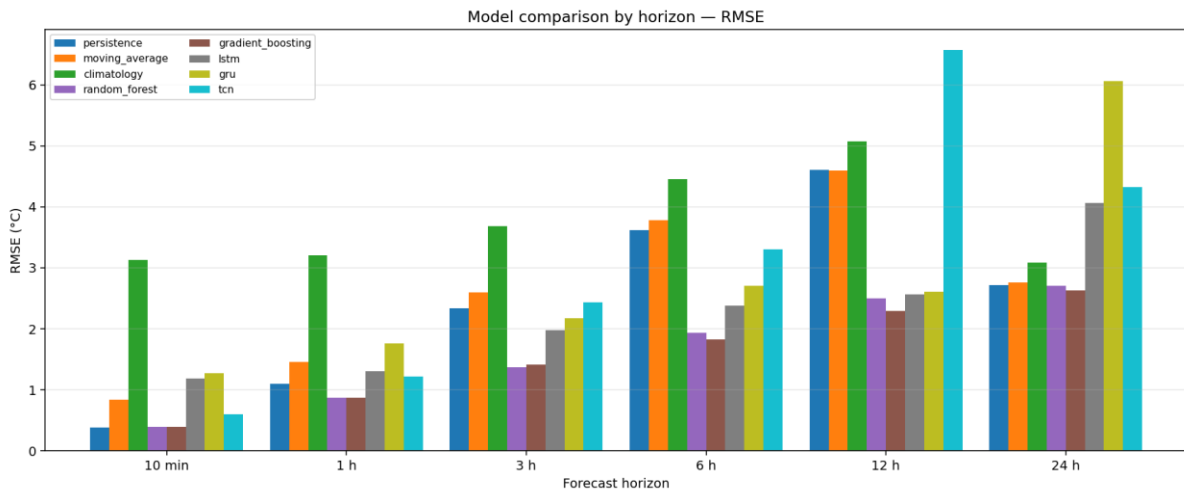
4.6 Οπτικοποίηση και Διαγνωστική Ανάλυση Αποτελεσμάτων

Για την καλύτερη ερμηνεία των αποτελεσμάτων δημιουργήθηκαν διαγράμματα σύγκρισης, πραγματικών έναντι προβλεπόμενων τιμών και διαγνωστικής ανάλυσης υπολοίπων. Τα διαγράμματα αυτά δεν λειτουργούν μόνο ως οπτική παρουσίαση των αποτελεσμάτων, αλλά βοηθούν στον εντοπισμό μοτίβων σφάλματος, αστάθειας μοντέλων και διαφορών μεταξύ χρονικών οριζόντων.



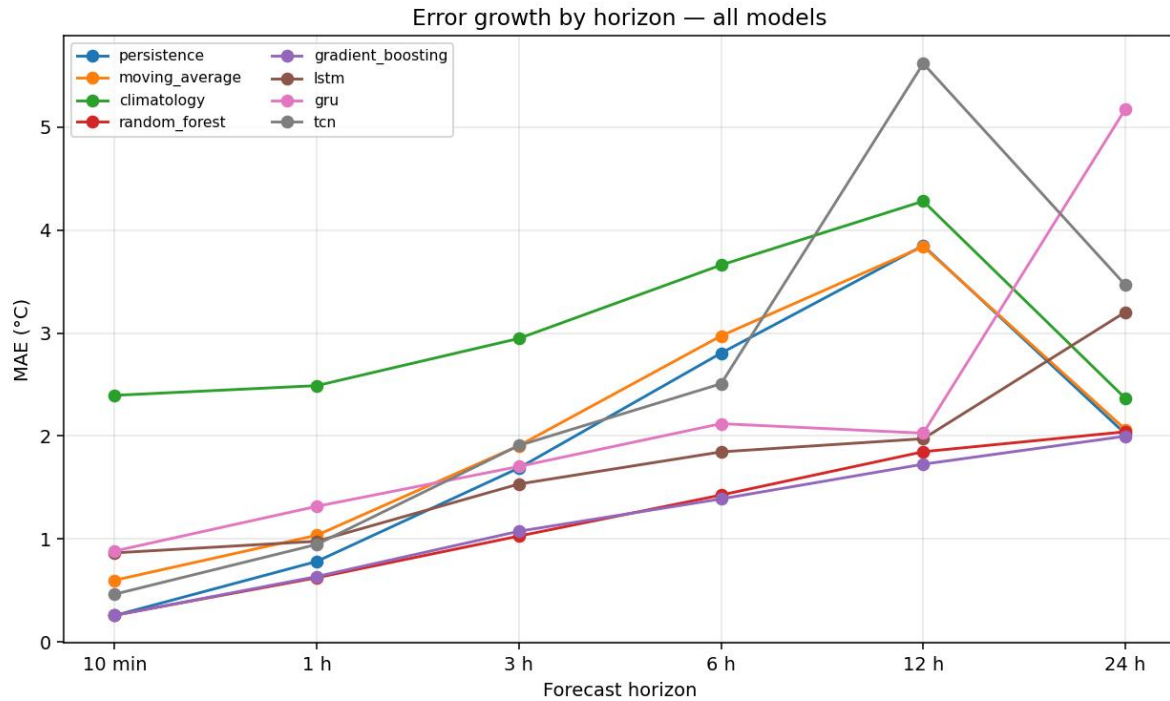
Σχήμα 4.5 - Διαγνωστική σύγκριση MAE όλων των μοντέλων ανά χρονικό ορίζοντα.

Τα συγκριτικά διαγράμματα MAE δείχνουν τη μεταβολή του μέσου απόλυτου σφάλματος ανά μοντέλο και ανά χρονικό ορίζοντα. Από αυτά προκύπτει ότι τα tree-based Machine Learning μοντέλα έχουν σταθερά καλύτερη συμπεριφορά στους ενδιάμεσους ορίζοντες, ενώ το persistence είναι ιδιαίτερα ανταγωνιστικό στα 10 λεπτά και στις 24 ώρες.



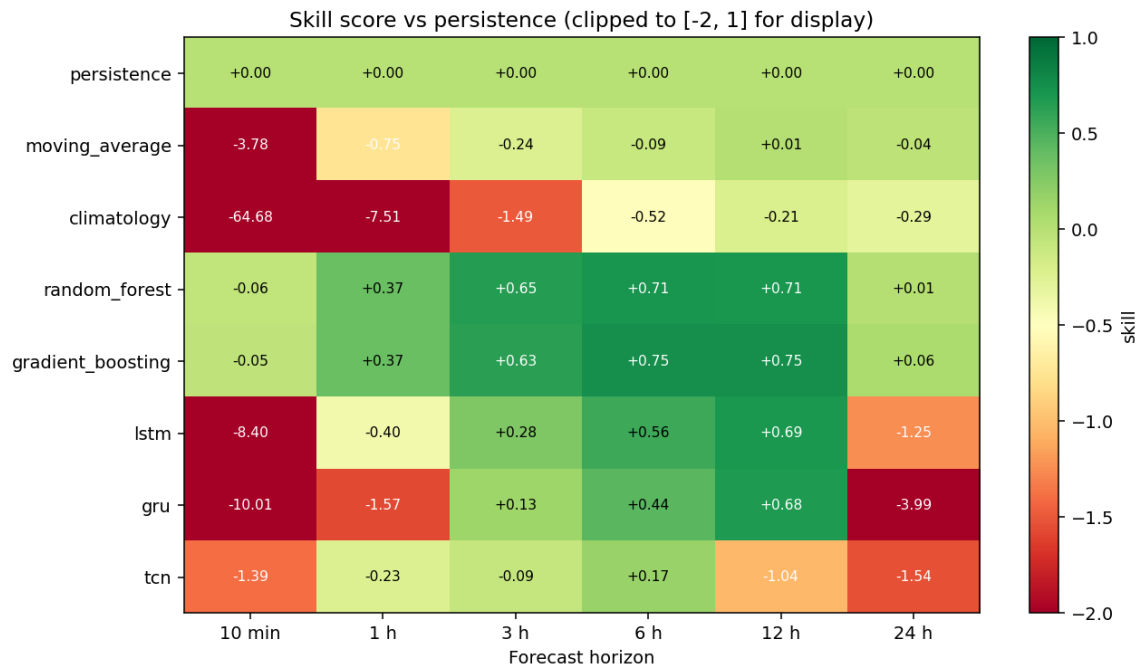
Σχήμα 4.6 - Σύγκριση RMSE όλων των μοντέλων ανά χρονικό ορίζοντα.

Η σύγκριση RMSE είναι ιδιαίτερα σημαντική επειδή η συγκεκριμένη μετρική δίνει μεγαλύτερη βαρύτητα στα μεγάλα σφάλματα. Στην περίπτωση τόσο της linear_regression όσο και του ridge, η μεγάλη διαφορά μεταξύ MAE και RMSE σε μεγαλύτερους ορίζοντες δείχνει ότι τα γραμμικά μοντέλα μπορούν να παράγουν ακραίες προβλέψεις. Το εύρημα αυτό υποστηρίζει την ερμηνεία ότι τα γραμμικά μοντέλα λειτουργούν κυρίως ως baselines και όχι ως τελικές πρακτικές επιλογές.



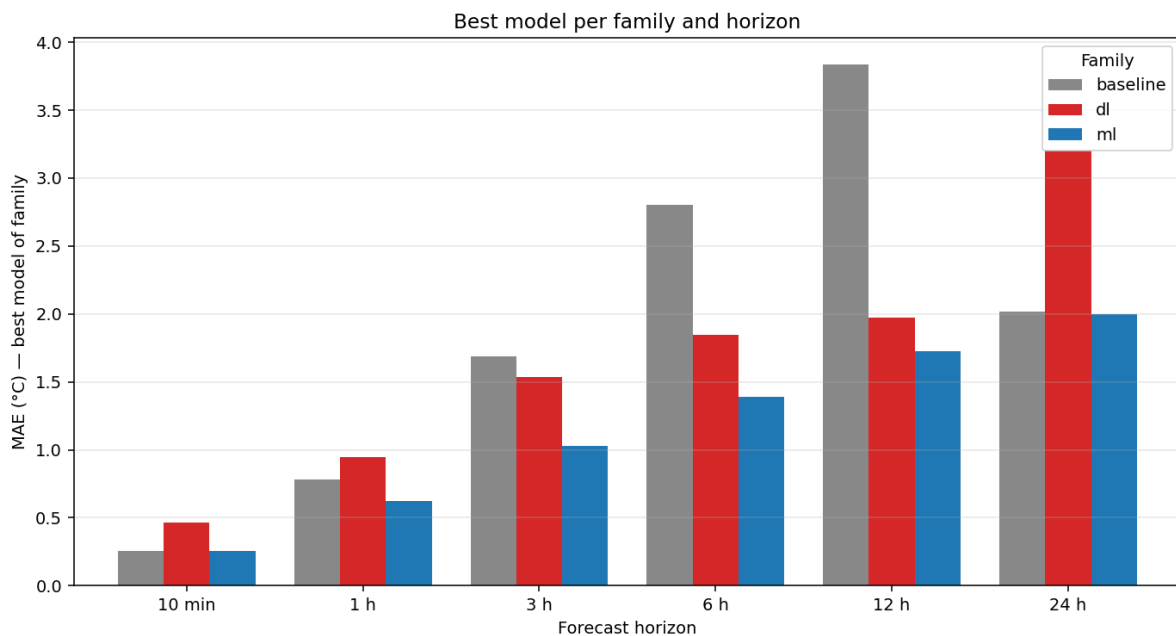
Σχήμα 4.7 - Μεταβολή MAE ανά χρονικό ορίζοντα για κάθε μοντέλο.

Το διάγραμμα μεταβολής του σφάλματος ανά χρονικό ορίζοντα δείχνει ότι η δυσκολία του προβλήματος αυξάνεται γενικά όσο απομακρύνεται ο χρονικός ορίζοντας από την τρέχουσα παρατήρηση. Ωστόσο, η σχέση αυτή δεν είναι πλήρως γραμμική. Το σφάλμα στις 24 ώρες δεν είναι απαραίτητα μεγαλύτερο από το σφάλμα στις 12 ώρες, επειδή η θερμοκρασία παρουσιάζει έντονο ημερήσιο κύκλο. Η τιμή στο t+24h αντιστοιχεί περίπου στην ίδια ώρα της επόμενης ημέρας, ενώ η τιμή στο t+12h βρίσκεται συχνά σε αντίθετη φάση του ημερήσιου κύκλου.



Σχήμα 4.8 - Heatmap persistence skill score ανά μοντέλο και χρονικό ορίζοντα.

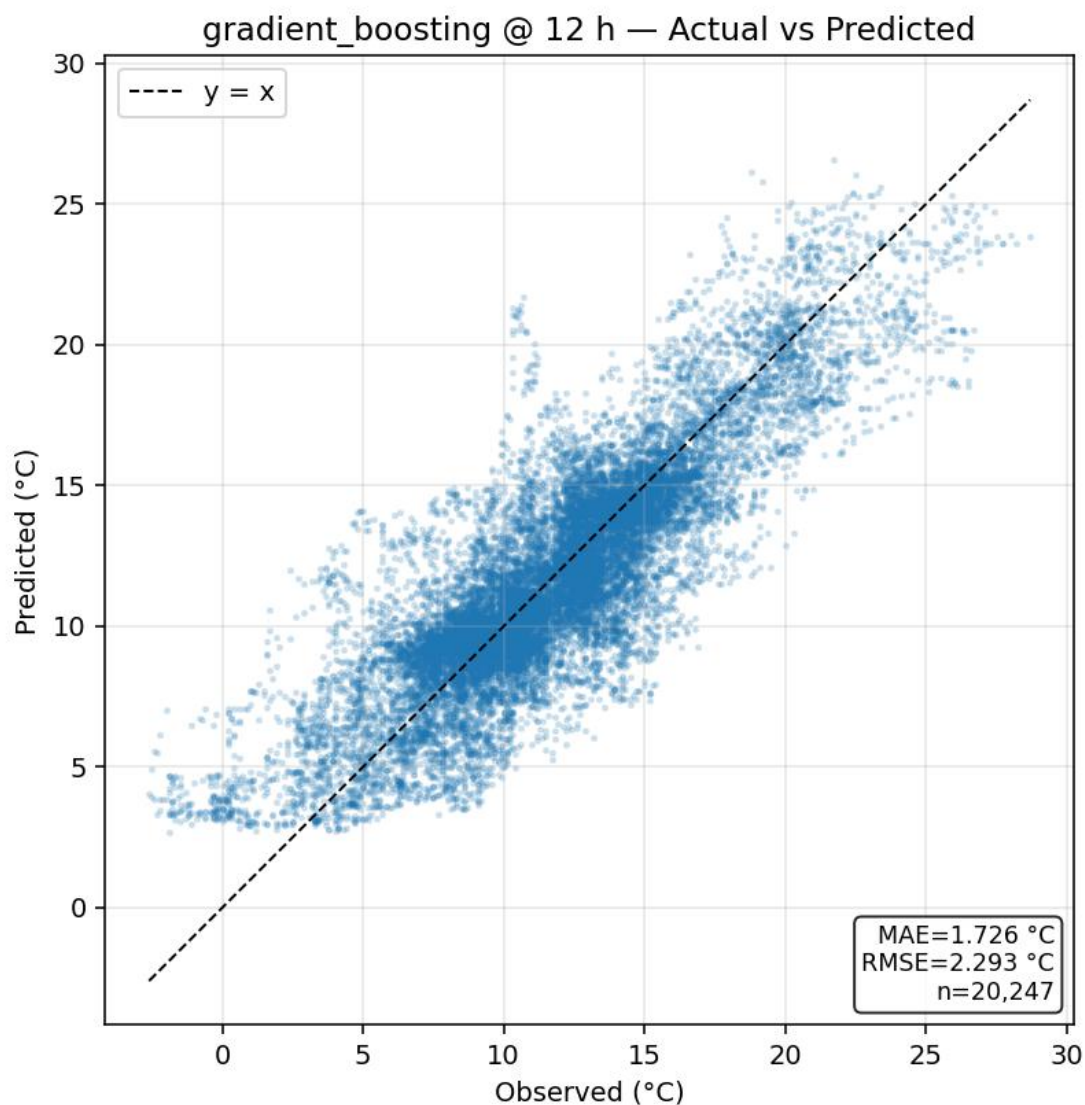
Για τη συνοπτική αξιολόγηση της βελτίωσης έναντι του persistence χρησιμοποιήθηκε heatmap του persistence skill score. Το διάγραμμα αυτό είναι ιδιαίτερα χρήσιμο, επειδή δείχνει όχι μόνο ποιο μοντέλο έχει χαμηλό σφάλμα, αλλά και αν η επίδοσή του είναι ουσιαστικά καλύτερη από ένα πολύ απλό και ισχυρό baseline. Η ερμηνεία του skill score επιβεβαιώνει ότι η σημαντικότερη βελτίωση εμφανίζεται στους ορίζοντες των 6 και 12 ωρών, όπου το Gradient Boosting επιτυγχάνει skill score περίπου 0.75. Αντίθετα, στον ορίζοντα των 24 ωρών το skill score μειώνεται σημαντικά, δείχνοντας ότι το καλύτερο μοντέλο παραμένει μόνο οριακά καλύτερο από το persistence.



Σχήμα 4.9 - Σύγκριση καλύτερου baseline, καλύτερου Machine Learning και καλύτερου Deep Learning μοντέλου ανά χρονικό ορίζοντα.

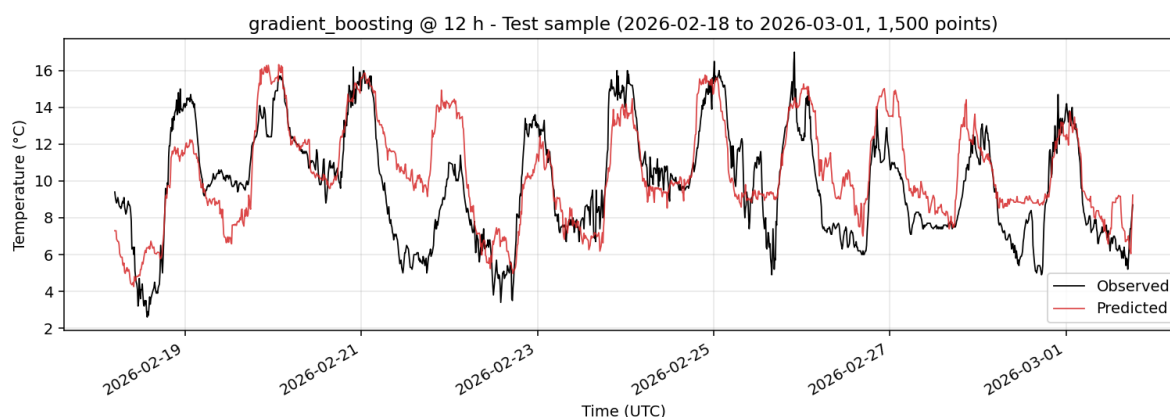
Το διάγραμμα best per family συνοψίζει την καλύτερη επίδοση κάθε οικογένειας μοντέλων. Η σύγκριση δείχνει ότι τα Machine Learning μοντέλα υπερέχουν στους περισσότερους χρονικούς ορίζοντες, ενώ τα Deep Learning μοντέλα δεν ξεπερνούν τα καλύτερα tree-based μοντέλα. Τα baseline μοντέλα παραμένουν ιδιαίτερα σημαντικά για την ερμηνεία, καθώς το persistence είναι πολύ ισχυρό στα 10 λεπτά και στις 24 ώρες.

Για τη διαγνωστική ανάλυση πραγματικών και προβλεπόμενων τιμών επιλέχθηκε ο ορίζοντας των 12 ωρών, επειδή συνδυάζει υψηλό persistence skill score και πρακτική σημασία για ενδοημερήσια πρόβλεψη.



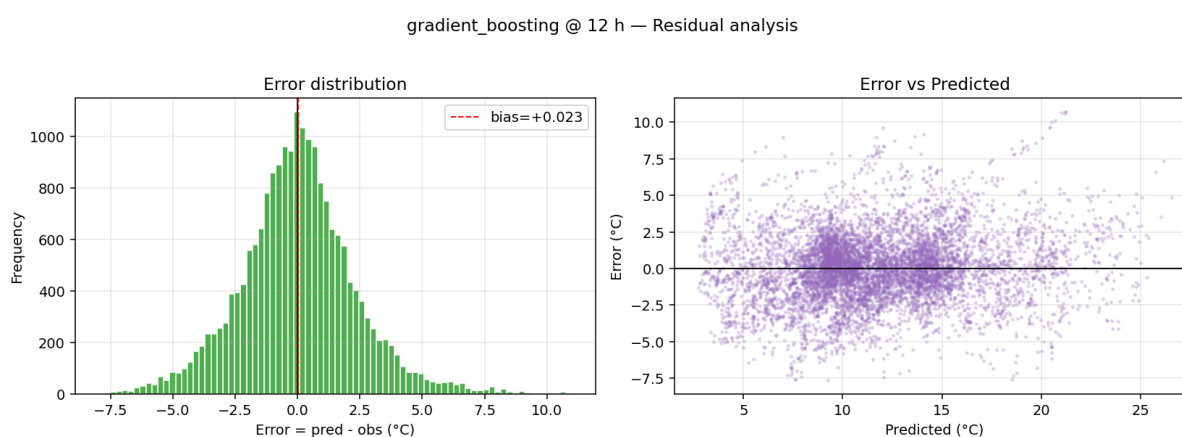
Σχήμα 4.10 - Scatter plot πραγματικών και προβλεπόμενων τιμών για το Gradient Boosting στον ορίζοντα των 12 ωρών.

Τα scatter plots πραγματικών και προβλεπόμενων τιμών επιτρέπουν την οπτική αξιολόγηση του βαθμού συμφωνίας μεταξύ πρόβλεψης και πραγματικής τιμής. Όσο πιο κοντά βρίσκονται τα σημεία στη διαγώνιο $y=x$, τόσο καλύτερη είναι η πρόβλεψη. Στα καλύτερα μοντέλα, όπως το Gradient Boosting στους ορίζοντες των 6 και 12 ωρών, τα σημεία συγκεντρώνονται πιο κοντά στη διαγώνιο, αν και εξακολουθούν να υπάρχουν αποκλίσεις σε περιόδους ταχύτερης μεταβολής της θερμοκρασίας.



Σχήμα 4.11 - Time-series plot πραγματικής και προβλεπόμενης θερμοκρασίας για το Gradient Boosting στον ορίζοντα των 12 ωρών.

Τα time-series plots δείχνουν την εξέλιξη των πραγματικών και προβλεπόμενων τιμών μέσα στον χρόνο. Τα διαγράμματα αυτά είναι χρήσιμα επειδή αποκαλύπτουν αν το μοντέλο ακολουθεί τη γενική τάση της θερμοκρασίας ή αν παρουσιάζει συστηματική καθυστέρηση, υπερεξομάλυνση ή αδυναμία αποτύπωσης απότομων μεταβολών. Τα καλύτερα Machine Learning μοντέλα ακολουθούν ικανοποιητικά τη συνολική μορφή της χρονοσειράς, αλλά τείνουν να εξομαλύνουν ορισμένες απότομες μεταβολές.



Σχήμα 4.12 - Residual diagnostics για το Gradient Boosting στον ορίζοντα των 12 ωρών.

Τέλος, τα residual plots χρησιμοποιήθηκαν για την ανάλυση των σφαλμάτων πρόβλεψης. Η κατανομή των residuals δίνει ένδειξη για το αν τα σφάλματα είναι περίπου συμμετρικά γύρω από το μηδέν, ενώ το διάγραμμα residuals έναντι προβλεπόμενων τιμών βοηθά στον εντοπισμό πιθανής ετεροσκεδαστικότητας. Η ανάλυση αυτή είναι σημαντική, επειδή ένα μοντέλο μπορεί να έχει χαμηλό μέσο σφάλμα αλλά να παρουσιάζει συστηματικά μεγαλύτερα σφάλματα σε συγκεκριμένες θερμοκρασιακές περιοχές.

Συνολικά, οι οπτικοποιήσεις επιβεβαιώνουν και ενισχύουν τα αριθμητικά ευρήματα. Τα Machine Learning μοντέλα, και ειδικά το Gradient Boosting, εμφανίζονται ως η πιο αξιόπιστη επιλογή για τους ενδιάμεσους χρονικούς ορίζοντες. Το persistence παραμένει εξαιρετικά ισχυρό σε πολύ μικρό ορίζοντα, ενώ η μείωση του skill score στις 24 ώρες υποδεικνύει ότι η διαθέσιμη πληροφορία από έναν μόνο τοπικό σταθμό έχει πρακτικό όριο.

4.7 Παραγόμενα Artifacts και Πρακτικά Αποτελέσματα Χρήσης του Λογισμικού

Η εφαρμογή του λογισμικού στην πλήρη πειραματική διαδικασία παράγαγε ένα οργανωμένο σύνολο αρχείων και αποτελεσμάτων, τα οποία χρησιμοποιήθηκαν για την αξιολόγηση των μοντέλων και τη συγγραφή της παρούσας εργασίας. Τα παραγόμενα artifacts δεν αποτελούν απλώς βοηθητικά αρχεία εκτέλεσης, αλλά τεκμηριώνουν τη διαδικασία του πειράματος και επιτρέπουν τον έλεγχο, τη σύγκριση και την αναπαραγωγή των αποτελεσμάτων.

Η αναπαραγωγή της κύριας πειραματικής εκτέλεσης πραγματοποιείται μέσω του αρχείου ρύθμισης `configs/default.yaml` και της εντολής:

```
python -m weather_forecasting_pipeline run-all --config
configs/default.yaml
```

Η εντολή αυτή εκτελεί διαδοχικά τα στάδια `ingest`, `preprocess`, `train` και `evaluate`. Τα ενδιάμεσα και τελικά `outputs` αποθηκεύονται στους φακέλους `data/interim/`, `data/processed/`, `artifacts/` και `runs/<run_id>/`. Η ύπαρξη ενιαίου `configuration file`, `split metadata`, `fitted scalers`, `metrics files` και `run snapshots` επιτρέπει τον έλεγχο της πειραματικής διαδικασίας και τη σύνδεση των αποτελεσμάτων του Κεφαλαίου 4 με συγκεκριμένη εκτέλεση του λογισμικού.

Τα βασικά artifacts περιλαμβάνουν τα `canonical` και `prepared datasets`, τα `supervised datasets` ανά χρονικό ορίζοντα, τα `split metadata`, τους `fitted scalers`, τα αποθηκευμένα μοντέλα, τα αρχεία προβλέψεων, τις μετρικές αξιολόγησης, τα διαγράμματα και τις συνοπτικές αναφορές. Η ύπαρξη αυτών των αρχείων επιτρέπει την πλήρη ιχνηλασιμότητα της πειραματικής διαδικασίας, από τα αρχικά `Weathercloud CSV` αρχεία έως τις τελικές μετρικές αξιολόγησης.

Τα `run snapshots` λειτουργούν ως `frozen` εκδόσεις των πειραματικών αποτελεσμάτων. Κάθε `snapshot` συγκεντρώνει τα `configuration files`, τα `datasets`, τα μοντέλα, τις προβλέψεις, τις μετρικές, τα `plots` και τις αναφορές που αντιστοιχούν σε συγκεκριμένη εκτέλεση. Αυτό επιτρέπει την επιστροφή σε μια προηγούμενη εκτέλεση χωρίς αμφιβολία για το ποια δεδομένα, ποια μοντέλα και ποιες ρυθμίσεις χρησιμοποιήθηκαν.

Τα αρχεία προβλέψεων περιλαμβάνουν τις πραγματικές και τις προβλεπόμενες τιμές για κάθε μοντέλο και χρονικό ορίζοντα. Τα αρχεία αυτά είναι απαραίτητα για την αναπαραγωγή των μετρικών, τη δημιουργία διαγραμμάτων και τη διαγνωστική ανάλυση των σφαλμάτων. Η αποθήκευση των `predictions` κάνει δυνατή την εκ νέου παραγωγή `scatter plots`, `time-series plots` και `residual diagnostics` χωρίς να απαιτείται επανεκπαίδευση των μοντέλων.

Ιδιαίτερα σημαντικά είναι τα `split metadata`, επειδή καταγράφουν τα χρονικά όρια των `training`, `validation` και `test sets` για κάθε χρονικό ορίζοντα. Με αυτόν τον τρόπο μπορεί να ελεγχθεί ότι ο διαχωρισμός έγινε χρονολογικά και ότι δεν υπήρξε τυχαία ανάμειξη των δεδομένων. Αντίστοιχα, οι `fitted scalers` αποθηκεύονται ώστε να είναι σαφές ότι η κλιμάκωση των χαρακτηριστικών έγινε με βάση το `training set` και στη συνέχεια εφαρμόστηκε στα `validation` και `test sets`.

Ορίζοντας	Train start	Train end	Test start	Test end	Train n	Validation n	Test n
10 λεπτά	2023-03-04 16:30 UTC	2025-06-16 13:30 UTC	2025-11-27 09:10 UTC	2026-05-16 10:40 UTC	95.366	20.436	20.436
1 ώρα	2023-03-04 16:30 UTC	2025-06-16 22:00 UTC	2025-11-27 12:50 UTC	2026-05-16 09:50 UTC	95.183	20.396	20.396
3 ώρες	2023-03-04 16:30 UTC	2025-06-17 06:10 UTC	2025-11-27 11:20 UTC	2026-05-16 07:50 UTC	94.936	20.343	20.343
6 ώρες	2023-03-04 16:30 UTC	2025-06-17 13:20 UTC	2025-11-27 07:20 UTC	2026-05-16 04:50 UTC	94.567	20.264	20.265
12 ώρες	2023-03-04 16:30 UTC	2025-06-18 01:40 UTC	2025-11-27 09:40 UTC	2026-05-15 22:50 UTC	94.499	20.250	20.250
24 ώρες	2023-03-04 16:30 UTC	2025-06-17 22:50 UTC	2025-11-26 23:50 UTC	2026-05-15 10:40 UTC	94.223	20.190	20.191

Πίνακας 4.9 - Ενδεικτικά χρονικά όρια splits ανά χρονικό ορίζοντα στο πείραμα αναφοράς.

Οι μετρικές συγκεντρώνονται σε ενιαίο αρχείο, ώστε να είναι δυνατή η σύγκριση όλων των μοντέλων κάτω από το ίδιο πειραματικό πλαίσιο. Το αρχείο metrics.csv περιλαμβάνει, για κάθε μοντέλο και χρονικό ορίζοντα, τις τιμές MAE, RMSE, MAPE, persistence skill score και το πλήθος των test samples. Με αυτόν τον τρόπο, η αξιολόγηση δεν βασίζεται σε επιμέρους χειροκίνητους υπολογισμούς, αλλά σε αναπαραγώγιμη έξοδο του pipeline.

Κατηγορία	Πλήθος / Περιγραφή
Model artifacts	66 αρχεία μοντέλων (11 μοντέλα × 6 ορίζοντες)
Predictions	66 αρχεία προβλέψεων
Comparison plots	5 συγκριτικά διαγράμματα
Actual vs predicted plots	48 διαγράμματα
Residual plots	24 διαγνωστικά διαγράμματα
Supervised datasets	6 αρχεία, ένα ανά χρονικό ορίζοντα
Split metadata	6 αρχεία, ένα ανά χρονικό ορίζοντα
Scalers	Feature και target scalers ανά ορίζοντα

Πίνακας 4.10 - Σύνοψη βασικών παραγόμενων αρχείων στο πείραμα αναφοράς.

Το πείραμα αναφοράς οδήγησε σε ένα οργανωμένο σύνολο αρχείων που καλύπτει όλα τα αξιολογημένα μοντέλα και χρονικούς ορίζοντες. Συγκεκριμένα, αποθηκεύτηκαν 66 αρχεία προβλέψεων (11 μοντέλα × 6 ορίζοντες), τα αντίστοιχα serialized μοντέλα και scalers, οι μετρικές αξιολόγησης σε ενιαίο CSV, καθώς και actual-vs-predicted, residual και comparison plots. Με αυτόν τον τρόπο, η εργασία τεκμηριώνει τόσο τα μοντέλα που τελικά προκρίνονται ως πιο κατάλληλα όσο και τα μοντέλα που αξιολογήθηκαν αλλά κρίθηκαν λιγότερο πρακτικά, υπό την ίδια πειραματική διάταξη και τα ίδια test sets.

Συνεπώς, το βασικό πρακτικό αποτέλεσμα της υλοποίησης είναι ότι η πειραματική διαδικασία δεν παραμένει άτυπη ή χειροκίνητη. Αντίθετα, μετατρέπεται σε οργανωμένη διαδικασία παραγωγής τεκμηριωμένων artifacts, τα οποία υποστηρίζουν τόσο την αξιολόγηση των μοντέλων όσο και την αναπαραγωγικότητα της εργασίας.

Κεφάλαιο 5. Συζήτηση - Συμπεράσματα - Μελλοντικές Επεκτάσεις

5.1 Ανακεφαλαίωση

Η παρούσα εργασία είχε ως στόχο την ανάπτυξη και αξιολόγηση ενός επαναλήψιμου πειραματικού πλαισίου για βραχυπρόθεσμη πρόβλεψη θερμοκρασίας σε τοπική κλίμακα, χρησιμοποιώντας αποκλειστικά ιστορικές παρατηρήσεις από έναν μεμονωμένο μετεωρολογικό σταθμό. Η προσέγγιση αυτή διαφοροποιείται από συστήματα που βασίζονται σε Numerical Weather Prediction outputs, reanalysis datasets, δορυφορικά δεδομένα ή παρατηρήσεις από πολλαπλούς σταθμούς. Στόχος ήταν να εξεταστεί τι μπορεί να επιτευχθεί όταν η διαθέσιμη πληροφορία περιορίζεται αυστηρά σε station-local observation-only δεδομένα.

Η ερευνητική προσέγγιση βασίστηκε στην κατασκευή ενός πλήρους πειραματικού pipeline. Το pipeline καλύπτει όλα τα στάδια της διαδικασίας, από την εισαγωγή των πρωτογενών Weathercloud CSV αρχείων έως την παραγωγή μετρικών αξιολόγησης, διαγραμμάτων, αρχείων προβλέψεων και run snapshots. Η προετοιμασία των μετεωρολογικών χρονοσειρών πραγματοποιήθηκε μέσω του MetDataPy, ενώ το παρόν λογισμικό ανέλαβε την πειραματική ενορχήστρωση, την εκπαίδευση των μοντέλων, την αξιολόγηση και τη διαχείριση των παραγόμενων artifacts.

Σε επίπεδο δεδομένων, η εργασία αξιοποίησε χρονοσειρές με χρονική ανάλυση 10 λεπτών από τον μετεωρολογικό σταθμό Weathercloud της Νέας Τρίγλιας. Η πειραματική αξιολόγηση επικεντρώθηκε στην πρόβλεψη της θερμοκρασίας αέρα (temp_c) για πολλαπλούς χρονικούς ορίζοντες: 10 λεπτά, 1 ώρα, 3 ώρες, 6 ώρες, 12 ώρες και 24 ώρες. Για κάθε ορίζοντα δημιουργήθηκε ξεχωριστό supervised dataset και κάθε μοντέλο εκπαιδεύτηκε ανεξάρτητα, ώστε η αξιολόγηση να είναι συγκρίσιμη και μεθοδολογικά καθαρή.

Στο πειραματικό σκέλος αξιολογήθηκαν τρεις βασικές οικογένειες μοντέλων: baseline models, παραδοσιακά μοντέλα Machine Learning και μοντέλα Deep Learning. Τα baseline models περιλάμβαναν persistence, moving average και climatology. Τα μοντέλα Machine Learning περιλάμβαναν linear regression, ridge regression, random forest, gradient boosting και support vector regression. Τα μοντέλα Deep Learning περιλάμβαναν LSTM, GRU και TCN.

Τα αποτελέσματα έδειξαν ότι η βραχυπρόθεσμη πρόβλεψη θερμοκρασίας από έναν μόνο τοπικό σταθμό είναι εφικτή και μπορεί να αποδώσει χρήσιμα αποτελέσματα, κυρίως στους ορίζοντες από 1 έως 12 ώρες. Στον ορίζοντα των 10 λεπτών, το persistence baseline παρέμεινε εξαιρετικά ισχυρό, ενώ το ridge πέτυχε οριακά καλύτερη αριθμητική επίδοση. Στους ορίζοντες της 1 και των 3 ωρών, το Random Forest πέτυχε την καλύτερη απόδοση, ενώ στους ορίζοντες των 6, 12 και 24 ωρών επικράτησε το Gradient Boosting.

Το σημαντικότερο ποσοτικό εύρημα ήταν ότι το gradient_boosting πέτυχε **skill score +0.745 στις 6 ώρες και +0.752 στις 12 ώρες** έναντι του persistence baseline, δηλαδή μείωση του τετραγωνικού σφάλματος κατά περίπου 75%. Στις 24 ώρες, η αντίστοιχη βελτίωση μειώνεται σε μόλις **+0.058**, υποδεικνύοντας ότι σε αυτόν τον ορίζοντα η observation-only πληροφορία τοπικού σταθμού πλησιάζει σε σημείο κορεσμού έναντι του ισχυρού ημερήσιου-κύκλου persistence. Το εύρημα αυτό δείχνει ότι τα engineered χαρακτηριστικά (lag features, rolling statistics, κυκλικές χρονικές μεταβλητές, QC flags και παράγωγα μετεωρολογικά χαρακτηριστικά) περιέχουν αξιοποιήσιμη πληροφορία για βραχυπρόθεσμη πρόβλεψη θερμοκρασίας στο διάστημα 1-12 ωρών, αλλά δεν επαρκούν για ουσιαστική βελτίωση στις 24 ώρες χωρίς εξωτερικές προγνωστικές εισόδους.

Παράλληλα, η εργασία ανέδειξε και τα όρια της observation-only προσέγγισης. Στον ορίζοντα των 24 ωρών, η βελτίωση έναντι του persistence μειώθηκε σημαντικά, με το καλύτερο μοντέλο να παραμένει μόνο οριακά καλύτερο από το baseline. Το αποτέλεσμα αυτό δείχνει ότι η πληροφορία που περιέχεται αποκλειστικά στις τοπικές παρατηρήσεις έχει πεπερασμένη προβλεπτική αξία και ότι για μεγαλύτερους χρονικούς ορίζοντες πιθανώς απαιτούνται πρόσθετες εξωτερικές πηγές πληροφορίας.

5.2 Συζήτηση και Συμπεράσματα

Τα αποτελέσματα της εργασίας οδηγούν σε τέσσερα βασικά συμπεράσματα. Το πρώτο αφορά τη δυνατότητα βραχυπρόθεσμης πρόβλεψης θερμοκρασίας από έναν μόνο μετεωρολογικό σταθμό. Τα αποτελέσματα δείχνουν ότι οι τοπικές ιστορικές παρατηρήσεις περιέχουν επαρκή πληροφορία για χρήσιμη πρόβλεψη σε χρονικούς ορίζοντες λίγων ωρών. Η επίδοση των μοντέλων Machine Learning στους ορίζοντες από 1 έως 12 ώρες δείχνει ότι το station-local observation-only setup είναι επιστημονικά και πρακτικά αξιοποιήσιμο.

Το δεύτερο συμπέρασμα αφορά τη σημασία των baseline μοντέλων. Το persistence baseline αποδείχθηκε ιδιαίτερα ισχυρό, ειδικά στον πολύ μικρό ορίζοντα των 10 λεπτών και στον ορίζοντα των 24 ωρών. Στα 10 λεπτά, η τελευταία διαθέσιμη παρατήρηση αποτελεί πολύ ισχυρή πρόβλεψη, επειδή η θερμοκρασία συνήθως μεταβάλλεται αργά σε τόσο σύντομο χρονικό διάστημα. Στις 24 ώρες, η ισχύς του persistence συνδέεται με την ημερήσια περιοδικότητα της θερμοκρασίας, καθώς η τιμή στο $t+24h$ αντιστοιχεί περίπου στην ίδια ώρα της επόμενης ημέρας.

Το τρίτο συμπέρασμα αφορά την υπεροχή των tree-based Machine Learning μοντέλων. Το Random Forest ήταν το καλύτερο μοντέλο στους ορίζοντες της 1 και των 3 ωρών, ενώ το Gradient Boosting ήταν το καλύτερο μοντέλο στους ορίζοντες των 6, 12 και 24 ωρών. Ιδιαίτερα στους ορίζοντες των 6 και 12 ωρών, το Gradient Boosting πέτυχε persistence skill score περίπου 0.75, γεγονός που δείχνει ουσιαστική βελτίωση έναντι ενός ισχυρού baseline. Το αποτέλεσμα αυτό αποτελεί ένα από τα πιο σημαντικά ευρήματα της εργασίας.

Το τέταρτο συμπέρασμα αφορά τη συμπεριφορά των μοντέλων Deep Learning. Παρότι τα LSTM, GRU και TCN είναι θεωρητικά κατάλληλα για χρονοσειρές, στο συγκεκριμένο πείραμα δεν ξεπέρασαν τα καλύτερα Machine Learning μοντέλα. Το LSTM ήταν το πιο σταθερό από τα Deep Learning μοντέλα, όμως παρέμεινε πίσω από το Gradient Boosting στους σημαντικότερους ορίζοντες. Το GRU και το TCN παρουσίασαν αστάθειες σε ορισμένες περιπτώσεις, ιδίως σε μεγαλύτερους χρονικούς ορίζοντες.

#	Εύρημα	Ποσοτική τεκμηρίωση	Σημασία
1	Η βραχυπρόθεσμη πρόβλεψη θερμοκρασίας από έναν μόνο τοπικό σταθμό είναι εφικτή.	Skill score έναντι persistence: +0.374 στη 1 h, +0.652 στις 3 h, +0.745 στις 6 h, +0.752 στις 12 h.	Το station-local observation-only setup είναι επιστημονικά αξιοποιήσιμο στους ορίζοντες 1-12 h.
2	Το persistence αποτελεί εξαιρετικά ισχυρό baseline, ειδικά στα 10 min και στις 24 h.	MAE 0.257 °C στα 10 min· επανέρχεται σε MAE 2.021 °C στις 24 h λόγω ημερήσιας περιοδικότητας.	Κάθε πιο σύνθετο μοντέλο πρέπει να αξιολογείται σε σχέση με το persistence, όχι μόνο σε απόλυτο MAE.
3	Τα tree-based ensemble μοντέλα υπερισχύουν συνολικά.	Random Forest: καλύτερο στα 1 h (MAE 0.624 °C) και στις 3 h (MAE 1.029 °C). Gradient Boosting:	Αποτελούν την πιο αποτελεσματική πρακτική

#	Εύρημα	Ποσοτική τεκμηρίωση	Σημασία
		καλύτερο στις 6, 12 και 24 h (MAE 1.389 / 1.726 / 1.998 °C).	επιλογή για το συγκεκριμένο πρόβλημα.
4	Τα Deep Learning μοντέλα δεν ξεπέρασαν τα tree-based ML σε κανέναν χρονικό ορίζοντα.	Καλύτερο DL (LSTM h12): MAE 1.974 °C, έναντι 1.726 °C του Gradient Boosting. Παρατηρήθηκαν αστάθειες (TCN h12: MAE 5.617 °C· GRU h24: MAE 5.175 °C).	Η πολυπλοκότητα του μοντέλου δεν συνεπάγεται καλύτερη απόδοση σε station-based, καλά δομημένα datasets.
5	Η observation-only προσέγγιση έχει σαφές πρακτικό όριο στις 24 h.	Skill score του καλύτερου μοντέλου στις 24 h: μόλις +0.058.	Σε μεγαλύτερους ορίζοντες απαιτούνται εξωτερικές προγνωστικές πηγές για ουσιαστική βελτίωση.

Πίνακας 5.1 - Κύρια ευρήματα της εργασίας.

Ιδιαίτερη προσοχή χρειάζεται στην ερμηνεία της σχέσης μεταξύ των οριζόντων 12 και 24 ωρών. Το γεγονός ότι το απόλυτο σφάλμα στις 24 ώρες δεν είναι απαραίτητα μεγαλύτερο από το αντίστοιχο σφάλμα στις 12 ώρες δεν σημαίνει ότι η 24ωρη πρόβλεψη είναι ευκολότερη με γενικό τρόπο. Η συμπεριφορά αυτή συνδέεται με την ημερήσια περιοδικότητα της θερμοκρασίας. Η τιμή στο t+24h αντιστοιχεί στην ίδια περίπου ώρα της επόμενης ημέρας, ενώ η τιμή στο t+12h βρίσκεται συχνά σε αντίθετη φάση του ημερήσιου κύκλου. Για τον λόγο αυτό, η σύγκριση πρέπει να γίνεται όχι μόνο με βάση το MAE, αλλά και με βάση τη βελτίωση έναντι του persistence.

Σε σχέση με τη βιβλιογραφία, τα αποτελέσματα της εργασίας είναι συνεπή με την παρατήρηση ότι τα Deep Learning μοντέλα είναι ιδιαίτερα ισχυρά σε προβλήματα με μεγάλα και πολυδιάστατα datasets, αλλά δεν αποτελούν πάντα την καλύτερη επιλογή σε μικρότερα, καλά δομημένα και station-based datasets. Η παρούσα εργασία δείχνει ότι για τοπική πρόβλεψη από έναν σταθμό, τα παραδοσιακά Machine Learning μοντέλα μπορούν να είναι όχι μόνο ανταγωνιστικά, αλλά και ανώτερα από πιο σύνθετες αρχιτεκτονικές.

Η σύγκριση με άλλες ερευνητικές προσπάθειες πρέπει επίσης να λάβει υπόψη τη φύση των εισόδων. Πολλές σύγχρονες μελέτες πρόγνωσης καιρού χρησιμοποιούν NWP δεδομένα, δορυφορικές παρατηρήσεις, reanalysis datasets ή πολλαπλούς σταθμούς. Αυτές οι πηγές αυξάνουν την πληροφοριακή πληρότητα του προβλήματος, αλλά παράλληλα μεταβάλλουν τη φύση της πρόβλεψης. Η παρούσα εργασία εξετάζει ένα αυστηρότερο και πιο περιορισμένο σενάριο, στο οποίο η πρόβλεψη βασίζεται αποκλειστικά σε τοπικές παρατηρήσεις. Επομένως, τα αποτελέσματα δεν πρέπει να συγκρίνονται απευθείας με συστήματα που χρησιμοποιούν εξωτερικές προγνωστικές εισόδους.

Συμπερασματικά, η εργασία καταλήγει ότι η καταλληλότερη προσέγγιση για το συγκεκριμένο πρόβλημα είναι η χρήση tree-based Machine Learning μοντέλων, με έμφαση στο Gradient Boosting για τους ορίζοντες των 6 και 12 ωρών. Τα baseline μοντέλα παραμένουν απαραίτητα για την ερμηνεία των αποτελεσμάτων, ενώ τα Deep Learning μοντέλα προσφέρουν χρήσιμο σημείο σύγκρισης αλλά δεν αποτελούν την καλύτερη πρακτική επιλογή στο συγκεκριμένο πειραματικό πλαίσιο.

5.3 Αξιοποίηση και Πρακτικές Προεκτάσεις της Έρευνας

Τα αποτελέσματα της εργασίας μπορούν να αξιοποιηθούν τόσο σε πρακτικό όσο και σε ερευνητικό επίπεδο. Η βασική πρακτική αξία της προσέγγισης είναι ότι δείχνει πως ένας τοπικός μετεωρολογικός σταθμός μπορεί να αποτελέσει επαρκή βάση για βραχυπρόθεσμη πρόβλεψη θερμοκρασίας, χωρίς την ανάγκη πρόσβασης σε σύνθετες εξωτερικές προγνωστικές πηγές. Αυτό είναι ιδιαίτερα σημαντικό για περιπτώσεις όπου οι διαθέσιμοι πόροι είναι περιορισμένοι ή όπου ζητείται ένα ελαφρύ, τοπικό και αυτοτελές forecasting system.

Σε πρακτικό επίπεδο, ένα τέτοιο σύστημα θα μπορούσε να χρησιμοποιηθεί για βραχυπρόθεσμη υποστήριξη αποφάσεων σε τοπικές εφαρμογές. Παραδείγματα περιλαμβάνουν μικρής κλίμακας αγροτικές εφαρμογές, τοπική παρακολούθηση περιβαλλοντικών συνθηκών, υποστήριξη μικρών ενεργειακών εφαρμογών ή απλή ενημέρωση για την αναμενόμενη θερμοκρασία των επόμενων ωρών. Η εργασία δεν ισχυρίζεται ότι αντικαθιστά επίσημα μετεωρολογικά συστήματα πρόγνωσης, αλλά δείχνει ότι υπάρχει αξιοποιήσιμο τοπικό σήμα στις ιστορικές παρατηρήσεις.

Ερευνητικά, η εργασία προσφέρει ένα επαναλήψιμο πλαίσιο σύγκρισης μοντέλων για station-based forecasting. Η ύπαρξη pipeline, configuration files, split metadata, predictions, metrics και plots επιτρέπει τη συστηματική επανάληψη των πειραμάτων με διαφορετικούς στόχους, διαφορετικές περιοχές ή διαφορετικές αρχιτεκτονικές μοντέλων. Αυτό καθιστά την υλοποίηση χρήσιμη βάση για μελλοντικές εργασίες που θα εξετάσουν περισσότερες μετεωρολογικές μεταβλητές ή περισσότερους σταθμούς.

Κατηγορία	Προέκταση	Σύντομη Περιγραφή
Πρακτική	Τοπικό σύστημα υποστήριξης αποφάσεων	Αξιοποίηση των μοντέλων για βραχυπρόθεσμη ενημέρωση σε ορίζοντες 1-12 ωρών σε εφαρμογές περιορισμένης κλίμακας.
Πρακτική	Μικρής κλίμακας αγροτικές εφαρμογές	Πρόβλεψη θερμοκρασίας για προγραμματισμό άρδευσης, προστασίας από παγετό και εργασιών πεδίου.
Πρακτική	Τοπική περιβαλλοντική παρακολούθηση	Συνδυασμός παρατηρήσεων σταθμού και βραχυπρόθεσμων προβλέψεων για παρακολούθηση τοπικών συνθηκών.
Πρακτική	Υποστήριξη μικρών ενεργειακών εφαρμογών	Εκτίμηση επόμενων ωρών για διαχείριση φορτίου, θέρμανσης/ψύξης ή αυτόνομων συστημάτων.
Ερευνητική	Επαναλήψιμο πλαίσιο σύγκρισης μοντέλων	Δυνατότητα επανάληψης του πειράματος για διαφορετικούς στόχους, σταθμούς ή αρχιτεκτονικές υπό κοινή μεθοδολογία.
Ερευνητική	Βάση για διεύρυνση σε πολλαπλούς στόχους	Δυνατότητα εναλλαγής μεταβλητής-στόχου διατηρώντας τη μεθοδολογία (διαχωρισμός, lags, rolling, μετρικές).
Εκπαιδευτική	Πλήρες παράδειγμα ML forecasting workflow	Καλύπτει προεπεξεργασία χρονοσειρών, feature engineering, χρονολογικά splits, train-only scaling, multi-horizon evaluation και διαγνωστική ανάλυση.
Επιχειρησιακή	Επέκταση σε real-time inference	Σύνδεση με ροή νέων παρατηρήσεων και παραγωγή προβλέψεων σε πραγματικό χρόνο με τα εκπαιδευμένα μοντέλα.

Κατηγορία	Προέκταση	Σύντομη Περιγραφή
Επιχειρησιακή	Διεπαφή χρήστη ή dashboard	Παρουσίαση τρέχουσας κατάστασης σταθμού, προβλέψεων ανά ορίζοντα και ιστορικού σφάλματος για μη τεχνικούς χρήστες.

Πίνακας 5.2 - Πρακτικές και ερευνητικές προεκτάσεις της εργασίας.

Επιπλέον, το λογισμικό μπορεί να χρησιμοποιηθεί εκπαιδευτικά, καθώς ενσωματώνει πολλά κρίσιμα στοιχεία ενός πραγματικού machine learning forecasting workflow: προεπεξεργασία χρονοσειρών, feature engineering, χρονολογικά splits, train-only scaling, baseline σύγκριση, multi-horizon evaluation, model artifacts και διαγνωστική ανάλυση αποτελεσμάτων. Με αυτόν τον τρόπο, μπορεί να αποτελέσει χρήσιμο παράδειγμα για φοιτητές ή ερευνητές που μελετούν εφαρμογές μηχανικής μάθησης σε χρονοσειρές.

Τέλος, η εργασία μπορεί να αξιοποιηθεί ως αρχικό βήμα για πιο ολοκληρωμένο σύστημα τοπικής πρόγνωσης. Με την προσθήκη scheduled execution, ενημέρωσης με νέα δεδομένα και απλής διεπαφής χρήστη, το pipeline θα μπορούσε να μετατραπεί από πειραματικό εργαλείο σε λειτουργικό σύστημα παραγωγής προβλέψεων. Ωστόσο, μια τέτοια επέκταση θα απαιτούσε πρόσθετο σχεδιασμό σε θέματα deployment, παρακολούθησης ποιότητας δεδομένων, επανεκπαίδευσης μοντέλων και παρουσίασης των προβλέψεων σε τελικό χρήστη.

5.4 Περιορισμοί και Μελλοντικές Επεκτάσεις

Παρότι η εργασία πέτυχε τον βασικό της στόχο, δηλαδή την ανάπτυξη και αξιολόγηση ενός observation-only πλαισίου βραχυπρόθεσμης πρόβλεψης από έναν τοπικό μετεωρολογικό σταθμό, υπάρχουν ορισμένοι περιορισμοί που πρέπει να ληφθούν υπόψη. Οι περιορισμοί αυτοί αφορούν τόσο τη φύση των δεδομένων όσο και το εύρος της μοντελοποίησης, την υπολογιστική διάσταση της υλοποίησης και τη δυνατότητα γενίκευσης των αποτελεσμάτων.

Ο πρώτος βασικός περιορισμός είναι ότι η εργασία βασίστηκε σε δεδομένα από έναν μόνο μετεωρολογικό σταθμό. Αυτό ήταν συνειδητή μεθοδολογική επιλογή, καθώς στόχος ήταν να μελετηθεί η πρόβλεψη σε αυστηρά τοπικό και single-station πλαίσιο. Ωστόσο, η επιλογή αυτή περιορίζει τη δυνατότητα γενίκευσης των αποτελεσμάτων σε άλλες γεωγραφικές περιοχές. Η συμπεριφορά της θερμοκρασίας στη Νέα Τρίγλια μπορεί να επηρεάζεται από τοπικά μικροκλιματικά χαρακτηριστικά, τα οποία δεν είναι απαραίτητα ίδια σε άλλους σταθμούς.

Ένας δεύτερος περιορισμός αφορά τον στόχο πρόβλεψης. Στο πλήρες πείραμα αναφοράς, η πρόβλεψη επικεντρώθηκε στη θερμοκρασία αέρα (temp_c). Άλλες μετεωρολογικές μεταβλητές, όπως η σχετική υγρασία, η βροχόπτωση, η ταχύτητα ανέμου ή η ατμοσφαιρική πίεση, χρησιμοποιήθηκαν ως είσοδοι αλλά δεν αποτέλεσαν ξεχωριστούς στόχους πρόβλεψης. Επομένως, τα συμπεράσματα της εργασίας αφορούν κυρίως την πρόβλεψη θερμοκρασίας και δεν πρέπει να γενικεύονται αυτόματα σε πλήρη πρόγνωση καιρού.

Ένας τρίτος περιορισμός σχετίζεται με την απουσία εξωτερικών προγνωστικών εισόδων. Η εργασία δεν χρησιμοποίησε Numerical Weather Prediction outputs, δορυφορικά δεδομένα, reanalysis δεδομένα ή παρατηρήσεις από γειτονικούς σταθμούς. Η επιλογή αυτή καθιστά το πρόβλημα πιο καθαρό μεθοδολογικά και επιτρέπει την αποτίμηση της καθαρής αξίας των τοπικών παρατηρήσεων. Ταυτόχρονα όμως περιορίζει την προβλεπτική ικανότητα σε μεγαλύτερους χρονικούς ορίζοντες, όπως φάνηκε από τη μικρή βελτίωση έναντι του persistence στις 24 ώρες.

Σε επίπεδο μοντελοποίησης, τα αποτελέσματα έδειξαν ότι τα Deep Learning μοντέλα δεν ξεπέρασαν τα καλύτερα Machine Learning μοντέλα. Αυτό αποτελεί εύρημα της εργασίας, αλλά και ένδειξη ότι η συγκεκριμένη διαμόρφωση DL μπορεί να μην είναι η βέλτιστη. Η απόδοση των DL μοντέλων ενδέχεται να βελτιωθεί με συστηματικότερο hyperparameter tuning, περισσότερα seeds, μεγαλύτερα training budgets, διαφορετικές αρχιτεκτονικές ή μεγαλύτερο ιστορικό παράθυρο εισόδου. Ωστόσο, τέτοιες επεκτάσεις θα πρέπει να γίνουν με προσοχή, ώστε να παραμείνει δίκαιη η σύγκριση με τα υπόλοιπα μοντέλα.

Υπολογιστικός και μεθοδολογικός περιορισμός παρατηρήθηκε και στην αξιολόγηση ορισμένων μοντέλων. Το Support Vector Regression αξιολογήθηκε στο πείραμα, αλλά είχε υψηλό υπολογιστικό κόστος και δεν επικράτησε σε κανέναν χρονικό ορίζοντα. Για τον λόγο αυτό κρίθηκε μη πρακτικό για μελλοντικές πλήρεις εκτελέσεις στο συγκεκριμένο dataset. Αντίστοιχα, η απλή linear regression παρουσίασε ακραίες τιμές RMSE σε μεγαλύτερους ορίζοντες, γεγονός που ανέδειξε την ανάγκη αξιολόγησης regularized linear μοντέλου, όπως Ridge regression. Το Ridge βελτίωσε την επίδοση στον ορίζοντα των 10 λεπτών, αλλά παρέμεινε μη ανταγωνιστικό σε μεγαλύτερους ορίζοντες λόγω υψηλών τιμών RMSE.

Περιορισμός	Φύση	Πιθανή μελλοντική επέκταση
Δεδομένα από έναν μόνο σταθμό (Νέα Τρίγλια)	Γενικευσιμότητα	Αξιολόγηση του pipeline σε περισσότερους σταθμούς και διαφορετικά μικροκλίματα.
Στόχος πρόβλεψης μόνο η θερμοκρασία αέρα	Εύρος μοντελοποίησης	Εκτέλεση αντίστοιχων πειραμάτων για σχετική υγρασία, ατμοσφαιρική πίεση, ταχύτητα/ρίπες ανέμου και ένταση βροχόπτωσης.
Απουσία εξωτερικών προγνωστικών εισόδων (NWP, reanalysis, δορυφορικά, γειτονικοί σταθμοί)	Προβλεπτικό εύρος	Διερεύνηση multi-source forecasting setup ως ξεχωριστή μεθοδολογική κατεύθυνση.
Offline αξιολόγηση με ιστορικό train/val/test split	Λειτουργικό σενάριο	Ανάπτυξη real-time inference και συνεχούς αξιολόγησης (drift monitoring, rolling re-evaluation).
DL μοντέλα: δεν ξεπέρασαν τα tree-based και εμφάνισαν αστάθειες σε ορισμένους ορίζοντες	Μοντελοποίηση	Συστηματικό hyperparameter tuning, διαφορετικά sequence lengths, dropout rates, learning-rate schedules, πολλαπλά seeds.
SVR: υψηλό υπολογιστικό κόστος, καμία επικράτηση σε ορίζοντα	Υπολογιστικός	Αποκλεισμός από μελλοντικές πλήρεις εκτελέσεις στο συγκεκριμένο πειραματικό πλαίσιο.
Linear regression: ακραίες τιμές RMSE σε μεγαλύτερους ορίζοντες (έως 13.863 °C στις 24 h)	Σταθερότητα	Προτίμηση regularized γραμμικών μοντέλων (π.χ. Ridge) ή διερεύνηση robust εναλλακτικών.
Ridge: αν και βελτίωσε τα 10 min (MAE 0.254 °C), παρέμεινε μη ανταγωνιστικό στους μεγαλύτερους ορίζοντες (RMSE 5.420-11.562 °C)	Σταθερότητα γραμμικών	Διερεύνηση εναλλακτικών feature engineering ή robust loss functions για μείωση των ακραίων προβλέψεων.
Συγκεκριμένη παραμετροποίηση και περιορισμένος αριθμός random seeds	Μεθοδολογικός	Πιο εκτεταμένο hyperparameter search και πολλαπλά seeds για στατιστικά πιο σταθερή εκτίμηση.

Περιορισμός	Φύση	Πιθανή μελλοντική επέκταση
Έλλειψη ερμηνευσιμότητας μοντέλων	Επιστημονική κατανόηση	Χρήση feature importances ή SHAP values για τα tree-based μοντέλα και σύνδεση με μετεωρολογικές σχέσεις.
Απουσία λειτουργίας παραγωγικής χρήσης	Λογισμικό	Ανάπτυξη pipeline inference, scheduled execution και dashboard για βραχυπρόθεσμες προβλέψεις σε τελικούς χρήστες.

Πίνακας 5.3 - Περιορισμοί της εργασίας και πιθανές μελλοντικές επεκτάσεις.

Με βάση τους παραπάνω περιορισμούς, προκύπτουν αρκετές μελλοντικές επεκτάσεις. Η πρώτη και πιο άμεση επέκταση είναι η εκτέλεση αντίστοιχων πειραμάτων για επιπλέον μετεωρολογικούς στόχους πρόβλεψης, όπως η σχετική υγρασία, η ατμοσφαιρική πίεση, η ταχύτητα ανέμου, οι ριπές ανέμου ή η ένταση βροχόπτωσης. Η υφιστάμενη υλοποίηση του pipeline υποστηρίζει ήδη μεθοδολογικά την εναλλαγή στόχου διατηρώντας τον ίδιο διαχωρισμό, τις ίδιες lag και rolling μεταβλητές και τις ίδιες μετρικές. Επομένως, τα supplementary αυτά πειράματα μπορούν να εκτελεστούν σε επόμενη φάση της εργασίας, ώστε να διαπιστωθεί αν τα συμπεράσματα που προέκυψαν για τη θερμοκρασία ισχύουν και για πιο ασταθείς ή πιο δύσκολες μεταβλητές.

Μια δεύτερη επέκταση είναι η αξιολόγηση του pipeline σε περισσότερους μετεωρολογικούς σταθμούς. Με τον τρόπο αυτό θα μπορούσε να εξεταστεί αν τα μοντέλα που αποδίδουν καλύτερα στη Νέα Τρίγλια διατηρούν την απόδοσή τους και σε άλλες περιοχές. Η σύγκριση διαφορετικών σταθμών θα βοηθούσε στην εκτίμηση της γενικευσιμότητας της προσέγγισης και στην κατανόηση του ρόλου των τοπικών μικροκλιματικών συνθηκών.

Μια τρίτη επέκταση αφορά την εισαγωγή εξωτερικών δεδομένων, όπως NWP outputs ή δεδομένα από γειτονικούς σταθμούς. Ωστόσο, μια τέτοια επέκταση θα αποτελούσε μεθοδολογική αλλαγή σε σχέση με την παρούσα εργασία, επειδή θα μετέτρεπε το πρόβλημα από αυστηρά observation-only και single-station σε πιο σύνθετο multi-source forecasting setup. Για τον λόγο αυτό, θα πρέπει να τεκμηριωθεί ξεχωριστά και να αξιολογηθεί ως νέα πειραματική κατεύθυνση.

Σε επίπεδο μοντέλων, θα μπορούσε να πραγματοποιηθεί πιο συστηματικό hyperparameter tuning για τα Random Forest, Gradient Boosting και Deep Learning μοντέλα. Ειδικά για τα DL μοντέλα, θα είχε ενδιαφέρον να εξεταστούν διαφορετικά sequence lengths, μεγαλύτερα training budgets, dropout rates, learning-rate schedules, batch sizes και αρχιτεκτονικές. Επιπλέον, η χρήση πολλαπλών random seeds θα επέτρεπε πιο σταθερή εκτίμηση της απόδοσης και της διακύμανσης των αποτελεσμάτων.

Μια επιπλέον ερευνητική κατεύθυνση είναι η χρήση μεθόδων explainability. Για τα tree-based μοντέλα, μπορούν να εξεταστούν feature importances ή SHAP values, ώστε να εντοπιστούν τα χαρακτηριστικά που συμβάλλουν περισσότερο στην πρόβλεψη. Αυτό θα βοηθούσε στην καλύτερη κατανόηση της συμπεριφοράς των μοντέλων και στη σύνδεση των αποτελεσμάτων με φυσικά ερμηνεύσιμες μετεωρολογικές σχέσεις.

Σε επίπεδο λογισμικού, μελλοντική επέκταση θα μπορούσε να είναι η ανάπτυξη λειτουργίας inference σε πραγματικό χρόνο. Το υπάρχον pipeline εκπαιδεύει και αξιολογεί μοντέλα, αλλά θα μπορούσε να επεκταθεί ώστε να διαβάζει νέες παρατηρήσεις από τον σταθμό, να εφαρμόζει

τους αποθηκευμένους scalars και να παράγει βραχυπρόθεσμες προβλέψεις με τα αποθηκευμένα μοντέλα. Μια τέτοια επέκταση θα μετέτρεπε το ερευνητικό pipeline σε πρακτικό σύστημα πρόγνωσης.

Τέλος, θα μπορούσε να αναπτυχθεί απλή διεπαφή χρήστη ή dashboard για την παρουσίαση των προβλέψεων. Ένα τέτοιο dashboard θα μπορούσε να εμφανίζει την τρέχουσα κατάσταση του σταθμού, τις προβλέψεις ανά χρονικό ορίζοντα, την αβεβαιότητα ή το ιστορικό σφάλματος του μοντέλου. Η προσθήκη διεπαφής θα έκανε το σύστημα πιο προσβάσιμο σε μη τεχνικούς χρήστες και θα διευκόλυνε την πρακτική του αξιοποίηση.

5.5 Τελικό Συμπέρασμα

Η παρούσα εργασία έδειξε ότι η βραχυπρόθεσμη πρόβλεψη θερμοκρασίας σε τοπική κλίμακα μπορεί να υποστηριχθεί αποτελεσματικά με χρήση αποκλειστικά ιστορικών παρατηρήσεων από έναν μεμονωμένο μετεωρολογικό σταθμό. Παρότι η προσέγγιση αυτή είναι περιορισμένη σε σχέση με συστήματα που αξιοποιούν NWP outputs, δορυφορικά δεδομένα, reanalysis datasets ή πολλαπλούς σταθμούς, τα αποτελέσματα έδειξαν ότι περιέχει επαρκή πληροφορία για χρήσιμες προβλέψεις σε χρονικούς ορίζοντες λίγων ωρών.

Το βασικό εύρημα είναι ότι τα παραδοσιακά μοντέλα Machine Learning, και ειδικότερα τα tree-based ensemble μοντέλα, αποτέλεσαν την πιο αποδοτική επιλογή στο συγκεκριμένο πειραματικό πλαίσιο. Το Random Forest και το Gradient Boosting αξιοποίησαν αποτελεσματικά τα lag features, τα rolling statistics, τις κυκλικές χρονικές μεταβλητές, τα QC flags και τα παράγωγα μετεωρολογικά χαρακτηριστικά, επιτυγχάνοντας καλύτερη απόδοση από τα baseline και τα Deep Learning μοντέλα στους περισσότερους χρονικούς ορίζοντες.

Παράλληλα, η εργασία ανέδειξε ότι η αυξημένη πολυπλοκότητα ενός μοντέλου δεν συνεπάγεται απαραίτητα καλύτερη απόδοση. Τα Deep Learning μοντέλα παρείχαν χρήσιμο σημείο σύγκρισης, αλλά δεν ξεπέρασαν τα καλύτερα Machine Learning μοντέλα. Το αποτέλεσμα αυτό είναι σημαντικό, επειδή υπογραμμίζει την ανάγκη επιλογής μοντέλων με βάση το συγκεκριμένο πρόβλημα, τα διαθέσιμα δεδομένα και το υπολογιστικό κόστος, και όχι μόνο με βάση τη θεωρητική ισχύ μιας μεθόδου.

Ένα ακόμη κεντρικό συμπέρασμα είναι ότι η observation-only πρόβλεψη έχει σαφές πρακτικό όριο. Η μεγαλύτερη βελτίωση έναντι του persistence παρατηρήθηκε στους ορίζοντες των 6 και 12 ωρών (skill +0.745 και +0.752 αντίστοιχα, δηλαδή περίπου 75% μείωση του τετραγωνικού σφάλματος), ενώ στις 24 ώρες η βελτίωση μειώθηκε σε **+0.058**, δηλαδή σε σχεδόν αμελητέα διαφορά έναντι του ημερήσιου-κύκλου persistence. Αυτό δείχνει ότι οι τοπικές ιστορικές παρατηρήσεις επαρκούν για χρήσιμη βραχυπρόθεσμη πρόβλεψη, ιδίως στο διάστημα 1-12 ωρών, αλλά δεν μπορούν να αντικαταστήσουν την πληροφορία που παρέχουν πιο σύνθετες πηγές (NWP, reanalysis, multi-station) για μεγαλύτερους χρονικούς ορίζοντες.

Η συνεισφορά της εργασίας δεν περιορίζεται στα αριθμητικά αποτελέσματα. Η υλοποίηση του λογισμικού παρέχει ένα επαναλήψιμο πειραματικό πλαίσιο, το οποίο οργανώνει τη διαδικασία από τα πρωτογενή δεδομένα έως την αξιολόγηση και την αρχειοθέτηση των αποτελεσμάτων. Η χρήση χρονολογικών splits, train-only scaling, MetDataPy-based preparation και structured artifacts ενισχύει τη μεθοδολογική αξιοπιστία της προσέγγισης.

Συνολικά, η εργασία προσφέρει μια τεκμηριωμένη απάντηση στο αρχικό ερευνητικό ερώτημα. Ένας μόνο τοπικός μετεωρολογικός σταθμός μπορεί να υποστηρίξει χρήσιμη βραχυπρόθεσμη πρόβλεψη θερμοκρασίας, ιδιαίτερα σε ορίζοντες έως 12 ώρες, εφόσον τα δεδομένα

προετοιμαστούν σωστά και αξιολογηθούν με αυστηρή μεθοδολογία. Ταυτόχρονα, τα αποτελέσματα δείχνουν με σαφήνεια τα όρια της προσέγγισης και δημιουργούν βάση για μελλοντικές επεκτάσεις σε περισσότερες μεταβλητές, περισσότερους σταθμούς και πιο ολοκληρωμένα συστήματα πρόγνωσης.

Βιβλιογραφία

- Abdulla, N., Demirci, M., & Ozdemir, S. (2022). Design and evaluation of adaptive deep learning models for weather forecasting. *Engineering Applications of Artificial Intelligence*, 116, 105440. doi:10.1016/j.engappai.2022.105440
- Alerskans, E., Nyborg, J., Birk, M., & Kaas, E. (2022). A transformer neural network for predicting near-surface temperature. *Meteorological Applications*, 29(5), e2098. doi:10.1002/met.2098
- Allen, A., Markou, S., Tebbutt, W., et al. (2025). End-to-end data-driven weather prediction. *Nature*, 641, 1172–1179. doi:10.1038/s41586-025-08897-0
- Bi, K., Xie, L., Zhang, H., et al. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619, 533–538. doi:10.1038/s41586-023-06185-3
- Dos Santos, R. P., Matos-Carvalho, J. P., & Leithardt, V. R. Q. (2025). Deep learning in time series forecasting with transformer models and RNNs. *PeerJ Computer Science*, 11, e3001. doi:10.7717/peerj-cs.3001
- Fang, W., Xue, Q., Shen, L., & Sheng, V. S. (2021). Survey on the application of deep learning in extreme weather prediction. *Atmosphere*, 12(6), 661. doi:10.3390/atmos12060661
- Hewage, P., Behera, A., Trovati, M., Pereira, E., Ghahremani, M., Palmieri, F., & Liu, Y. (2020). Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station. *Soft Computing*, 24(20), 16453–16482. doi:10.1007/s00500-020-04954-0
- Kartas, K. (2026). MetDataPy: A Python library for reproducible preparation of meteorological time-series data. *SSRN*. doi:10.2139/ssrn.6523260
- Kochkov, D., Yuval, J., Langmore, I., et al. (2024). Neural general circulation models for weather and climate. *Nature*, 632, 1060–1066. doi:10.1038/s41586-024-07744-y
- Lam, R., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, 382, 1416–1421. doi:10.1126/science.adi2336
- Li, B., & Qian, Y. (2024). *Weather prediction using CNN-LSTM for time series analysis: A case study on Delhi temperature data*. arXiv:2409.09414 [cs.LG]. doi:10.48550/arXiv.2409.09414
- Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194), 20200209. doi:10.1098/rsta.2020.0209
- Lu, G., Ou, Y., Wang, Z., Qu, Y., Xia, Y., Tang, D., Kotenko, I., & Li, W. (2025). A survey of deep learning for time series forecasting: Theories, datasets, and state-of-the-art techniques. *Computers, Materials & Continua*, 85(2), 2403–2441. doi:10.32604/cmc.2025.068024
- Nguyen, T., Shah, R., Bansal, H., Arcomano, T., Maulik, R., Kotamarthi, V., Foster, I., Madireddy, S., & Grover, A. (2024). Scaling transformer neural networks for skillful and reliable medium-range weather forecasting. *arXiv:2312.03876*. doi:10.48550/arXiv.2312.03876

- Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., Hassanzadeh, P., Kashinath, K., & Anandkumar, A. (2022). FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. *arXiv:2202.11214*. doi:10.48550/arXiv.2202.11214
- Price, I., Sanchez-Gonzalez, A., Alet, F., et al. (2025). Probabilistic weather forecasting with machine learning. *Nature*, 637, 84–90. doi:10.1038/s41586-024-08252-9
- Rahman, M. S., Tumpa, F. A., Islam, M. S., Al Arabi, A., Hossain, M. S. B., & Haque, M. S. U. (2024). Comparative evaluation of weather forecasting using machine learning models. *arXiv preprint arXiv:2402.01206*. doi:10.48550/arXiv.2402.01206
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W.-C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *arXiv preprint arXiv:1506.04214*. doi:10.48550/arXiv.1506.04214
- Sun, J., Qian, Y., Sun, Y., et al. (2026). A single-station hourly multi-step forecasting model for sea fog based on ResNet and LSTM networks. *Journal of the Meteorological Society of Japan*, 104(1). doi:10.1007/s44394-025-00004-1
- Xu, P., Zheng, X., Gao, T., et al. (2025). An artificial intelligence-based limited area model for forecasting of surface meteorological variables. *Communications Earth & Environment*, 6, 372. doi:10.1038/s43247-025-02347-5
- Zhang, H., Liu, Y., Zhang, C., & Li, N. (2025). Machine learning methods for weather forecasting: A survey. *Atmosphere*, 16(1), 82. doi:10.3390/atmos16010082

Παράρτημα Α – Αναπαραγωγή Πειραματικής Εκτέλεσης

Το παρόν παράρτημα συνοψίζει τις βασικές εντολές εγκατάστασης, ελέγχου και εκτέλεσης του πειραματικού pipeline που χρησιμοποιήθηκε για την παραγωγή των αποτελεσμάτων του Κεφαλαίου 4.

Repository: <https://github.com/kkartas/local-weather-forecasting-ml>

Βασική εγκατάσταση:

```
python -m venv .venv
.\.venv\Scripts\Activate.ps1
python -m pip install --upgrade pip
pip install -r requirements.txt
pip install -e .
```

Έλεγχος tests:

```
python -m pytest
```

Κύρια πειραματική εκτέλεση:

```
python -m weather_forecasting_pipeline run-all --config configs/default.yaml
```

Επιμέρους στάδια:

```
python -m weather_forecasting_pipeline ingest --config configs/default.yaml
```

```
python -m weather_forecasting_pipeline preprocess --config configs/default.yaml
```

```
python -m weather_forecasting_pipeline train --config configs/default.yaml
```

```
python -m weather_forecasting_pipeline evaluate --config configs/default.yaml
```