



ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Γεωχωρική Ανάλυση και Μοντελοποίηση Κινδύνου
Αγροτοδοασικών Πυρκαγιών μέσω Τεχνητής Νοημοσύνης**



Του φοιτητή:
Ιωάννη Δεληβέη
Αριθμός Μητρώου: 133989

Επιβλέπων:
Παναγιώτης Αδαμίδης
Καθηγητής

Θεσσαλονίκη, Ιούνιος 2026

Τίτλος Δ.Ε.: Γεωχωρική Ανάλυση και Μοντελοποίηση Κινδύνου Αγροτοδοασικών Πυρκαγιών
μέσω Τεχνητής Νοημοσύνης

Κωδικός Δ.Ε.: 22859

Ονοματεπώνυμο φοιτητή: Ιωάννης Δεληβέης

Ονοματεπώνυμο εισηγητή: Παναγιώτης Αδαμίδης

Ημερομηνία ανάληψης Δ.Ε.: 13-03-2025

Ημερομηνία περάτωσης Δ.Ε.: 30-05-2026

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως πτυχιακή εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Ιωάννη Δεληβέη, που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της πτυχιακή εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος,

Αφιέρωση

"Αφιερώνω αυτή την εργασία στην μητέρα μου, που δεν πρόλαβε να με δει να την τελειώνω."

Ευχαριστίες

"Ευχαριστώ από καρδιάς τον πατέρα μου, Γιώργο, και την γυναίκα μου, Χρύσα, γιατί πίστεψαν σε εμένα μέχρι το τέλος και στάθηκαν δίπλα μου, σε όλες τις δύσκολες στιγμές.

Εκφράζω, επίσης, την ευγνωμοσύνη μου στον επιβλέποντα καθηγητή μου, κ. Αδαμίδη Παναγιώτη, για την καλή θέληση που έδειξε, την καθοδήγηση και την επιστημονική γνώση που προσέφερε. Χωρίς την βοήθειά του, η ολοκλήρωση αυτού του έργου δεν θα ήταν δυνατή."

Πρόλογος

Η απόφαση να μελετήσω συγκεκριμένο θέμα στηρίχθηκε σε τρία κίνητρα. Το πρώτο είναι το ισχυρό αίσθημα ευθύνης προς την κοινωνία. Από τότε που θυμάμαι, οι πυρκαγιές αποτελούν ένα ζήτημα μείζονος σημασίας για την Ελλάδα. Δεν πρόκειται απλά για μία αφηρημένη, μακρινή απειλή· είναι ένα θέμα προσωπικής ασφάλειας που ενδέχεται να αντιμετωπίσουμε όλοι, εφόσον βρεθούμε στην ύπαιθρο για κάποιο διάστημα. Φαίνεται δε, πως κάθε λίγα χρόνια, προκύπτει κάποιο σοβαρό επεισόδιο μεγάλης κλίμακας, που προκαλεί εκτεταμένες καταστροφές σε κοινότητες, στην πρωτογενή παραγωγή και στο οικοσύστημα. Αυτό που κάνει την κατάσταση ακόμα πιο ανησυχητική και επείγουσα, είναι οι τραγωδίες που έχουνε χαραχθεί στην πρόσφατη κοινή μνήμη. Μετρώντας 2 πολύνεκρες περιόδους πυρκαγιών τα τελευταία 20 χρόνια, τα καλοκαίρια θυμίζουν πλέον παιχνίδι ρωσικής ρουλέτας. Με μία διαφορά: ακόμα και οι άδαιοι κύλινδροι, επιφέρουν μεγάλο υλικό και περιβαλλοντικό τίμημα.

Το δεύτερο κίνητρο είναι το έντονο, έμφυτο προσωπικό ενδιαφέρον για το φυσικό περιβάλλον. Οι διεργασίες που διέπουν τα οικοσυστήματα και οι αλληλεπιδράσεις μεταξύ της βιολογίας και της γεωγραφίας, παρέχουν μία τεράστια ποικιλία από καταστάσεις, η κάθε μία από τις οποίες περιγράφει ένα μικρό τμήμα του κόσμου μας. Ακόμη και στην πιο καταστροφική τους εκδοχή, τα φυσικά φαινόμενα έχουν λεπτές πτυχές που αξίζει να αναλογιστούμε.

Τέλος, είχα την επιθυμία να εφαρμόσω την εμπειρία και τεχνογνωσία που αποκόμισα κατά την επαγγελματική και προσωπική μου πορεία: Αφενός με τη δημιουργία εργαλείων γεωπληροφορικής και τεχνητής νοημοσύνης σε επαγγελματικό επίπεδο, αφετέρου, με την διαίσθηση για τις αγροτοδοσικές πυρκαγιές, την οποία καλλιέργησα μέσω της οικογενειακής ενασχόλησης με την ελαιοκομία. Θεωρώ πως, με αυτά τα μέσα, είμαι σε θέση να συνεισφέρω στην έρευνα και την ανάπτυξη πάνω σε αυτόν τον κρίσιμο τομέα.

Abstract

Wildfires are a significant environmental hazard in Greece, where the mediterranean climate, local vegetation and human activity create favourable conditions for the outbreak and spread of fire. This thesis investigates the creation of a complete system, which would facilitate their analysis and modelling, with the aim of generating practical risk estimates.

Literature review indicated dry vegetation as a requirement for the outbreak of a wildfire, and human influences as the main cause of wildfires in Greece. In addition, an array of meteorological, geographical and social factors, that can act as indicators of the state of fuels or the probability of ignition, were identified.

A pilot for the system was designed and developed successfully for the Heraklion prefecture in Crete. Implemented functionality includes automated collection and processing of pyric factors, spatiotemporal analysis of historic fire events and conditions with Machine Learning algorithms and a live geospatial presentation of risk levels.

The chosen risk estimation model generates results 3,66 times better than the ones of a random classifier, with pronounced seasonal deviation. Although this proves the system's ability to capture certain statistical patterns in the study area, overall prediction precision remained low, across all tested models. This leaves the system in a prototype state, however, many paths for future research remain available.

Περίληψη

Οι αγροτοδοασικές πυρκαγιές είναι σημαντικός περιβαλλοντικός κίνδυνος στην Ελλάδα, όπου το μεσογειακό κλίμα, η τοπική βλάστηση και η ανθρώπινη δραστηριότητα δημιουργούν ευνοϊκές συνθήκες για την πρόκληση και τη διάδοση της φωτιάς. Αυτή η μελέτη πραγματεύεται την δημιουργία ενός ολοκληρωμένου συστήματος, που θα διευκολύνει την ανάλυση και την μοντελοποίησή τους, με τελικό στόχο την πρακτική εκτίμηση κινδύνου.

Η βιβλιογραφική έρευνα που έγινε κατέδειξε την ξερή βλάστηση ως προϋπόθεση για το ξέσπασμα αγροτοδοασικής πυρκαγιάς, και τις ανθρωπογενείς επιρροές ως την κύρια αιτία εκδήλωσής τους στην Ελλάδα. Επιπλέον, εντοπίστηκε ένα πλήθος μετεωρολογικών, γεωγραφικών και κοινωνικών παραγόντων που μπορούν να λειτουργήσουν ως ενδείξεις για την κατάσταση της καύσιμης ύλης και την πιθανότητα ανάφλεξης.

Το σύστημα σχεδιάστηκε και αναπτύχθηκε πιλοτικά για την Π.Ε Ηρακλείου με επιτυχία. Στις υλοποιημένες δυνατότητές περιλαμβάνεται η αυτόματη συλλογή και επεξεργασία των πυρικών παραγόντων, η χωροχρονική ανάλυσή των ιστορικών συμβάντων και συνθηκών με χρήση αλγόριθμων Μηχανικής Μάθησης, και η ζωντανή γεωχωρική παρουσίαση των επιπέδων κινδύνου.

Το επιλεγμένο μοντέλο εκτίμησης κινδύνου παράγει αποτελέσματα 3,66 φορές καλύτερα από αυτά ενός τυχαίου κατηγοριοποιητή, με έντονες εποχικές διακυμάνσεις. Αν και αυτό αποδεικνύει την ικανότητα του συστήματος να αποτυπώσει συγκεκριμένα στατιστικά μοτίβα στην περιοχή μελέτης, η ακρίβεια των προβλέψεων παρέμεινε χαμηλή, για όλα τα δοκιμασμένα μοντέλα. Για αυτόν τον λόγο, το σύστημα παραμένει σε πρωτότυπο στάδιο, αλλά με πολλές διαθέσιμες επιλογές για μελλοντική έρευνα.

Περιεχόμενα

1	Εισαγωγή	1
1.1	Φαινόμενο Πυρκαγιών	1
1.2	Διατύπωση Ζητήματος Μελέτης	1
1.3	Σκοπός Μελέτης	2
1.3.1	Ερευνητικός Στόχος	2
1.3.2	Τεχνικός Στόχος	3
1.4	Ερευνητικά Ερωτήματα	3
1.5	Σημασία της Μελέτης	3
1.6	Περιοχή Μελέτης	4
1.7	Μεθοδολογική Προσέγγιση	5
1.8	Πλαίσιο Μελέτης	5
1.9	Δομή Εργασίας	6
2	Θεωρητικό Υπόβαθρο	7
2.1	Επισκόπηση Κεφαλαίου	7
2.2	Θερμοδυναμική Δασικών Πυρκαγιών	7
2.3	Συνθήκες Εκδήλωσης Δασικών Πυρκαγιών	7
2.3.1	Κατάσταση Βλάστησης	7
2.3.2	Μετεωρολογικές Συνθήκες	8
2.3.3	Μορφολογία Εδάφους	9
2.4	Εναύσματα Δασικών Πυρκαγιών	9
2.4.1	Φυσικά Φαινόμενα	9
2.4.2	Ανθρώπινες Δραστηριότητες	9
2.5	Αλληλεπίδραση Βλάστησης-Καιρικών Συνθηκών	10
2.6	Ο Ανθρώπινος Παράγοντας στις Δασικές Πυρκαγιές	11
2.7	Μελέτες Δασικών Πυρκαγιών	12
2.8	Χαρακτηριστικά Δασικών Πυρκαγιών στην Ελλάδα	13
2.8.1	Χωρική Κατανομή Πυρκαγιών στην Ελλάδα	13
2.8.2	Χρονική Κατανομή Πυρκαγιών στην Ελλάδα	14
2.8.3	Αίτια Πυρκαγιών στην Ελλάδα	15
2.9	Μοντελοποίηση Πυρκαγιών	15
2.9.1	Προσεγγίσεις Μοντελοποίησης Πυρκαγιών	15
2.9.2	Ιδιαιτερότητες Εκτίμησης Κινδύνου	16
2.10	Σύνοψη Κεφαλαίου	16

3	Μεθοδολογία	17
3.1	Επισκόπηση Κεφαλαίου	17
3.2	Διαδικασία Έρευνας και Ανάπτυξης	17
3.2.1	Ροή Έρευνας και Εργασιών	17
3.3	Εντοπισμός και Διαλογή Δεδομένων	18
3.3.1	Αρχείο Δασικών Πυρκαγιών	18
3.3.2	Πρωτογενή Μετεωρολογικά Δεδομένα	19
3.3.3	Πηγές Γεωχωρικών Δεδομένων	20
3.3.4	Κριτήρια Καταλληλότητας Δεδομένων	21
3.4	Σχεδιασμός Βάσει Βιβλιογραφίας	21
3.4.1	Παράγοντες Κινδύνου Εκδήλωσης Πυρκαγιών στην Σχεδίαση και Ανάπτυξη	21
3.4.2	Επίδραση Σύγχρονων Προσεγγίσεων Εκτίμησης Κινδύνου Πυρκαγιών	21
3.4.3	Υιοθέτηση Υπαρχουσών Προσεγγίσεων για Κλίμακες Κινδύνου Πυρκαγιών	22
3.5	Προκαταρκτική Μελέτη Σκοπιμότητας	22
3.5.1	Μελέτη και Ανάλυση των Αρχείων Δεδομένων	22
3.5.2	Αρχική Διαμόρφωση Χαρακτηριστικών	25
3.5.3	Προκαταρκτικός Έλεγχος Απόδοσης Μοντέλων	26
3.5.4	Αρχικά Ευρήματα	28
3.6	Μεθοδολογία Διαχείρισης των Δεδομένων	28
3.6.1	Συλλογή και Αποθήκευση	28
3.6.2	Καθαρισμός και Εναρμόνιση	29
3.6.3	Χρονική και Χωρική Κβαντοποίηση	29
3.6.4	Χρονική και Χωρική Συσχέτιση	31
3.6.5	Δημιουργία Χαρακτηριστικών	31
3.6.6	Προετοιμασία Δεδομένων Εκπαίδευσης/Πρόβλεψης	33
3.7	Μοντελοποίηση Εκτίμησης Κινδύνου	34
3.7.1	Ορισμός Προβλήματος Πρόβλεψης	34
3.7.2	Μετρικές Αξιολόγησης	35
3.7.3	Στρατηγική Επιλογής Μοντέλων	37
3.7.4	Διαμόρφωση Διαδικασίας Εκπαίδευσης	37
3.7.5	Προσέγγιση Ρύθμισης Υπερπαραμέτρων	38
3.7.6	Ορισμός Κατωφλιού και Δείκτη Κινδύνου	39
3.8	Μεθοδολογία Αξιολόγησης Μοντέλων	40
3.9	Αρχές Σχεδίασης του Συστήματος	41
3.9.1	Αρθρωτότητα και Επεκτασιμότητα	41
3.9.2	Ευελιξία στην Παραμετροποίηση	42
3.9.3	Φορητότητα	42
3.9.4	Ζωντανή Εκτίμηση Κινδύνου	42
3.9.5	Συνεχής Βελτίωση	42
3.10	Μεθοδολογικοί Περιορισμοί και Αδυναμίες	43
3.10.1	Ελλειπείς Παράγοντες Εκδήλωσης Πυρκαγιών	43
3.10.2	Αδυναμίες Μεθόδου Μοντελοποίησης Πυρκαγιών	44
3.10.3	Προβλήματα Συνόλου Δεδομένων	44
3.10.4	Επιχειρησιακές Αδυναμίες	44
3.11	Σύνοψη Κεφαλαίου	45

4	Σχεδιασμός και Υλοποίηση	46
4.1	Επισκόπηση Κεφαλαίου	46
4.2	Απαιτήσεις Συστήματος	46
4.2.1	Πλαίσιο Υλοποίησης	46
4.2.2	Λειτουργικές Απαιτήσεις	47
4.2.3	Μη Λειτουργικές Απαιτήσεις	48
4.3	Αρχιτεκτονική Συστήματος	48
4.3.1	Περιγραφή Συστήματος	49
4.3.2	Επιμέρους Υποσυστήματα	49
4.3.3	Ροή Δεδομένων	49
4.4	Οντότητες Πεδίου Εφαρμογής	51
4.4.1	Κύριες Οντότητες	51
4.4.2	Αλληλεπιδράσεις Οντοτήτων	52
4.5	Πρωτόκολλα Συστήματος	53
4.5.1	Μονάδες Μέτρησης	53
4.5.2	Χωροχρονική Αποτύπωση	53
4.6	Σχεδίαση Βάσης Δεδομένων	54
4.6.1	Επιλογή Μοντέλου Μόνιμης Αποθήκευσης	54
4.6.2	Τεχνολογίες Βάσης Δεδομένων	54
4.6.3	Σχήμα Βάσης	54
4.6.4	Αρχειοθέτηση Δεδομένων	55
4.6.5	Βελτιστοποίηση Ερωτημάτων	55
4.6.6	Διεπαφή Βάσης Δεδομένων	56
4.7	Ροή Επεξεργασίας - ETL	57
4.7.1	Υποσύστημα Συλλογής Δεδομένων	57
4.7.2	Υποσύστημα Επεξεργασίας	58
4.8	Υποσύστημα Εκπαίδευσης	59
4.8.1	Παράμετροι Εκπαίδευσης	59
4.8.2	Δημιουργία Συνόλου Εκπαίδευσης	60
4.8.3	Ροή Εκπαίδευσης	60
4.8.4	Αξιολόγηση Μοντέλου	61
4.8.5	Αποθήκευση Μοντέλου	61
4.9	Υποσύστημα Πρόβλεψης	62
4.10	Εκτέλεση Εργασιών	62
4.10.1	Διαχείριση Ουράς Εργασιών	62
4.10.2	Παραγωγοί Εργασιών	62
4.10.3	Ροή Εκτέλεσης Εργασιών	63
4.10.4	Διαχείριση Σφαλμάτων	63
4.11	Διαδικτυακή Διεπαφή	64
4.11.1	Διαδικτυακές Τεχνολογίες	64
4.11.2	Υλοποίηση REST API	64
4.12	Διεπαφή Χρήστη	64
4.12.1	Οπτικοποίηση Κινδύνου	64
4.12.2	Τεχνολογίες Γεωχωρικής Απεικόνισης	66
4.13	Υποσύστημα Καταγραφής	66

4.14	Εγκατάσταση και Ανάπτυξη Συστήματος	67
4.14.1	Υπηρεσίες σε Κοντέινερ	67
4.14.2	Οργάνωση Ανάπτυξης με Docker	67
4.14.3	Παραμετροποίηση Περιβάλλοντος Υπηρεσιών	68
4.14.4	Διασύνδεση Υπηρεσιών	68
4.14.5	Βήματα Εγκατάσταση Συστήματος	68
4.15	Σενάρια Εκτέλεσης	69
4.16	Έλεγχος και Δοκιμές	69
4.16.1	Τοπικές Δοκιμές	69
4.16.2	Δοκιμές Ανάπτυξης Συστήματος	70
4.17	Περιορισμοί και Αδυναμίες Υλοποίησης	70
4.17.1	Περιορισμοί Συστήματος	70
4.17.2	Συμβιβασμοί στην Σχεδίαση	72
4.18	Μελλοντική Υλοποίηση	72
4.18.1	Υπολειπόμενη Λειτουργικότητα	72
4.18.2	Μελέτη Συστήματος Παραγωγής	75
4.19	Σύνοψη Κεφαλαίου	75
5	Αποτελέσματα	76
5.1	Επισκόπηση Κεφαλαίου	76
5.2	Μετρικές Ροής Εργασιών	76
5.3	Αποτίμηση Διαδικασίας Εκπαίδευσης	77
5.3.1	Περιγραφή Συνόλου Εκπαίδευσης	77
5.3.2	Μετρικές Εκπαίδευσης	77
5.4	Στατιστικά Εκδήλωσης Πυρκαγιών	78
5.4.1	Περιγραφικά Στατιστικά Συμβάντων	78
5.4.2	Στατιστικά Επικίνδυνων Συνθηκών	79
5.5	Αποτελέσματα Μοντέλων	86
5.5.1	Αποτελέσματα Συγκριτικών Δοκιμών Μοντέλων	86
5.5.2	Στατιστική Σημασία Χαρακτηριστικών	87
5.5.3	Αποτελέσματα Δοκιμών Παραμετροποίησης	88
5.5.4	Ανάλυση Προβλέψεων	89
6	Συζήτηση	92
6.1	Ερμηνεία Δεδομένων	92
6.1.1	Ερμηνεία Χωρικής και Χρονικής Κατανομής	92
6.1.2	Ερμηνεία Καιρικών Συνθηκών	93
6.2	Ανασκόπηση Πειραμάτων Μοντελοποίησης	95
6.2.1	Καταλληλότητα Αλγόριθμων	95
6.2.2	Επίδραση Παραμέτρων	95
6.2.3	Ερμηνεία Προβλέψεων Μοντέλων	95
6.2.4	Αποτίμηση Αποτελεσμάτων	97
6.3	Αποτίμηση Υλοποίησης	97
6.4	Σύγκριση με Άλλες Μελέτες	98

7	Συμπεράσματα	100
7.1	Αντίκτυπος της Μελέτης	100
7.2	Περιορισμοί	101
7.3	Μελλοντική Έρευνα	101
7.4	Επίλογος	102

Κατάλογος σχημάτων

1.1	Κύριοι Γεωγραφικοί Σχηματισμοί Περιοχής Μελέτης Υπόβαθρο: © 2017-2026 Αποκεντρωμένη Διοίκηση Κρήτης, Διεύθυνση Πληροφορικής και Επικοινωνιών. .	4
2.1	Φρύγανα, χαρακτηριστική χαμηλή βλάστηση της περιοχής μελέτης Photo: Krzysztof Ziarnik, Kenraiz, Phrygana near Elafonissi Beach, SW Crete, CC BY-SA 4.0.	8
2.2	Σύνολο Πυρκαγιών ανά Έτος (Κρήτη 2000 - 2025)	11
2.3	Σύνολο Καμμένης Έκτασης ανά Έτος (Κρήτη 2000 - 2025)	11
2.4	Κατανομή Μεγέθους - Συχνότητας Πυρκαγιών (Κρήτη 2000 - 2025)	12
2.5	Χρονικό Μεγάλων Πυρκαγιών (Κρήτη 2000 - 2025)	14
3.1	Γεωχωρική Κατανομή Πυρκαγιών (Κρήτη 2020 - 2025). Υπόβαθρο: © OpenStreetMap contributors, available under the Open Database License.	19
3.2	DEM Κρήτης από το δίκτυο Copernicus [50]	20
3.3	Μέσος Αριθμός Πυρκαγιών ανά Μήνα (Κρήτη 2000 - 2025)	23
3.4	Μέση Καμμένη Έκταση ανά Μήνα (Κρήτη 2000 - 2025)	24
3.5	Μέσος Υετός Ανά Σταθμό και Μήνα (Κρήτη 2015 - 2025)	24
3.6	Μέση Θερμοκρασία Ανά Σταθμό κατά την Δεκαετία 2015 - 2025	25
3.7	Συνολικός Αριθμός Πυρκαγιών ανά Ώρα Εκδήλωσης (Κρήτη 2000 - 2025)	26
3.8	Συνολικός Αριθμός Πυρκαγιών ανά Ημέρα της Εβδομάδας (Κρήτη 2000 - 2025) .	27
3.9	Συνολικός Αριθμός Πυρκαγιών ανά Μήνα και Ημέρα (Κρήτη 2000 - 2025)	27
3.10	Πλέγμα Διαχωρισμού Voronoi βάσει των Σταθμών της Περιοχής Μελέτης Υπόβαθρο: © OpenStreetMap contributors, available under the Open Database License. .	30
3.11	Πρότυπη Μοντελοποίηση Εκδήλωσης Πυρκαγιών	40
4.1	Επισκόπηση Υψηλού Επιπέδου	48
4.2	Αρχιτεκτονική Πρωτότυπου Συστήματος	50
4.3	Διάγραμμα Οντοτήτων-Σχέσεων	52
4.4	Οπτικοποίηση Εκτίμησης Κινδύνου με Γεωχωρική Ανάλυση 10χμ. Υπόβαθρο: © OpenStreetMap contributors, available under the Open Database License.	65
4.5	Οπτικοποίηση Εκτίμησης Κινδύνου με Γεωχωρική Ανάλυση 1χμ Υπόβαθρο: © OpenStreetMap contributors, available under the Open Database License.	71
5.1	Κατανομή Αριθμού Πυρκαγιών ανά Ώρα της Ημέρας (Περιοχή Μελέτης 2020 - 2025)	78
5.2	Κατανομή Αριθμού Πυρκαγιών ανά Ημέρα της Εβδομάδας (Περιοχή Μελέτης 2020 - 2025)	78

5.3	Κατανομή Αριθμού Πυρκαγιών ανά Μήνα και Ημερολογιακό Έτος (Περιοχή Μελέτης 2020 - 2025)	79
5.4	Κατανομή Αριθμού Πυρκαγιών ανά Τομέα (Περιοχή Μελέτης 2020 - 2025) Υπόβαθρο: © OpenStreetMap contributors, available under the Open Database License.	80
5.5	Κατανομή Τιμών Θερμοκρασίας (Περιοχή Μελέτης 2020 - 2025)	81
5.6	Κατανομή Τιμών Υγρασίας (Περιοχή Μελέτης 2020 - 2025)	82
5.7	Αναλογία Επικίνδυνων Στιγμιοτύπων ανά Επίπεδο Θερμοκρασίας και Μήνα (Περιοχή Μελέτης 2020 - 2025)	82
5.8	Αναλογία Επικίνδυνων Στιγμιοτύπων ανά Επίπεδο Υγρασίας και Μήνα (Περιοχή Μελέτης 2020 - 2025)	83
5.9	Αναλογία Επικίνδυνων Στιγμιοτύπων ανά Συνδυασμό Επιπέδου Υγρασίας και Θερμοκρασίας (Περιοχή Μελέτης 2020 - 2025)	83
5.10	Αναλογία Επικίνδυνων Στιγμιοτύπων ανά Συνδυασμό Επιπέδου Υγρασίας και Βροχόπτωσης Μήνα (Περιοχή Μελέτης 2020 - 2025)	84
5.11	Αναλογία Επικίνδυνων Στιγμιοτύπων ανά Συνδυασμό Επιπέδου Υγρασίας και Βροχόπτωσης Εξαμήνου (Περιοχή Μελέτης 2020 - 2025)	84
5.12	Αναλογία Επικίνδυνων Στιγμιοτύπων ανά Επίπεδο Ηλιακής Ακτινοβολίας και Περιόδου της Ημέρας (Περιοχή Μελέτης 2020 - 2025)	85
5.13	Αναλογία Επικίνδυνων Στιγμιοτύπων ανά Επίπεδο Ατμοσφαιρικής Πίεσης και Υψόμετρο (Περιοχή Μελέτης 2020 - 2025)	85
5.14	Κατανομή Εσφαλμένων Αρνητικών Προβλέψεων ανά επίπεδο Ηλιακής Ακτινοβολίας	90
5.15	Κατανομή Εσφαλμένων Αρνητικών Προβλέψεων ανά επίπεδο Θερμοκρασίας	91
5.16	Κατανομή Εσφαλμένων Αρνητικών Προβλέψεων ανά επίπεδο Υγρασίας	91
6.1	Χρήσεις Γής, Βλάστηση Υπόβαθρο: © © 2017-2026 Αποκεντρωμένη Διοίκηση Κρήτης, Διεύθυνση Πληροφορικής και Επικοινωνιών.	93
6.2	Photo: Freideriki Dionysopoulou / Wikimedia Commons, Burning olive branches after olive harvesting, CC BY-SA 3.0.	94

Κατάλογος πινάκων

4.1	Οντότητες Πεδίου Εφαρμογής για Πρωτότυπο Σύστημα	51
4.2	Μονάδες Μέτρησης Συστήματος	53
4.3	Διαθέσιμη Παραμετροποίηση Μοντέλων	60
4.4	Στάδια Υλοποίησης Υποσυστημάτων	73
5.1	Χαρακτηριστικά Πρόβλεψης	77
5.2	Αποτύπωμα Συνόλου Δεδομένων για Διαφορετικούς Τρόπους Διαχωρισμού	77
5.3	Χρόνος Διεκπεραίωσης (s) Βημάτων Εκπαίδευσης	77
5.4	Σύγκριση Μετρικών Απόδοσης Εξεταζόμενων Μοντέλων	86
5.5	Συντελεστές Διακύμανσης Μετρικών μη-Ντετερμινιστικών Αλγορίθμων	86
5.6	Βελτιστοποιημένες Παράμετροι Επιλεγμένου Αλγορίθμου XGB, Βιβλιοθήκη Python XGBoost	87
5.7	Πίνακας Σύγκρισης για κατηγοριοποίηση επικίνδυνων συνθηκών με XGB	87
5.8	Κέρδος Πληροφορίας Χαρακτηριστικών Αλγόριθμου XGB	88
5.9	Δοκιμές XGB με Διαταραχές	88
5.10	Δοκιμές XGB ανά Περίοδο του Έτους	89
5.11	Αποτελέσματα Επιπέδων Κινδύνου με XGB	89
5.12	Αποτελέσματα Επιπέδων Κινδύνου με XGB (Θερμοί Μήνες)	90

Συντομογραφίες

FMC	Fuel Moisture Content
DFMC	Dead Fuel Moisture Content
WUI	Wildland-Urban Interface
FWI	Fire Weather Index
ML	Machine Learning
SVM	Support Vector Machine
RF	Random Forest
LR	Logistic Regression
DL	Deep Learning
RNN	Recursive Neural Networks
CNN	Convolutional Neural Networks
DEM	Digital Elevation Model
PR-AUC	Precision-Recall Area Under the Curve
ROC-AUC	Area Under the Receiver Operating Characteristic Curve
XGB	Extreme Gradient Boosting / XGBoost
CV	Cross Validation
ETL	Extract Transform Load
API	Application Programming Interface
REST	Representational State Transfer
DBMS	Database Management System
SQL	Structured Query Language
MV	Materialized View
CRUD	Create Read Update Delete
CSV	Comma Separated Values
ORM	Object Relational Model
HTTP	Hypertext Transfer Protocol
URL	Unique Resource Locator
GIS	Geographic Information System
MVT	Mapbox Vector Tiles
VM	Virtual Machine
UI	User Interface
PF	Polynomial Features
TP	True Positive
FN	False Negative
FP	False Positive
Π.Ε	Περιφερειακή Ενότητα
ΒΔ	Βάση Δεδομένων

Κεφάλαιο 1

Εισαγωγή

1.1 Φαινόμενο Πυρκαγιών

Οι αγροτοδοσικές πυρκαγιές έχουν έντονες περιβαλλοντικές, κοινωνικές και οικονομικές επιπτώσεις. Στην Μεσόγειο, και κατ' επέκταση στην Ελλάδα, συγκαταλέγονται στις πιο κρίσιμες περιβαλλοντικές απειλές, καθότι συνιστούν άμεσο κίνδυνο για την ασφάλεια και την ευημερία των ανθρώπων. Η απώλεια περιουσίας και κατοικίας αγγίζουν το κοινό αίσθημα και τα συμβάντα πυρκαγιών καλύπτονται εκτενώς στα μέσα ενημέρωσης, ειδικά όταν εξαπλώνονται σε κατοικημένες περιοχές. Οι εξελίξεις αναφορικά με τέτοια περιστατικά ακολουθούνται στενά από την κοινωνία, που συμμερίζεται την αγωνία των πληγέντων. Τα παραπάνω πολλαπλασιάζονται, όταν η ίδια η ανθρώπινη ζωή βρίσκεται σε κίνδυνο. Εάν, τελικά, μία εκτός ελέγχου πυρκαγιά καταλήξει σε τραγωδία, θα λάβει, αναπόφευκτα, και πολιτικές διαστάσεις. Δυστυχώς, τα τελευταία χρόνια υπήρξε ένας αριθμός συμβάντων, που ταιριάζουν σε ανησυχητικό βαθμό με την παραπάνω περιγραφή και φέρνουν το φαινόμενο στο προσκήνιο της επικαιρότητας. Πλέον, η πρόληψη και αντιμετώπιση των ανεξέλεγκτων πυρκαγιών έχει εξελιχθεί σε μία πρόκληση μείζονος σημασίας για την πολιτεία [1][2][3][4].

Οι συνθήκες που διέπουν τα μεσογειακά περιβάλλοντα, ήταν ανέκαθεν ιδιαίτερα ευνοϊκές για τις πυρκαγιές: Τα ζεστά και ξηρά καλοκαίρια, οι περίοδοι έντονης ξηρασίας, η εύφλεκτη βλάστηση και οι περιοδικοί άνεμοι συντελούν στην εκδήλωση πυρκαγιών που, ενίοτε, τίθενται εκτός ελέγχου και εξαπλώνονται με ραγδαίους και απρόβλεπτους ρυθμούς [5][6]. Ταυτόχρονα, πολλές ανθρώπινες δραστηριότητες λαμβάνουν χώρα σε περιοχές που περικλείονται από άγρια βλάστηση και αυξάνουν σημαντικά τις πιθανότητες απρόβλεπτης ανάφλεξης. Η συνεχής αστικοποίηση σπρώχνει αυτές τις δραστηριότητες όλο και πιο κοντά σε άγριες εκτάσεις και, συγχρόνως, δημιουργεί ευάλωτα περιαστικά περιβάλλοντα, όπου οι αγροτοδοσικές πυρκαγιές αποβαίνουν καταστροφικές [7].

1.2 Διατύπωση Ζητήματος Μελέτης

Η χρονική κατανομή των αγροτοδοσικών πυρκαγιών δεν είναι ομοιόμορφη· ο όρος "εποχή πυρκαγιών" ή "αντιπυρική περίοδος" χρησιμοποιείται συχνά όταν γίνεται αναφορά στους μήνες του χρόνου, κατά τους οποίους το φαινόμενο είναι εντονότερο [8][9]. Το ίδιο ισχύει και για την χωρική κατανομή· κάποιες περιοχές αποτελούν "ζεστές ζώνες" και πλήττονται επανειλημμένα, ενώ άλλες όχι [2][10]. Αυτές οι διακυμάνσεις οφείλονται σε παράγοντες που χαρακτηρίζουν ειδικά τον

τόπο ή/και την περίοδο, όπως η κατάσταση της βλάστησης, οι ατμοσφαιρικές συνθήκες οι χρήσεις γης και η τοπολογία κλπ. [11].

Ιστορικά, η εκτίμηση κινδύνου γινόταν με εμπειρικές μεθόδους βασισμένες σε κανόνες προερχόμενους από γενίκευση ανεξάρτητων παραγόντων [12]. Καθώς, όμως, τα συστήματα που οδηγούν την εκδήλωση και διασπορά των πυρκαγιών είναι εξαιρετικά περίπλοκα, η πλήρης αποτύπωση των σχέσεων μεταξύ συνθηκών και κινδύνου παραμένει δύσκολη. Οι σύγχρονες μελέτες ερευνούν την αξιοποίηση πλήθους περιβαλλοντικών και ανθρωπογενών παραγόντων για την παρακολούθηση του φαινομένου των πυρκαγιών, με στόχο την εκτίμηση του κινδύνου κατά τόπους και ανά περίοδο [11][12].

Η μοντελοποίηση της πρόκλησης πυρκαγιών σε μία συγκεκριμένη γεωγραφική περιοχή μπορεί να γίνει αναλύοντας τα μοτίβα των ιστορικών περιστατικών, σε σχέση με τις συνθήκες που επικρατούσαν στο σημείο κατά το διάστημα εκδήλωσης. Τα συμπεράσματα της ανάλυσης μπορούν να προβληθούν στις τρέχουσες συνθήκες με έναν βαθμό αξιοπιστίας ανάλογο του βάθους της αποτύπωσης [13][11][12][14][4].

Υπάρχει η ανάγκη της καλύτερης κατανόησης της έκφρασης του φαινομένου σε διαφορετικές γεωγραφικές περιοχές, ώστε να είναι δυνατή η έγκυρη και έγκαιρη προειδοποίηση των τοπικών αρχών και των κατοίκων. Για αυτόν τον λόγο, απαιτούνται συστήματα, που μπορούν να φέρουν εις πέρας τις διαδικασίες που απαιτούνται για την επιτυχή μοντελοποίηση, από την αρχή, μέχρι το τέλος.

1.3 Σκοπός Μελέτης

Ο στόχος της παρούσας μελέτης είναι η πρωτοτύπηση, για μία συγκεκριμένη περιοχή μελέτης, ενός Συστήματος Ζωντανής Εκτίμησης Κινδύνου Εκδήλωσης Αγροτοδοσικών Πυρκαγιών: του Aristaeos. Η επίτευξη αυτού του στόχου μπορεί να χωριστεί σε ερευνητικό και τεχνικό σκέλος.

1.3.1 Ερευνητικός Στόχος

Η μελέτη της δυναμικής των πυρκαγιών είναι ένας κλάδος των περιβαλλοντικών επιστημών με έντονο διεπιστημονικό χαρακτήρα· στην περιγραφή του φαινομένου συμβάλλουν τα πεδία της φυσικής, της οικολογίας, της μετεωρολογίας και της κοινωνιολογίας. Δεδομένου ότι το ζήτημα της μελέτης είναι τόσο πολύπλευρο, κρίνεται σκόπιμο να αφιερωθεί ένα σημαντικό τμήμα της έρευνας στην δημιουργία μία σύγχρονης και επιστημονικά τεκμηριωμένης προσέγγισης μοντελοποίησης. Αυτό συνεπάγεται την εύρεση και εξέταση των συνθηκών που συντελούν στην εκδήλωση πυρκαγιών, δηλαδή των πυρικών παραγόντων. Η εξέταση θα γίνει σε χωροχρονικό επίπεδο, βάσει των πτυχών του ζητήματος, που διατυπώθηκαν στην ενότητα 1.2.

Επιπλέον, οι σημαντικές διακυμάνσεις στην κατά τόπους εκδήλωση των πυρκαγιών, ωθούν την μελέτη να γίνει τοπικοποιημένα. Πράγματι, συνηθίζεται οι έρευνες σε αυτόν τον τομέα να έχουν την μορφή εμπειρικής μελέτης συγκεκριμένης περίπτωσης, επικεντρωμένη σε συγκεκριμένη περιοχή [15] [16][13]. Επομένως, η ανάλυση επικεντρώνεται στην συσχέτιση του ιστορικού των πυρκαγιών με τις επικρατούσες συνθήκες, εντός της περιοχής ενδιαφέροντος.

Τελικός στόχος της έρευνας είναι η σύνθεση μίας μεθόδου μοντελοποίησης, που θα εκφράζει σε ικανοποιητικό βαθμό την δυναμική του φαινομένου, εντός της επιλεγμένης περιοχής.

1.3.2 Τεχνικός Στόχος

Η διερεύνηση των ζητημάτων της μοντελοποίησης εκτιμάται πως μπορεί να διευκολυνθεί σημαντικά, με την χρήση οργανωμένων πρωτοκόλλων ανάλυσης και τυποποιημένων ρουτίνων εκτέλεσης των εργασιών. Επιπλέον, η υιοθέτηση του Aristaeos από υποψήφιους χρήστες, όπως οργανισμούς και την τοπική αυτοδιοίκηση, θα διευκολυνθεί εφόσον πληρούνται συγκεκριμένες λειτουργικές προϋποθέσεις.

Για το πρακτικό κομμάτι της μελέτης απαιτείται η σχεδίαση και υλοποίηση μίας ολοκληρωμένης υποδομής έρευνας και ανάπτυξης. Αυτή η υποδομή, θα χρησιμοποιηθεί πάνω στην περιοχή μελέτης, για την διενέργεια των ερευνητικών πειραμάτων και, συνάμα, θα είναι σε θέση να αξιοποιήσει άμεσα τα αποτελέσματα, στην διαχείριση των πυρκαγιών σε πρακτικό επίπεδο.

1.4 Ερευνητικά Ερωτήματα

Τα θέματα προς διερεύνηση είναι τα εξής:

1. Η συλλογή διαθέσιμων ιστορικών αρχείων δεδομένων για πυρικούς παράγοντες και περιστατικά.
2. Η ανάλυση της χωρικής και χρονικής κατανομής των πυρκαγιών, καθώς και σχέσεών τους με τους επιλεγμένους παράγοντες κινδύνου.
3. Η ενσωμάτωση ετερογενών μετεωρολογικών, χωρικών και ανθρωπογενών μεταβλητών σε ένα συγκροτημένο σύνολο χαρακτηριστικών πρόβλεψης.
4. Η υλοποίηση ολοκληρωμένων πρωτοκόλλων για εκπαίδευση, αξιολόγηση και δοκιμή μοντέλων.
5. Η δημιουργία και σύγκριση απόδοσης επιλεγμένων μοντέλων εκτίμησης.
6. Η υλοποίηση ολοκληρωμένων πρωτοκόλλων για σχεδόν-ζωντανή συλλογή μετρήσεων, επεξεργασία δεδομένων και εκτίμηση επιπέδου κινδύνου.
7. Η ανάπτυξη ενός ενιαίου, πρακτικού και βιώσιμου συστήματος που θα υποστηρίζει την έγκαιρη ενημέρωση.

1.5 Σημασία της Μελέτης

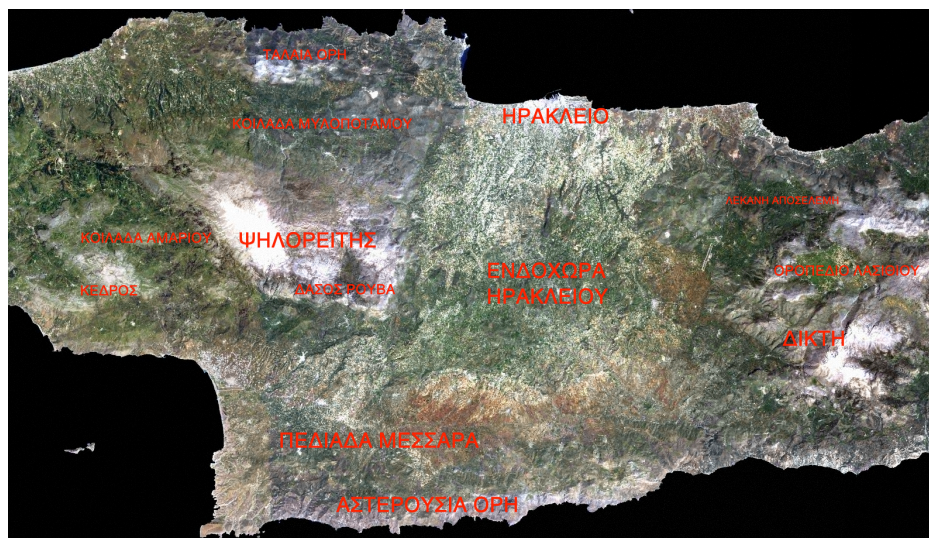
Η βιβλιογραφία υποδεικνύει δυσκολίες στην μοντελοποίηση των πυρκαγιών που χρήζουν διερεύνησης. Ένα ολοκληρωμένο σύστημα ανάλυσης θα επέτρεπε στους ερευνητές να αναπτύσσουν, να δοκιμάζουν και να εφαρμόζουν με μεγάλη ευχέρεια μοντέλα εκτίμησης κινδύνου για λογαριασμό τοπικών φορέων.

Η Ελλάδα, κατ' ανάγκη, συμμετέχει ενεργά στην έρευνα εκτίμησης κινδύνου πυρκαγιών. Εντούτοις, οι υπάρχουσες έρευνες στον ελληνικό χώρο βασίζονται κυρίως σε κανόνες και στατιστικές μεθόδους και έχουν πιο γενικό χαρακτήρα. Οι μελέτες πάνω σε συγκεκριμένες περιοχές, που ενσωματώνουν σύγχρονα εργαλεία της τεχνητής νοημοσύνης και παρέχουν συγκεκριμένα αριθμητικά αποτελέσματα, είναι ακόμα λίγες.

Έργα, όπως αυτά της ομάδας FLAME του FireWeather.eu [17], αποσκοπούν στην διερεύνηση της μοντελοποίησης των πυρκαγιών και την πρακτική εφαρμογή των αποτελεσμάτων σε συνθήκες πεδίου. Η επιτυχία τέτοιων συστημάτων σε συγκεκριμένες γεωγραφικές περιοχές μπορεί να δράσει ως πρότυπο για την ευρύτερη υιοθέτηση τους σαν επικουρικά εργαλεία στην πολιτική ασφαλείας.

Η προτεινόμενη πρωτοβουλία μπορεί να ενταχθεί σε ένα ευρύτερο τεχνολογικό οικοσύστημα στην υπηρεσία των πολιτών, προς την επίτευξη ενός απώτερου οράματος· την δημοκρατοποίηση της τεχνολογίας, της επιστήμης και της ασφάλειας, μέσω της διάχυσης τεχνολογικών μέσων, τα οποία σέβονται και εξυπηρετούν τους ανθρώπους.

1.6 Περιοχή Μελέτης



Σχήμα 1.1: Κύριοι Γεωγραφικοί Σχηματισμοί Περιοχής Μελέτης **Υπόβαθρο**: © 2017-2026 Αποκεντρωμένη Διοίκηση Κρήτης, Διεύθυνση Πληροφορικής και Επικοινωνιών.

Το Aristaeos θα αναπτυχθεί πιλοτικά για την έκταση, στο περίπου, της Π.Ε Ηρακλείου Κρήτης. Η γεωγραφία της περιοχής (Σχ. 1.1) κυριαρχείται από ένα λοφώδες τοπίο που καλύπτει το κεντρικό και βόρειο τμήμα, καθώς και από τις οροσειρές του Ψηλορείτη και της Δίκτης, στα δυτικά και ανατολικά του νομού αντίστοιχα. Ο Ψηλορείτης σχηματίζει κοιλάδες με τα Τάλαιο Όρη στον βορρά και το Όρος Κέδρος στον νότο, ενώ στην μέση της Δίκτης βρίσκεται το Οροπέδιο Λασιθίου. Στο νότιο τμήμα της περιοχής, βρίσκεται η πεδιάδα της Μεσσαράς που περικλείεται στον νότο από τα Αστερούσια όρη.

Όπως μπορεί να γίνει αντιληπτό, ανάμεσα στα βουνά και τους λόφους της περιοχής, επικρατεί μεγάλη ποικιλία γεωγραφικών στοιχείων, τα οποία συνθέτουν ένα εξαιρετικά ετερογενές ανάγλυφο. Το κλίμα που επικρατεί είναι μεσογειακό με θερμά και ξηρά καλοκαίρια, ήπιους χειμώνες και κατά τόπους εποχικές βροχοπτώσεις. Επιπλέον, η γεωγραφία της περιοχής αλληλεπιδρά με την ατμόσφαιρα, με αποτέλεσμα να προκύπτουν πολλά μικροκλίματα, χαρακτηριστικά στο κάθε μέρος. Η άγρια βλάστηση του νησιού είναι κυρίως χαμηλή (φρύγανα) και θαμνώδης [18][19].

Η αστική περιοχή του Ηρακλείου απλώνεται στο βόρειο τμήμα και αφήνει σημαντικό αποτύπωμα στον περίγυρό της. Διάσπαρτα χωριά βρίσκονται στα πεδινά και ημιορεινά μέρη της ενδοχώ-

ρας του νομού και των κοιλάδων. Το μεγαλύτερο τμήμα του εδάφους αξιοποιείται για καλλιέργεια ελίων και αμπελιών, ενώ οι λιγότερο γόνιμοι λόφοι χρησιμοποιούνται για βοσκή και κτηνοτροφία. Το τουριστικό ενδιαφέρον για τις δραστηριότητες, το φυσικό κάλος και τον πολιτισμικό και ιστορικό πλούτο της περιοχής είναι αυξημένο, με τις βόρειες ακτές να φιλοξενούν τα περισσότερα τουριστικά θέρετρα [19][20].

1.7 Μεθοδολογική Προσέγγιση

Η οργάνωση της μεθοδολογίας που ακολουθήθηκε, έγινε βάσει των συμπερασμάτων της βιβλιογραφικής ανασκόπησης που προηγήθηκε· οι συχνά αναφερόμενοι παράγοντες κινδύνου συλλέχθηκαν. Άλλα στοιχεία, που δεν ήταν άμεσα διαθέσιμα, προσεγγίστηκαν με μηχανική χαρακτηριστικών. Τα διαμορφωμένα χαρακτηριστικά πρόβλεψης καλύπτουν, άμεσα ή έμμεσα, ένα εύρος μετεωρολογικών, χωρικών και ανθρωπογενών συνθηκών.

Σε διάφορα στάδια της μελέτης πραγματοποιήθηκε στατιστική ανάλυση του ιστορικού αρχείου. Με αυτόν τον τρόπο, έγινε διασταύρωση των μοτίβων εκδήλωσης πυρκαγιών της περιοχής ενδιαφέροντος, με τα καλά κατανοητά μοτίβα που έχουν καταγραφεί σε προηγούμενες μελέτες στην ευρύτερη περιοχή.

Στο σύνολό της, η μεθοδολογία περιλαμβάνει την συλλογή δεδομένων, την επεξεργασία τους, την ανάλυσή τους, την δημιουργία μοντέλων πρόβλεψης με εκπαίδευση στα επεξεργασμένα δεδομένα, την αξιολόγηση της απόδοσης των παραγόμενων μοντέλων και την χρήση τους για την συνεχή εκτίμηση του επιπέδου κινδύνου.

1.8 Πλαίσιο Μελέτης

Αυτή η μελέτη επικεντρώνεται στην εκτίμηση της πιθανότητας εκδήλωσης πυρκαγιών, και δεν αποπειράται να αναλύσει την διάρκεια και την εξάπλωση τους με λεπτομέρεια. Η δεύτερη είναι μία ξεχωριστή προσέγγιση του προβλήματος, που έχει την δική της θέση στην διαχείριση των πυρκαγιών.

Επιπλέον, η έρευνα που διενεργήθηκε, περιορίζεται στην μοντελοποίηση των πυρκαγιών στην περιοχή μελέτης με τα διαθέσιμα δεδομένα. Αυτό σημαίνει πως τα συμπεράσματα για την αποτελεσματικότητα του έργου αφορούν αποκλειστικά αυτή την περιοχή και δεν θα έχει αποσαφηνιστεί η δυνατότητα γενίκευσής τους με το πέρασ αυτής της μελέτης.

Όπως θα φανεί και κατά την έρευνα, οι μηχανισμοί που προκαλούν τις πυρκαγιές είναι στοχαστικοί [21][22][12]· μία ανεξέλεγκτη φωτιά μπορεί να προκληθεί από τις μη γνώσιμες ενέργειες ενός μοναδικού ατόμου. Επίσης, δεν υπάρχουν δεδομένα που να περιγράφουν λεπτομερώς τις συνθήκες σε κάθε σημείο και χρονική στιγμή. Για αυτούς τους λόγους, το σύστημα κάνει προβλέψεις ενδεικτικές του επιπέδου κινδύνου και δεν "μαντεύει" που και πότε θα εκδηλωθεί η επόμενη πυρκαγιά.

1.9 Δομή Εργασίας

Η εργασία οργανώθηκε στα εξής κεφάλαια:

Κεφάλαιο 1, που παρουσιάζει το γενικότερο πλαίσιο και τον σκοπό της μελέτης.

Κεφάλαιο 2, στο οποίο συλλέγονται τα ευρήματα της βιβλιογραφικής έρευνας πάνω στις πυρκαγιές και τις μεθόδους μοντελοποίησης.

Κεφάλαιο 3, στο οποίο παρατίθεται η μεθοδολογική προσέγγιση διαχείρισης και επεξεργασίας των δεδομένων για την εξαγωγή των ζητούμενων αποτελεσμάτων.

Κεφάλαιο 4, στο οποίο περιγράφεται η δομή και οι τεχνικές λεπτομέρειες του υλοποιημένου συστήματος.

Κεφάλαιο 5, όπου αναγράφονται και σχολιάζονται τα αποτελέσματα της στατιστικής ανάλυσης και της μέτρησης των επιδόσεων του συστήματος και των παραγόμενων προβλέψεων.

Κεφάλαιο 6, όπου γίνεται απολογισμός των αποτελεσμάτων και προσπάθεια ερμηνείας τους.

Κεφάλαιο 7, που κλείνει την μελέτη με μία αποτίμηση του έργου στο σύνολο του.

Κεφάλαιο 2

Θεωρητικό Υπόβαθρο

2.1 Επισκόπηση Κεφαλαίου

Στο κεφάλαιο αυτό θα παρατεθεί η βιβλιογραφική ανασκόπηση που έγινε πάνω στο φαινόμενο των δασικών πυρκαγιών, ξεκινώντας από τις βασικές αρχές και τους κύριους παράγοντες εκδήλωσης και εξάπλωσής τους. Στη συνέχεια θα αναλυθούν οι πτυχές του φαινομένου, ειδικότερα στον μεσογειακό και σύγχρονο ελλαδικό χώρο. Το κεφάλαιο θα κλείσει με τις πρόσφατες μελέτες πάνω στην μοντελοποίηση των πυρκαγιών και την εκτίμηση πυρικής επικινδυνότητας.

2.2 Θερμοδυναμική Δασικών Πυρκαγιών

Τα οικοσυστήματα, όπως τα δάση και οι θαμνώδεις λόφοι, δρουν σαν τεράστιες αποθήκες ενέργειας, όπου η βλάστηση μετατρέπει την ηλιακή ακτινοβολία σε δυναμική χημική ενέργεια στους δεσμούς μεταξύ των θεμελιωδών δομικών στοιχείων των φυτών, όπως οι δεσμοί της κυτταρίνης ή της λιγνίνης [23]. Αυτή η αποθηκευμένη ενέργεια καταναλώνεται σταδιακά από χορτοφάγα ζώα ή μικροοργανισμούς, και διασκορπίζεται στο περιβάλλον. Όταν όμως οι συνθήκες είναι κατάλληλες, μπορεί να εκδηλωθεί καύση, μία αλυσιδωτή χημική αντίδραση οξειδωσης των υδρογονανθράκων, από τους οποίους αποτελείται η βλάστηση, με το οξυγόνο της ατμόσφαιρας. Αυτό έχει ως αποτέλεσμα την ταχύτατη απελευθέρωση μεγάλων ποσοτήτων της αποθηκευμένης ενέργειας μέσω θερμικής ακτινοβολίας, το φαινόμενο της πυρκαγιάς [21][22][4].

2.3 Συνθήκες Εκδήλωσης Δασικών Πυρκαγιών

2.3.1 Κατάσταση Βλάστησης

Καθώς η βλάστηση υπάρχει σε μεγάλη αφθονία στα περισσότερα οικοσυστήματα, αποτελεί και την κύρια καύσιμη ύλη για τις δασικές πυρκαγιές. Για να μπορεί να εξαπλωθεί η φωτιά και να εξελιχθεί σε πυρκαγιά, απαιτείται αφθονία και συνέχεια των καυσίμων, που συνήθως ταυτίζεται με τον όγκο της βλάστησης, σε συνδυασμό με τον τύπο του εδάφους [22].

Η υγρασία στην βλάστηση περιορίζει σε μεγάλο βαθμό την καύση, καθώς απαιτούνται μεγάλα ποσά ενέργειας για την εξάτμιση του νερού. Ως εκ τούτου, πολλές έρευνες θεωρούν την υγρασία της καύσιμης ύλης (*Fuel Moisture Content - FMC*), ως την πιο σημαντική ένδειξη, αναφο-

ρικά με το πόσο εύαλωτο σε πυρκαγιά είναι ένα οικοσύστημα, σε κάθε δεδομένη χρονική στιγμή [24][22][25][15][26][12].

Η διακίνηση της υγρασίας στο οικοσύστημα καθορίζεται σε μεγάλο βαθμό από την βιολογία των φυτών. Τα περισσότερα είδη, αποθηκεύουν νερό στους ιστούς τους όταν υπάρχει σε αφθονία, έτσι ώστε να μπορούν να συνεχίζουν να επιβιώνουν και να φωτοσυνθέτουν, ακόμα και όταν οι συνθήκες δεν είναι ευνοϊκές. Συνεπώς, το ποσοστό της υγρασίας στην βλάστηση εξαρτάται από το είδος του φυτού, αλλά και από τα αποθέματα που έχουν χρησιμοποιηθεί [27].

Ο τύπος της βλάστησης ενδέχεται να έχει διάφορα δομικά και χημικά χαρακτηριστικά που επηρεάζουν την συμπεριφορά των πυρκαγιών. Πολλά κωνοφόρα είδη, για παράδειγμα, εκλύουν εύφλεκτες ρητίνες που προάγουν την καύση και, συγχρόνως, λόγω των κώνων και των ξερών πευκοβελόνων που εναποθέτουν στο έδαφος, δημιουργούν συνέχεια καύσιμης ύλης, που διευκολύνει την εξάπλωση της φωτιάς [28].

Ο καίριος ρόλος που διαδραματίζουν οι πυρκαγιές στα μεσογειακά οικοσυστήματα φαίνεται από συγκεκριμένες εξελικτικές προσαρμογές που παρατηρούνται σε πολλά είδη βλάστησης της Μεσογείου. Η χαμηλή βλάστηση στα ξηρά κλίματα, όπως τα Φρύγανα (Φωτ. 2.1), έχει εξελιχθεί για να ευδοκιμεί στην ξηρασία, καίγεται εξαιρετικά εύκολα και ανακάμπτει γρήγορα από πυρκαγιές. Συχνά, μάλιστα, καταλαμβάνει καμένες εκτάσεις που πρότερα φιλοξενούσαν άλλου είδους βλάστηση, με αποτέλεσμα να αλλάζει τον τύπο του εδάφους, κάνοντας το ακόμα πιο εύφλεκτο [22]. Κάποια κωνοφόρα, δε, φύονται ευκολότερα έπειτα από δασικές πυρκαγιές (*βραδυχωρία*) [29].



Σχήμα 2.1: Φρύγανα, χαρακτηριστική χαμηλή βλάστηση της περιοχής μελέτης Photo: Krzysztof Ziarnik, Kenraiz, Phrygana near Elafonissi Beach, SW Crete, CC BY-SA 4.0.

2.3.2 Μετεωρολογικές Συνθήκες

Η θερμοκρασία του περιβάλλοντος αποτελεί σημαντική παράμετρο της διαδικασίας ανάφλεξης, καθώς επηρεάζει την θερμοκρασία των καυσίμων υλών. Όταν θερμαίνονται τα καύσιμα, το κατώφλι

της ενέργειας που απαιτείται για να ξεκινήσει η καύση μειώνεται, άρα αναφλέγονται πιο εύκολα [22][15].

Όταν ο αέρας είναι κορεσμένος με υγρασία, προάγεται η διατήρηση της υγρασίας στις καύσιμες ύλες και η ανάφλεξη είναι πολύ πιο δύσκολη. Η ατμοσφαιρική υγρασία αποτελεί σημαντική ένδειξη για το πόσο εύφλεκτο είναι το περιβάλλον [30][24].

Ο άνεμος, αφενός παρέχει μία συνεχή ροή οξυγόνου στην καύση, αφετέρου, στέλνει καυτό αέρα και φλεγόμενα υλικά στην γειτονική βλάστηση, με αποτέλεσμα να προκαλείται εκ νέου ανάφλεξη και να διαδίδεται η φωτιά. Κατά τις δραστηριότητες που περιλαμβάνουν φλόγες και διεξάγονται με την παρουσία ανέμου, ενδέχεται η φωτιά να τεθεί εκτός ελέγχου, μέσω της διάδοσης της σε προσκείμενους χώρους και αντικείμενα. Η ισχύς και διεύθυνση του ανέμου θεωρούνται οι πιο σημαντικοί παράγοντες διάδοσης μίας πυρκαγιάς και, συνεπώς, της έκτασης και ταχύτητας εξάπλωσης της και διαδραματίζουν σημαντικό ρόλο στην πρόκληση πυρκαγιών λόγω αμέλειας [16][22][24][30].

2.3.3 Μορφολογία Εδάφους

Το οξυγόνο είναι απαραίτητο για την καύση. Όσο αυξάνεται το υψόμετρο, η περιεκτικότητα της ατμόσφαιρας σε οξυγόνο μειώνεται, με αποτέλεσμα η δυνατότητα για ανάφλεξη και διατήρηση της καύσης να ελαττώνεται [15]. Επιπλέον, η κάλυψη βλάστησης τείνει να αραιώνει σε περιοχές που βρίσκονται σε μεγάλο ύψος [19]. Επομένως, με την άνοδο του υψόμετρου περιορίζονται οι πυρκαγιές.

Ο ζεστός αέρας έχει την τάση να ανεβαίνει. Εάν το έδαφος έχει κλίση, αυτό κάνει την διάδοση της φωτιάς πιο εύκολη από βλάστηση στο χαμηλό επίπεδο σε αυτήν στα ψηλότερα επίπεδα [15]. Η κλίση του εδάφους είναι συμπληρωματικός παράγοντας για την εξάπλωση της πυρκαγιάς.

2.4 Εναύσματα Δασικών Πυρκαγιών

Η βλάστηση αυτή καθαυτή, όσο ξερή και αν είναι, δεν μπορεί σε καμία περίπτωση, κάτω από συνήθεις συνθήκες, να φτάσει την θερμοκρασία ανάφλεξης (150-200 °C για λεπτή βλάστηση, και 280-500 °C για ξυλώδη υλικά), ώστε να προκληθεί αυτανάφλεξη και να ξεκινήσει καύση [31].

2.4.1 Φυσικά Φαινόμενα

Οι κεραυνοί μπορούν στιγμιαία να ανεβάσουν την θερμοκρασία σε ακραίες τιμές, και, εφόσον η εκκένωση γίνει κοντά σε καύσιμη ύλη, να προκαλέσουν ανάφλεξη. Οι στεγνοί κεραυνοί κατά τις ηλεκτρικές καταιγίδες, είναι η μόνη στατιστικά σημαντική αιτία πρόκλησης δασικών πυρκαγιών, έξω από την ανθρώπινη επιρροή, και είναι ο κυρίαρχος παράγοντας ανάφλεξης σε αχανείς και απομονωμένες περιοχές. Η ηφαιστειακή δραστηριότητα είναι μία ακόμη αιτία πρόκλησης πυρκαγιάς, αλλά η παρουσία της σε καθημερινό επίπεδο είναι μηδαμινή [5].

2.4.2 Ανθρώπινες Δραστηριότητες

Από την Νεολιθική εποχή και έπειτα, ο αντίκτυπος των ανθρώπινων ενεργειών στα δασικά οικοσυστήματα μεγάλωνε σταδιακά, με αποτέλεσμα την αυξημένη συχνότητα δασικών πυρκαγιών σε σχέση με το παρελθόν [5]. Οι ανθρωπογενείς πυρκαγιές μπορεί να είναι ακούσιες, ειδικά κατά την

διάρκεια αγροτικών εργασιών, όπως κάψιμο κλαδιών ή καλάμων. Μπορεί, επίσης, να οφείλονται σε αμέλεια, όπως κατά την διαχείριση φωτιάς για μαγειρική και διασκέδαση, από πυροτεχνήματα ή λόγω ρίψης τσιγάρων κλπ. Στην περίπτωση που η πυρκαγιά προκλήθηκε εκούσια, πρόκειται για εμπρησμό, ο οποίος εκτελείται με οικονομικά κίνητρα σαν τον καθαρισμό εκτάσεων, από δόλο ή ακόμα και λόγω ψυχολογικών διαταραχών [32][5].

Η ανθρώπινη δραστηριότητα μπορεί να προκαλέσει πυρκαγιές και με έμμεσο τρόπο. Σε περίπτωση βλάβης ηλεκτρολογικού ή μηχανικού εξοπλισμού, ενδέχεται να εκδηλωθούν θερμοί σπινθήρες, που δύνανται να προκαλέσουν ανάφλεξη. Σε πολύ σπάνιες περιπτώσεις, γυάλινα απορρίμματα μπορεί να επικεντρώσουν και να συσσωρεύσουν την ηλιακή ακτινοβολία με τέτοιο τρόπο, ώστε η θερμοκρασία να ανεβεί σε επικίνδυνα επίπεδα. Το ίδιο μπορεί να συμβεί και με συγκεκριμένα οργανικά απορρίμματα, λόγω των υψηλών θερμοκρασιών και των χημικών που απελευθερώνουν ορισμένες διεργασίες ζύμωσης των μικροοργανισμών [32].

2.5 Αλληλεπίδραση Βλάστησης-Καιρικών Συνθηκών

Οι μετεωρολογικές συνθήκες που επικρατούν, σε συνάρτηση με τον χρόνο, έχουν επίδραση στην βλάστηση. Λόγω της εξάρτησης των φυτών από το διαθέσιμο νερό, ο πρόσφατος υετός είναι καθοριστικός για την ανάπτυξη και την υγεία της χλωρίδας [33][27]. Η υγρασία, η θερμοκρασία, ο άνεμος και η ηλιακή ακτινοβολία, επίσης μεταβάλλουν σταδιακά την κατάσταση της βλάστησης, με τρόπο που επηρεάζει τον βαθμό ευφλεκτότητάς της [30][24].

Τα φυτά καταναλώνουν τα αποθέματα υγρασίας τους, κατά την διάρκεια των περιόδων με χαμηλό υετό, με αποτέλεσμα ο δείκτης FMC να πέφτει. Αυτό καθιστά τις παρατεταμένες ξηρασίες, όπως το μεσογειακό καλοκαίρι, εξαιρετικά πρόσφορες για εκδήλωση πυρκαγιών [30][24]. Διάφορα είδη βλάστησης αντιδρούν διαφορετικά στην ξηρασία: Τα μεγάλα δέντρα έχουν πρόσβαση σε βαθύτερα υπόγεια ύδατα, ενώ τα χόρτα παραδίδονται στην ξηρασία και εξασφαλίζουν την συνέχεια τους με σπόρους [33][27].

Η νεκρή βλάστηση αποτελείται από μονοετή χόρτα και πόες, πολυετή φυτά που έχουν κλείσει τον κύκλο ζωής τους, αλλά και τμήματα φυτών που έχουν απορριφθεί, όπως νεκρά φύλλα και κλαδιά. Η υγρασία της νεκρής ύλης έχει ραγδαίες διακυμάνσεις, ανάλογα με τις συνθήκες του περιβάλλοντος· σε συνθήκες ξηρασίας, αυτά τα υλικά αποξηραίνονται απότομα. Το ποσοστό υγρασίας της νεκρής βλάστησης (*Dead Fuel Moisture Content - DFMC*) είναι σημαντικός δείκτης της ευαισθησίας του οικοσυστήματος σε πυρκαγιές [34][24][12].

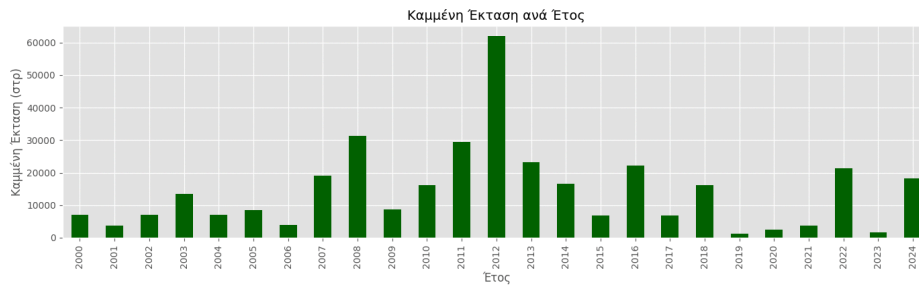
Είναι φυσικό επακόλουθο ότι οι καιρικές συνθήκες των προηγούμενων εβδομάδων, και ειδικά η ανομβρία, συμβάλλουν στην εκδήλωση και την διάδοση δασικών πυρκαγιών.

Σε βάθος μηνών, ακόμα και μία ήπια, αλλά παρατεταμένη έλλειψη βροχοπτώσεων μπορεί να ρίξει τους δείκτες FMC και DFMC σε κρίσιμα χαμηλά επίπεδα. Παρομοίως, ένας συνδυασμός μίας περιόδου έντονων βροχοπτώσεων, που ακολουθείται από μία έντονη ξηρασία, δημιουργεί ευνοϊκές συνθήκες για μεγάλες πυρκαγιές, καθώς, η βλάστηση που αναπτύσσεται την περίοδο των βροχών, ξεραίνεται απότομα [24][22][15].

Οι κλιματικές τάσεις του τελευταίου αιώνα δείχνουν άνοδο της μέσης θερμοκρασίας και μεταβολή των μοτίβων βροχοπτώσεων και αυτό αναμένεται να επιμηκύνει τις περιόδους υψηλού κινδύνου για πυρκαγιές [25][35].



Σχήμα 2.2: Σύνολο Πυρκαγιών ανά Έτος (Κρήτη 2000 - 2025)



Σχήμα 2.3: Σύνολο Καμμένης Έκτασης ανά Έτος (Κρήτη 2000 - 2025)

2.6 Ο Ανθρώπινος Παράγοντας στις Δασικές Πυρκαγιές

Αν και είναι αδύνατον να προβλεφθούν επακριβώς οι λεπτομερείς κινήσεις των ατόμων, υπάρχουν κάποιες μετρικές που υποδηλώνουν αυξημένη κινητικότητα και, επομένως, αυξημένες πιθανότητες εκδήλωσης πυρκαγιάς. Οι περιοχές που είναι εύκολα προσβάσιμες ή βρίσκονται κοντά σε κάποια κοινότητα είναι πιθανότερο να συγκεντρώνουν άτομα και να φιλοξενούν την ανθρώπινη δραστηριότητα που ενδέχεται να προκαλέσει ανάφλεξη. Επομένως, η απόσταση από τον δρόμο και η απόσταση από οικισμό μπορούν να αξιοποιηθούν σαν έμμεσος δείκτης της πιθανής κινητικότητας [36].

Στο ίδιο μήκος, η χρήση γης και η κάλυψη του εδάφους, πέρα από την πυκνότητα και το είδος των καύσιμων υλών, υποδηλώνουν την συχνότητα και το επίπεδο των αλληλεπιδράσεων μεταξύ των ανθρώπων και του περιβάλλοντος [7]. Οι περισσότερες υπαίθριες δραστηριότητες τείνουν να αυξομειώνονται ανάλογα με την ώρα της ημέρας και την περίοδο του χρόνου, διότι συνηθίζεται να προτιμώνται συγκεκριμένες συνθήκες για την εκτέλεση συγκεκριμένων εργασιών, παρομοίως και για εξορμήσεις με σκοπό την ψυχαγωγία [32][5].

Η ζώνη γύρω από αστικά κέντρα, όπου τα οικοσυστήματα με υψηλή συγκέντρωση καύσιμης ύλης εφάπτονται χώρων ανθρώπινης δραστηριότητας, συχνά αποκαλείται Δασική-Αστική Διεπαφή (*Wilderness-Urban Interface - WUI*) στην μελέτη των πυρκαγιών. Αυτές οι ζώνες, αφενός αντιμετωπίζουν πυρκαγιές με μεγάλη συχνότητα και, αφετέρου, συνθέτουν ένα περιβάλλον υψηλού κινδύνου για τους ανθρώπους που κατοικούν και δραστηριοποιούνται σε αυτές [37][7][5]. Οι επιπτώσεις μίας μεμονωμένης πυρκαγιάς που τίθεται εκτός ελέγχου σε περιοχή WUI ενδέχεται να είναι δραματικές, όπως έγινε στην Αττική το 2018 [3][10].

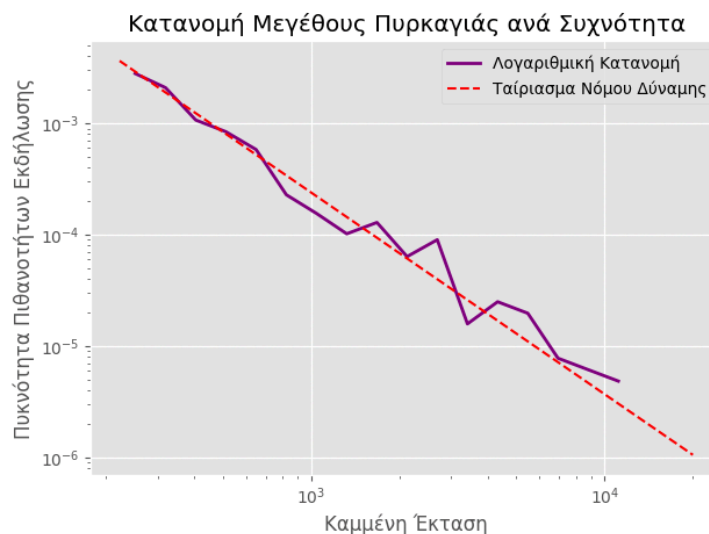
Στην ύπαιθρο, καθώς και ένα μεγάλο μέρος της WUI, οι αγροτικές, δασοκομικές και κτηνοτροφικές δραστηριότητες, όπως η βοσκή χορτοφάγων ζώων, συμβάλλουν στον έλεγχο της καύσιμης

ύλης. Οι μεταβολές στην χρήση γης και η εγκατάλειψη της υπαίθρου και των επαγγελματικών ενασχολήσεων σε αυτήν, εντείνουν την συγκέντρωση και συνέχεια της βλάστησης, ειδικά όταν δεν υπάρχει μέριμνα για την διαχείριση της με εναλλακτικούς τρόπους [38][5][2]. Αυτοί οι κοινωνικο-οικονομικοί παράγοντες αυξάνουν την συχνότητα και την έκταση των πυρκαγιών σε περιοχές που είναι ήδη ιδιαίτερα ευάλωτες.

Η ανάδειξη των δασικών πυρκαγιών σαν μάστιγα των ημιαστικών περιοχών, καθώς και η ανάπτυξη των περιβαλλοντικών επιστημών, προήγαγαν την υπεύθυνη διαχείριση των καύσιμων υλών στις ευαίσθητες περιοχές, οδηγώντας σε θεσμοθέτηση δράσεων αποτροπής: Χάραγμα αντιπυρικών ζωνών, καθαρισμό της βλάστησης, απαγόρευση εργασιών με φωτιά και καύση κατά τις ημέρες υψηλού ρίσκου, ακόμα και πρόκληση εσκεμμένων ελεγχόμενων πυρκαγιών για κατανάλωση της καύσιμης ύλης [1][4][25]. Οι πολιτικές πρόληψης και το επίπεδο ευαισθητοποίησης και ενημέρωσης ανά περιοχή επηρεάζουν τα μοτίβα εκδήλωσης πυρκαγιών.

2.7 Μελέτες Δασικών Πυρκαγιών

Οι μηχανισμοί που διέπουν τα δάση και τις δασικές πυρκαγιές, ταιριάζουν με τα μοτίβα που διέπουν την θεωρία της Αυτοοργανωμένης Κρισιμότητας, ως συστήματα που συγκεντρώνουν ενέργεια και τείνουν προς μία εύθραυστη ισορροπία. Σε αυτήν την κατάσταση, ένα μικρο συμβάν μπορεί να επιφέρει δραστική απελευθέρωση ενέργειας. Αυτό οδηγεί σε εκδήλωση μεγάλων πυρκαγιών ανά διαστήματα, ενώ η κατανομή της καμένης έκτασης ως προς την συχνότητα εκδήλωσης ακολουθεί νόμο δύναμης, με τις μεγάλες σε έκταση πυρκαγιές να είναι σπανιότερες (Σχ. 2.4).



Σχήμα 2.4: Κατανομή Μεγέθους - Συχνότητας Πυρκαγιών (Κρήτη 2000 - 2025)

Σε αυτήν την προσέγγιση, οι Andreas Y. Troumbis κα. διεξήγαγαν στατιστική ανάλυση, που συνιστά ότι η αστικοποίηση, η μεταβολή της χρήσης γης, οι κλιματικές ανωμαλίες και τα εκλογικά έτη στην Ελλάδα εξηγούν σε μεγάλο βαθμό την κατά καιρούς απόκλιση των στατιστικών για τις πυρκαγιές. Τα αποτελέσματα αποκαλύπτουν μοτίβα σε βάθος ετών καθώς και βαθύτερες κοινωνι-

κές, οικονομικές και πολιτικές πτυχές του φαινομένου, με την ομάδα να καταλήγει σε κριτική της πολιτικής κατάστασης στην Ελλάδα ως σημαντικό παράγοντα επιδείνωσης του φαινομένου [4].

Τον τελευταίο αιώνα στην Ελλάδα, η ετήσια μέση καμένη έκταση έχει τετραπλασιαστεί [6] και πολλές μελέτες προβλέπουν περαιτέρω αύξηση των επικινδυνών ημερών, της συχνότητας των πυρκαγιών και της καμένης έκτασης, μέχρι το 2100, σε όλες τις περιοχές της Ελλάδας [25][35].

Οι οικονομικές επιπτώσεις των δασικών πυρκαγιών στην Ελλάδα είναι δύσκολο να αποτιμηθούν με ένα ενιαίο μέγεθος, επειδή οι διαθέσιμες εκτιμήσεις συχνά συνδυάζουν άμεσες και έμμεσες ζημιές ή εντάσσουν τις πυρκαγιές σε ευρύτερα ακραία κλιματικά επεισόδια. Ενδεικτικά, σύμφωνα με την Eurostat, η οικονομική ζημιά των ακραίων κλιματικών επεισοδίων στην Ελλάδα, κυμαίνεται στα 50€ κατά κεφαλήν το έτος. Ένα σημαντικό ποσό από αυτά αναλογεί στις δασικές πυρκαγιές, με κόστος που ενδέχεται να ανέρχεται σε εκατοντάδες εκατομμύρια [39][40].

Λαμβάνοντας υπ' όψιν τις πολιτικές και οικονομικές επιπτώσεις του ζητήματος, είναι αναμενόμενο πως η δημιουργία αποτελεσματικών κλιμάκων επικινδυνότητας αποτελεί κύριο στόχο πολυάριθμων μελετών· τελικός σκοπός, η μείωση των ανεξέλεγκτων πυρκαγιών μέσω της έγκαιρης προετοιμασίας και ενημέρωσης και η βελτίωση της ασφάλειας των περιοχών που είναι ευάλωτες στις δασικές πυρκαγιές.

Το FWI (fire weather index) είναι ένας δείκτης επικινδυνότητας πυρκαγιάς, σε κλίμακα από πολύ χαμηλό έως ακραίο κίνδυνο, που βασίζεται αποκλειστικά στις πρόσφατες καιρικές συνθήκες. Έχει αναπτυχθεί στην Βόρεια Αμερική και εφαρμόζεται σε διάφορες χώρες της Ευρώπης και παγκοσμίως [25][20]. Προηγούμενες μελέτες έδειξαν πως έχει περιορισμένη αποτελεσματικότητα στις περιοχές της Ελλάδας, ενώ άλλες μελέτες προτείνουν μετατροπές που θα το καταστήσουν κατάλληλο για της συνθήκες του Ελληνικού χώρου [30][41][17]. Παραδείγματα πρόσφατων μελετών πάνω σε μεθόδους παραγωγής χαρτών επικινδυνότητας με την χρήση μετεωρολογικών, τοπογραφικών και δημογραφικών δεδομένων, όπως η γεωχωρική ανάλυση μεγάλων πυρκαγιών των τελευταίων ετών στην Χίο των Nektaria Adaktylou κα, δίνουν αισιόδοξα σημάδια για την επιχειρησιακή δυνατότητα συστημάτων εκτίμησης κινδύνου [15].

2.8 Χαρακτηριστικά Δασικών Πυρκαγιών στην Ελλάδα

2.8.1 Χωρική Κατανομή Πυρκαγιών στην Ελλάδα

Οι πυρκαγιές στην Ελλάδα δεν έχουν ομοιόμορφη χωρική κατανομή, αλλά τείνουν να συγκεντρώνονται σε συγκεκριμένες ευάλωτες περιοχές. Οι Paragiannaki κα. χρησιμοποιώντας το αρχείο του πυροσβεστικού σώματος, διαχώρισαν 34.120 συμβάντα δασικών πυρκαγιών για την περίοδο 2000 - 2016, με την Ηλία, την Λακωνία και την Εύβοια να κατέχουν το 28% της καμένης έκτασης. Αυτά τα στοιχεία υποδεικνύουν πως οι τάσεις των πυρκαγιών στην Ελλάδα διαμορφώνονται από τις τοπικές ιδιομορφίες στην βλάστηση, τα μικροκλίματα και την χρήση της γης [41].

Τα στοιχεία για την Αττική υποστηρίζουν αυτήν την ερμηνεία: Πολύχρονες καταγραφές για την περίοδο 1977 - 2024 δείχνουν επανειλημμένες πυρκαγιές, με το 45% της καμένης έκτασης να αποτελείται από πευκοδάσος, και το 18.5% από θαμνώδεις φρυγανικές εκτάσεις [10]. Αυτές οι παρατηρήσεις συνθέτουν ένα μωσαϊκό περιφερειακών περιοχών και μεσογειακής βλάστησης που είναι ιδιαίτερα επιρρεπές στις πυρκαγιές λόγω της ανάμειξης των φυσικών τάσεων και των ανθρωπογενών πιέσεων.

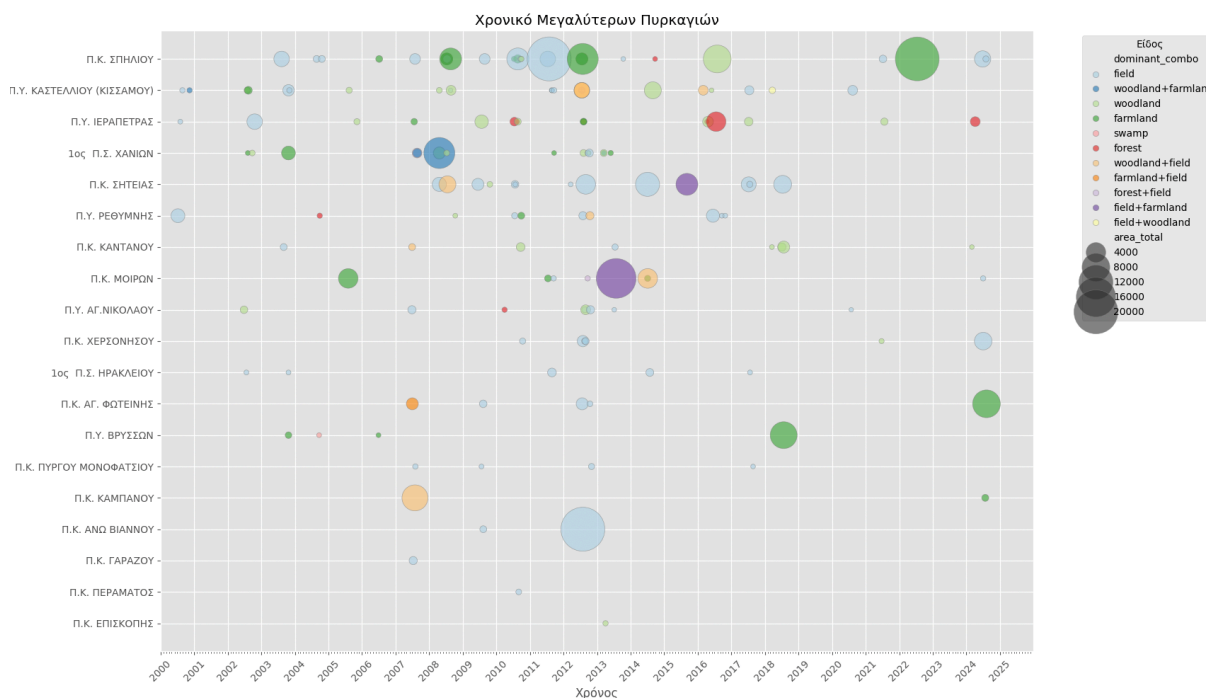
2.8.2 Χρονική Κατανομή Πυρκαγιών στην Ελλάδα

Η χρονική κατανομή των δασικών πυρκαγιών στην Ελλάδα είναι στενά συνδεδεμένη με την ξηρασία και την σύμπραξη καιρικών συνθηκών που υποστηρίζουν τις πυρκαγιές. Οι Dimitrakopoulos κα. έδειξαν πως, κατά την περίοδο 1961 - 1997, το πλήθος και το μέγεθος των πυρκαγιών είχαν θετική συσχέτιση με την ξηρασία, ενώ οι παρατεταμένες ξηρασίες, όπως αυτές μετά το 1978, έχουν ισχυρότερη επίδραση. Τα αποτελέσματα αυτής της μελέτης δείχνουν τον αντίκτυπο της ξηρασίας στην συχνότητα εκδήλωσης και την έκταση των δασικών πυρκαγιών, ακόμα και σε περιοχές που, σε άλλες περιόδους, δεν παρουσιάζουν έντονη τάση [9].

Μελέτες πάνω στις τραυματικές πυρκαγιές του 2007 στην Πελοπόννησο ερμήνευσαν την εξαιρετικά δύσκολη χρονιά ως αποτέλεσμα της μεγάλης συγκέντρωσης καύσιμων υλών, σε συνδυασμό με τις καιρικές συνθήκες, οι οποίες αποδείχθηκαν πρόσφορες για την εξάπλωση των πυρκαγιών [2].

Η μετεωρολογική ανάλυση των Papavasileiou και Giannaros με Μηχανική Μάθηση (Self-Organizing Maps) υποστηρίζει το μοντέλο των ακραίων καιρικών επεισοδίων. Τα πορίσματα περιλαμβάνουν την κατάδειξη 5 διακριτών καιρικών μοτίβων που σχετίζονται με την εκδήλωση των πυρκαγιών. Από αυτά, 3 δείχνουν ξεκάθαρη σύνδεση μεταξύ της έντασης των περιστατικών και ασυνήθιστων μοτίβων υψηλής θερμοκρασίας, ξηρασίας, ισχυρού ανέμου και ατμοσφαιρικής αστάθειας. Τα άλλα 2 φανερώνουν αρνητική σχέση μεταξύ της εκδήλωσης των πυρκαγιών και την ύπαρξη υψηλής υγρασίας, νηνεμίας, χαμηλής ατμοσφαιρικής πίεσης και χαμηλών θερμοκρασιών [42].

Στην Ελλάδα, τα πυρρικά φαινόμενα μπορούν να γίνουν αντιληπτά ως ένα σύστημα, όπου η παρατεταμένη ξηρασία, οι υψηλές θερμοκρασίες, και η διαθέσιμη καύσιμη ύλη συνωμοτούν, ώστε οι αναφλέξεις να εξελίσσονται σε εκτενείς και καταστρεπτικές πυρκαγιές.



Σχήμα 2.5: Χρονικό Μεγάλων Πυρκαγιών (Κρήτη 2000 - 2025)

2.8.3 Αίτια Πυρκαγιών στην Ελλάδα

Σχετικά με τα αίτια των πυρκαγιών στην Ελλάδα, οι έρευνες, όπως η επισκόπηση ειδικών στις δασικές πυρκαγιές των Fantina Tedim κα., αλλά και τα διαθέσιμα στατιστικά στοιχεία, φανερώνουν ξεκάθαρα την κυριαρχία των ανθρωπογενών παραγόντων. Συγκεκριμένα, η αμέλεια ακολουθούμενη από τον εμπρησμό εκτιμάται ότι ευθύνονται για την συντριπτική πλειοψηφία των συμβάντων, έως και 95% σύμφωνα με κάποιες εκτιμήσεις, ενώ οι κεραυνοί και άλλες αιτίες, όπως οι βλάβες ηλεκτρολογικού εξοπλισμού, έχουν συγκριτικά μικρή μερίδα [43][32].

Οι αστικές διεπαφές και η εγγύς ύπαιθρος στην Ελλάδα, είναι χώροι όπου η αλληλεπίδραση με δυνητικά καύσιμες ύλες είναι συχνές. Αυτές οι αλληλεπιδράσεις, ενίοτε περιλαμβάνουν καταστάσεις όπου είναι εύκολο η φωτιά να τεθεί εκτός ελέγχου, λόγω της ζέστης και της ξηρασίας, αλλά και της παραμελημένης και παρατημένης βλάστησης [7].

2.9 Μοντελοποίηση Πυρκαγιών

Στις πρόσφατες μελέτες, το πρόβλημα της μοντελοποίησης των δασικών πυρκαγιών έχει δύο βασικές παραλλαγές, την εκτίμηση πιθανότητας εκδήλωσης πυρκαγιάς, όπου ο στόχος είναι ο εντοπισμός έναρξης πυρκαγιάς εντός ενός ορίζοντα πρόβλεψης, και την εκτίμηση ευαλωτότητας, όπου ο στόχος εξετάζει εάν μία περιοχή επηρεάστηκε ή όχι από δασική πυρκαγιά.

2.9.1 Προσεγγίσεις Μοντελοποίησης Πυρκαγιών

Σε μεθοδολογικό επίπεδο, η βιβλιογραφική ανασκόπηση των Xu κα. παρουσιάζει το φάσμα των προσεγγίσεων ως εξής: Τα μοντέλα βασισμένα σε συνδυασμό εμπειρικών στοιχείων με μαθηματικές και φυσικές μεθόδους έχουν την πιο μακροχρόνια ιστορία χρήσης, αλλά και περιορισμένη αποτελεσματικότητα. Τα στατιστικά μοντέλα, τα οποία κάνουν διερεύνηση σχέσεων μεταξύ των δασικών πυρκαγιών και των παραγόντων κινδύνου με παλινδρόμηση, παρέχουν ερμηνεύσιμα αποτελέσματα, ωστόσο, δυσκολεύονται να ανταποκριθούν στις μη γραμμικές σχέσεις που διέπουν το φαινόμενο. Στην στοχαστικότητα των συστημάτων εκδήλωσης πυρκαγιών ανταποκρίνονται καλύτερα τα μοντέλα Μηχανικής Μάθησης (*Machine Learning - ML*), όπως οι αλγόριθμοι Τυχαίου Δάσους (*Random Forest - RF*), Μηχανές Υποστήριξης Διανυσμάτων (*Support Vector Machines - SVM*) και Μπεζιανές (*Bayesian*) μέθοδοι, αποτυπώνοντας τις μη γραμμικές σχέσεις από ιστορικά δεδομένα [12][11].

Η καταλληλότητα των μεθόδων ML υποστηρίζεται και πειραματικά: οι Mishra κα. χρησιμοποιούν 19 γεωγραφικούς, μετεωρολογικούς, βιοφυσικούς και ανθρωπογενείς παράγοντες πρόβλεψης και κάνουν σύγκριση μεταξύ SVM και RF σε γεωχωρική εκτίμηση ευαλωτότητας σε πυρκαγιά, με τον RF να παρουσιάζει καλύτερη απόδοση [44]. Παρομοίως, οι Moghim και Mehrabi συγκρίνουν την Λογιστική Παλινδρόμηση (*Logistic Regression - LR*) με τον RF για την εκτίμηση ευαλωτότητας σε δύο ανόμοιες περιοχές με την χρήση 11 καιρικών, τοπολογικών και γεωγραφικών χαρακτηριστικών και βρίσκουν τον RF να υπερέχει [45]. Αντίθετα, οι Gholami κα., που επιλέγουν να διαμορφώσουν το πρόβλημα σαν ερώτημα εκτίμησης προγνωστικών εκδήλωσης πυρκαγιών, πάνω σε ένα πλέγμα κελιών πλευράς 500 μ, βάσει ιστορικών στοιχείων προηγούμενων ετών, βρίσκουν τον LR να υπερέχει σε σχέση με πιο περίπλοκα μη γραμμικά μοντέλα [13].

Τα μοντέλα Βαθειας Μάθησης (*Deep Learning - DL*) έχουν την δυνατότητα εξειδίκευσης σε πεδία σχετικά με την πρόγνωση πυρκαγιών όπως οι καιρικές χρονοσειρές, η κατηγοριοποίηση δο-

ρυφορικών εικόνων και η χωροχρονική ανάλυση εκδήλωσης πυρκαγιών. Με αυτά τα δομημένα δεδομένα, οι αλγόριθμοι DL μπορούν να εξάγουν αυτόματα τους συντελεστές πρόκλησης πυρκαγιών. Η χρήση της DL στην πρόβλεψη πυρκαγιών είναι η πιο τρέχουσα τάση στο αντικείμενο και έχει πολλές ανοιχτές προκλήσεις, όπως η αποτελεσματική σχεδίαση των δεδομένων, η δυσκολία κλιμάκωσης των μοντέλων και το πρωτόκολλο αξιολόγησης [12].

2.9.2 Ιδιαιτερότητες Εκτίμησης Κινδύνου

Καθώς οι πυρκαγιές εκδηλώνονται σποραδικά, η μοντελοποίησή τους είναι μία μορφή πρόβλεψης σπάνιου γεγονότος ή ανωμαλίας, με αποτέλεσμα τα σύνολα δεδομένων που προκύπτουν να έχουν πολύ μεγάλη ανισορροπία σήματος φωτιάς - μη φωτιάς [12]. Για αυτόν τον λόγο, πολλές φορές απαιτούνται τεχνικές εξισορρόπησης κατά την εκπαίδευση των μοντέλων, με σκοπό την αντιστάθμιση της ανισορροπίας σήματος. Η χρήση αυτών των τεχνικών, έχει ως παρενέργεια την μείωση της κατανοητότητας και χρησιμότητας των πιθανοτήτων εκτίμησης που δίνουν τα μοντέλα ως αποτέλεσμα. Πολλές μελέτες επιλέγουν να επανορθώσουν για αυτό μεταφράζοντας τις πιθανότητες σε ένα διατεταγμένο σύνολο κλάσεων ή κλίμακα κινδύνου, με σκοπό να παρέχουν ερμηνεύσιμα αποτελέσματα, κοντά στις επιχειρησιακές ανάγκες που επιτάσσουν οι πυρκαγιές [15][41]. Η μετατροπή από πιθανότητα σε κατηγορία επικινδυνότητας γίνεται με συσταδοποίηση, όπως η μέθοδος Jenks, κατώφλια, ή άλλες τεχνικές διακριτοποίησης [12].

2.10 Σύνοψη Κεφαλαίου

Η βαθύτερη κατανόηση των μηχανισμών που διέπουν την εύθραυστη ισορροπία των οικοσυστημάτων και του ρόλου που διαδραματίζουν οι πυρκαγιές σε αυτά, έχει οδηγήσει σε αλλαγή της φιλοσοφίας αντιμετώπισης των πυρκαγιών, από την ένθερμη αντιδραστική καταστολή, σε μελετημένη πρόληψη μέσω της παρακολούθησης και διαχείρισης των καύσιμων υλών. Σε αυτό το κλίμα, η έρευνα πάνω στην εκτίμηση κινδύνου έρχεται στο προσκήνιο, με την μελέτη των πυρικών αίτιων και μοτίβων να αποσκοπεί στην ελαχιστοποίηση του αντίκτυπου των πυρκαγιών στους ανθρώπους μέσω της έγκαιρης και έγκυρης προειδοποίησης.

Κεφάλαιο 3

Μεθοδολογία

3.1 Επισκόπηση Κεφαλαίου

Το παρόν κεφάλαιο περιγράφει την Μεθοδολογία που ακολουθήθηκε για την δημιουργία και την αξιολόγηση του προτεινόμενου Συστήματος Εκτίμησης Κινδύνου Δασικών Πυρκαγιών. Παρακάτω θα αναλυθούν με λεπτομέρεια τα βήματα διερεύνησης και υλοποίησης, η σειρά με την οποία ακολουθήθηκαν, πως το κάθε στάδιο επηρέασε τα επόμενα, αλλά και πως οι αρχές και οι στόχοι της μελέτης καθόρισαν την πορεία της.

Μεταξύ άλλων, θα παρουσιαστεί η προσπάθεια εύρεσης κατάλληλων δεδομένων, οι περιορισμοί που τέθηκαν από αυτά, η διαχείριση τους για την εξαγωγή πληροφορίας, τα επίκεντρα της βιβλιογραφικής έρευνας και πως τα ευρήματα που προέκυψαν συνέβαλλαν στην σχεδίαση του συστήματος, και, τέλος, η στρατηγική μοντελοποίησης του προβλήματος, συμπεριλαμβανομένων των μεθόδων αξιολόγησης της απόδοσης.

3.2 Διαδικασία Έρευνας και Ανάπτυξης

3.2.1 Ροή Έρευνας και Εργασιών

Η σειρά με την οποία ακολουθήθηκαν τα βήματα εκπόνησης της παρούσας μελέτης, διαμορφώθηκε από την δύσκολη πραγματικότητα που έχει να αντιμετωπίσει στην Ελλάδα ένα ανεξάρτητο έργο με επίκεντρο τα δεδομένα. Η εύρεση ταιριαστών και ποιοτικών συνόλων δεδομένων που να ανταποκρίνονται στο αντικείμενο και την εμβέλεια του έργου ήταν υψίστης σημασίας και η έλλειψη των κατάλληλων πόρων σε αυτόν τον τομέα θα ήταν ένα σημαντικό κώλυμα για το εγχείρημα. Ως εκ τούτου, πριν καν προταθεί προς διερεύνηση το θέμα της μελέτης, διασφαλίστηκε ότι τα προαπαιτούμενα για την μελέτη σύνολα δεδομένων υπάρχουν και είναι διαθέσιμα για αυτόν τον σκοπό. Έτσι, τέθηκαν ξεκάθαρα πλαίσια όσον αφορά την περιοχή και το βάθος της ανάλυσης, από τα ίδια τα δεδομένα που ήταν δυνατό να αποκτηθούν για τους σκοπούς του έργου.

Τα επόμενα στάδια οργανώθηκαν με τέτοιο τρόπο, ώστε κάθε βήμα να επωφελείται από ό,τι χτίστηκε στα προηγούμενα. Η βιβλιογραφική έρευνα πραγματοποιήθηκε στα αρχικά στάδια της εκπόνησης, με σκοπό την διαμόρφωση ισχυρού θεωρητικού υπόβαθρου, πάνω στο οποίο κινήθηκε η υλοποίηση του συστήματος. Η αναζήτηση στην βιβλιογραφία επικεντρώθηκε, αλλά δεν περιορίστηκε, σε περιοχές με παρόμοιο κλίμα και γεωγραφία σε σχέση με την περιοχής μελέτης. Ιδιαίτερη

έμφαση δόθηκε σε συστήματα που αξιοποιούν δεδομένα σαν αυτά που συλλέχθηκαν για τους σκοπούς της μελέτης.

Πριν την υλοποίηση, έγινε μία προκαταρκτική μελέτη σκοπιμότητας, βασισμένη στα συμπεράσματα που κατέδειξε η διερεύνηση της βιβλιογραφίας. Αρχικά, τα δεδομένα που συλλέχθηκαν πέρασαν από μία σύντομη διαδικασία ανάλυσης και οπτικοποίησης, ώστε να επιτευχθεί ένας επαρκής βαθμός κατανόησης και εξοικείωσης με την περιοχή ενδιαφέροντος και τα μοτίβα που την διέπουν. Σε αυτό το σημείο, έγιναν αντιληπτά τα δυνατά και τα αδύνατα σημεία των διαθέσιμων στοιχείων και εξετάστηκαν τρόποι αξιοποίησής τους, ώστε να προκύψει ένα ξεκάθαρο πλάνο διαχείρισης και επεξεργασίας. Ακολούθησε μία σειρά από δοκιμές αλγόριθμων ML, με σκοπό την εξέταση της καταλληλότητας των δεδομένων, την αξιολόγηση διαφόρων προσεγγίσεων μηχανικής χαρακτηριστικών και την εκτίμηση της βιωσιμότητας του έργου υπό τους υπάρχοντες περιορισμούς.

Η διαδικασία της υλοποίησης, η οποία ήταν και το πιο μεγάλο σε διάρκεια τμήμα της μελέτης, αποσκοπούσε στην ανέγερση της κατάλληλης υποδομής για να πλακισωθούν τα μοντέλα εκτίμησης πυρικού κινδύνου. Αφενός, λοιπόν, η υλοποίηση ενός ολοκληρωμένου συστήματος αποτελεί το κύριο μέλημα της παρούσας μελέτης. Αφετέρου, εφόσον ολοκληρωθεί σε ικανοποιητικό βαθμό μία τέτοια υποδομή, αναμένεται να διευκολύνει, να επιταχύνει και να τυποποιήσει τις διεργασίες συλλογής και επεξεργασίας δεδομένων, δημιουργίας και αξιολόγησης μοντέλων και εξαγωγής προβλέψεων σε σχεδόν ζωντανό ρυθμό.

Έχοντας έτοιμο τον σκελετό του Aristaeos, το τελικό στάδιο ήταν η επιστροφή στον πυρήνα του συστήματος και η χρήση της υποδομής για την συστηματική αξιολόγηση της απόδοσης επιλεγμένων αλγορίθμων, βάσει μετρικών, ώστε να επιτύχουν τα καλύτερα δυνατά αποτελέσματα για τα υπάρχοντα δεδομένα.

Λαμβάνοντας υπ' όψιν την ευστοχία των μοντέλων εκτίμησης στις δοκιμές και την γενικότερη πρακτικότητα του συστήματος, είναι δυνατό να κριθεί εάν είναι σκόπιμο να προχωρήσει το έργο σε επόμενο στάδιο. Και τον μακροπρόθεσμο στόχο ενσωμάτωσης σε ένα ευρύτερο οικοσύστημα εργαλείων προστασίας και ενημέρωσης.

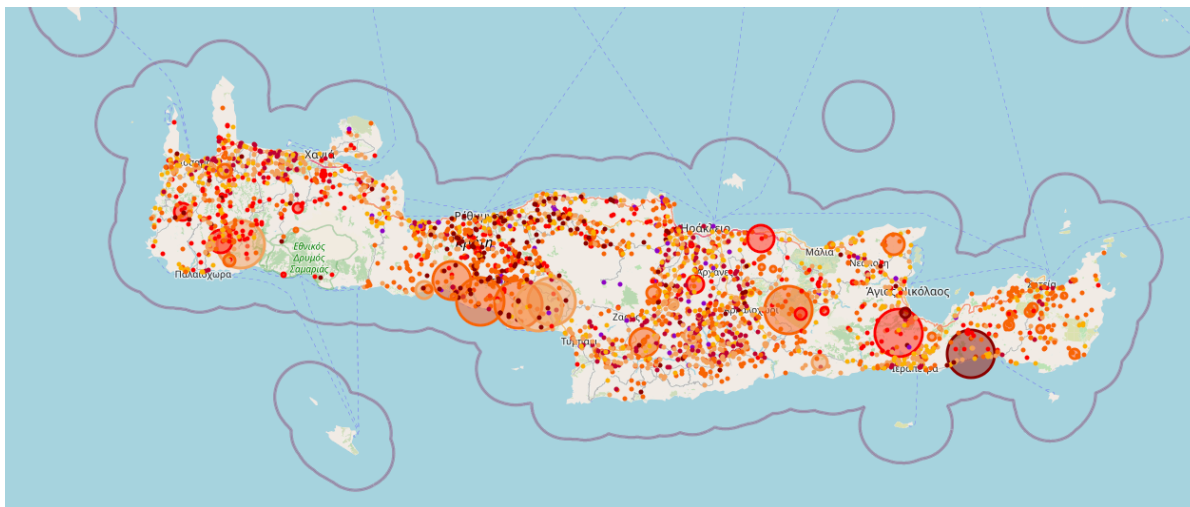
3.3 Εντοπισμός και Διαλογή Δεδομένων

3.3.1 Αρχείο Δασικών Πυρκαγιών

Εφόσον ο στόχος των προβλέψεων του συστήματος αφορά τα συμβάντα πυρκαγιών, ο πρώτος και πιο σημαντικός πόρος που πρέπει να διασφαλιστεί είναι ένα αρχείο καταγραφής συμβάντων δασικών πυρκαγιών. Ελλείψει αυτού, η μοντελοποίηση του προβλήματος με χρήση ML είναι αδύνατη.

Το αρχείο αγροτοδασικών πυρκαγιών, που διατηρεί και έχει δημοσιεύσει το Πυροσβεστικό Σώμα Ελλάδος, παρέχει χρονικά και χωρικά στοιχεία, για όλες τις δασικές πυρκαγιές στις οποίες χρειάστηκε επέμβαση της πυροσβεστικής, από το 2000 έως και το 2024 και σε όλη την Ελληνική επικράτεια [46]. Περιλαμβάνεται χρόνος καταγραφής συμβάντος με ακρίβεια λεπτού για όλες τις εγγραφές, ενώ για αρκετές δίνεται και ώρα λήξης. Όσον αφορά την τοποθεσία, για τις πυρκαγιές από το 2020 και έπειτα, παρέχονται ακριβείς γεωχωρικές συντεταγμένες, ενώ ο προσδιορισμός τοποθεσιών των πυρκαγιών πριν το 2020 γίνεται μέσω οδού ή τοπωνυμίου, γεγονός που κάνει την χωρική ευθυγράμμιση δύσκολη.

Για κάθε πυρκαγιά παρέχεται η καμένη έκταση ανά τύπο κάλυψης εδάφους, ενώ για αυτές από



Σχήμα 3.1: Γεωχωρική Κατανομή Πυρκαγιών (Κρήτη 2020 - 2025). **Υπόβαθρο:**© OpenStreetMap contributors, available under the Open Database License.

το 2011 και έπειτα, έχουν καταγραφεί και τα πυροσβεστικά μέσα που κινητοποιήθηκαν για την κατάσβεσή τους. Το αρχείο στο σύνολο του περιέχει 251.150 καταγεγραμμένα συμβάντα πυρκαγιών και παρέχονται σε μορφή xls ανά χρονολογία.

3.3.2 Πρωτογενή Μετεωρολογικά Δεδομένα

Η αναζήτηση κατάλληλου αρχείου μετεωρολογικών δεδομένων συνάντησε πολλά εμπόδια. Οι καιρικές μετρήσεις θεωρούνται πόρος με οικονομική σημασία για διάφορους τομείς του εμπορίου και της παραγωγής και, επιπλέον, η συλλογή τους απαιτεί την εγκατάσταση, προστασία και συντήρηση εξειδικευμένου εξοπλισμού σε συγκεκριμένες τοποθεσίες. Συνεπώς, η εύρεση ελεύθερων δεδομένων που να ανταποκρίνονται στα κριτήρια και τις απαιτήσεις του έργου αποδείχθηκε δύσκολη.

Από τα διαθέσιμα ελληνικά μετεωρολογικά αρχεία που εξετάστηκαν, το ελεύθερο σύνολο καιρικών δεδομένων της Αποκεντρωμένης Διοίκησης Κρήτης κρίθηκε να έχει τον μεγαλύτερο βαθμό πληρότητας, σύμφωνα με τα κριτήρια αξιολόγησης που ορίστηκαν. Αφορά 11 επιστημονικούς μετεωρολογικούς σταθμούς σε διάφορα σημεία της Κρήτης. Οι σταθμοί κατέγραφαν ημωριαίες, ακριβείς μετρήσεις βασικών μετεωρολογικών τιμών, συμπεριλαμβανομένης της θερμοκρασίας του αέρα, της υγρασίας, του ανέμου, του υετού ημώρου, ηλιακής ακτινοβολίας, ρυθμού εξάτμισης και βαρομετρικής πίεσης [47]. Οι μετρήσεις καλύπτουν το διάστημα 2015 - 2025, δηλαδή μία περίοδο 10 ετών, με ελάχιστες βραχυχρόνιες διακοπές. Από τους 11 σταθμούς, επιλέχθηκαν 8, με τρόπο που να επιτυγχάνει την μεγαλύτερη δυνατή κάλυψη κάποιας γεωγραφικής περιοχής.

Η επιλεγμένη γεωγραφική περιοχή, έκτασης περίπου 4.000 χμ², η οποία εκτίνεται στο μεγαλύτερο μέρος του νομού Ηρακλείου και το ανατολικό ήμισυ του νομού Ρεθύμνου, ορίστηκε ως η περιοχή μελέτης.

Αυτή η διαλογή, αύξησε σημαντικά την πυκνότητα σταθμών μετρήσεων από το αραιό 1 σταθμό/768χμ², όπως προκύπτει από την επιφάνεια της Κρήτης (8.450χμ²) δια το πλήθος των σταθμών, σε 1 σταθμό/500χμ². Έτσι, καλύπτονται καλύτερα τα μικροκλίματα που σχηματίζει η

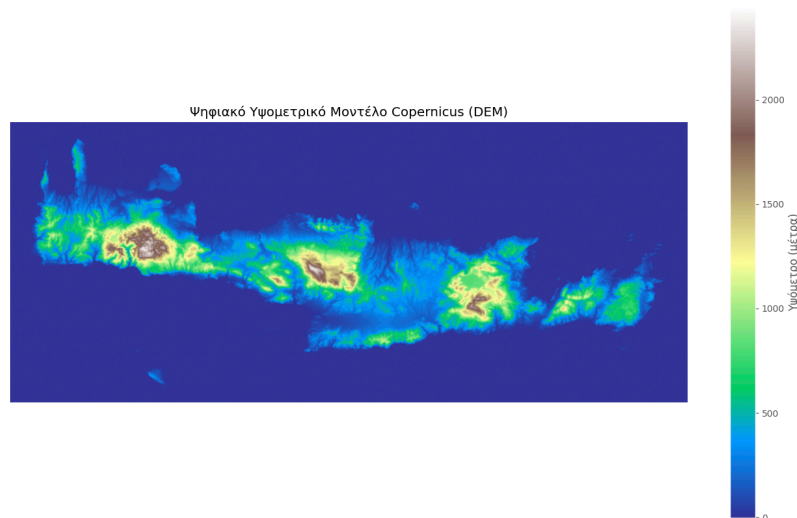
γεωμορφολογία της περιοχής μελέτης. Τα αρχεία αυτά είναι ανοικτά και διαθέσιμα σε μορφή CSV ανά σταθμό.

Για τους σκοπούς της δοκιμής του συστήματος ζωντανής συλλογής και πρόβλεψης, ελλείπει εναλλακτικής, χρησιμοποιήθηκε η ζωντανή διαδικτυακή υπηρεσία OpenWeatherMap, η οποία παρέχει Προγραμματιστική Διεπαφή (*Application Programming Interface - API*) με ζωντανές εκτιμήσεις καιρικών συνθηκών [48], συμπεριλαμβανομένης της θερμοκρασίας, υγρασίας, ατμοσφαιρικής πίεσης, αέρα και νετού ανά ώρα. Αυτές οι εκτιμήσεις καλύπτουν όλον τον κόσμο και προκύπτουν από ένα δίκτυο διαφόρων πηγών. Έπειτα από αντιπαραβολή με τις ζωντανές μετρήσεις των μετεωρολογικών σταθμών του δικτύου meteo [49], διαπιστώθηκε πως δεν μπορούν να θεωρηθούν αξιόπιστες για την περιοχή έρευνας και την ανάλυση της παρούσας μελέτης. Οι κλήσεις σε αυτό το API λειτουργούν ως υποκατάστατο ενός πυκνού δικτύου φυσικών σταθμών ακριβείας, με δυνατότητα λήψης ζωντανών μετρήσεων, παρόμοιο με αυτό που αξιοποιήθηκε για το χτίσιμο του αρχείου από όπου προήλθαν τα δεδομένα εκπαίδευσης.

3.3.3 Πηγές Γεωχωρικών Δεδομένων

Το εύρος γεωχωρικών δεδομένων που ενσωματώθηκαν στα χαρακτηριστικά πρόβλεψης παρέμεινε περιορισμένο. Το ψηφιακό υψομετρικό μοντέλο (*Digital Elevation Model - DEM*), είναι ένα raster, η γεωμετρία του οποίου αντιστοιχίζεται με γεωγραφικές συντεταγμένες και η τιμή κάθε ειconoστοιχείου (*pixel*) αντιπροσωπεύει το υψόμετρο της συγκεκριμένης περιοχής. Σε αυτήν την μελέτη, ενσωματώθηκε το DEM της Κρήτης από το δίκτυο Copernicus [50] σε ανάλυση 1pixel/100μ που είναι διαθέσιμο σε μορφή GeoTIFF.

Ο ρυθμός με τον οποίο οι γεωχωρικές στρώσεις, όπως το DEM, επιβάλλεται να ανανεώνονται για να παραμείνουν έγκυρες, εξαρτάται από το τι αναπαριστούν· για παράδειγμα, μία στρώση που αφορά την πυκνότητα της βλάστησης απαιτεί πολύ πιο συχνή ανανέωση από ένα υψομετρικό μοντέλο.



Σχήμα 3.2: DEM Κρήτης από το δίκτυο Copernicus [50]

3.3.4 Κριτήρια Καταλληλότητας Δεδομένων

Όλα τα δεδομένα επιβάλλεται να έχουν υψηλή ακρίβεια και πληρότητα για να είναι δυνατή η επιτυχής διερεύνηση του λεπτού προβλήματος που πραγματεύεται η μελέτη. Επίσης πρέπει, προφανώς, να είναι διαθέσιμα, ώστε να είναι αξιοποιήσιμα σε αυτό το στάδιο του έργου.

Τα αυστηρά κριτήρια καταλληλότητας για το αρχείο καιρικών συνθηκών αντικατοπτρίζουν τα απαιτητικά χαρακτηριστικά του προτεινόμενου συστήματος. Οι μετρήσεις πρέπει να σχηματίζουν χρονοσειρά, ένα σχήμα που προϋποθέτει περιοδικές εγγραφές σε σταθερό και σχετικά γρήγορο ρυθμό, με περιορισμένα κενά λειτουργίας. Οι σταθμοί πρέπει να παρέχουν επιστημονική ακρίβεια σε ένα εύρος καιρικών συνθηκών που επιτρέπει την μοντελοποίηση της εκδήλωσης πυρκαγιών. Ταυτόχρονα, η πυκνότητα των σταθμών πρέπει να είναι επαρκής ώστε να είναι έγκυρες οι μετρήσεις για την περιοχή που καλύπτουν.

Από το αρχείο των πυρκαγιών απαιτείται πολύ υψηλός βαθμός πληρότητας, ώστε, όταν θα γίνει η διασταύρωση με τις συνθήκες, τα μοτίβα που οδηγούν σε πυρκαγιές να είναι όσο το δυνατόν πιο ολοκληρωμένα και φανερά. Συνάμα, πρέπει να υπάρχει επαρκές πλήθος συμβάντων για την ορθή λειτουργία των αλγορίθμων. Επιπλέον, είναι απαραίτητη και η ακριβής χρονική και χωρική τοποθέτηση τους, διαφορετικά, η σύγχυση που θα προκληθεί, είναι πιθανό να στρεβλώσει τα μοτίβα και να υποβαθμίσει την απόδοση.

3.4 Σχεδιασμός Βάσει Βιβλιογραφίας

3.4.1 Παράγοντες Κινδύνου Εκδήλωσης Πυρκαγιών στην Σχεδίαση και Ανάπτυξη

Από την μελέτη σύγχρονων ερευνών πάνω στους μηχανισμούς των αγροτοδοασικών πυρκαγιών, έγινε αντιληπτό πως υπάρχουν τρεις προϋποθέσεις για την εκδήλωσή τους: Επαρκής δυναμική ενέργεια στο σύστημα, συνθήκες που διευκολύνουν την καύση, και ένα έναυσμα. Αυτή η διαπίστωση, κατηύθυνε την μοντελοποίηση προς την αναζήτηση τιμών και μεταβλητών που συνδέονται με αυτούς τους παράγοντες. Ο όγκος και οι συνθήκες των καύσιμων υλών εμφανίστηκαν με μεγάλη συνέπεια σε διάφορες έρευνες, για αυτό έπρεπε να προβλεφθούν τρόποι προσέγγισης και αναπαράστασης της κατάστασης της καύσιμης ύλης στο σύστημα, όπως με γεωχωρική δεδομένα, ή συνδυαστικές σχέσεις και συναρτήσεις.

Η εφαρμογή του FWI στην Ελλάδα δεν πέτυχε στον ίδιο βαθμό με άλλες χώρες [30][41]. Ένα πιθανό αίτιο για αυτό είναι η μη συμπερίληψη των ανθρώπινων παραγόντων, κάτι που ενδεχομένως δεν επηρεάζει τόσο την ικανότητα μοντελοποίησης στις αχανείς εκτάσεις της Β. Αμερικής. Εκεί, οι περισσότερες πυρκαγιές προκαλούνται από κεραυνούς [5]. Για την Ελλάδα, όπου την συντριπτική πρωτιά στα εναύσματα κατέχουν ανθρωπογενείς αιτίες [43], ένα κύριο μέλημα της μοντελοποίησης ήταν η εύρεση στοιχείων, τα οποία προδίδουν την εκτέλεση των δραστηριοτήτων που ενέχουν τον κίνδυνο της ανεξέλεγκτης φωτιάς.

3.4.2 Επίδραση Σύγχρονων Προσεγγίσεων Εκτίμησης Κινδύνου Πυρκαγιών

Η διάκριση μεταξύ της εκτίμησης πιθανότητας εκδήλωσης πυρκαγιάς και της εκτίμησης ευαλωτότητας κάποιας περιοχής σε πυρκαγιά έχει σημασία σε επιχειρησιακό επίπεδο. Ένα σύστημα που στοχεύει στην πιθανότητα εκδήλωσης των πυρκαγιών, προσαρμόζεται αβίαστα σε μηχανισμούς

επιφυλακής και έγκαιρης προειδοποίησης, ενώ αυτό που στοχεύει στην παθογένεια μίας περιοχής μπορεί να λειτουργήσει σαν υπενθύμιση για προληπτικά μέτρα [12]. Σε αυτό το επίπεδο, λήφθηκε υπ' όψιν η απρόβλεπτη φύση των δασικών πυρκαγιών στον ελληνικό χώρο, όπου μικρά συμβάντα ενδέχεται να εξελιχθούν σε πύρινες λαίλαπες. Γνωρίζοντας αυτό, επιλέχθηκε συνειδητά ως προσέγγιση μοντελοποίησης η απλή εκδήλωση πυρκαγιάς, δηλαδή η εκτίμηση της πιθανότητας ανάφλεξης ανά πάσα στιγμή.

Η γενικότερη τάση των ερευνών μοντελοποίησης να ενσωματώνουν πολλές διαφορετικές κατηγορίες πυρικών παραγόντων ως χαρακτηριστικά πρόβλεψης, έκανε ξεκάθαρη την ανάγκη δημιουργίας των κατάλληλων υποδομών και μεθόδων που θα επιτρέψουν την αξιοποίηση πολλών διαφορετικών πηγών δεδομένων για αυτά τα χαρακτηριστικά. Παρότι το εύρος των παραγόντων πρόβλεψης που χρησιμοποιήθηκαν στα πρότυπα μοντέλα αυτής της μελέτης ήταν σχετικά περιορισμένο, η ύπαρξη αυτών των υποδομών καθιστά την μελλοντική δοκιμή και ενσωμάτωσή τους απλή.

3.4.3 Υιοθέτηση Υπαρχουσών Προσεγγίσεων για Κλίμακες Κινδύνου Πυρκαγιών

Η παραγωγή ενός ερμηνεύσιμου και κατανοητού δείκτη επικινδυνότητας είναι εξαιρετικά σημαντική για την πρακτική εφαρμογή των συστημάτων μοντελοποίησης. Τέτοιες κλίμακες είναι οικίες στου θεσμούς δημόσιας προστασίας και κατανοητές από τους πολίτες, και, επιπλέον, αποκρύπτουν την αριθμητική τιμή της πιθανότητας, η οποία μπορεί να βρίσκεται σε μία αυθαίρετη κλίμακα, ανάλογα με τις λεπτομέρειες των δεδομένων και της μοντελοποίησης. Η παραγωγή κλίμακας επικινδυνότητας αποτελεί μέρος του στόχου του συστήματος, αν και η μέθοδος διακριτοποίησης θα παραμείνει απλή.

3.5 Προκαταρκτική Μελέτη Σκοπιμότητας

3.5.1 Μελέτη και Ανάλυση των Αρχείων Δεδομένων

Χαρακτηριστικά Αρχείου Δεδομένων

Το πρώτο σκέλος της μελέτης των δεδομένων έχει καθαρά λειτουργικό χαρακτήρα. Η εφαρμογή των κριτηρίων καταλληλότητας των δεδομένων περιόρισε το ωφέλιμο όγκο των στοιχείων και την έκταση της μελέτης. Στον κάθετο άξονα, ένας μεγάλος αριθμός εγγραφών απορρίφθηκε διότι δεν ανταποκρινόταν σε λειτουργικές προϋποθέσεις: Η εξαγωγή της τοποθεσίας από το λεκτικό τοπωνύμιο, αν και πραγματοποιήσιμη, δεν θεωρήθηκε πρακτική λύση, λόγω της περιπλοκότητας και της αστάθειας που θα προκαλούσε στα δεδομένα. Αυτό σημαίνει πως οι πυρκαγιές χωρίς ακριβείς συντεταγμένες, υποσύνολο που περιλαμβάνει κάθε καταγραφή πριν το 2020, απορρίπτονται αυτόματα. Δεδομένου αυτού, το διάστημα ενδιαφέροντος περιορίστηκε στις πρόσφατες καταγραφές και μετρήσεις από την έναρξη του 2020 μέχρι και το τέλος του 2024. Φιλτράροντας τις καταγραφές που τοποθετήθηκαν εκτός της περιοχής ενδιαφέροντος, σκόπιμα ή εκ παραδρομής, απομένουν 1.344 περιπτώσεις καταγεγραμμένων πυρκαγιών.

Οι καταγραφές μετεωρολογικών τιμών για τους 8 σταθμούς εντός της περιοχής πραγματοποιούνται ανά 30 λεπτά, καθ' όλη τη διάρκεια του έτους. Για το διάστημα ενδιαφέροντος των 5 ετών που

προέκυψε από τη διαδικασία διαλογής των πυρκαγιών, ο αναμενόμενος αριθμός καταγραφών υπολογίζεται βάσει του 3.1 στα 700800 ξεχωριστά στιγμιότυπα.

$$N_{\text{exp}} = 2 \cdot 24 \cdot 365 \cdot 5 \cdot 8 \quad (3.1)$$

Έπειτα από την αφαίρεση των διαστημάτων διακοπής λειτουργίας, τα οποία εντοπίζονται από την απουσία όλων ή σχεδόν όλων των μετρήσιμων τιμών, απομένουν $N_{\text{valid}} = 686913$ ωφέλιμες εγγραφές.

Με βάση τα παραπάνω, ο χρόνος λειτουργίας (*uptime*) των σταθμών εκτιμάται από τη σχέση (3.2).

$$U = \frac{N_{\text{valid}}}{N_{\text{exp}}} \cdot 100 = \frac{686913}{700800} \cdot 100 \approx 98.02\% \quad (3.2)$$

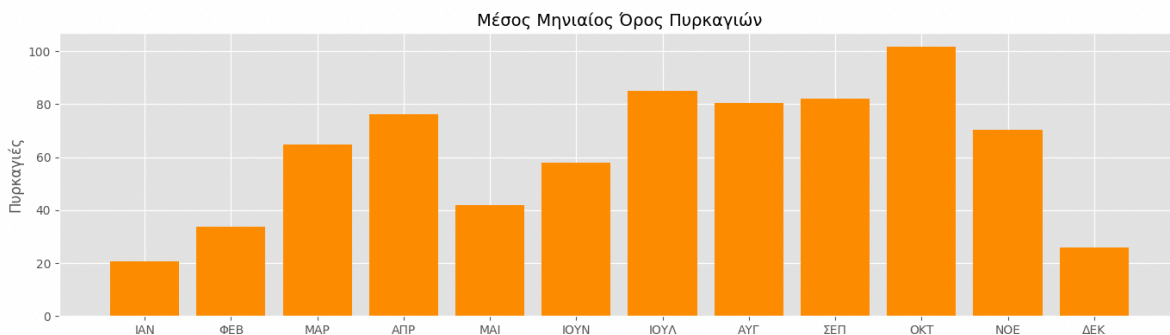
Στον κάθετο άξονα, μετά από καταμέτρηση της απουσίας τιμών ανά μετρήσιμες μεταβλητές, έγινε αντιληπτό, ότι ένα πολύ μεγάλο ποσοστό των καταγραφών δεν είχαν μετρήσεις για τον άνεμο ή τον ρυθμό εξάτμισης, πιθανώς λόγω ελαττωματικών οργάνων ή συστημικού λάθους, για αυτό και αυτές οι μεταβλητές αυτές απορρίπτονται από την αρχική μοντελοποίηση.

Ο έλεγχος για ακραίες τιμές (*outliers*), που μπορεί να προέκυψαν από σφάλματα οργάνων, πραγματοποιήθηκε με χρήση συναρτήσεων ελάχιστων και μέγιστων στις τιμές των καιρικών μεταβλητών, και έδειξε πως το αρχείο μετρήσεων είναι φαινομενικά καθαρό.

Το αρχείο πυρκαγιών και μετεωρολογικών μετρήσεων αποτελεί τον πυρήνα της μοντελοποίησης του προβλήματος.

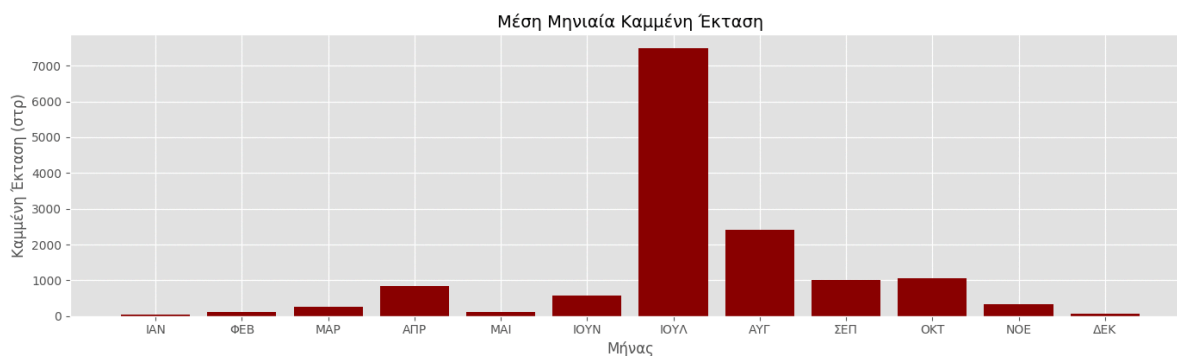
Ανάλυση Αρχείου Δεδομένων

Το δεύτερο σκέλος της ανάλυσης των δεδομένων αποσκοπεί στην βαθύτερη κατανόηση των μοτίβων που διέπουν το αρχείο, και, συνάμα, στην υποστήριξη της γενικότερης βιβλιογραφικής αναζήτησης πάνω στις δασικές πυρκαγιές στην περιοχή της μελέτης, αντλώντας συμπεράσματα από τα ίδια τα πρωτογενή δεδομένα που την αφορούν.



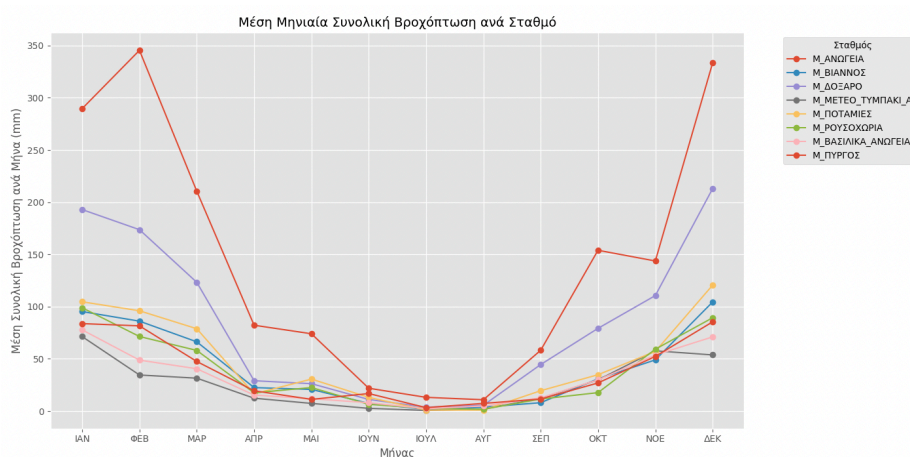
Σχήμα 3.3: Μέσος Αριθμός Πυρκαγιών ανά Μήνα (Κρήτη 2000 - 2025)

Για τον σκοπό αυτό εφαρμόστηκαν μέθοδοι ανάλυσης και οπτικοποίησης δεδομένων πάνω στο αρχείο, η επιλογή των οποίων έγινε βάσει των τάσεων που παρατηρήθηκαν σε παρόμοιες μελέτες. Για την οπτικοποίηση χρησιμοποιήθηκε η βιβλιοθήκη matplotlib της Python. Έγινε κατανομή των συμβάντων πυρκαγιών, υετού και θερμοκρασίας ανά μήνα (Σχ. 3.3, 3.5, 3.6) για την διερεύνηση



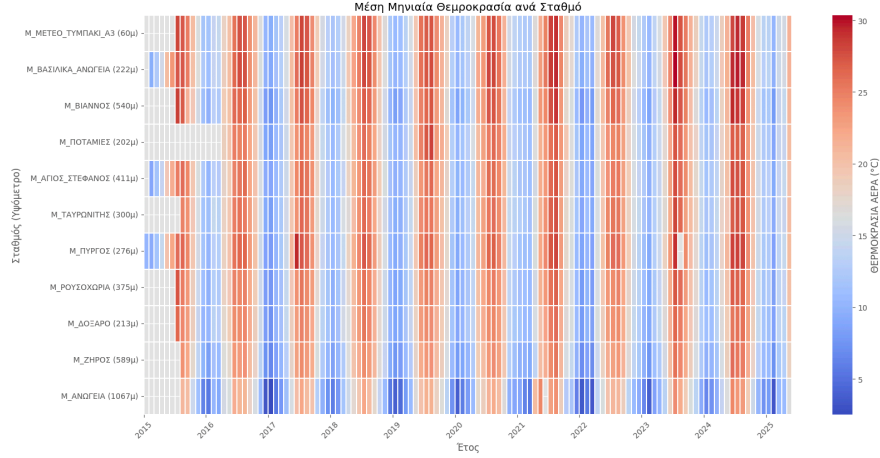
Σχήμα 3.4: Μέση Καμμένη Έκταση ανά Μήνα (Κρήτη 2000 - 2025)

απλών εποχικών μοτίβων. Εδώ σημειώθηκε πως ο σταθμός ΑΝΩΓΕΙΑ παρουσίαζε μεγάλη απόκλιση από τους υπόλοιπους, σημειώνοντας σημαντικά χαμηλότερες θερμοκρασίες λόγω υψηλού υψόμετρου. Ο εντοπισμός ζωνών υψηλής επικινδυνότητας έγινε με γεωχωρική οπτικοποίηση πυρκαγιών και καμμένης έκτασης (Σχ. 3.1).



Σχήμα 3.5: Μέσος Υετός Ανά Σταθμό και Μήνα (Κρήτη 2015 - 2025)

Εμβαθύνοντας στην ανάλυση: Το μέγεθος της καμμένης έκτασης κατανεμημένο ανά χρονολογία (Σχ. 2.5) υποδεικνύει μοτίβα πολυετούς παροδικότητας και μεταβολές στις πολιτικές πρόληψης και καταπολέμησης των πυρκαγιών. Εξετάζοντας την κατανομή των συμβάντων σε σχέση με την περίοδο του έτους (Σχ. 3.3) και την περίοδο της ημέρας (Σχ. 3.7) εντοπίζουμε ενδείξεις της περιοδικής ανθρώπινης δραστηριότητας που παρέχει το έναυσμα στην πλειοψηφία των περιστατικών. Ο διαμοιρασμός των περιστατικών ανά τις μέρες της εβδομάδας (Σχ. 3.8) παρουσιάζει ελαφρώς αυξημένη εκδήλωση τα Σαββατοκύριακα, αλλά, όταν η μέρα της εβδομάδας αναλύεται σε συνδυασμό με τον μήνα (Σχ. 3.9), φαίνεται η επιρροή των εποχικών δραστηριοτήτων.



Σχήμα 3.6: Μέση Θερμοκρασία Ανά Σταθμό κατά την Δεκαετία 2015 - 2025

3.5.2 Αρχική Διαμόρφωση Χαρακτηριστικών

Δεδομένης της ανάλυσης και της έρευνας που έγινε, προτάσσονται δύο υποθέσεις που, εάν επιβεβαιωθούν, ενισχύουν την δυνητική ικανότητα των διαθέσιμων δεδομένων να αποτυπώνουν τις συνθήκες που οδηγούν σε πυρκαγιές και να μοντελοποιούν το πρόβλημα της εκδήλωσης πυρκαγιών:

- **Ο προγενέστερος υετός αντανakλά την κατάσταση της βλάστησης:** Δεδομένων των αλληλεπιδράσεων μεταξύ των καιρικών συνθηκών, συναρτήσει του χρόνου, και της βλάστησης, οι οποίες διατυπώθηκαν στην ενότητα 2.5, θεωρείται ότι για μία περιοχή s και μία χρονική στιγμή t , το FMC και ο όγκος της βλάστησης μπορούν να περιγραφούν εννοιολογικά από τις σχέσεις (3.3) και (3.4) αντίστοιχα. Προτείνεται ότι οι τιμές των μεγεθών αυτών επηρεάζονται θετικά, κυρίως από τον προγενέστερο υετό που καταγράφηκε στην περιοχή s κατά ένα σύνολο χρονικών διαστημάτων που προηγούνται της στιγμής t .

$$H(s, t) = f(P_1(s, t), P_2(s, t), \dots, P_n(s, t)), \quad (3.3)$$

$$V(s, t) = g(Q_1(s, t), Q_2(s, t), \dots, Q_m(s, t)), \quad (3.4)$$

Κάθε όρος $P_i(s, t)$ της (3.3) εκφράζει τον αθροιστικό υετό στην περιοχή s κατά το προγενέστερο βραχυπρόθεσμο χρονικό ορίζοντα D_i . Κάθε όρος $Q_j(s, t)$ της (3.4) επίσης εκφράζει τον αθροιστικό υετό στην περιοχή s εντός του προγενέστερου μεσοπρόθεσμου χρονικού ορίζοντα E_j .

Στο πλαίσιο της παρούσας μελέτης, θεωρείται ότι οι υπόλοιποι παράγοντες που ενδέχεται να επηρεάζουν τις σχέσεις αυτές είτε παραμένουν σταθεροί, όπως ο τύπος και η κάλυψη του εδάφους, είτε δεν λαμβάνονται υπόψη, όπως οι πολιτικές προληπτικού καθαρισμού.

- **Η ανθρώπινη δραστηριότητα μπορεί να προσεγγιστεί μέσω απλών χρονικών παραγόντων:** Όπως ειπώθηκε στις ενότητες 2.4.2 και 2.6, ο ενεργός ανθρώπινος παράγοντας κυριαρχεί στις αιτίες εκδήλωσης των πυρκαγιών. Εντούτοις, η ανθρώπινη παρουσία δεν είναι

άμεσα διαθέσιμη ως μετρήσιμο μέγεθος στο αρχείο δεδομένων, άρα, κρίνεται αναγκαία η κατά προσέγγιση εκτίμησή της με βάση τα διαθέσιμα στοιχεία.

Για κάθε περιοδικό χρονικό διάστημα T , η φάση μιας χρονικής στιγμής t εντός αυτού μπορεί να αναπαρασταθεί μέσω της τριγωνομετρικής κωδικοποίησης (3.5)

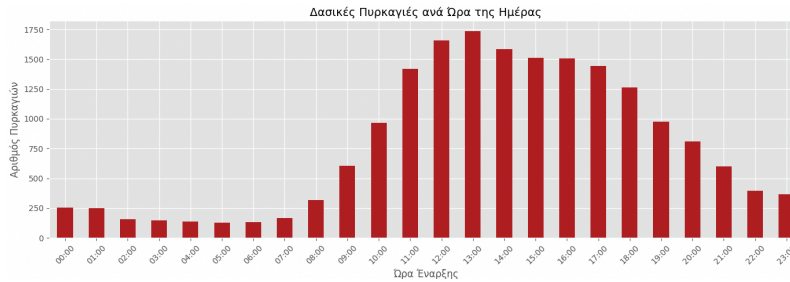
$$\sin\left(\frac{2\pi t}{T}\right), \cos\left(\frac{2\pi t}{T}\right). \quad (3.5)$$

Εφόσον η ανθρώπινη δραστηριότητα εμφανίζει περιοδικότητα σε διαφορετικές χρονικές κλίμακες, όπως εντός του ημερολογιακού έτους, της εβδομάδας και της ημέρας, θεωρείται ότι μπορεί να περιγραφεί από μία σχέση $H(t)$, η τιμή της οποίας αντιστοιχεί στην εκτιμώμενη ένταση της ανθρώπινης δραστηριότητας κατά τη χρονική στιγμή t . Η σχέση αυτή υποτίθεται ότι εξαρτάται από τις τριγωνομετρικές συνιστώσες που αντιστοιχούν στη φάση των παραπάνω περιοδικών κύκλων.

Συνεπώς, σε εννοιολογικό επίπεδο, η εκτιμώμενη ανθρώπινη δραστηριότητα H αποδίδεται από την σχέση (3.6)

$$H(t) = h\left(\sin\left(\frac{2\pi t}{T_y}\right), \cos\left(\frac{2\pi t}{T_y}\right), \sin\left(\frac{2\pi t}{T_w}\right), \cos\left(\frac{2\pi t}{T_w}\right), \sin\left(\frac{2\pi t}{T_d}\right), \cos\left(\frac{2\pi t}{T_d}\right)\right) \quad (3.6)$$

όπου T_y , T_w και T_d αντιστοιχούν στις περιόδους του έτους, της εβδομάδας και της ημέρας αντίστοιχα.

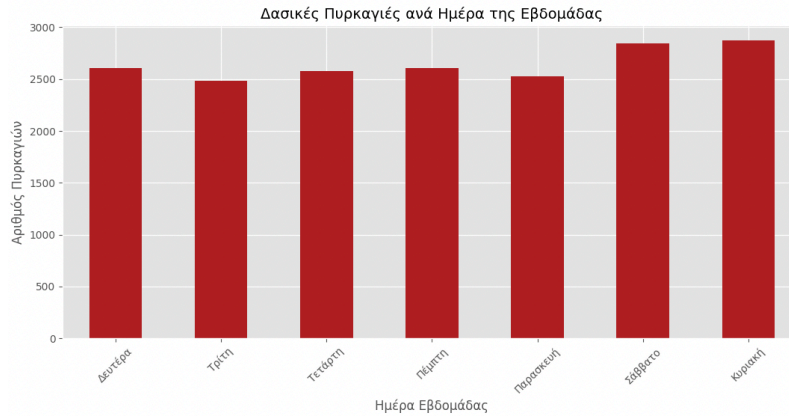


Σχήμα 3.7: Συνολικός Αριθμός Πυρκαγιών ανά Ώρα Εκδήλωσης (Κρήτη 2000 - 2025)

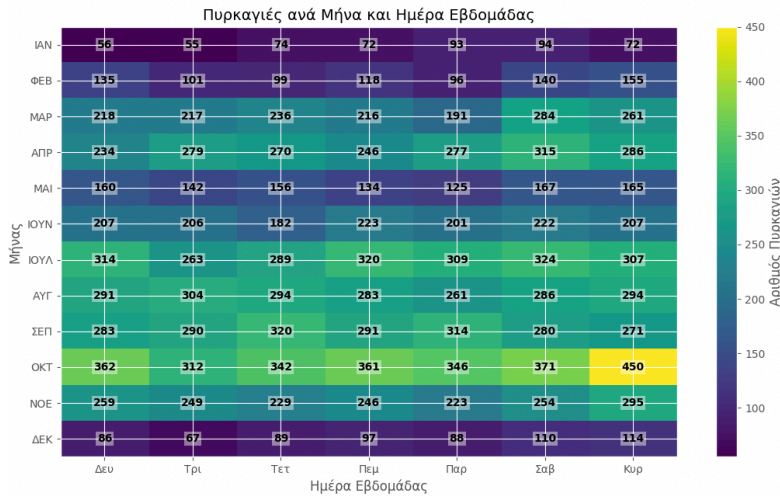
Για να δοκιμαστούν αυτές οι υποθέσεις, εντάσσεται στα χαρακτηριστικά πρόβλεψης η αθροιστική βροχόπτωση για τα διαστήματα 1 ημέρας, 7 ημερών, 30 ημερών, 60 ημερών, 90 ημερών και 180 ημερών, καθώς και το ημίτονο και συνημίτονο της φάσης της ημέρας, της εβδομάδας, και του έτους.

3.5.3 Προκαταρκτικός Έλεγχος Απόδοσης Μοντέλων

Οι δοκιμές των αλγόριθμων πρόβλεψης ξεκίνησαν στο πλαίσιο της Μηχανικής Χαρακτηριστικών με σκοπό την εκτίμηση του αντίκτυπου των διαφόρων παραμέτρων. Πολύ σύντομα, έγινε ξεκάθαρο ότι το σήμα πυρκαγιών ως προς το σύνολο των δεδομένων είναι εξαιρετικά μικρό. Η έντονα δυσανάλογη κατανομή των κλάσεων, με ποσοστό 0.195% για την κλάση <<Πυρκαγιά>>



Σχήμα 3.8: Συνολικός Αριθμός Πυρκαγιών ανά Ημέρα της Εβδομάδας (Κρήτη 2000 - 2025)



Σχήμα 3.9: Συνολικός Αριθμός Πυρκαγιών ανά Μήνα και Ημέρα (Κρήτη 2000 - 2025)

και 99.805% για την κλάση <<Όχι Πυρκαγιά>>, κυριαρχεί σε τέτοιο βαθμό, ώστε να δυσχεραίνει σημαντικά τη διαδικασία εκπαίδευσης ενός πιθανοτικού μοντέλου. Αυτό εξομαλύνθηκε σε μεγάλο βαθμό όταν το πρόβλημα του συστήματος ορίστηκε ξεκάθαρα ως ο εντοπισμός συνθηκών που οδηγούν σε πυρκαγιά τις ακόλουθες ώρες. Με την μεταβολή του στόχου από Πυρκαγιά Τώρα σε Πυρκαγιά μέσα στις επόμενες 3 ώρες, αυξήθηκε το σήμα σε ένα πιο έντονο 1.17%.

Ο αλγόριθμος LR υπολογίζει τις συνιστώσες των μεταβλητών, οι οποίες μπορούν να αξιοποιηθούν ως δείκτης σημαντικότητας για την κάθε μεταβλητή. Υπολογίζοντας την ισχυρή αρνητική συσχέτιση μεταξύ της πρόσφατης βροχόπτωσης και της πρόκλησης των πυρκαγιών, ενισχύθηκε πειραματικά η υπόθεση πως, μέσω της βραχυπρόθεσμης συγκεντρωτικής βροχόπτωσης, μπορεί να εξαχθεί συμπέρασμα για την υγρασία της βλάστησης. Εξετάζοντας τις συνιστώσες του συγκεντρωτικού νετού στα μεσοπρόθεσμα διαστήματα 1 μέχρι 6 μηνών, ενισχύεται πειραματικά η υπόθεση πως η εκδήλωση πυρκαγιών διέπεται από εποχικά μοτίβα βροχόπτωσης.

Επιπλέον, η περιορισμένη απόδοση του γραμμικού αλγόριθμου LR φανερώνει τον μη γραμμικό χαρακτήρα των σχέσεων του συνόλου δεδομένων, τύπος προβλήματος που τείνει να δυσκολεύει

αυτόν τον αλγόριθμο.

Προχωρώντας στον αλγόριθμο RF, οι δοκιμές έδειξαν την θετική επίδραση των τριγωνομετρικών χαρακτηριστικών φάσης στην ακρίβεια και την ευστοχία των προβλέψεων και με αυτόν τον τρόπο ενισχύει την υπόθεση ότι δρουν ως έμμεσοι δείκτες για την ανθρώπινη δραστηριότητα στο γενικό πλαίσιο της μελέτης δασικών πυρκαγιών. Αυτά τα χαρακτηριστικά βελτίωσαν την ποιότητα των προβλέψεων και ξεχωριστά, αλλά και συνδυαστικά μεταξύ τους, προσφέροντας έτσι μία μέθοδο μοντελοποίησης των περίπλοκων μοτίβων της ανθρώπινης δραστηριότητας. Όσον αφορά την προβλεπτική δύναμη του RF, τα πειραματικά αποτελέσματα δεν έφτασαν κοντά στους απαιτούμενους στόχους, είχαν όμως θετικές ενδείξεις, που σε συνδυασμό με την βιβλιογραφική έρευνα, ήδη έδειξαν σε αυτό το σημείο την δυνητική ικανότητα του αλγορίθμου να ανταποκριθεί στο πρόβλημα.

3.5.4 Αρχικά Ευρήματα

Τα ευρήματα της προκαταρκτικής μελέτης ήταν καίρια για την διαμόρφωση της τελικής μεθοδολογικής προσέγγισης και την πορεία της υλοποίησης. Μέσω των δοκιμών εντός της δύσκολης πραγματικότητας των ακατέργαστων δεδομένων για την περιοχή μελέτης, ήταν δυνατό να γίνουν οι κατάλληλες προετοιμασίες για την συλλογή και την επεξεργασία των δεδομένων και να προσαρμοστεί η στρατηγική μοντελοποίησης, ώστε να είναι πιο κατάλληλη για την παραγωγή χρήσιμων αποτελεσμάτων.

Διαμορφώθηκαν υποθέσεις αναφορικά με δύο κύριους παράγοντες για την εκδήλωση πυρκαγιών και αποτυπώθηκαν με κάποια επιτυχία, δείχνοντας πειραματικά την θετική επίδραση τους στο σύστημα. Τα ευρήματα που σχετίζονται με την προγνωστική ικανότητα των μοντέλων, έδειξαν ότι προβλέπουν πάνω από το αναμενόμενο ρυθμό και, επομένως, δίνουν το πράσινο φως για την σκοπιμότητα του συστήματος και την υλοποίηση.

3.6 Μεθοδολογία Διαχείρισης των Δεδομένων

3.6.1 Συλλογή και Αποθήκευση

Τα αρχεία των ιστορικών μετρήσεων και τα αρχεία γεωχωρικών δεδομένων καταχωρούνται αυτόματα στο σύστημα κατά το αρχικό στήσιμο ή μπορούν να εισαχθούν σε μεταγενέστερο χρόνο. Η τακτική συλλογή των μετρήσεων των σταθμών/αισθητήρων γίνεται χρονοδρομολογημένα, με έναν σταθερό ρυθμό c , το διάστημα του οποίου είναι παράμετρος του συστήματος, και για το οποίο επιλέχθηκε $c = 30min$, σε συμφωνία με τα διαθέσιμα πρωτογενή δεδομένα. Τα δεδομένα που συλλέγονται περιοδικά χρησιμοποιούνται για την εκτίμηση κινδύνου και έπειτα αποθηκεύονται, ώστε να χτιστεί ένα αρχείο μετρήσεων.

Η αποθήκευση διατηρεί τον πρωτογενή χαρακτήρα των μετρήσεων, χωρίς καμία επεξεργασία ώστε να υπάρξει ευελιξία στην μετέπειτα διαχείρισή τους. Αποθηκεύονται, δηλαδή, η χρονική στιγμή t και ο σταθμός μέτρησης, από τα οποία μπορεί να προκύψει ένα στιγμιότυπο $w(s, t)$ όπου s ένας τομέας της περιοχής μελέτης που είναι στην εμβέλεια του σταθμού μέτρησης. Επιπλέον, καταχωρείται το όνομα της μετρήσιμης μεταβλητής καθώς και η τιμή που καταγράφηκε από τα όργανα. Οι επαναλαμβανόμενες περιοδικές μετρήσεις σχηματίζουν μία χρονοσειρά, ένα σχήμα που παρέχει σειρά και συνέχεια μεταξύ των εγγραφών, χαρακτηριστικά που είναι αξιοποιήσιμα κατά

την ανάλυση. Με την χρήση των ανεπεξέργαστων αποθηκευμένων μετρήσεων, μπορεί ανά πάσα στιγμή να γίνει μία εκτίμηση για οποιοδήποτε στιγμιότυπο w , εντός του διαστήματος και του χώρου λειτουργίας του συστήματος.

3.6.2 Καθαρισμός και Εναρμόνιση

Η ομοιομορφία των μετεωρολογικών σταθμών δεν είναι δεδομένη, παρεκκλίσεις μπορεί να προκύψουν από διαφορετικούς πάροχους, ή λόγω αλλαγών και αναβαθμίσεων. Στις προδιαγραφές του συστήματος περιλαμβάνονται και οι μονάδες μέτρησης που χρησιμοποιούνται για τον κάθε τύπο μεταβλητής και, βάσει αυτών, γίνονται οι κατάλληλες μετατροπές. Έτσι, έχοντας γνώση του ονόματος της μεταβλητής και από ποιόν σταθμό προέρχεται, οι μετρήσεις μπορούν να συγκριθούν και να ταυτιστούν ανεξάρτητα από την πηγή τους και τις μεταξύ τους διαφορές ως προς την καταγραφή.

Το αρχείο πυρκαγιών φιλτράρεται με βάση την καμμένη έκταση. Οι πυρκαγιές που λαμβάνουν χώρα σε χωματερές, βάσει του τύπου της καμμένης έκτασης, απορρίπτονται, καθώς αυτές οι ζώνες δεν υπακούν στα ίδια μοτίβα FMC που κατευθύνουν την μοντελοποίηση των αγροτοδοσικών πυρκαγιών. Επίσης, τα συμβάντα που δεν έχουν καθόλου καταγεγραμμένη καμμένη έκταση απορρίπτονται ως ενδεχόμενοι λάθος συναγερμοί ή άλλου είδους περιστατικά.

3.6.3 Χρονική και Χωρική Κβαντοποίηση

Τα μοντέλα πιθανοτήτων που χρησιμοποιεί το σύστημα λειτουργούν επί σαφώς ορισμένων στιγμιότυπων. Για τον λόγο αυτό, καθίσταται αναγκαία η κβαντοποίηση του αρχείου μετρήσεων ως προς τον χώρο και τον χρόνο, ώστε κάθε στιγμιότυπο w να αντιστοιχεί σε μία γεωγραφική περιοχή s και σε ένα χρονικό τμήμα t .

Ειδικότερα, κάθε χρονικό τμήμα t ορίζεται από μία χρονική στιγμή έναρξης t_{start} και μία χρονική στιγμή λήξης t_{end} , ώστε $t = [t_{\text{start}}, t_{\text{end}})$. Η τιμή της αρχής του t δίνεται από την σχέση (3.7), ενώ η τιμή της λήξης του από την σχέση (3.8)

$$t_{\text{start}} = nc \quad (3.7)$$

$$t_{\text{end}} = nc + c = (n + 1)c, \quad (3.8)$$

όπου $n \in \mathbb{N}$ και c είναι το χρονικό βήμα καταγραφής των μετρήσεων.

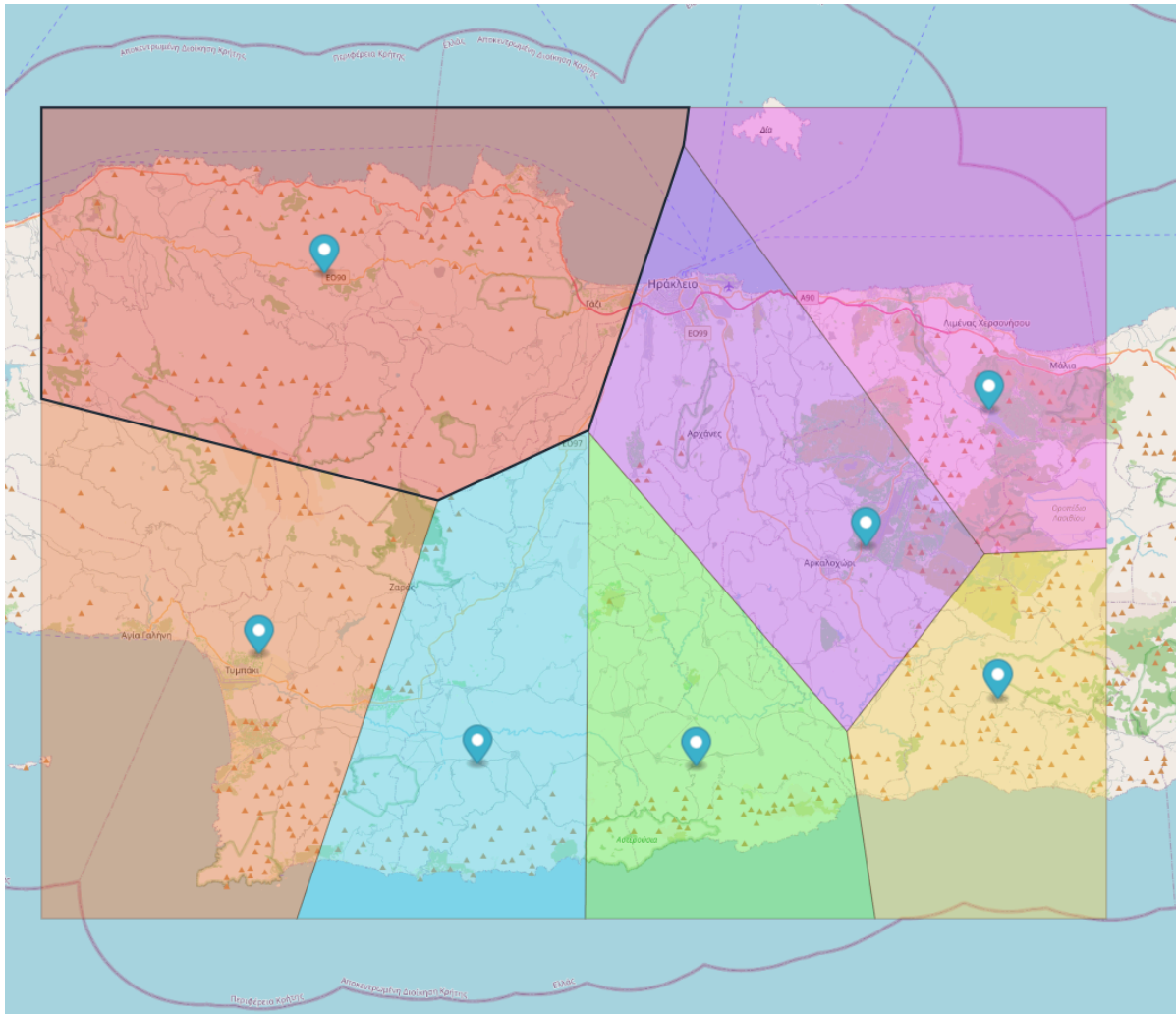
Αντίστοιχα, κάθε χωρικός τομέας s αποτελεί υποσύνολο της συνολικής περιοχής μελέτης S , δηλαδή $s \subseteq S$.

Κατά συνέπεια, κάθε στιγμιότυπο w μπορεί να θεωρηθεί ως ένα διακριτό ζεύγος χωρικής και χρονικής αναφοράς με μορφή $w(s, t)$.

Ο διαχωρισμός της S σε υποπεριοχές (*tassellation*), εφόσον γίνει ορθά, επιτρέπει μεγαλύτερη ακρίβεια στην εκτίμηση των συνθηκών πάνω στις εκτάσεις των επιμέρους τομέων, σε σχέση με εκτιμήσεις για την συνολική έκταση της S και, επομένως, πιο στοχευμένες εκτιμήσεις. Η εφαρμογή μίας μεθόδου διαχωρισμού G πάνω στην περιοχή S , έχει ως αποτέλεσμα την παραγωγή ενό σύνολου τομέων $\{s_1^{(G)}, s_2^{(G)}, \dots, s_n^{(G)}\}$, για το οποίο ισχύει η σχέση (3.9)

$$S = \bigcup_{i=1}^n s_i^{(G)} \quad (3.9)$$

Για την παρούσα μελέτη, η μέθοδος Voronoi επιλέχθηκε για τον διαχωρισμό της περιοχής ενδιαφέροντος, σύμφωνα με την οποία ο κάθε τομέας διαμορφώνεται από τον χώρο που είναι εγγύτερος σε κάθε δεδομένο σημείο. Στο πλέγμα που προκύπτει χρησιμοποιώντας τους σταθμούς ως τα κεντρικά σημεία άγκυρες, ο κάθε τομέας αποτελείται από τον χώρο που περιτριγυρίζει τον αντίστοιχο σταθμό (Σχ. 3.10). Άλλες μέθοδοι που υποστηρίζονται από το σύστημα είναι ο διαχωρισμός σε όμοια τετράγωνα (*rasterisation*) ή εξάγωνα, μεταβλητής ανάλυσης r για μεγαλύτερη λεπτομέρεια, αλλά και η χρήση προσαρμοσμένων περιοχών αυθαίρετου αριθμού κορυφών και μεγέθους για την καλύτερη αποτύπωση των μικροκλιμάτων και ιδιομορφιών της κάθε περιοχής. Σε κάθε περίπτωση, οι τομείς που προκύπτουν από τον διαχωρισμό δεν πρέπει να έχουν καμία επικάλυψη, δηλαδή να ισχύει $s_i^{(G)} \cap s_j^{(G)} = \emptyset, \forall i \neq j$.



Σχήμα 3.10: Πλέγμα Διαχωρισμού Voronoi βάσει των Σταθμών της Περιοχής Μελέτης **Υπόβαθρο:**© OpenStreetMap contributors, available under the Open Database License.

3.6.4 Χρονική και Χωρική Συσχέτιση

Μία από τις πιο απλές μεθόδους ανάθεσης τιμών για καιρικές συνθήκες $X^{(W)}$ σε κάθε τομέα s είναι η απευθείας ανάθεση των καιρικών συνθηκών του εγγύτερου σταθμού W^* , όπως αποτυπώνεται στην σχέση (3.10). Η απόσταση μεταξύ σταθμού και τομέα υπολογίζεται βάσει της ευκλείδειας απόστασης ανάμεσα στο γεωχωρικό κεντροειδές C_s του τομέα και τον εκάστοτε σταθμό, πάνω στο γεωγραφικό σύστημα συντεταγμένων, σύμφωνα με την σχέση (3.11).

$$X_s^{(W)} = X_{W^*(s)}^{(W)}, W^*(s) = \arg \min d(s, W) \quad (3.10)$$

$$d((C_x, C_y), (W_x, W_y)) = \sqrt{(C_x - W_x)^2 + (C_y - W_y)^2} \quad (3.11)$$

Εφόσον ο διαχωρισμός έγινε με την μέθοδο Voronoi, ο εγγύτερος σταθμός είναι και αυτός από τον οποίο προέκυψε ο τομέας. Εναλλακτικές μέθοδοι όπως η παρεμβολή καιρικών συνθηκών είναι περίπλοκες και μπορούν να εφαρμοστούν, εφόσον μελετηθούν επαρκώς σε μεταγενέστερο στάδιο.

Δεδομένων των σαφώς ορισμένων στιγμιοτύπων $w(s, t)$, η αντιστοίχιση με τις πυρκαγιές, οι οποίες έχουν ήδη φιλτραριστεί ώστε η τοποθεσία τους να εμπίπτει στην περιοχή μελέτης S , συνιστά διαδικασία γεωχωρικής και χρονικής ευθυγράμμισης.

Εφόσον κάθε πυρκαγιά αναπαρίσταται από ένα ζεύγος συντεταγμένων, δηλαδή από ένα σημείο στον χώρο, και δεδομένου ότι δεν υπάρχει επικάλυψη μεταξύ των χωρικών τομέων της περιοχής μελέτης, κάθε τοποθεσία πυρκαγιάς που ανήκει στην περιοχή S αντιστοιχίζεται σε έναν και μόνο έναν τομέα s .

Αντίστοιχα, ως προς τον χρόνο, κάθε πυρκαγιά αντιστοιχίζεται στο χρονικό τμήμα t για το οποίο ισχύει $t_{\text{start}} \leq t_{\text{fire}} < t_{\text{end}}$.

3.6.5 Δημιουργία Χαρακτηριστικών

Για την μοντελοποίηση των δασικών πυρκαγιών μπορούν να αξιοποιηθούν διαφορετικά χαρακτηριστικά, οι τιμές των οποίων θα πρέπει να περιγράφουν επαρκώς τις συνθήκες που σχετίζονται με την εκδήλωση πυρκαγιάς σε ένα δεδομένο στιγμιότυπο $w(s, t)$. Ο τρόπος υπολογισμού του κάθε χαρακτηριστικού διαφέρει ανάλογα με τον τύπο του. Για τους σκοπούς της οργάνωσης και συντονισμένης λειτουργίας, ορίζεται χάρτης που περιγράφει το κάθε χαρακτηριστικό, τον τύπο του και σε ποιες μεταβλητές αντιστοιχίζεται.

Καιρικές Συνθήκες

Πολλά αξιοποιήσιμα χαρακτηριστικά μπορούν να παραχθούν απευθείας από της καιρικές μεταβλητές που μετράνε οι μετεωρολογικοί σταθμοί. Πυρικοί παράγοντες, όπως η τρέχουσα θερμοκρασία και υγρασία της ατμόσφαιρας, είναι άμεσα διαθέσιμοι από τις πρωτογενείς μετεωρολογικές μετρήσεις χωρίς επιπλέον επεξεργασία. Έχουν προβλεφθεί και συνθήκες για πιο περίπλοκες σχέσεις συνδυασμού ή μετατροπής καιρικών μεταβλητών, εφόσον απαιτηθούν από τις ανάγκες της μοντελοποίησης.

Χρονικά Παράθυρα και Συσσωρευτικά Χαρακτηριστικά

Η ανάλυση χρονοσειρών μέσω χρονικών παραθύρων επιτρέπει την αναπαράσταση σχέσεων χρονικής εξάρτησης, δηλαδή του τρόπου με τον οποίο οι συνθήκες που επικρατούν πριν από τη

χρονική στιγμή t του τρέχοντος στιγμιότυπου επηρεάζουν την παρούσα κατάσταση. Με τον τρόπο αυτό, οι τιμές μιας μεταβλητής μπορούν να εξεταστούν σε συνάρτηση με την πρόσφατη χρονική τους εξέλιξη.

Έστω μεταβλητή v , l μήκος χρονικού ορίζοντα συσσωρευτικού χαρακτηριστικού και τελεστής συσσώρευσης A . Τότε, για έναν χωρικό τομέα s και ένα χρονικό τμήμα t , το παράγωγο χαρακτηριστικό προκύπτει από την εφαρμογή του τελεστή A στο χρονικό παράθυρο μήκους l , πριν από το t . Αυτό περιλαμβάνει τα $n = l/c$ προηγούμενα στιγμιότυπα, όπου c το χρονικό μήκος των στιγμιότυπων, ο ρυθμός λήψης μετρήσεων σε αυτήν την περίπτωση. Η γενικευμένη συσσώρευση φαίνεται στην σχέση (3.12).

$$x_{s,t}^{(v,l,A)} = A(v_{s,t-1}, v_{s,t-2}, \dots, v_{s,t-n}). \quad (3.12)$$

Η παραπάνω διατύπωση καλύπτει διαφορετικές μορφές συσσώρευσης σε παράθυρα, όπως αθροιστική, μέσου όρου, ελάχιστου ή μέγιστου, και επιτρέπει την αποτύπωση διαφορετικών όψεων της χρονικής συμπεριφοράς της μεταβλητής.

Στην ειδική περίπτωση κατά την οποία ο τελεστής A είναι το άθροισμα, η συσσωρευτική τιμή της μεταβλητής v στο παράθυρο μήκους l υπολογίζεται βάσει του (3.13)

$$x_{s,t}^{(v,l,\Sigma)} = \sum_{k=1}^l v_{s,t-k}. \quad (3.13)$$

Η ποσότητα αυτή εκφράζει τη συνολική επίδραση των προηγούμενων τιμών της μεταβλητής v στον τομέα s κατά τα n προηγούμενα στιγμιότυπα. Η συσσώρευση μπορεί να χρησιμοποιηθεί για την περιγραφή χρονικών σχέσεων που δεν είναι άμεσα ορατές από την τρέχουσα μόνο τιμή.

Με αυτόν τον τρόπο, δοκιμάζεται η πρώτη υπόθεση αναφορικά με την βροχόπτωση σε μία σειρά από διαστήματα που προηγούνται του t , με σκοπό την προσέγγιση της υγρασίας και του όγκου της βλάστησης. Πέρα από την άθροιση, η ανάλυση σε χρονικά παράθυρα παράγει και άλλα αποτελέσματα που ενδέχεται να έχουν ενδιαφέρον, όπως την μέση ή την μέγιστη τιμή, μ χ της υγρασίας, το προηγούμενο χρονικό διάστημα.

Τιμές Φάσης

Οι τιμές φάσης περιόδων που χρησιμοποιούνται για τον έλεγχο της δεύτερης μεθοδολογικής υπόθεσης, σχετικά με την ανθρώπινη δραστηριότητα, προκύπτουν από τη χρονική στιγμή έναρξης t_{start} του στιγμιότυπου, τη διάρκεια T της περιόδου και από μία χρονική στιγμή άγκυρας t_0 , η οποία ορίζεται ως αρχή της φάσης. Παραδείγματα t_0 είναι η ώρα 00:00 για την ημέρα η 1η Ιανουαρίου ώρα 00:00 για το έτος και η Δευτέρα ώρα 00:00 για την εβδομάδα.

Για κάθε περίοδο T , η φάση του στιγμιότυπου μπορεί να βρεθεί από την σχέση (3.14) και, στην συνέχεια, να κωδικοποιηθεί με το ζεύγος των τριγωνομετρικών τιμών της σύμφωνα με το (3.15).

$$\theta_T(t) = \frac{2\pi}{T} ((t_{\text{start}} - t_0) \bmod T), \quad (3.14)$$

$$(\sin(\theta_T(t)), \cos(\theta_T(t))). \quad (3.15)$$

Το ημίτονο και το συνημίτονο της φάσης, όταν χρησιμοποιούνται από κοινού, αποδίδουν μονοσήμαντα τη θέση της χρονικής στιγμής t_{start} εντός του περιοδικού κύκλου.

Γεωχωρικές Τιμές

Εκτός από τις καιρικές συνθήκες, κάθε χωρικός τομέας s μπορεί να χαρακτηρίζεται και από στατικά ή βραδέως μεταβαλλόμενα γεωγραφικά γνωρίσματα, τα οποία αντλούνται από τις γεωχωρικές στρώσεις που διατηρεί το σύστημα. Τα γνωρίσματα αυτά μπορεί να προέρχονται είτε από στρώσεις τύπου *raster* είτε από διανυσματικές στρώσεις (*vector layers*).

Για μία στρώση *raster* R , θεωρείται ότι κάθε τομέας s αντιστοιχεί σε ένα σύνολο εικονοστοιχείων $P(s)$, δηλαδή στα εικονοστοιχεία της στρώσης που εμπίπτουν εντός του s . Τότε, ένα χωρικό χαρακτηριστικό που παράγεται από τη στρώση R ορίζεται ως η εφαρμογή ενός τελεστή συσσώρευσης A επί των τιμών των εικονοστοιχείων αυτών, όπως αποτυπώνεται στην σχέση συσσώρευσης στοιχείων raster (3.16).

$$x_s^{(R,A)} = A(\{R(p) : p \in P(s)\}). \quad (3.16)$$

Ανάλογα με τη φύση του γεωγραφικού γνωρίσματος, ο τελεστής A μπορεί να είναι, ενδεικτικά, ο μέσος όρος, το ελάχιστο, το μέγιστο ή η τυπική απόκλιση. Με τον τρόπο αυτό, μία συνεχής γεωχωρική επιφάνεια μετατρέπεται σε ένα διακριτό χαρακτηριστικό ανά τομέα s .

Αντίστοιχα, για μία διανυσματική στρώση V , το χαρακτηριστικό ενός τομέα s προκύπτει από τα γεωμετρικά στοιχεία της στρώσης που τέμνουν ή περιέχονται στον τομέα αυτό. Αν $F(s)$ είναι το σύνολο των στοιχείων της στρώσης V που σχετίζονται χωρικά με τον τομέα s , τότε ένα χωρικό χαρακτηριστικό μπορεί να υπολογιστεί από την σχέση συσσώρευσης διανυσματικών στοιχείων (3.17),

$$x_s^{(V,A)} = A(\{g(f, s) : f \in F(s)\}), \quad (3.17)$$

όπου $g(f, s)$ είναι μία συνάρτηση που αποδίδει τη συνεισφορά του στοιχείου f στον τομέα s , όπως για παράδειγμα παρουσία ή απουσία, πλήθος στοιχείων, συνολικό μήκος γραμμικών στοιχείων, συνολικό εμβαδόν πολυγώνων ή απόσταση από το πλησιέστερο στοιχείο.

Συνεπώς, τόσο οι *raster* όσο και οι διανυσματικές στρώσεις μπορούν να μετασχηματιστούν σε διακριτά χωρικά χαρακτηριστικά ανά τομέα s , μέσω κατάλληλης επιλογής χωρικής σχέσης και τελεστή συσσώρευσης.

Ιστορικό Πυρκαγιών

Αν και δεν χρησιμοποιήθηκαν στην μοντελοποίηση για αυτήν την μελέτη, δεδομένης της κρίσιμης ισορροπίας της καύσιμης ύλης που αναπτύσσεται και καταναλώνεται, είναι δυνατό να εξαχθούν χρήσιμες πληροφορίες για την κατάσταση του εδάφους μέσω της ανάλυσης του αρχείου πυρκαγιών που διατηρεί το σύστημα. Ειδικά το μέγεθος της καμμένης έκτασης και το διάστημα που μεσολάβησε από μεγάλη πυρκαγιά σε κάποιον τομέα, μπορούν να συνεισφέρουν στην εκτίμηση του όγκου της βλάστησης και να υποδείξουν μοτίβα σε βάθος ετών. Η πρόκληση είναι ότι χρειάζεται σχεδόν ζωντανή ενημέρωση του αρχείου των πυρκαγιών, για να αξιοποιηθούν τέτοια χαρακτηριστικά σε ένα σύστημα επιχειρησιακού επιπέδου.

3.6.6 Προετοιμασία Δεδομένων Εκπαίδευσης/Πρόβλεψης

Η προετοιμασία των δεδομένων είναι σχεδόν η ίδια, ανεξάρτητα με το αν πρόκειται για εκατομμύρια εγγραφές για την εκπαίδευση του μοντέλου, ή απλά μία εγγραφή για τα πιο πρόσφατα προγνωστικά. Όταν οι εγγραφές είναι πολλές, η ίδια διαδικασία γίνεται για κάθε χρονική περίοδο t σε κάθε τομέα s , ώστε να δημιουργηθούν όλα τα ξεχωριστά στιγμιότυπα.

1. Δημιουργία των χαρακτηριστικών που προκύπτουν από αντιστοίχιση με καιρικές συνθήκες του εγγύτερου σταθμού ή παρεμβολή. Για όποια περίπτωση αντιστοίχισης δεν βρεθεί καταγεγραμμένη μεταβλητή, το χαρακτηριστικό παραμένει κενό.
2. Δημιουργία των συσσωρευτικών χαρακτηριστικών από αντιστοίχιση με καιρική μεταβλητή και εφαρμογή του αντίστοιχου τελεστή συσώρευσης A επί του παραθύρου w του κάθε χαρακτηριστικού από τις τιμές του αρχείο καταγραφών.
3. Δημιουργία των χαρακτηριστικών φάσης από την περίοδο φάσης T και την χρονική στιγμή t_{start} .
4. Δημιουργία των γεωγραφικών χαρακτηριστικών με εφαρμογή του αντίστοιχου τελεστή συσώρευσης A επί το υποσύνολο της γεωχωρικής στρώσης που εμπίπτει στον τομέα s .
5. Εφαρμογή βημάτων προετοιμασίας, που απαιτούνται από το συγκεκριμένο μοντέλο που χρησιμοποιείται. Εδώ γίνεται η διαχείριση των κενων τιμών, η κανονικοποίηση και άλλες μετατροπές, ανάλογα με τις ανάγκες του εκάστοτε αλγόριθμου.

3.7 Μοντελοποίηση Εκτίμησης Κινδύνου

3.7.1 Ορισμός Προβλήματος Πρόβλεψης

Το πρόβλημα που εξετάζεται στην παρούσα μελέτη είναι η εκτίμηση του κινδύνου εκδήλωσης πυρκαγιάς, ανεξάρτητα από την τελική έκταση που μπορεί να λάβει το συμβάν. Υπό την παραδοχή ότι κάθε περιστατικό που απαιτεί επέμβαση του Πυροσβεστικού Σώματος συνιστά συμβάν εκτός ελέγχου και, συνεπώς, ουσιώδη κίνδυνο, το ζητούμενο περιορίζεται στην πρόβλεψη της εκδήλωσης ή μη πυρκαγιάς.

Σε τυπική μορφή, το πρόβλημα διατυπώνεται ως πρόβλημα δυαδικής κατηγοριοποίησης με δύο κλάσεις, δηλαδή $f \ll \text{φωτιά} \gg$ και $f' \ll \text{όχι φωτιά} \gg$. Για κάθε στιγμιότυπο $w(s, t)$, και δεδομένου του συνόλου των μετρήσιμων χαρακτηριστικών $\mathbf{x}_{s,t} = (x_{1,s,t}, x_{2,s,t}, \dots, x_{m,s,t})$ επιδιώκεται η εκτίμηση της πιθανότητας $P(f | \mathbf{x}_{s,t})$, η οποία αντιστοιχεί στην πιθανότητα εκδήλωσης πυρκαγιάς στον χωρικό τομέα s κατά το χρονικό τμήμα t .

Η τελική πρόβλεψη προκύπτει μέσω κατωφλίου απόφασης a , σύμφωνα με τον κανόνα (3.18)

$$\hat{y}_{s,t} = \begin{cases} f, & P(f | \mathbf{x}_{s,t}) > a, \\ nf, & P(f | \mathbf{x}_{s,t}) \leq a. \end{cases} \quad (3.18)$$

Λαμβάνοντας υπόψη λειτουργικούς και περιβαλλοντικούς περιορισμούς, ο στόχος της πρόβλεψης μπορεί να μετασχηματιστεί σε πιο πρακτική μορφή. Να αλλάξει από την εκδήλωση πυρκαγιάς στο χρονικό τμήμα t στην εκδήλωση πυρκαγιάς εντός του χρονικού ορίζοντα $[t, t + h)$, όπου h είναι ο χρονικός ορίζοντας πρόβλεψης από την παρούσα χρονική στιγμή.

Κατά συνέπεια, για κάθε στιγμιότυπο $w(s, t)$, και δεδομένου του συνόλου των διαθέσιμων χαρακτηριστικών $\mathbf{x}_{s,t}$ επιδιώκεται πλέον η εκτίμηση της πιθανότητας $P(f_h | \mathbf{x}_{s,t})$, όπου f_h δηλώνει την εκδήλωση πυρκαγιάς στον χωρικό τομέα s εντός του χρονικού διαστήματος $[t, t + h)$.

Με άλλα λόγια, το πρόβλημα λαμβάνει τη μορφή (3.19)

$$(x_{1,s,t}, x_{2,s,t}, \dots, x_{n,s,t}) \rightarrow P(f_h). \quad (3.19)$$

Επομένως, ο στόχος των εκτιμήσεων των μοντέλων είναι "*Θα ξεκινήσει πυρκαγιά εντός των επόμενων h λεπτών*". Με αυτόν τον τρόπο, η μοντελοποίηση στρέφεται προς τον εντοπισμό των επικίνδυνων συνθηκών αντί για την απλή πρόβλεψη της εκδήλωσης πυρκαγιάς.

Ο στόχος ορίστηκε με αυτόν τον τρόπο διότι, από περιβαλλοντική άποψη, οι βραχυπρόθεσμες ακραίες καιρικές συνθήκες έχουν καθυστερημένο και συσσωρευτικό αντίκτυπο στην καύσιμη ύλη, καθώς οι υψηλές θερμοκρασίες και η χαμηλή υγρασία οδηγούν την βλάστηση σε κρίσιμα εύφλεκτη κατάσταση μετά από κάποιο διάστημα έκθεσης. Σε επιχειρησιακό επίπεδο, ένα σύστημα εκτίμησης κινδύνου είναι πιο χρήσιμο όταν δίνει ένα παράθυρο ενέργειας στους χρήστες του. Επιπροσθέτως, τα προκαταρκτικά πειράματα έδειξαν ότι αυτός ο ορισμός στόχου αποφέρει πιο σταθερές προβλέψεις.

Σημειώνεται ότι η μέθοδος πρόβλεψης έχει στη διάθεσή της αποκλειστικά τις τιμές των χαρακτηριστικών που αντιστοιχούν στο συγκεκριμένο s και t . Επομένως, η πρόβλεψη για τον χρονικό ορίζοντα h βασίζεται μόνο στην πληροφορία που είναι διαθέσιμη κατά τη χρονική στιγμή αναφοράς και δεν προϋποθέτει γνώση των μελλοντικών τιμών των μεταβλητών εντός του διαστήματος ορίζοντα πρόβλεψης.

Η τιμή h πρέπει να παρέχει έναν ικανοποιητικό συμβιβασμό επιχειρησιακής χρησιμότητας και περιβαλλοντικής πληρότητας. Με αυτήν την λογική επιλέχθηκαν οι 3 ώρες σαν ένας βραχυπρόθεσμος ορίζοντας συσσωρευμένου καιρικού φόρτου και χρήσιμο παράθυρο ενέργειας. Ο ορίζοντας h είναι μία βασική παράμετρος της μοντελοποίησης που κάνει το σύστημα και, επομένως, είναι προσαρμόσιμος, αλλά η πλειοψηφία των πειραμάτων θα γίνει με ορίζοντα κινδύνου 3 ωρών.

3.7.2 Μετρικές Αξιολόγησης

Ο βασικός στόχος των συστημάτων εκτίμησης κινδύνου, όπως το παρόν, είναι να εξάγουν προγνωστικά έγκυρα και ακριβή, και από αυτό προκύπτει το κριτήριο αποτελεσματικότητας του συστήματος: Η βαρύτητα δίνεται στην επιτυχία της ανίχνευσης των επικίνδυνων συνθηκών, διατηρώντας, όμως, όσο το δυνατόν υψηλότερο δείκτη ακρίβειας. Με άλλα λόγια, η βελτιστοποίηση ανίχνευσης του στόχου, σε συνδυασμό με τον περιορισμό των λάθος συναγερμών.

Οι εκτιμήσεις των μοντέλων μπορούν να συγκριθούν με τα επιβεβαιωμένα συμβάντα (*Ground Truth*) ώστε να είναι δυνατόν να αξιολογηθούν· από αυτήν την σύγκριση προκύπτουν οι 4 περιπτώσεις ενός πίνακα σύγχυσης (*Confusion Matrix*). Οι προβλέψεις που συμπίπτουν με την πραγματικότητα είναι τα Ορθώς Θετικά (*True Positive - TP*) και είναι η πιο άμεση δήλωση επίτευξης του στόχου. Στις περιπτώσεις που οι προβλέψεις δεν ανταποκρίνονται στην αλήθεια που προκύπτουν δύο καταστάσεις, κάθε μία με ξεχωριστό επιχειρησιακό αντίκτυπο. Στην περίπτωση που η πρόβλεψη ήταν Εσφαλμένα Θετική (*False Positive - FP*), έχει προβλεφθεί πυρκαγιά που δεν εκδηλώθηκε. Για ένα περίπλοκο φαινόμενο όπως οι πυρκαγιές, η ύπαρξη FP είναι αναμενόμενη, ωστόσο, ένας πολύ μεγάλος αριθμός FP αντιστοιχεί με δυσανάλογο αριθμό λάθος συναγερμών, γεγονός που προκαλεί ενόχληση και ενδεχόμενη σπατάλη πόρων. Αντίθετα, η περίπτωση της Εσφαλμένα Αρνητικής εκτίμησης (*False Negative - FN*) είναι δυνητικά καταστροφική, διότι υποδηλώνει μία πυρκαγιά που δεν προβλέφθηκε. Εάν οι εκτιμήσεις αξιοποιούνται έμπρακτα σε επιχειρησιακό επίπεδο, ένα ατυχές FN μπορεί να έχει πραγματικές συνέπειες σε περιουσίες και ανθρώπινες ζωές. Είναι ξεκάθαρο πως υπάρχει μία ασυμμετρία κόστους μεταξύ των καταστάσεων.

Δεδομένου του κριτηρίου αποτελεσματικότητας του συστήματος, η απόδοσή του μπορεί να μετρηθεί με καλά ορισμένες μεθόδους. Απευθείας από τις τιμές του πίνακα σύγχυσης προκύπτει η Ανάκληση (*Recall*) και η Ακρίβεια (*Precision*). Η πρώτη υποδεικνύει την ευστοχία των προβλέ-

ψεων, ήτοι το ποσοστό επικίνδυνων συνθηκών που εντοπίστηκαν, και είναι ευαίσθητη στα FN. Η δεύτερη δείχνει το ποσοστό των έγκυρων προβλέψεων για επικίνδυνες συνθήκες και εξαρτάται από τα FP. Η βελτίωση του εντοπισμού που σηματοδοτείται από την αύξηση της ανάκλησης σε συνδυασμό με την διατήρηση ικανοποιητικής ακρίβειας, είναι το κεντρικό μοτίβο της βελτιστοποίησης του συστήματος.

Ο γενικός τύπος του F-Score (3.20) δίνει τον αρμονικό μέσο της ακρίβειας και της ανάκλησης, επιτυγχάνοντας έτσι μία ισορροπία μεταξύ των δύο μετρικών. Η τιμή b είναι κλειδί για την ρύθμιση αυτής της ισορροπίας, προσδίδοντας βάρος στην Ανάκληση, άρα και τον εντοπισμό όσο περισσότερων πυρκαγιών γίνεται. Η μείωση της, από την άλλη, στρέφει το βάρος στην ακρίβεια και, επομένως, την ελαχιστοποίηση των λάθος συναγερμών. Δεδομένης της ασυμμετρίας του κόστους μεταξύ των δύο μετρικών, η τιμή 2 ($F2Score$) χρησιμοποιείται για την ρύθμιση του κατωφλιού πρόβλεψης. Αυτό υποδηλώνει συγκρατημένο ζήλο για την επιτυχή ανίχνευση πυρκαγιών, ώστε να μην είναι υπερβολικά πολυάριθμοι οι λάθος συναγερμοί.

$$F_b = \frac{(1 + b^2) \cdot \text{Precision} \cdot \text{Recall}}{b^2 \cdot \text{Precision} + \text{Recall}} = \frac{5 \cdot \text{Precision} \cdot \text{Recall}}{4 \cdot \text{Precision} + \text{Recall}} \quad (3.20)$$

Ο δείκτης Περιοχής Κάτωθεν της Καμπύλης Ακρίβειας - Ανάκλησης (PR-AUC) είναι μία ολιστική μετρική, η οποία αποτυπώνει την δυνατότητα ενός μοντέλου να βρίσκει με ευστοχία και ακρίβεια την θετική κλάση σε όλο το εύρος των κατωφλιών. Προκύπτει εξετάζοντας τα αποτελέσματα των προβλέψεων ενός μοντέλου πάνω σε ένα δοκιμαστικό σύνολο, και ολοκληρώνοντας για την ακρίβεια προς την ανάκληση, σε όλο το μήκος των πιθανοτήτων, από το 0 μέχρι το 1 σύμφωνα με την σχέση (3.21) $R = \text{Recall}$.

$$\text{PR-AUC} = \int_0^1 \text{Precision}(R) dR \quad (3.21)$$

Η αύξησή του, υποδηλώνει βελτίωση της ικανότητας του μοντέλου να εντοπίζει τον στόχο χωρίς να προκαλεί λάθος συναγερμούς. Λόγω της ανεξαρτησίας του από το κατώφλι απόφασης και της έμφασης που δίνει στην θετική κλάση με την χρήση της Ανάκλησης και της Ακρίβειας, συνιστάται για προβλήματα με δυσανάλογα μικρό πλήθος θετικών στιγμιότυπων.

Καθώς το PR-AUC εξαρτάται από τα στιγμιότυπα θετικής κλάσης, είναι σκόπιμο να το συγκρίνουμε σε σχέση με την επικράτηση (*Prevalence*) των θετικών ενδείξεων κινδύνου στο σύνολο των στιγμιότυπων (3.22). Ένας ταξινομητής που προβλέπει τυχαία, αναμένεται να πετύχει ακρίβεια παρόμοια με την επικράτηση της κλάσης. Μπορεί, επομένως, να εξακριβωθεί, πόσο καλύτερα ένα μοντέλο διαχωρίζει τον κίνδυνο σε σχέση με έναν τυχαίο ταξινομητή, υπολογίζοντας την ανύψωση (*Lift*) για το PR-AUC (3.23)

$$\text{Prevalence}^+ = \frac{N^+}{N} \quad (3.22)$$

$$\text{Lift}_{\text{PR-AUC}} = \frac{\text{PR-AUC}}{\text{Prevalence}^+} \quad (3.23)$$

Η λογαριθμική απώλεια (log loss), υποδεικνύει την ποιότητα των πιθανοτήτων που εκτιμάει το μοντέλο, τιμωρώντας τις λάθος προβλέψεις που γίνονται με υψηλή σιγουριά. Ο υπολογισμός της πάνω σε ένα σύνολο αποτελεσμάτων γίνεται με την μέση τιμή των λογαρίθμων των υπολογισμένων πιθανοτήτων $P(f)$, σε σχέση με τις πραγματικές τιμές y , όπως φαίνεται στο (3.24).

$$\text{LogLoss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(P(f | \mathbf{x}_i)) + (1 - y_i) \log(1 - P(f | \mathbf{x}_i))] \quad (3.24)$$

Η χρήση του λογαρίθμου έχει ως αποτέλεσμα οι μεγάλες αποκλίσεις μεταξύ υπολογισμένης πιθανότητας και πραγματικής τιμής να λαμβάνουν πολύ μεγαλύτερη ποινή, προκαλώντας μεγαλύτερη απώλεια. Είναι επιχειρησιακά σκόπιμο να σημάνει και το αντίστοιχο επίπεδο συναγερμού, όταν προβλέπεται η κλάση κινδύνου με υψηλή πιθανότητα, δηλαδή σιγουριά. Επομένως, ο έλεγχος της αυτοπεποίθησης του μοντέλου είναι επιθυμητός για συστήματα με επίπεδα κινδύνου και μπορεί να γίνει με log loss.

3.7.3 Στρατηγική Επιλογής Μοντέλων

Εφόσον το πρόβλημα αφορά κατηγοριοποίηση με πιθανότητες, οι κατηγοριοποιητές που υπολογίζουν πιθανότητες είναι η φυσική επιλογή για τον πυρήνα του μοντέλου πρόβλεψης. Οι αλγόριθμοι, πάνω στους οποίους χτίστηκαν τα μοντέλα, επιλέχθηκαν με γνώμονα την δυνατότητά τους να διαχειρίζονται τις ιδιαιτερότητες του κύριου συνόλου δεδομένων, ήτοι, ακραία κατανομή κλάσεων και ενδεχόμενες μη γραμμικές σχέσεις. Επιπλέον, για αυτό το στάδιο του έργου προτιμήθηκαν καλά τεκμηριωμένοι αλγόριθμοι με αποδεδειγμένα αποτελέσματα και επαρκές υλικό. Ο RF επιλέχθηκε με σημαντική υποστήριξη από την βιβλιογραφική έρευνα σε συναφή συστήματα και ο XGBoost [51] λόγω της συμβατότητάς του με τα κριτήρια. Ο LR προστέθηκε σαν έλεγχος γραμμικών σχέσεων και επίδρασης χαρακτηριστικών.

Κάθε μοντέλο έχει κάποιες βασικές ιδιότητες, με πρώτη και κυριότερη τον τύπο αλγόριθμου στον οποίο βασίζεται. Άλλες ιδιότητες είναι το σύνολο χαρακτηριστικών που χρησιμοποιεί για τις εκτιμήσεις, συγκεκριμένα βήματα προετοιμασίας όπως η κανονικοποίηση, ο χειρισμός κενών τιμών, η ισορροπία κλάσεων και ο ορίζοντας του στόχου. Επιπλέον, οι αλγόριθμοι έχουν ένα πλήθος Υπερπαραμέτρων εκπαίδευσης που έχουν επίδραση στην αποτελεσματικότητα του παραγόμενου μοντέλου.

Αναφορικά με τα βήματα προετοιμασίας του κάθε μοντέλου, τα μοντέλα βασισμένα στον LR και τον RF συμπεριλαμβάνουν και συμπλήρωση κενών τιμών με την διάμεση τιμή του χαρακτηριστικού, για καλύτερη σταθερότητα. Ο XGB, σύμφωνα με τους οδηγούς [52], δεν απαιτεί τέτοια διαχείριση. Επιπλέον, ο τρόπος διαχείρισης της άνισης κατανομής κλάσεων είναι βασικό ερώτημα της έρευνας και, για αυτό, εξετάζονται μοντέλα με ειδικό χειρισμό της κατανομής των κλάσεων, όπως υποδειγματοληψία και χωρίς ειδικό χειρισμό [12]. Τα μοντέλα που βασίζονται σε LR συμπεριλαμβάνουν κανονικοποίηση για πιο ερμηνεύσιμους συντελεστές, μετατροπή δύναμης που συστήνεται για μη γραμμικά προβλήματα ή πολυωνμικά χαρακτηριστικά (*Polynomial Features - PF*), με τα οποία εξετάζονται οι παράγοντες πρόβλεψης συνδυαστικά, αλλά αυξάνεται δραματικά ο χρόνος εκτέλεσης.

3.7.4 Διαμόρφωση Διαδικασίας Εκπαίδευσης

Τα ιστορικά δεδομένα, αφότου έχουν τακτοποιηθεί σε στιγμιότυπα με τον τρόπο που περιγράφει η ενότητα 3.6.4, υπόκεινται επεξεργασία για να μετατραπούν στα χαρακτηριστικά μοντελοποίησης (ενότητα 3.6.6). Αυτά αποτελούν και την είσοδο για τα μοντέλα πρόβλεψης. Η ένδειξη πυρκαγιάς προστίθεται με ερώτημα γεωχωρικής και χρονικής σύνδεσης και απλώνεται στα προηγούμενα στιγμιότυπα, σύμφωνα με τον ορισμό του στόχου (ενότητα 3.7.1). Τα στιγμιότυπα με την ένδειξη <<Πυρκαγιά τις επόμενες 3 ώρες>> είναι τα θετικά δείγματα και τα υπόλοιπα είναι τα αρνητικά.

Πριν την εκπαίδευση, το σύνολο δεδομένων χωρίζεται σε ένα μεγαλύτερο σύνολο εκπαίδευσης, πάνω στο οποίο θα προσαρμοστεί το μοντέλο πρόβλεψης, και ένα μικρότερο σύνολο δοκιμής στο

οποίο δεν έχει πρόσβαση κατά την εκπαίδευση. Η εμπειρία των Gholami κα. έδειξε πως ο τυχαίος διαμοιρασμός των στιγμιότυπων σε σύνολο εκπαίδευσης και δοκιμής παράγει πολύ αισιόδοξα αποτελέσματα και, αντ' αυτού, προτείνουν χρονικό διαχωρισμό [13]. Επιπλέον, εφόσον εφαρμόζεται ορίζοντας κινδύνου, τα στιγμιότυπα πριν την έναρξη της πυρκαγιάς στα οποία ανατέθηκε ένδειξη κινδύνου, έχουν συσχέτιση μεταξύ τους, και ο διαχωρισμός τους σε διαφορετικά σύνολα θα σημαίνει διαρροή πληροφορίας μεταξύ των συνόλων. Για να αποφευχθεί αυτό, ο διαχωρισμός γίνεται σε ημερολογιακά έτη, ξεκινώντας από την 1 Ιανουαρίου.

Στις πραγματικές συνθήκες, το μοντέλα θα πρέπει να βασιστούν αποκλειστικά σε ιστορικά δεδομένα, αλλά οι προβλέψεις τους αφορούν τρέχουσες συνθήκες. Ένας καλός τρόπος για να προσομοιωθεί αυτή η κατάσταση, είναι να χρησιμοποιηθούν τα πρώτα σε χρονολογική σειρά έτη για την εκπαίδευση των μοντέλων, και να δοκιμαστούν στα τελευταία. Έτσι, τα αποτελέσματα αναμένεται να είναι πιο αντιπροσωπευτικά [13].

Διαμορφώνοντας με αυτόν τον τρόπο την διαδικασία, διασφαλίζεται πως κάθε στιγμιότυπο του συνόλου εκπαίδευσης περιέχει μόνο:

- Τις συνθήκες που επικρατούσαν στον τομέα του κατά την αρχή του χρονικού του διαστήματος.
- Εάν προκλήθηκε πυρκαγιά τις επόμενες 3 ώρες, που σημαίνει ότι οι συνθήκες αποδείχθηκαν επικίνδυνες.

και για τα στιγμιότυπα του συνόλου δοκιμής ισχύει ότι:

- Περιέχουν μόνο τις συνθήκες που επικρατούσαν στον τομέα του κατά την αρχή του χρονικού του διαστήματος.
- Δεν περιέχουν κάποια πληροφορία σχετικά με το εάν αυτές οι συνθήκες αποδείχθηκαν επικίνδυνες.
- Δεν σχετίζονται με κανένα συμβάν που να υπάρχει στο σύνολο εκπαίδευσης.

3.7.5 Προσέγγιση Ρύθμισης Υπερπαραμέτρων

Για την καλύτερη αξιοποίηση των βασικών αλγορίθμων, έγινε ρύθμιση των υπερπαραμέτρων του κάθε μοντέλου εξετάζοντας τυχαία δείγματα από τον χώρο των πιθανών τιμών των υπερπαραμέτρων. Ο χώρος πιθανών τιμών διαμορφώνεται από το (εμπειρικά) αποδεκτό εύρος τιμών κάθε υπαρπαραμέτρου. Οι αλγόριθμοι που δέχονται περισσότερες παραμέτρους, όπως ο XGB, μπορεί να υποφέρουν από την εκθετική αύξηση της έκτασης του χώρου τιμών. Για αυτό, ο αριθμός των εξεταζόμενων συνόλων τιμών για τις υπερπαραμέτρους του κάθε μοντέλου, αυξάνεται εκθετικά με το πλήθος των παραμέτρων που δέχεται ο αλγόριθμος που χρησιμοποιεί.

Τα δείγματα αξιολογήθηκαν με κριτήριο την ελαχιστοποίηση της λογαριθμικής απώλειας σε δοκιμές Cross Validation - CV. Οι ομάδες της διαδικασίας CV διαχωρίζονται χρονικά, με το κάθε ημερολογιακό έτος του συνόλου εκπαίδευσης να αποτελεί και μία ομάδα. Το σετ υπερπαραμέτρων με την καλύτερη απόδοση χρησιμοποιείται για την δημιουργία του μοντέλου και τις δοκιμές.

Καθώς αυτή η αναζήτηση είναι μακράν το πιο χρονοβόρο στάδιο των δοκιμών, η επιλογή των εξεταζόμενων συνόλων τιμών υπερπαραμέτρων γίνεται με τυχαία δειγματοληψία από το καθορισμένο εύρος τιμών των μεταβλητών. Αυτή η τεχνική προτιμήθηκε, παρότι δεν είναι τόσο εξονυ-

χιστική όσο η αναζήτηση πλέγματος, διότι είναι σημαντικά ταχύτερη και θεωρήθηκε επαρκής για αυτή τη μελέτη.

3.7.6 Ορισμός Κατωφλιού και Δείκτη Κινδύνου

Οι οδηγοί χρήσης αλγορίθμων πιθανοτήτων, όπως ο οδηγός του scikit-learn, συχνά συνιστούν την ρύθμιση του κατωφλιού, ανάλογα με τους επιχειρησιακούς στόχους καθώς και σε περιπτώσεις άνισης κατανομής κλάσεων [53].

Για αυτήν την μελέτη, η επιλογή της τιμής a στην συνάρτηση απόφασης 3.18 γίνεται βάσει δοκιμών με την μέθοδο CV, στο σύνολο εκπαίδευσης με διαχωρισμό ανά έτος, όπως συμβαίνει και για την δοκιμή τιμών υπερπαραμέτρων 3.7.5. Με αυτόν τον τρόπο, διασφαλίζεται η ανεξαρτησία των αποτελεσμάτων, που προκύπτουν από την εφαρμογή των επιλεγμένων κατωφλιών, στις πιθανότητες που προβλέφθηκαν στο σύνολο δοκιμών. Εάν υπάρχουν τιμές που ικανοποιούν τον ζητούμενο στόχο για 80% επιτυχία στην πρόβλεψη, δηλαδή Ανάκληση 0.8, και μέχρι 4 λάθος συναγερμοί για κάθε επιτυχημένη πρόβλεψη, ήτοι 0.2 ακρίβεια, τότε το κατώφλι επιλέγεται αποκλειστικά από αυτές. Η τιμή που επιλέγεται είναι αυτή που καταγράφει το υψηλότερο F2Score στις προβλέψεις.

Η πρόβλεψη με χρήση μίας μόνο τιμής κατωφλιού είναι χρήσιμη για τη σύγκριση της απόδοσης των μοντέλων, αλλά δεν επαρκεί για τον επιχειρησιακό στόχο της παραγωγής μίας εύχρηστης κλίμακας επικινδυνότητας [30][14]. Στην περίπτωση ενός πιθανοτικού μοντέλου, τα επίπεδα κινδύνου μπορούν να συσχετιστούν με την εκτιμώμενη πιθανότητα εκδήλωσης πυρκαγιάς $P(f)$. Υψηλότερες τιμές αυτής της πιθανότητας υποδηλώνουν μεγαλύτερη βεβαιότητα του μοντέλου ως προς το ενδεχόμενο εκδήλωσης πυρκαγιάς και, συνεπώς, υψηλότερο επίπεδο κινδύνου [41].

Έστω

$$p = P(f), \quad p \in [0, 1], \quad (3.25)$$

και έστω μία κλίμακα επικινδυνότητας με n επίπεδα. Ορίζονται $n - 1$ κατώφλια

$$0 = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_{n-1} < \tau_n = 1. \quad (3.26)$$

Το επίπεδο κινδύνου $r(p)$ αποδίδεται τότε σύμφωνα με τον κανόνα

$$r(p) = k \iff \tau_{k-1} \leq p < \tau_k, \quad k = 1, 2, \dots, n. \quad (3.27)$$

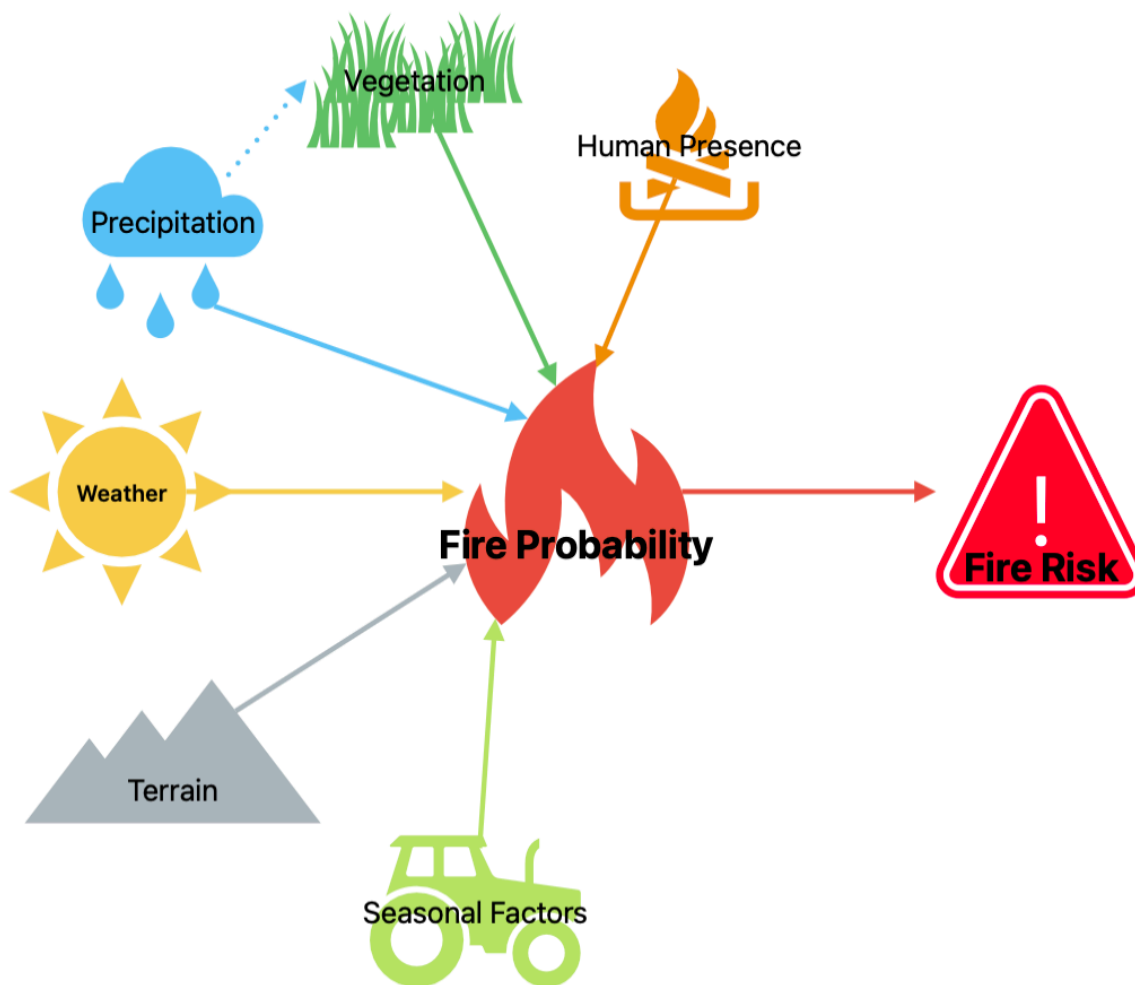
Για την τελευταία κλάση, το διάστημα μπορεί να οριστεί ως

$$r(p) = n \iff \tau_{n-1} \leq p \leq 1. \quad (3.28)$$

Ένας απλός τρόπος να ρυθμιστούν οι τιμές των κατωφλιών είναι από την αναμενόμενη ακρίβεια των προβλέψεων. Για παράδειγμα, ένα a_i που αναμένεται να δώσει $\text{Precision}_i > 0.5$, μπορεί να θεωρηθεί πολύ υψηλό ρίσκο, καθώς μία στις δύο περιπτώσεις που θα προκύψει, αναμένεται να εκδηλωθεί πυρκαγιά. Εφαρμόζοντας αυτή τη λογική, αντιστοιχίζεται κάθε κατώφλι a_i με ένα αντίστοιχο Precision_i ώστε να ισχύει η ανίσωση (3.29).

$$\text{Precision}_1 < \text{Precision}_2 < \dots < \text{Precision}_{n-1} \quad (3.29)$$

Εφαρμόζοντας περιορισμούς ακρίβειας στα αποτελέσματα του συνόλου εκπαίδευσης, μπορούν να εξαχθούν οι τιμές κατωφλιών που ικανοποιούν το κάθε επίπεδο ακρίβειας. Από αυτές τις τιμές,



Σχήμα 3.11: Πρότυπη Μοντελοποίηση Εκδήλωσης Πυρκαγιών

επιλέγεται αυτή με την καλύτερη απόδοση F2Score για να εκπροσωπήσει το αντίστοιχο επίπεδο κινδύνου.

Η χρήση ακρίβειας για τα κατώφλια των αποτελεσμάτων προϋποθέτει πως το μοντέλο μπορεί να διαχωρίσει καλά ένα τμήμα των στιγμιότυπων, ώστε να βρεθούν κατώφλια που επιτυγχάνουν την ακρίβεια του επιπέδου. Αυτό συνήθως δεν ισχύει όταν το μοντέλο είναι ακόμα σε αρχικό στάδιο. Ένας εναλλακτικός τρόπος υπολογισμού των κατωφλίων απόφασης, που χρησιμοποιείται ως εφεδρικός, είναι με την χρήση του FScore, με διάφορες τιμές b : Σε κάθε επίπεδο κινδύνου αντιστοιχεί μία τιμή b . Το κατώφλι στο κάθε επίπεδο κινδύνου ορίζεται ώστε το FScore για το αντίστοιχο b να είναι το μεγαλύτερο δυνατό. Όσο μεγαλώνει η τιμή του b , τόσο μειώνεται το κατώφλι πρόβλεψης που θα προκύψει, για αύξηση των επιτυχημένων ανιχνεύσεων και των λάθος συναγερμών.

3.8 Μεθοδολογία Αξιολόγησης Μοντέλων

Στο πρώτο στάδιο δοκιμάστηκαν όλα τα υποψήφια μοντέλα στο σύνολο των χαρακτηριστικών. Για κάθε μοντέλο έγινε η βελτιστοποίηση των υπερπαραμέτρων και ο υπολογισμός των κατωφλίων

και ακολούθησε η διαδικασία εκπαίδευσης με χρονολογικό διαχωρισμό, ώστε να μην υπάρξει σε κανένα σημείο διαρροή πληροφορίας μεταξύ συνόλου ελέγχου και εκπαίδευσης. Η διαδικασία εκπαίδευσης επαναλήφθηκε για 20 τυχαίες καταστάσεις στους μη ντετερμινιστικούς αλγόριθμους για την πληρότητα του αποτελέσματος.

Η ισορροπία μεταξύ ακρίβειας και ανάκλησης είναι σημαντική για την επίτευξη των στόχων του συστήματος, για αυτό το PR-AUC χρησιμοποιείται ως ο κύριος δείκτης απόδοσης των μοντέλων. Επειδή η ακριβής τιμή της πιθανότητας έχει επίσης σημασία για το σύστημα, καθώς συμβάλει στην ακριβέστερη αξιολόγηση επιπέδου κινδύνου, η λογαριθμική απώλεια χρησιμοποιείται συμπληρωματικά. Αυτές οι δύο μετρικές είναι ανεξάρτητες από το κατώφλι απόφασης και διαμορφώνουν ένα ολιστικό πλαίσιο αξιολόγησης.

Αναφορικά με την αξιολόγηση των προβλέψεων, τα δύο σκέλη του κριτηρίου αποτελεσματικότητας του συστήματος, πρώτη παράγραφος 3.7.2, είναι αντιθετικά. Οι προσπάθειες αύξησης του ενός επιφέρουν μείωση του άλλου. Η επιθυμητή ισορροπία μπορεί να επιτευχθεί μέσω δοκιμών και ρυθμίσεων της τιμής b του FScore.

Αυτοί με τα καλύτερα αποτελέσματα προχώρησαν σε δοκιμές με μέθοδο διαταραχών (*Perturbations*): Μέσω μικρών αλλαγών στην παραμετροποίηση, όπως την αφαίρεση χαρακτηριστικών πρόβλεψης, ή την μεταβολή του ορίζοντα στόχου, μπορεί να διαπιστωθεί κατά πόσο συγκεκριμένα χαρακτηριστικά και ρυθμίσεις είναι ωφέλιμα και να δοκιμαστεί η σταθερότητα του συστήματος.

Τα αποτελέσματα του τελικού μοντέλου με την καλύτερη απόδοση αναλύθηκαν ανά περίπτωση. Δόθηκε έμφαση στις περιπτώσεις όπου δεν προβλέφθηκε επιτυχώς μία πυρκαγιά, δηλαδή στα FN, ώστε να διαπιστωθεί σε ποια είδη στιγμιοτύπων έχει επιτυχία και σε ποιες καταστάσεις υστερεί.

3.9 Αρχές Σχεδίασης του Συστήματος

Στα πλαίσια αυτής της μελέτης θεωρείται δεδομένο ότι δεν θα βρεθεί ο βέλτιστος, ή ακόμα και κάποιος ιδιαίτερα αποτελεσματικός τρόπος μοντελοποίησης της εκτίμησης κινδύνου εκδήλωσης πυρκαγιάς. Για αυτό ένα κύριο μέλημα της έρευνας είναι η σχεδίαση ενός συστήματος που θα είναι σε θέση να υποστηρίξει την έρευνα, δοκιμή και εφαρμογή ποικίλων προσεγγίσεων μοντελοποίησης για οποιονδήποτε φορέα επιθυμεί να το υιοθετήσει.

3.9.1 Αρθρωτότητα και Επεκτασιμότητα

Η αρθρωτή σχεδίαση και υλοποίηση είναι προϋπόθεση για το μέλλον του συστήματος. Με αυτήν την αρχή, εφαρμόζεται ξεκάθαρος διαχωρισμός μεταξύ των διαφόρων υπηρεσιών, ώστε η κάθε μία να εξυπηρετεί τους στόχους της ανεξάρτητα με την υλοποίηση των άλλων. Σε χαμηλότερο επίπεδο, τα επιμέρους υποσυστήματα, τα οποία εκτελούν συγκεκριμένες διεργασίες, πρέπει να έχουν ξεκάθαρα όρια, ώστε να αποτελούν ανεξάρτητα μέλη ενός αρθρωτού συνόλου από εργαλεία, που όλα μαζί φέρνουν εις πέρας τις ζητούμενες εργασίες. Ακολουθώντας αυτό το μοτίβο, το σύστημα παραμένει ανοιχτό στην σταδιακή αναβάθμιση και ενσωμάτωση νέας λειτουργικότητας κατά τις ανάγκες και απαιτήσεις που θα προκύψουν.

3.9.2 Ευελιξία στην Παραμετροποίηση

Για να διευκολυνθεί η τρέχουσα αλλά και οποιεσδήποτε μελλοντικές προσεγγίσεις στην μοντελοποίηση, είναι αναγκαίο τα υποσυστήματα να σχεδιαστούν έτσι, ώστε να είναι σε θέση να υποστηρίξουν οποιεσδήποτε παραμέτρους μοντελοποίησης ενδέχεται να χρειαστούν. Βασικές τιμές όπως η συχνότητα μετρήσεων, η χωρική ανάλυση των τομέων και η μέθοδος διαχωρισμού των περιοχών είναι προσαρμόσιμες, και, επιπροσθέτως, υπάρχει πρόβλεψη για χρήση εναλλακτικών μεθόδων σε σημεία του συστήματος όπου αυτό είναι σκόπιμο, όπως η παρεμβολή των καιρικών συνθηκών. Παρέχεται πλήρης ελευθερία και στην επιλογή χαρακτηριστικών, με την υποστήριξη πολλών τύπων παραμέτρων σε κάθε συνδυασμό που θα εξυπηρετούσε την μέθοδο μοντελοποίησης.

3.9.3 Φορητότητα

Η δυνατότητα του Aristaeos να τεθεί σε λειτουργία με ελάχιστη ρύθμιση, για οποιονδήποτε οργανισμό και περιοχή διαθέτει σταθμούς και ιστορικό, είναι ένα βασικό αξίωμα του συστήματος. Η εφαρμογή είναι ανά πάσα στιγμή έτοιμη να τεθεί σε λειτουργία, και, αφότου έχουν καταχωρηθεί οι σταθμοί μετρήσεων, να αρχίσει να καταχωρεί δεδομένα. Εάν έχει καταχωρηθεί εξ αρχής ή έχει δημιουργηθεί κατά την λειτουργία ένα αρχείο μετρήσεων, το μοντέλο πρόβλεψης μπορεί να εκπαιδευτεί και να παράγει τα προγνωστικά του. Αυτό επιτυγχάνεται σχεδιάζοντας γενικευμένες διαδικασίες, όπου οι λεπτομέρειες, όπως η περιοχή μελέτης ή οι πηγές των δεδομένων, λειτουργούν ως παράμετροι. Επιπλέον, ορίζονται τα πρότυπα που ακολουθεί το σύστημα αλλά και οι απαραίτητες στρώσεις μετάφρασης, ώστε τα εξωτερικά στοιχεία να εναρμονίζονται με τα πρωτόκολλα του συστήματος.

3.9.4 Ζωντανή Εκτίμηση Κινδύνου

Η δυνατότητα ζωντανών ή σχεδόν ζωντανών προγνωστικών παρουσιάζει ιδιαίτερο ενδιαφέρον για τα συστήματα έγκαιρης προειδοποίησης [26]. Ο στόχος της περιοδικής ανάκτησης μετρήσεων από τους σταθμούς και αναπροσαρμογής των εκτιμήσεων ανά τακτά χρονικά διαστήματα, είναι μία δυναμική και άμεσα αξιοποιήσιμη ένδειξης εγρήγορσης για τις αρμόδιες υπηρεσίες και τους πολίτες. Η συχνή αναπροσαρμογή αναμένεται να αντικατοπτρίζει καλύτερα την πραγματικότητα στο πεδίο που διαμορφώνουν οι μεταβαλλόμενες περιβαλλοντικές συνθήκες, αλλά και να εντοπίζει και να προβάλλει τις ιδιάζουσες και οριακές περιπτώσεις, που έχουν αυξημένη πιθανότητα να οδηγήσουν σε πυρκαγιά.

3.9.5 Συνεχής Βελτίωση

Η προβολή συμπερασμάτων που προέκυψαν από ανάλυση μακροχρόνιων αρχείων δεδομένων σε τρέχουσες συνθήκες, ενέχει την παγίδα της Μετατόπισης Πλαισίου (Context Drift), όπου τα μοτίβα πάνω στα οποία εκπαιδεύτηκε ένα σύστημα ML έχουν μεταλλαχθεί ή δεν ισχύουν πλέον. Τα μεταβαλλόμενα κλιματικά δεδομένα [25][24][9] θα μπορούσαν να καταστήσουν αναποτελεσματικό ένα μοντέλο που για περίοδο κάποιον ετών είχε επιτυχία. Η ιδιότητα που έχει το σύστημα να εμπλουτίζει το αρχείο του με την συνεχή καταγραφή νέου υλικού κατά την διάρκεια της λειτουργίας του, αναμένεται να επιτρέψει τα μοντέλα να επανεκπαιδεύονται στα μεταβαλλόμενα κλιματικά

μοτίβα. Έτσι, είναι δυνατό να διατηρούν ή ακόμα και να βελτιώσουν την απόδοση τους παρά την πάροδο του χρόνου.

3.10 Μεθοδολογικοί Περιορισμοί και Αδυναμίες

Αν και το αντικείμενο και οι στόχοι της παρούσας μελέτης χαρακτηρίζονται από φιλοδοξία, το εύρος της είναι περιορισμένο, με αποτέλεσμα να επιβάλλονται κάποιοι συμβιβασμοί και περιορισμοί στην μεθοδολογία. Τα προβλεπόμενα κενά στην μοντελοποίηση είναι και ο λόγος που ένα μεγάλο μέρος των εργασιών αφιερώθηκε στο χτίσιμο ενός συστήματος που θα υποστηρίξει επακόλουθες δοκιμές με νέες, πιο στιβαρές προσεγγίσεις μοντελοποίησης.

3.10.1 Ελλειπείς Παράγοντες Εκδήλωσης Πυρκαγιών

Παρότι οι αρχές που καθοδήγησαν την μοντελοποίηση είναι καλά τεκμηριωμένες από την βιβλιογραφία και τις μελέτες πάνω στις δασικές πυρκαγιές, αυτό δεν ισχύει για τις μεθόδους έμμεσης προσέγγισης παραγόντων που χρησιμοποιήθηκαν. Οι δύο υποθέσεις που έγιναν για να προχωρήσει η μοντελοποίηση με τα διαθέσιμα δεδομένα, είναι πολύ πιθανό να αγγίζουν μόνο ένα μικρό μέρος των διεργασιών που επιχειρούν να αποτυπώσουν και δεν έχουν επιβεβαιωθεί πειραματικά για την περιοχή και τους σκοπούς της μελέτης. Έτσι, ενώ η υγρασία και ο όγκος της βλάστησης, καθώς και η ανθρώπινη κινητικότητα έχουν αποδεδειγμένη σύνδεση με την εκδήλωση δασικών πυρκαγιών, παραμένουν ανοιχτά ερωτήματα για τον τρόπο με τον οποίο αναπαριστώνται στην πρότυπη μοντελοποίηση. Ο τρόπος προσέγγισης της κατάστασης της βλάστησης αποκλειστικά μέσω της βροχόπτωσης, αποδεικνύεται εξαιρετικά απλοϊκός σε σχέση με αντίστοιχες τεχνικές που παρουσιάζονται στην βιβλιογραφία οι οποίες συνδυάζουν πλήθος καιρικών μεταβλητών και θερμοδυναμικών συναρτήσεων [12].

Έπειτα, αναγνωρίζεται πως το σύνολο των χαρακτηριστικών που χρησιμοποιήθηκαν για την μοντελοποίηση δεν αντιπροσωπεύουν επαρκώς την "εξίσωση των πυρκαγιών". Πράγματι, μεταβλητές που αναγνωρίζονται σε μεγάλο βαθμό ως παράγοντες που επηρεάζουν τις πυρκαγιές έχουν αγνοηθεί. Αυτό συνέβη με τον άνεμο, διότι οι μετρήσεις των οργάνων στα πρωτογενή δεδομένα ήταν προβληματικές, αλλά και με γεωχωρικές μεταβλητές όπως η κάλυψη του εδάφους και η χρήση γής. Είναι αναμενόμενο πως οι μεταβλητές που αξιοποιήθηκαν από την μελέτη δεν επαρκούν από μόνες τους για την δημιουργία ικανοποιητικού μοντέλου. Ειδικά η απουσία ενός πυρικού παράγοντα τόσο σημαντικού όσο ο άνεμος σίγουρα θα επιφέρει πλήγμα στην τελική απόδοση.

Επιπροσθέτως, εκτιμάται ότι η γεωχωρική ανάλυση της περιοχής μελέτης έκτασης 4.000 χμ² σε 8 υποπεριοχές είναι πενιχρή για την αποτελεσματική διερεύνηση των χαρακτηριστικών της εκάστοτε πυρκαγιάς. Εφόσον αναγνωριστεί πως η τοπολογία της Κρήτης υποστηρίζει τα χαρακτηριστικά μικροκλίματα των Μεσογειακών νήσων, είναι πολύ πιθανό οι μετρήσεις ενός μετεωρολογικού σταθμού να είναι ανακριβείς για τις τιμές σε μία περιοχή 50 χμ μακριά. Η χαμηλή γεωχωρική ανάλυση, καθιστά την χρήση γεωχωρικών χαρακτηριστικών λιγότερο επωφελή, καθώς τιμές όπως το ποσοστό καλλιεργήσιμης γης, η καμμένη έκταση τα προηγούμενα έτη, ή η πυκνότητα υποδομών, χάνουν την δυνατότητα να περιγράψουν λεπτομερώς την κατάσταση του εδάφους στον τομέα. Τα κελιά πλευράς 500 μ που επέλεξαν οι Gholami κα. είναι ένας πολύ πιο αναλυτικός [13], αλλά και υπολογιστικά απαιτητικός τρόπος. Αν και είναι διαθέσιμη στο σύστημα η δυνατότητα διαχωρισμού του εδάφους σε παρόμοια ανάλυση, προτιμήθηκε ένας χωρισμός με πολύ μικρότερο αριθμό

τομέων, για τους σκοπούς της αρχικής μοντελοποίησης.

3.10.2 Αδυναμίες Μεθόδου Μοντελοποίησης Πυρκαγιών

Στον πυρήνα της, η γενικότερη μεθοδολογία μοντελοποίησης που προβλέπεται για το σύστημα είναι σχετικά απλή: Εφαρμογή μηχανικής χαρακτηριστικών σε ιστορικό αρχείο, χρήση του συνόλου δεδομένων για εκπαίδευση αλγόριθμου ML, παραγωγή χαρακτηριστικών από ζωντανές μετρήσεις και μετάφραση της προβλεπόμενης πιθανότητας εκδήλωσης σε επίπεδο κινδύνου. Μέσα σε αυτό το πλαίσιο παρέχονται αυξημένες δυνατότητες προσαρμογής από το σύστημα, αλλά το ίδιο το πλαίσιο ενδέχεται να είναι περιοριστικό. Η μεθοδολογία αυτή σχεδιάστηκε βάσει παρόμοιων μελετών των προηγούμενων ετών, αλλά μπορεί να βρίσκεται ήδη στα όρια των δυνατών επιδόσεων της· κάποιες μελέτες δείχνουν πως τα περίπλοκα μη-γραμμικά μοντέλα είναι επιρρεπή σε υπερπροσαρμογή (overfitting) στα δεδομένα εκπαίδευσης [13] ή πως δεν ανταποκρίνονται επαρκώς στις ανάγκες χωρικής μοντελοποίησης [12].

Ίσως, η μεθοδολογία με το υψηλότερο ταβάνι σε αυτό το πρόβλημα να βρίσκεται στις τεχνικές DL, τις οποίες δεν υποστηρίζει πλήρως η παρούσα υλοποίηση. Εναλλακτικά, είναι πιθανό να επωφελείται η διαδικασία από συγκεκριμένα επιπλέον ενδιάμεσα βήματα, όπως η συσταδοποίηση των πυρκαγιών σε κατηγορίες, ή η εφαρμογή συνδυαστικών μεθόδων. Αυτά, βέβαια, θα απαιτούσαν επιπλέον μελέτη και σχεδίαση.

3.10.3 Προβλήματα Συνόλου Δεδομένων

Αν και οι πηγές των πρωτογενών δεδομένων θεωρήθηκαν αξιόπιστες και διενεργήθηκαν βασικοί έλεγχοι εγκυρότητας κατά την ανάλυση του αρχείου 3.5.1, ενδέχεται να ενέχουν παγίδες που δεν προβλέφθηκαν στην μεθοδολογία. Ήδη εντοπίστηκε η ελλαττωματική καταγραφή μετρήσεων ανέμου και εξάτμισης, η απουσία των οποίων προβλέπεται να δημιουργήσει πρόβλημα. Μία σειρά λάθος μετρήσεων ή κακών και άσχετων καταγραφών που διέφυγε της προσοχής, μπορεί να καλύψει με θόρυβο τα βαθύτερα μοτίβα και να επηρεάσει αρνητικά την δυνατότητα των μοντέλων ML.

Ιδιαίτερο ενδιαφέρον παρουσιάζει ο αριθμός των πυρκαγιών που καταμετρούν οι Papagiannaki κα., ο οποίος διαφέρει σημαντικά από αυτόν που υπολογίζεται στο αρχείο της μελέτης, παρότι η πηγή είναι η ίδια και εφαρμόζονται φίλτρα παρόμοια με αυτά που περιγράφονται, γεγονός που υποδηλώνει παρέκκλιση στην μεθοδολογία φιλτραρίσματος των πυρκαγιών [41].

Επιπλέον, τα πέντε έτη μετρήσεων που χρησιμοποιούνται για την εκπαίδευση εκ πρώτης όψεως φαίνονται σαν μία μεγάλη περίοδος, αλλά αφότου γίνουν οι διαχωρισμοί που απαιτούνται για την ορθή εκπαίδευση των μοντέλων, είναι πιθανό να αποδειχθεί ότι δεν επαρκούν. Ακόμα και εάν γίνει επιτυχώς η μετατροπή των τοποθεσιών σε συντεταγμένες για τις πυρκαγιές παλαιότερες του 2020, η μεταφορά ενός μοντέλου, που έχει εκπαιδευτεί με πολύ προγενέστερα δεδομένα στις τωρινές και μελλοντικές συνθήκες, είναι επίφοβη λόγω context drift, όπως εξηγήθηκε στην υποενότητα 3.9.5.

3.10.4 Επιχειρησιακές Αδυναμίες

Το προτεινόμενο σύστημα λειτουργεί βάσει της εκπαίδευσης στα μοτίβα μίας γεωγραφικής περιοχής. Επομένως, υπάρχουν αμφιβολίες και ως προς την αποτελεσματικότητα της μεταφοράς των εκπαιδευμένων μοντέλων αυτούσιων σε άλλες περιοχές [45], ειδικά αυτές όπου επικρατούν διαφορετικές κλιματικές συνθήκες και ανθρώπινες συνήθειες.

Για μία νέα περιοχή, εφόσον δεν υπάρχει έτοιμο μοντέλο που να ταιριάζει σε αυτήν, η τοποθέτηση του Aristaeos στην παραγωγή απαιτεί και την ύπαρξη ιστορικού αρχείου, συν ένα σύνολο πηγών με σχεδόν ζωντανή ροή δεδομένων, συμπεριλαμβανομένου δικτύου μετεωρολογικών σταθμών με επαρκή πυκνότητα. Αυτό ενδέχεται να αποδειχθεί εξαιρετικά περίπλοκο, όχι μόνο λόγω της διαχείρισης των σταθμών, αλλά και λόγω της ανάγκης χρήσης γεωχωρικών στρωμάτων που απεικονίζουν ταχέως μεταβαλλόμενες συνθήκες. Τα τελευταία ενδέχεται να είναι απρόσιτα ή απλά να μην υφίστανται.

Κατά την εξαγωγή κατωφλιών απόφασης, η τιμή της απαιτούμενης ακρίβειας (ή του b) για κάθε επίπεδο κινδύνου είναι στατική και αυθαίρετη. Τα αποτελέσματα επιπέδου κινδύνου που προέρχονται δεν ανταποκρίνονται και καλά ορισμένες επιχειρησιακές ανάγκες ή μαθηματικά τεκμηριωμένες μεθόδους.

3.11 Σύνοψη Κεφαλαίου

Η μεθοδολογία που ακολουθήθηκε είχε τον απώτερο σκοπό της γεφύρωσης της επιστημονικής γνώσης με το πληροφοριακό σύστημα. Μελετώντας τις εξελίξεις και εφαρμόζοντας τα συμπεράσματα των προηγούμενων ερευνών, η παρούσα μελέτη αποσκοπεί κυρίως στην προετοιμασία του εδάφους για την συνεχή και σταθερή πρόοδο και βελτίωση. Με αυτή την προοπτική, η πρότυπη μοντελοποίηση της εκδήλωσης πυρκαγιών για αυτήν την μελέτη παραμένει συνειδητά απλή, ενώ η υλοποίηση της υποδομής του Aristaeos δημιουργήθηκε για να φιλοξενήσει αυτήν την πρόοδο και να της δώσει επιχειρησιακή υπόσταση.

Κεφάλαιο 4

Σχεδιασμός και Υλοποίηση

4.1 Επισκόπηση Κεφαλαίου

Σε αυτό το κεφάλαιο θα παρουσιαστεί η σχεδίαση και η υλοποίηση του προτεινόμενου συστήματος Aristaeos για εποπτεία και εκτίμηση κινδύνου δασικών πυρκαγιών. Αφού παρατεθούν οι κύριες αποφάσεις μηχανικής λογισμικού, θα περιγραφούν τα τμήματα του συστήματος, η ροή των δεδομένων, η ενσωμάτωση ML, η διαδικασία ανάπτυξης, εγκατάστασης και διάθεσης. Τέλος, θα γίνει αναφορά στους περιορισμούς της προσέγγισης υλοποίησης καθώς και στην λειτουργικότητα που ακόμη απαιτείται για ένα πλήρως παραγωγικό σύστημα.

4.2 Απαιτήσεις Συστήματος

4.2.1 Πλαίσιο Υλοποίησης

Για τους σκοπούς της πρωτοτύπησης, κάποια τμήματα του προτεινόμενου συστήματος έχουν υλοποιηθεί σε μεγάλο βαθμό, και κάποια άλλα σε μικρότερο. Η μεγαλύτερη βάση δόθηκε σε λειτουργικά μέρη που υποστηρίζουν την διαδικασία της σχεδόν-ζωντανής εκτίμησης κινδύνου. Αυτό το σύνολο περιλαμβάνει τα υποσυστήματα που ενσωματώνουν ML, ήτοι, το υποσύστημα εκπαίδευσης και το υποσύστημα πρόβλεψης, το υποσύστημα προεπεξεργασίας δεδομένων, το οποίο δημιουργεί τα χαρακτηριστικά εισόδου που είναι απαραίτητα για την λειτουργία των μοντέλων ML, η γραμμή ανάκτησης, μετατροπής και καταχώρησης δεδομένων (Extract Transform Load - ETL Pipeline), η οποία είναι αναγκαία για την ενημέρωση των προβλέψεων και την επέκταση του αρχείου, και, φυσικά η βάση δεδομένων - ΒΔ για την αποθήκευση δεδομένων και παραμέτρων. Τα παραπάνω τμήματα έχουν ολοκληρωμένη την βασική λειτουργικότητα και χρήζουν μόνο προσθήκης συμπληρωματικών λειτουργιών, βελτίωσης χειρισμού ειδικών περιπτώσεων και βελτιστοποιήσεων στην απόδοση.

Η ανάπτυξη των υπόλοιπων υποσυστημάτων έχει προχωρήσει σε πολύ μικρότερο βαθμό, καθώς η συνεισφορά τους στην διαδικασία εκτίμησης κινδύνου δεν είναι άμεση. Ο σκελετός του υποσυστήματος επιτήρησης είναι λειτουργικός: Εξυπηρετεί πλήρως την ροή ETL και το υποσύστημα προβλέψεων και είναι δυνατόν να αναπτυχθεί και για τα υπόλοιπα τμήματα, για κάθε ανάγκη επιτήρησης λειτουργίας του συστήματος που θα προκύψει.

Η διαδικτυακή προγραμματιστική διεπαφή (*web API*) και η διαδικτυακή διεπαφή χρήστη

(*frontend*), αν και προορίζονται να χρησιμοποιηθούν ως πλήρες οικοσύστημα διαχείρισης του συστήματος, εντούτοις, στην πρωτότυπη υλοποίηση, διαθέτουν μόνο τους μηχανισμούς που επιτρέπουν τον διαμοιρασμό και την παρουσίαση των εκτιμήσεων.

Η διαδικτύωση και η ασφάλεια είναι καίριες για το τελικό στάδιο του Aristaeos ως παραγωγική υπηρεσία, αλλά βρίσκονται εκτός του πλαισίου μελέτης. Τα φυσικά όργανα μετρήσεων θεωρούνται εξωτερική πηγή δεδομένων, άρα, η εγκατάσταση και η συντήρησή τους είναι πέρα από το αντικείμενο της μελέτης.

4.2.2 Λειτουργικές Απαιτήσεις

Οι λειτουργικές απαιτήσεις, δηλαδή οι κύριες δυνατότητες που πρέπει να παρέχει το σύστημα, καθορίζονται από την φύση του ως ένα εργαλείο εκτίμησης κινδύνου καθοδηγούμενο από δεδομένα. Οι απαιτήσεις μπορούν να χωριστούν βάσει του ρόλου τους σε αυτήν την αρχιτεκτονική.

Απαιτήσεις Συστήματος Καθοδηγούμενου από Δεδομένα

Το σύστημα μελέτης πυρκαγιών που βασίζεται σε δεδομένα πρέπει να υποστηρίζει:

- Καταχώρηση και αρχειοθέτηση Ιστορικών Δεδομένων Πυρκαγιών.
- Περιοδική ανάκτηση, αποθήκευση και αρχειοθέτηση καιρικών μετρήσεων από μετεωρολογικούς σταθμούς ή APIs.
- Ανάκτηση, αποθήκευση, αρχειοθέτηση και ανανέωση γεωχωρικών στρώσεων δεδομένων.
- Την δημιουργία των απαιτούμενων χαρακτηριστικών εισόδου που έχουν ανάγκη τα μοντέλα ML, μέσω της εναρμόνισης, επεξεργασίας και χωροχρονικής αντιστοίχισης των συνθηκών μεταξύ τους και με το αρχείο των πυρκαγιών.

Απαιτήσεις Εκτίμησης Κινδύνου

Το σύστημα που αποσκοπεί στην πρακτική εκτίμηση κινδύνου πρέπει να υποστηρίζει:

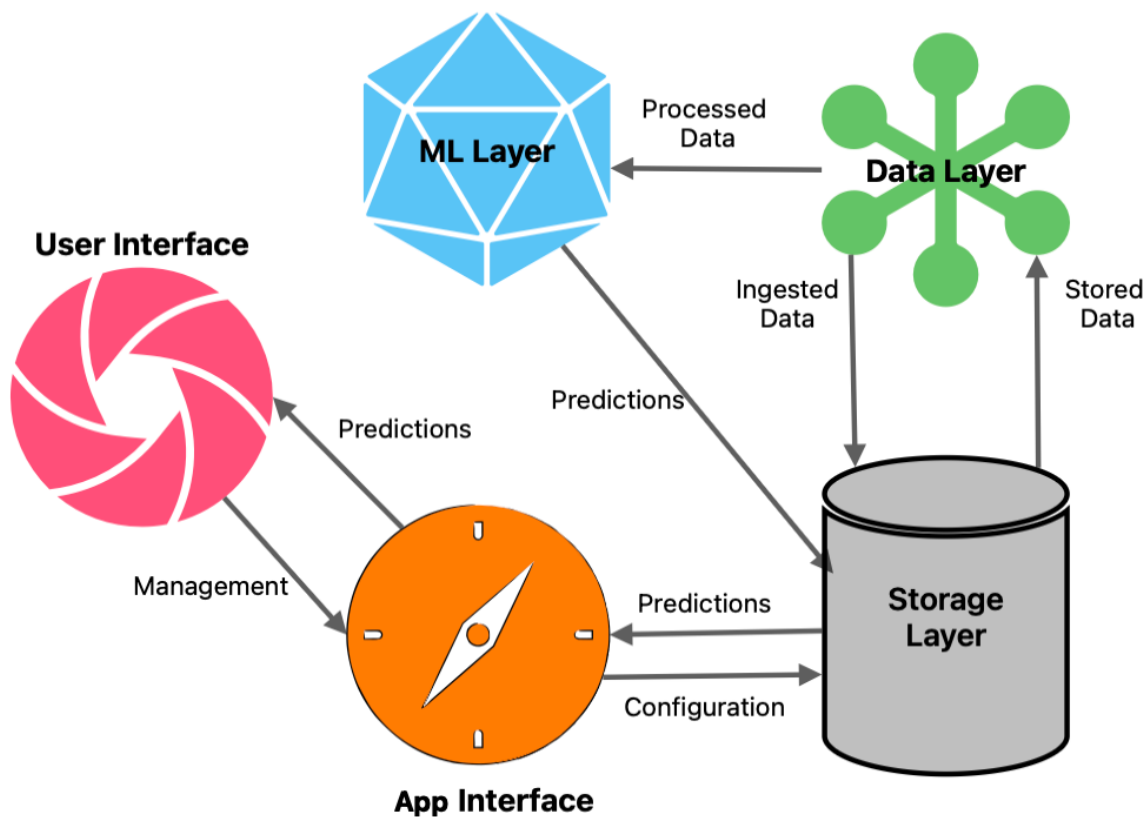
- Εκπαίδευση, αξιολόγηση και χρήση μοντέλων ML για την πρόβλεψη του κινδύνου πυρκαγιάς.
- Παραγωγή πιθανοτήτων και επιπέδου κινδύνου πυρκαγιάς για δεδομένες χρονικές στιγμές και τομείς.
- Διάθεση των προβλέψεων σε σχεδόν-ζωντανό (εντός λίγων λεπτών) επίπεδο μέσω διαδικτυακού API.
- Παρουσίαση του κινδύνου σε μορφή αντιληπτή από τους χρήστες.

4.2.3 Μη Λειτουργικές Απαιτήσεις

Στις μη λειτουργικές απαιτήσεις εντάσσονται προδιαγραφές που διασφαλίζουν την εύρυθμη και αποτελεσματική λειτουργία και συντήρηση του συστήματος:

- Αποτελεσματική διαχείριση μεγάλου όγκου δεδομένων.
- Καταγραφή σφαλμάτων εκτέλεσης και ποιότητας μετρήσεων και ενημέρωση των χειριστών και των χρηστών.
- Υποστήριξη βασικής επιτήρησης εκτέλεσης εργασιών.
- Αναπαραγωγιμότητα των εργασιών, συμπεριλαμβανομένης της εγκατάστασης.
- Ρυθμιζόμενες μεταβλητές περιβάλλοντος και εκτέλεσης των εργασιών.
- Αρθρωτότητα για ευκολία αναβάθμισης και διαχείρισης των υποσυστημάτων.

4.3 Αρχιτεκτονική Συστήματος



Σχήμα 4.1: Επισκόπηση Υψηλού Επιπέδου

4.3.1 Περιγραφή Συστήματος

Στο υψηλότερο επίπεδο, το σύστημα αποτελείται από λίγα εννοιολογικά στρώματα, που υλοποιούνται σε διακριτές υπηρεσίες, οι οποίες επικοινωνούν μεταξύ τους για την επίτευξη των στόχων του συστήματος, όπως φαίνεται στο Σχ. 4.1. Τα στρώματα-υπηρεσίες, είναι διαμορφωμένα με τέτοιο τρόπο, ώστε η τεχνολογίες υλοποίησης τους να μην επηρεάζουν τη λειτουργία των άλλων υπηρεσιών, εφόσον τα δεδομένα διακινούνται μεταξύ τους με τον προβλεπόμενο τρόπο. Με αυτήν την σχεδίαση είναι δυνατόν οι υπηρεσίες που συνθέτουν το σύστημα να συντηρούνται, αναβαθμίζονται και επεκτείνονται ανεξάρτητα.

Κομβικό για το σύστημα είναι το στρώμα αποθήκευσης, στο οποίο καταχωρείται το αρχείο της εφαρμογής και, από το οποίο, πραγματοποιείται ανάκτηση δεδομένων από τις άλλες υπηρεσίες. Το στρώμα δεδομένων οργανώνει και πραγματοποιεί την συλλογή πρωτογενών δεδομένων από τις πηγές που αξιοποιεί το σύστημα, καθώς και την επεξεργασία τους, ενώ το στρώμα ML χειρίζεται την εκπαίδευση και χρήση των μοντέλων πρόβλεψης. Η διεπαφή της εφαρμογής παρέχει πρόσβαση σε επιλεγμένες δυνατότητες και πόρους του συστήματος σε εξωτερικούς πράκτορες, όπως η διεπαφή του χρήστη, μέσω της οποίας ο χρήστης βλέπει τα αποτελέσματα και ρυθμίζει την λειτουργία.

4.3.2 Επιμέρους Υποσυστήματα

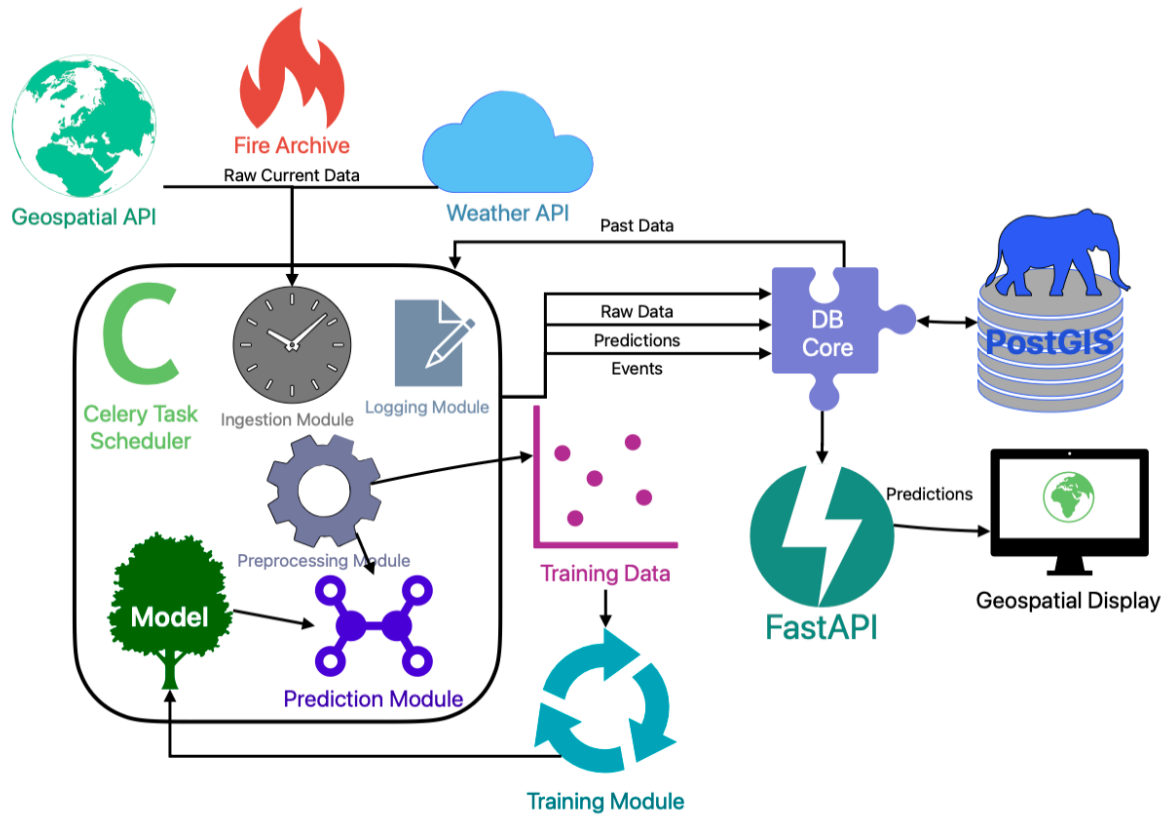
Σε λειτουργικό επίπεδο, οι υπηρεσίες αποτελούνται από εξειδικευμένα υποσυστήματα, σε κάθε ένα από τα οποία ανατίθεται η εκτέλεση συγκεκριμένου τμήματος της λειτουργικότητας του συστήματος, όπως απεικονίζει το Σχ.4.2. Για το στρώμα αποθήκευσης λειτουργεί σχεσιακή ΒΔ, με την οποία τα υπόλοιπα υποσυστήματα αλληλεπιδρούν μέσω μίας διεπαφής, η οποία παρέχει υπορουτίνες για αποθήκευση και ανάκτηση. Στο στρώμα των δεδομένων, υπάρχει το υποσύστημα ανάκτησης, το οποίο παραλαμβάνει και αποθηκεύει δεδομένα από εξωτερικές πηγές, και το υποσύστημα επεξεργασίας δεδομένων, που δημιουργεί τα χαρακτηριστικά που χρησιμοποιούνται για τις εκτιμήσεις. Όσον αφορά το στρώμα ML, η εκπαίδευση των μοντέλων πρόβλεψης εκτελείται μέσω του υποσυστήματος εκπαίδευσης και η παραγωγή των προβλέψεων γίνεται από το υποσύστημα πρόβλεψης, με χρήση των εκπαιδευμένων μοντέλων.

Ένα υποσύστημα χρονοδρομολόγησης καλεί σε σταθερό ρυθμό τα υποσυστήματα ανάκτησης, επεξεργασίας και πρόβλεψης, ώστε να πραγματοποιείται η περιοδική ανανέωση των προγνωστικών. Στο στρώμα διεπαφής των δεδομένων λειτουργεί εφαρμογή διαδικτυακού API, ενώ στο στρώμα διεπαφής του χρήστη παρέχεται στατική σελίδα γεωχωρικής οπτικοποίησης.

4.3.3 Ροή Δεδομένων

Η πορεία των δεδομένων ξεκινάει εκτός των ορίων του πληροφοριακού συστήματος, στα διάφορα μέσα παραγωγής πρωτογενών μετρήσεων, όπως μετεωρολογικούς σταθμούς, αισθητήρες, δορυφόρους και ανθρώπινες καταγραφές. Οι μηχανισμοί ανάκτησης του συστήματος αναλαμβάνουν να συγκεντρώσουν τα πρωτογενή δεδομένα από τα κανάλια στα οποία διατίθενται και να τα μεταφέρουν στους χώρους αποθήκευσης, όπου καταχωρούνται και, τελικά, αρχειοθετούνται μόνιμα. Αφού τα πρωτογενή δεδομένα περιέλθουν στο υποσύστημα επεξεργασίας, μετασχηματίζονται σε μορφή που είναι χρήσιμη για το στρώμα ML και παραδίδονται είτε στο υποσύστημα εκπαίδευσης για την παραγωγή μοντέλων ML, είτε στο υποσύστημα πρόβλεψης.

Το υποσύστημα πρόβλεψης αξιοποιεί τα δεδομένα που παραλαμβάνει για την παραγωγή της



Σχήμα 4.2: Αρχιτεκτονική Πρωτότυπου Συστήματος

πληροφορίας, δηλαδή των προβλέψεων. Οι προβλέψεις, με την σειρά τους, κατευθύνονται στο στρώμα αποθήκευσης, όπου και παραμένουν για ένα διάστημα. Το διαδικτυακό API παρέχει ένα κανάλι μεταξύ των τελευταίων αποθηκευμένων προβλέψεων και των εξωτερικών χρηστών, όπως οι εφαρμογές γεωχωρικής οπτικοποίησης ή έγκαιρης προειδοποίησης.

4.4 Οντότητες Πεδίου Εφαρμογής

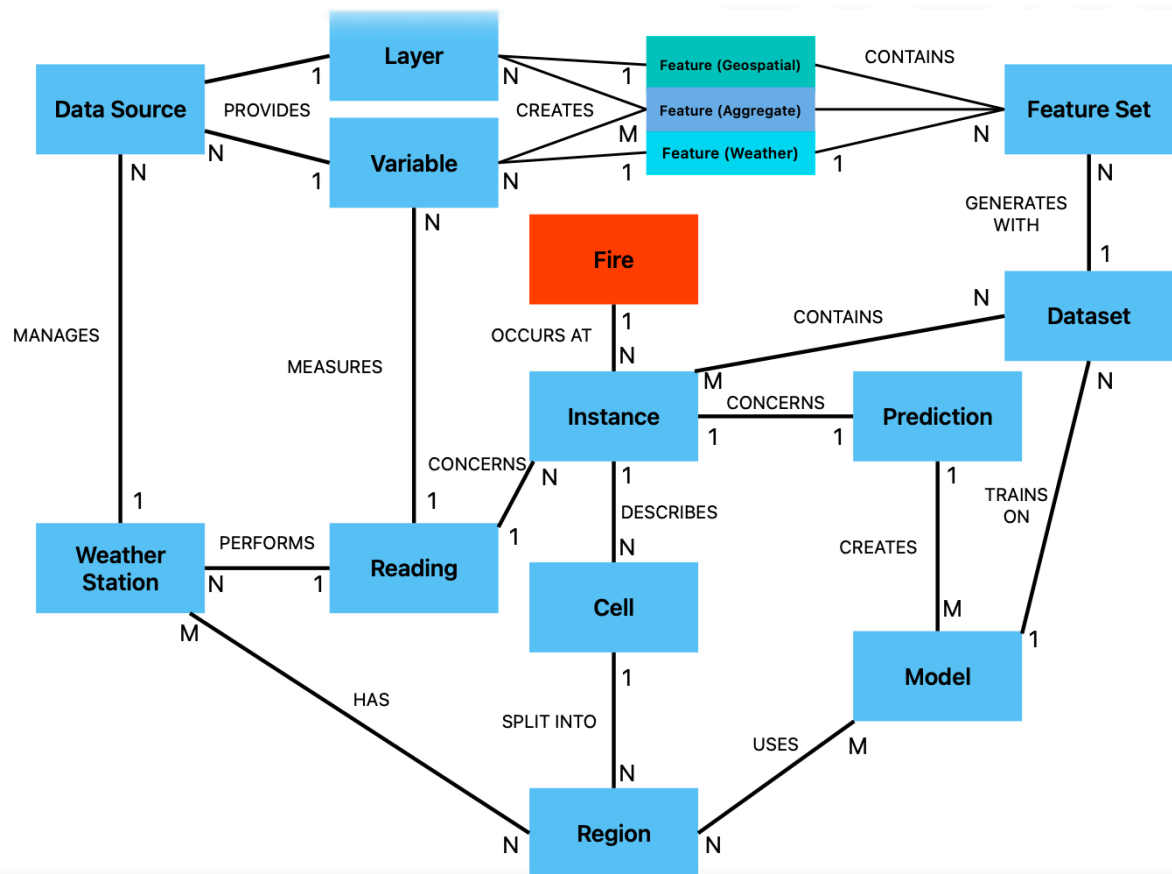
4.4.1 Κύριες Οντότητες

Πίνακας 4.1: Οντότητες Πεδίου Εφαρμογής για Πρωτότυπο Σύστημα

Οντότητα	Περιγραφή
Πυρκαγιές	Το αντικείμενο ενδιαφέροντος του συστήματος, αντιπροσωπεύει ένα καταγεγραμμένο συμβάν ανεξέλεγκτης φωτιάς με σημείο και ώρα έναρξης.
Περιοχές	Γεωγραφικές εκτάσεις για τις οποίες το σύστημα εξετάζει την εκδήλωση πυρκαγιών.
Μεταβλητές	Κατηγορίες πρωτογενών δεδομένων που αφορούν τις μεταβαλλόμενες συνθήκες που μπορούν να μετρηθούν σε μία τοποθεσία κατά μία χρονική στιγμή.
Πηγές Δεδομένων	Οι δίοδοι πρωτογενών δεδομένων όπου το σύστημα μπορεί ανά πάσα στιγμή να κάνει ερώτημα για τις τρέχουσες συνθήκες.
Σταθμοί	Γεωγραφικά σημεία από τα οποία προέρχονται μετρήσεις μετεωρολογικών μεταβλητών για την τρέχουσα χρονική στιγμή.
Μέτρηση	Η καταγραφή για μία μετεωρολογική μεταβλητή σε μία δεδομένη χρονική στιγμή.
Γεωχωρική Στρώση	Ψηφιακή αποτύπωση ενός συγκεκριμένου γεωγραφικού παράγοντα πάνω σε μία γεωγραφική περιοχή, με την οποία το σύστημα υπολογίζει γεωγραφικές μεταβλητές.
Πλέγμα	Ο τρόπος διαχωρισμού της περιοχής ενδιαφέροντος σε υποπεριοχές.
Κελί	Γεωγραφική έκταση που προκύπτει από τον διαχωρισμό μία περιοχής και χρησιμοποιούνται για προσδιορισμό της τοποθεσίας.
Χαρακτηριστικά	Κατηγορίες δεδομένων εισόδου για τις μεθόδους εκτίμησης. Υπολογίζονται κυρίως βάσει μεταβλητών για μία δεδομένη τοποθεσία και χρονική περίοδο
Στιγμιότυπο	Θεμελιώδης μονάδα χωροχρονικού προσδιορισμού που χαρακτηρίζεται από μία τοποθεσία και μία χρονική περίοδο.
Σύνολα Δεδομένων	Μία συλλογή από στιγμιότυπα με χρονική συνέχεια, για καθορισμένες τοποθεσίες και με τιμές για επιλεγμένα χαρακτηριστικά, που συνοδεύονται από ένδειξη αναφορικά με την εκδήλωση πυρκαγιάς.
Μοντέλα	Παραμετροποιησιμες ρουτίνες παραγωγής προβλέψεων βάσει συνθηκών. Τα μοντέλα ML εκπαιδεύονται με σύνολα δεδομένων.
Προβλέψεις	Ένδειξη για την εκτίμηση κινδύνου πυρκαγιάς σε ένα δεδομένο στιγμιότυπο, βασισμένη στις πιθανότητες που υπολογίζει ένα μοντέλο για αυτό το στιγμιότυπο.

4.4.2 Αλληλεπιδράσεις Οντοτήτων

Οι οντότητες πεδίου συνδέονται μεταξύ τους σε ένα πλέγμα συσχετίσεων, όπως αποτυπώνεται στο διάγραμμα οντοτήτων-σχέσεων 4.3. Οι διασυνδέσεις μεταξύ οντοτήτων και το είδος των σχέσεων καθοδηγούν την σχεδίαση και υλοποίηση του συστήματος. Παρότι στην παρούσα μελέτη



Σχήμα 4.3: Διάγραμμα Οντοτήτων-Σχέσεων

εξετάζεται μόνο μία περιοχή, το σύστημα έχει την δυνατότητα ταυτόχρονης διαχείρισης πολλών περιοχών, με κάθε μία από τις οποίες να έχει πολλούς διαφορετικούς μετεωρολογικούς σταθμούς, πολλά μοντέλα που δημιουργούν ξεχωριστές προβλέψεις για αυτήν και πολλά κελιά· το κάθε κελί προκύπτει από την εφαρμογή ενός πλέγματος διαχωρισμού στην περιοχή. Ένα στιγμιότυπο αντιστοιχεί με ένα κελί και ένα χρονικό διάστημα, αλλά για ένα κελί θα προκύψουν πολλά στιγμιότυπα σε βάθος χρόνου. Οι μετεωρολογικοί σταθμοί μπορεί να ανήκουν σε πολλές περιοχές και να παράγουν πολλές μετρήσεις και μεταβλητές, όμως ο κάθε σταθμός ανήκει μόνο σε μία πηγή δεδομένων. Κάθε μέτρηση αναφέρεται μόνο σε μία μεταβλητή, και κάθε μεταβλητή προέρχεται μόνο από μία πηγή, ακόμα και όταν διάφορες πηγές προσφέρουν μετρήσεις που αναφέρονται εννοιολογικά στον ίδιο παράγοντα, αυτές θα θεωρηθούν ως ξεχωριστές μεταβλητές. Κάθε γεωχωρική στρώση επίσης προέρχεται από μία μόνο πηγή.

Τα χαρακτηριστικά δημιουργούνται από την επεξεργασία των μεταβλητών και των στρώσεων με ποικίλους τρόπους: Πολλές μεταβλητές μπορεί να αποδίδουν το ίδιο χαρακτηριστικό, ένα χαρακτηριστικό μπορεί να αποδίδεται από πολλές μεταβλητές και μία μεταβλητή μπορεί να αποδίδει

πολλά χαρακτηριστικά. Τα γεωχωρικά χαρακτηριστικά, συγκεκριμένα, αναφέρονται σε ένα κελί και προέρχονται από μία ή περισσότερες στρώσεις, αλλά τα καιρικά χαρακτηριστικά διαμοιράζονται στα κελιά με ανάθεση ή παρεμβολή. Ένα σύνολο δεδομένων παράγεται για συγκεκριμένο πλέγμα και σύνολο χαρακτηριστικών, περιέχει πολλές πυρκαγιές, πολλά στιγμιότυπα, και μπορεί να χρησιμοποιείται από πολλά μοντέλα, ένα μοντέλο, όμως, χρησιμοποιεί μόνο ένα σύνολο δεδομένων. Μία πρόβλεψη παράγεται από ένα μοντέλο και για ένα στιγμιότυπο, άρα αφορά ένα μόνο κελί.

4.5 Πρωτόκολλα Συστήματος

4.5.1 Μονάδες Μέτρησης

Για να είναι προβλέψιμη η συμπεριφορά του συστήματος και να παραμένει αναπαράξιμη και κατανοητή η λειτουργία του, ορίστηκαν συγκεκριμένες μονάδες μέτρησης που εκφράζουν συγκεκριμένες κατηγορίες μεγεθών στο σύστημα, όπως φαίνεται στον πίνακα 4.2. Η επιλογή των μονάδων μέτρησης έγινε βάσει των συνηθειών και προτύπων των ελληνικών φορέων.

Πίνακας 4.2: Μονάδες Μέτρησης Συστήματος

Μέγεθος	Μονάδα Μέτρησης
Χρονικό Διάστημα	Λεπτά (<i>min</i>).
Αποστάσεις	Μέτρα (<i>m</i>).
Θερμοκρασία	Βαθμοί Κελσίου ($^{\circ}\text{C}$).
Ταχύτητα	Μέτρα ανά Δευτερόλεπτο (<i>m/s</i>).
Γωνίες	Μοίρες ($^{\circ}$).
Υετός	Χιλιοστά Βροχής (<i>mm</i>).
Πίεση	Εκατοπασκάλ (<i>hPa</i>).
Ακτινοβολία	Watt ανά Τετραγωνικό Μέτρο (W/m^2).

4.5.2 Χωροχρονική Αποτύπωση

Όπως συζητήθηκε στις ενότητες 3.6.3 και 3.6.1, οι συνθήκες και οι προβλέψεις για το σύστημα, χωρίζονται σε στιγμιότυπα $w(s, t)$ συγκεκριμένου τομέα και χρονικού διαστήματος. Οι καιρικές μετρήσεις χαρακτηρίζονται από το την στιγμή τέλος του χρονικού διαστήματος στο οποίο αναφέρονται, και τον σταθμό από όπου προέρχονται. Η διάρκεια των χρονικών διαστημάτων των στιγμιότυπων είναι σταθερή και ταυτίζεται με τον ρυθμό εκτέλεσης μετρήσεων c .

Οι χρονικές στιγμές καταγράφονται στο σύστημα ως χρονικά αποτυπώματα με ακρίβεια λεπτών και ένδειξη χρονικής ζώνης Συντονισμένης Παγκόσμιας Ώρας (*UTC*). Τα λεπτά της ώρας θεωρήθηκαν ως ο κατάλληλος βαθμός χρονικής ακρίβειας για τους στόχους του συστήματος.

Οι περιοχές και τα κελιά είναι έγκυρα πολύγωνα, είναι δηλαδή κλειστά και αποτελούνται από ευθείες που δεν τέμνονται, ενώ οι σταθμοί και οι πυρκαγιές θεωρούνται σημεία. Τα γεωχωρικά στοιχεία περιγράφονται με συντεταγμένες στο Παγκόσμιο Γεωδαιτικό Σύστημα Αναφοράς WGS 84 (*EPSG:4326*).

4.6 Σχεδίαση Βάσης Δεδομένων

4.6.1 Επιλογή Μοντέλου Μόνιμης Αποθήκευσης

Για την μόνιμη αποθήκευση του αρχείου δεδομένων, καθώς και των επιλογών παραμετροποίησης των λειτουργιών του συστήματος, επιλέχθηκε μοντέλο τύπου σχεσιακής βάσης δεδομένων SQL. Αυτή η επιλογή έγινε με κριτήριο συγκεκριμένα πλεονεκτήματα που παρέχουν οι σχεσιακές ΒΔ [54], σε σχέση με εναλλακτικές τεχνολογίες αποθήκευσης για συστήματα δεδομένων, όπως οι NoSQL ΒΔ ή το Object Storage:

- Δικλείδες ασφαλείας ως προς την ακεραιότητα των δεδομένων που εξασφαλίζουν την ομαλή εκτέλεση των εργασιών, όπως η χρήση προκαθορισμένου σχήματος, η επιβολή περιορισμών και η διαχείριση των συναλλαγών.
- Παροχή εργαλείων και επεκτάσεων για χειρισμό εξειδικευμένων τύπων δεδομένων και βελτίωση του χρόνου αναζήτησης και ανάκτησης.
- Πλούσια γλώσσα ερωτημάτων SQL για αναζήτηση και ανάλυση της πληροφορίας.
- Καλά κατανοητή δομή που μεταφράζει αβίαστα τις σχέσεις των οντοτήτων και τις λειτουργικές απαιτήσεις.

Δεδομένου ότι η μελέτη πραγματοποιείται ένα πρωτότυπο συστήματος, η αυστηρότητα που προσφέρουν οι σχεσιακές βάσεις είναι επιθυμητή για την ορθή κατεύθυνση και υλοποίηση. Ταυτόχρονα, οι εξειδικευμένες δυνατότητες των συστημάτων σχεσιακών ΒΔ κρίνονται ωφέλιμες για ένα έργο γεωχωρικής ανάλυσης με προοπτική διαχείρισης σχετικά μεγάλου όγκου δεδομένων.

4.6.2 Τεχνολογίες Βάσης Δεδομένων

Η PostgreSQL [55], σαν δωρεάν, ελεύθερο, σταθερό και εξαιρετικά πλούσιο σύστημα διαχείρισης ΒΔ (DBMS), επιλέχθηκε για την υλοποίηση του πρωτότυπου συστήματος. Το κύριο προτέρημα αυτής της τεχνολογίας για το έργο είναι οι επιλογές επεκτασιμότητας που υπάρχουν: Αφενός, η επέκταση PostGIS [56] επιτρέπει την άμεση επεξεργασία χωρικών τύπων δεδομένων, διευκολύνει την καταχώρηση γεωχωρικών στρώσεων, την διαχείριση συστημάτων γεωαναφοράς, παρέχει χρήσιμες χωρικές συναρτήσεις και επιτυγχάνει αύξηση της ταχύτητας διεκπεραίωσης των χωρικών ερωτημάτων. Αφετέρου, η επέκταση TimescaleDB [57] διαθέτει δυνατότητες για την διαχείριση, την ανάλυση και την εκτέλεση ερωτημάτων σε δεδομένα με χρονική συνέχεια, όπως οι μετεωρολογικές μετρήσεις. Η χρήση αυτών των επεκτάσεων είναι σχεδόν αυτονόητη για ένα έργο που εξετάζει γεωχωρικά δεδομένα και χρονοσειρές.

4.6.3 Σχήμα Βάσης

Το σχήμα της βάσης ακολουθεί σε μεγάλο βαθμό τις οντότητες του πεδίου εφαρμογής, όπως περιγράφηκαν στην ενότητα 4.4. Οι σχέσεις μεταξύ οντοτήτων σε επίπεδο βάσης αποδόθηκαν με ξένα κλειδιά, οι σχέσεις ένα-προς-ένα αποδόθηκαν με περιορισμούς μοναδικότητας στα ξένα κλειδιά και οι σχέσεις πολλά-προς-πολλά αποδόθηκαν με χρήση ενδιάμεσων πινάκων. Συγκεκριμένα, για τις περίπλοκες σχέσεις μεταξύ μεταβλητών και χαρακτηριστικών χρησιμοποιήθηκαν πίνακες

χάρτες, για καιρικά και για γεωχωρικά χαρακτηριστικά, που, πέρα από την σχέση μεταξύ μεταβλητής και χαρακτηριστικού, καταχωρούν και το είδος της σχέσης, διευκρινίζοντας έτσι την μέθοδο επεξεργασίας.

Οι περιορισμοί στις τιμές των εγγραφών (*Constraints*), χρησιμοποιήθηκαν μεθοδικά, ώστε να διασφαλίζεται η ακεραιότητα των δεδομένων, κυρίως η μοναδικότητα, σε επιλεγμένα σημεία, όπου κριθηκε πως αυτό έχει νόημα, όπως ονόματα, εκδόσεις και σχέσεις ξένων κλειδιών.

Το σχήμα δημιουργήθηκε σε υψηλό βαθμό κανονικοποίησης, έτσι ώστε κάθε στοιχείο στην βάση να έχει διακριτή θέση και ρόλο και, επομένως, η διαχείριση της να γίνεται αβασάνιστα. Οι προβολές (*Views*), προκατασκευασμένα ερωτήματα SQL, χρησιμοποιούνται για να κάνουν πιο πρακτική την αναζήτηση στοιχείων που απαιτείται συχνά να βρίσκονται μαζί, αλλά, λόγω της κανονικοποίησης, βρίσκονται σε ξεχωριστούς πίνακες και πρέπει να χρησιμοποιηθούν αλληπάληλες εντολές JOIN για να συγκεντρωθούν.

Για την υποστήριξη ποικίλων καιρικών μετρήσεων από διάφορες μεταβαλλόμενες πηγές, επιλέχθηκε μακριά μορφή πίνακα (*long-form table*) για την καταχώρηση των καιρικών συνθηκών. Η κάθε εγγραφή περιέχει, φυσικά, την καταγεγραμμένη τιμή, καθώς και το τέλος της περιόδου αναφοράς της, τον σταθμό καταγραφής και την μετρήσιμη μεταβλητή, σαν σύνθετο κύριο κλειδί.

4.6.4 Αρχειοθέτηση Δεδομένων

Για να ξεκινήσει το σύστημα την παραγωγή εκτιμήσεων, απαιτείται ένα ιστορικό αρχείο, πάνω στο οποίο θα εκπαιδευτούν τα μοντέλα πρόβλεψης. Επομένως, για να τεθεί άμεσα σε λειτουργία, τα ιστορικά αρχεία τοποθετούνται στη ΒΔ κατά την εγκατάσταση του συστήματος, στον τομέα των αρχειοθετημένων καταγραφών. Κατά την λειτουργία του συστήματος, μεγάλοι όγκοι πρωτογενών δεδομένων αντλούνται περιοδικά και αποθηκεύονται για την εκτίμηση του τρέχοντος κινδύνου. Ως επακόλουθο, το αρχείο καταγραφών εμπλουτίζεται και μπορεί ανά πάσα στιγμή να αξιοποιηθεί για εκ νέου εκπαίδευση, επαλήθευση ή σύγκριση απόδοσης μοντέλων.

Η προσέγγιση που ακολουθείται στην καταχώρηση των μετρήσεων, είναι να διατηρούνται τα δεδομένα αυτούσια, ακριβώς όπως αυτά παραλαμβάνονται από τις πηγές. Με αυτόν τον τρόπο, παραμένουν διαθέσιμα σε διαφορετικές προσεγγίσεις επεξεργασίας. Για εξοικονόμηση αποθηκευτικού χώρου, οι επεξεργασμένες τιμές και οι προβλέψεις δεν διατηρούνται μόνιμα, αλλά υπάρχει ανά πάσα στιγμή η δυνατότητα επαναδημιουργίας τους, από την ανάλυση του αρχείου πρωτογενών δεδομένων.

4.6.5 Βελτιστοποίηση Ερωτημάτων

Καθώς το σύστημα στοχεύει στην παροχή "σχεδόν-ζωντανής" εκτίμησης κινδύνου, η παραγωγή των προβλέψεων από την στιγμή ανάκτησης των πιο πρόσφατων μετρήσεων, έως και την εξαγωγή των προβλέψεων, δεν πρέπει να υπερβαίνει τα λίγα λεπτά. Για την παραγωγή των χαρακτηριστικών εισόδου των μοντέλων πρόβλεψης απαιτείται η εκτέλεση χωρικών υπολογισμών και η ανάκτηση πρόσφατων δεδομένων στο αρχείο, το οποίο ενδέχεται να περιέχει εκατομμύρια εγγραφές και, με αυτόν τον τρόπο, να επιβραδυνθεί η διαδικασία.

Βελτιστοποίηση Χωρικών Ερωτημάτων με PostGIS

Οι μέθοδοι που παρέχει η επέκταση PostGIS ήδη έρχονται βελτιστοποιημένες για επίτευξη μικρότερου δυνατού χρόνου διεκπεραίωσης και, για αυτόν τον λόγο, η συντριπτική πλειοψηφία των χωρικών ερωτημάτων, δηλαδή υπολογισμοί απόστασης, έλεγχοι τομής ή εμπερίεξης και ανάθεση χωρικών μεταβλητών, διεκπεραιώνονται εντός της ΒΔ. Επιπροσθέτως, η PostGIS επιτρέπει την δημιουργία ευρετηρίων (*indexes*) σε χωρικές στήλες που επιτρέπουν στα χωρικά ερωτήματα να παραβλέψουν μεγάλο μέρος των εγγραφών κατά την σάρωση των πινάκων και να μειώσουν περαιτέρω τον χρόνο εκτέλεσης.

Βελτιστοποίηση Διαχείρισης Χρονοσειρών με TimescaleDB

Οι χρονοσειρές μετρήσεων που αξιοποιεί το σύστημα, αποτελούνται από συνεχώς αυξανόμενο αριθμό εγγραφών. Βάσει του υπολογισμού 3.1 και τον διαχωρισμό της κάθε μέτρησης στις μετρήσιμες μεταβλητές είναι φανερό ότι ο αριθμός των εγγραφών οι εγγραφές αγγίζει την κλίμακα των εκατομμυρίων σε μικρό χρονικό διάστημα.

Η TimescaleDB εξειδικεύεται στην διαχείριση δεδομένων με χρονική συνέχεια: Οι εγγραφές των χρονοσειρών χωρίζονται σε υποπίνακες (*chunking*), βάσει της χρονικής ετικέτας τους, και τα χρονικά ερωτήματα προς την συγκεκριμένη χρονοσειρά κατευθύνονται αυτόματα προς το κατάλληλο υποπίνακα, ώστε να αποφεύγεται η σάρωση ολόκληρης της χρονοσειράς. Επιπλέον, παρέχει την δυνατότητα περιοδικής, αυτόματης συσσώρευσης τιμών με έξυπνο τρόπο. Η στιγμή εκτέλεσης αυτής της αυτόματης συσσώρευσης μπορεί να μετατοπιστεί στην νεκρή χρονική περίοδο ενδιάμεσα των ανανεώσεων των προβλέψεων. Με αυτόν τον τρόπο, είναι δυνατόν να προετοιμαστούν τα συσσωρευτικά καιρικά χαρακτηριστικά, όπως ο υετός των προηγούμενων ημερών, και να αποφευχθεί η επιβάρυνση από την εκτέλεση χρονοβόρων ερωτημάτων συσσώρευσης.

Προϋπολογισμός Αποτελεσμάτων με Υλοποιημένες Προβολές

Οι υλοποιημένες προβολές (*Materialized Views - MV*) χρησιμοποιούνται για την αποθήκευση των αποτελεσμάτων εξαιρετικά χρονοβόρων ερωτημάτων, ώστε το σύστημα να μπορεί να κάνει απλή ανάκτηση των προϋπολογισμένων αποτελεσμάτων. Μπορούν να εφαρμοστούν συνδυαστικά με τις βελτιστοποιήσεις των επεκτάσεων, όπως για την προετοιμασία της αυτόματης συσσώρευσης καιρικών τιμών ή για την καταχώρηση των αποστάσεων μεταξύ κελιών και σταθμών. Ο ρυθμός ανανέωσης των MV εξαρτάται από την συχνότητα μεταβολής των δεδομένων που εμπεριέχουν, αλλά και τις απαιτήσεις επικαιρότητας της πληροφορίας που περιέχουν. Για παράδειγμα, οι αποστάσεις δημιουργούνται όταν δημιουργείται νέο πλέγμα διαχωρισμού και χρειάζεται να ανανεώνονται με την καταχώρηση νέου σταθμού, ενώ, η βροχόπτωση του τελευταίου εικοσιτετράωρου, πρέπει να ανανεώνεται σε κάθε κύκλο μετρήσεων.

4.6.6 Διεπαφή Βάσης Δεδομένων

Η δημιουργία και παροχή διεπαφής ΒΔ για χρήση από τα υποσυστήματα του Aristaeos τυποποιεί τα κανάλια δεδομένων από και προς την ΒΔ, συγκεντρώνει τις λειτουργίες ανάκτησης και καταγραφής σε ένα τμήμα του έργου, όπου είναι εύκολα διαχειρίσιμες, και περιορίζει τις περιττές επαναλήψεις κώδικα. Η διεπαφή έγινε σε γλώσσα Python σε συμφωνία με τις υπηρεσίες του συστήματος.

Σύνδεση με την Βάση Δεδομένων

Η τυποποιημένες υπορουτίνες σύνδεσης στην βάση διασφαλίζουν την σταθερότητα και προβλεψιμότητα των εργασιών ανάκτησης και αποθήκευσης. Αξιοποιώντας την σύνταξη της Python για ορισμό πλαισίων εκτέλεσης (εντολή `with`), η σύνδεση στην βάση πραγματοποιείται συνοπτικά, με μία γραμμή κώδικα. Επιπλέον, η διαχείριση των συναλλαγών γίνεται αυτόματα, στέλνοντας σήμα `commit` μετά την εκτέλεση των ερωτημάτων ή `rollback` σε περίπτωση σφάλματος και με το αυτόματο κλείσιμο του καναλιού σύνδεσης με το πέρας του πλαισίου εκτέλεσης που ορίστηκε. Αυτή η ροή είναι ασφαλής, λειτουργική και εξαιρετικά βολική για τον προγραμματιστή.

Έχουν υλοποιηθεί υπορουτίνες σύνδεσης στην βάση με δύο διαφορετικούς οδηγούς (*drivers*): Αρχικά μία σύγχρονη σύνδεση με τον οδηγό `psycopg2`, ιδανική για προγράμματα που εκτελούνται σειριακά, όπως η συλλογή και η επεξεργασία των δεδομένων και η δημιουργία των προβλέψεων, καθώς η παρακώλυση της εκτέλεσης των εντολών που ακολουθούν ένα ερώτημα είναι επιθυμητή. Έπειτα, υλοποιήθηκε και υπορουτίνα ασύγχρονης σύνδεσης με τον οδηγό `asyncpg`, για χρήση από διαδικτυακό API που εξυπηρετεί πολλούς ταυτόχρονους χρήστες. Σε αυτήν την περίπτωση, η δυνατότητα επικαλυμμένης εκτέλεσης ερωτημάτων βελτιώνει σημαντικά την εμπειρία των χρηστών, χωρίς να επηρεάζει την ακεραιότητα της εκτέλεσης των εργασιών.

Εκτέλεση Ερωτημάτων

Το μεγαλύτερο τμήμα των υπορουτίνων καταχώρησης, ανάκτησης, ενημέρωσης και διαγραφής (*CRUD*), κάνουν χρήση Αντικειμενοστραφούς Σχεσιακού Μοντέλου (*ORM*) για σταθερότητα και ασφάλεια στην εκτέλεση των ερωτημάτων, ταχύτερη πρωτοτύπηση και πιο κατανοήσιμο περιεχόμενο. Συγκεκριμένες περιπτώσεις, όπου η χρήση του *ORM* δεν ήταν πρακτική ή εφικτή, έγιναν με δυναμική SQL. Μία από αυτές τις περιπτώσεις είναι η καταγραφή καιρικών συνθηκών, όπου διαμορφώθηκε ερώτημα Καταχώρησης-Ανανέωσης (*Upsert*) εγγραφής, ώστε να αποφευχθούν οι διπλοεγγραφές σε περίπτωση επανεκτέλεσης της διεργασίας για το ίδιο στιγμιότυπο.

4.7 Ροή Επεξεργασίας - ETL

Το στρώμα δεδομένων του συστήματος υλοποιείται από μία ροή Ανάκτησης-Μετατροπής-Φόρτωσης (*ETL Pipeline*), στα πλαίσια της οποίας πραγματοποιείται το μεγαλύτερο μέρος των διεργασιών διαχείρισης δεδομένων. Σε πρακτικό επίπεδο, η ροή ETL λαμβάνει την μορφή μίας ρουτίνας Python που μπορεί να εκτελεστεί οποτεδήποτε υπάρχει η ανάγκη για ανάκτηση, επεξεργασία και καταχώρηση των ζητούμενων δεδομένων.

4.7.1 Υποσύστημα Συλλογής Δεδομένων

Οι δύο κύριες κατηγορίες πηγών για την ανάκτηση δεδομένων είναι οι πηγές μετεωρολογικών μετρήσεων και οι πηγές γεωχωρικών στρώσεων. Το υποσύστημα συλλογής δεδομένων θα λάβει τα πιο πρόσφατα γεωχωρικά δεδομένα για ένα γεωχωρικό στοιχείο μόνο εάν έχει την έχει περάσει το διάστημα που θεωρούνται έγκυρα τα δεδομένα του συγκεκριμένου στοιχείου. Για τις μετεωρολογικές μετρήσεις, θα ερωτηθούν όλοι οι σταθμοί που ανήκουν στις περιοχές αρμοδιότητας του συστήματος.

Η ερωτήσεις στις πηγές δεδομένων γίνονται μέσω πρωτοκόλλου HTTP με την βιβλιοθήκη requests της Python, χωρίς να αποκλείεται η υλοποίηση εναλλακτικών μεθόδων ανάκτησης στο μέλλον, όπως web sockets για τους αισθητήρες που δεν παρέχουν διαδικτυακό API. Η επεξεργασία των ανακτημένων δεδομένων περιορίζεται στον εντοπισμό των ζητούμενων μεταβλητών εντός του πλαισίου απάντησης (response), και διαφέρει ελαφρώς ανά πηγή, ανάλογα με την δομή που ακολουθείται από τον κάθε πάροχο.

Η απαντήσεις από μετεωρολογικούς σταθμούς αναμένονται σε μορφή JSON, ενώ τα γεωχωρικά στρώματα σε μορφή GeoJSON για τα διανυσματικά δεδομένα και GeoTIFF για τα rasters. Τα πρωτογενή δεδομένα αποστέλλονται στην ΒΔ για καταχώρηση, μέσω των διαθέσιμων υπορουτίων της διεπαφής ΒΔ και επιστρέφονται σαν αποτέλεσμα της εργασίας συλλογής δεδομένων για χρήση από άλλα υποσυστήματα.

4.7.2 Υποσύστημα Επεξεργασίας

Το υποσύστημα επεξεργασίας αποτελείται από εργαλεία για την δημιουργία των ζητούμενων χαρακτηριστικών πρόβλεψης από τα πρωτογενή δεδομένα. Τα εργαλεία αυτά λαμβάνουν την μορφή ανεξάρτητων κλάσεων Python, με την κάθε μία από αυτές να έχει συγκεκριμένη αρμοδιότητα. Ο χειρισμός και οι μετατροπές στα στιγμιότυπα γίνονται πάνω σε δομή δεδομένων DataFrame χρησιμοποιώντας την βιβλιοθήκη polars [58], η οποία παρέχει ταχεία και αποδοτική, από άποψη μνήμης, εκτέλεση μεταμορφώσεων και συνενώσεων σε σύνολα του ενός μέχρι και εκατομμυρίων στιγμιότυπων.

Μηχανική Χαρακτηριστικών

Σε αυτό το υποσύστημα υλοποιείται η μεθοδολογία μηχανικής χαρακτηριστικών, μέσω της οποίας τα πρωτογενή δεδομένα μετασχηματίζονται σε διαφορετικές μορφές για την αποκάλυψη πληροφορίας που δεν ήταν προσβάσιμη από την αρχική μορφή τους. Ο οδηγός για αυτούς τους μετασχηματισμούς είναι οι χάρτες αντιστοίχισης μεταβλητών-χαρακτηριστικών που περιγράφονται στην τελευταία παράγραφο της ενότητας 4.4.2. Η δημιουργία των καιρικών και γεωχωρικών χαρακτηριστικών γίνεται βάσει του είδους της σχέσης μεταξύ χαρακτηριστικού/ών και μεταβλητής/ών.

Τα χαρακτηριστικά που αξιοποιούνται σε αυτήν την μελέτη προκύπτουν κυρίως από σχέσεις που υποδηλώνουν απλή αντιστοίχιση με μία καιρική μεταβλητή ή σχέσεις συσσώρευσης εντός ενός δεδομένου χρονικού παραθύρου πριν το στιγμιότυπο, ή εντός του τομέα του στιγμιότυπου. Τα χαρακτηριστικά φάσης, ή αλλιώς κυκλικά χαρακτηριστικά, υπολογίζονται για την στιγμή τέλους του χρονικού διαστήματος του στιγμιότυπου. Οι μέθοδοι υπολογισμού περιγράφηκαν λεπτομερώς στην ενότητα 3.6.5.

Τα καιρικά χαρακτηριστικά συσσώρευσης που απαιτούν χρονοβόρους υπολογισμούς, λαμβάνονται έτοιμα και προϋπολογισμένα από την ΒΔ. Το ίδιο ισχύει και για τα γεωχωρικά χαρακτηριστικά που υπολογίζονται εκ νέου για κάθε κελί κατά την εισαγωγή ή ανανέωση της αντίστοιχης στρώσης που τα παράγει και παραμένουν αποθηκευμένα σε πίνακα αντιστοίχισης στην ΒΔ, από όπου και ανακτούνται έτοιμα, ώστε να μην πρέπει να εκτελούνται οι χρονοβόρες γεωχωρικές αντιστοιχίσεις σε κάθε κύκλο εκτέλεσης.

Εναρμόνιση Μεταβλητών-Χαρακτηριστικών

Όλα τα χαρακτηριστικά ακολουθούν τις μονάδες μέτρησης του συστήματος που βρίσκονται στον πίνακα 4.2, αλλά δεν είναι απαραίτητο πως τα πρωτογενή δεδομένα από διάφορες πηγές θα ακολουθούν τα ίδια πρωτόκολλα. Για τον σκοπό αυτό δημιουργήθηκε εργαλείο μετατροπής μονάδων μέτρησης που μεταφράζει την τιμή της μεταβλητής, βάσει της μονάδας μέτρησης που ακολουθεί η πηγή της, στην μονάδα μέτρησης με την οποία εκφράζεται το χαρακτηριστικό στο οποίο αντιστοιχίζεται.

Χωρική Ανάθεση Χαρακτηριστικών

Η ανάθεση των γεωχωρικών χαρακτηριστικών στα κελιά γίνεται άμεσα, με χρήση χωρικών ερωτημάτων και γεωσυσσορευτικών συναρτήσεων ανάμεσα στις γεωχωρικές στρώσεις και την γεωμετρία του κελιού. Τα χαρακτηριστικά φάσης είναι ανεξάρτητα από τους χωρικούς παράγοντες. Όσον αφορά τα καιρικά χαρακτηριστικά, αυτά ανατίθενται στα κελιά από ένα εργαλείο προσέγγισης καιρικών συνθηκών ανά περιοχή, το οποίο, για τους σκοπούς της μελέτης έχει υλοποιημένη μόνο την ανάθεση από κοντινότερο γείτονα, όπως περιγράφεται στην πρώτη παράγραφο την ενότητας 3.6.4.

Εφόσον έχουν διασταυρωθεί τα κελιά με τις καιρικές, γεωχωρικές και χρονικές συνθήκες, το σύνολο χαρακτηριστικών πρόβλεψης είναι έτοιμο να προχωρήσει στο στρώμα ML.

4.8 Υποσύστημα Εκπαίδευσης

Το υποσύστημα εκπαίδευσης αποτελείται από μία κλάση εκπαιδευτή, η οποία αναλαμβάνει την δημιουργία ενός μοντέλου ML, που δημιουργείται και παραμετροποιείται από τον διαχειριστή του συστήματος. Το μεγαλύτερο μέρος των μεθόδων ML υλοποιήθηκε με χρήση της βιβλιοθήκης scikit-learn στην Python [59].

4.8.1 Παράμετροι Εκπαίδευσης

Οι παράμετροι του μοντέλου βρίσκονται στο σύστημα αποθήκευσης και καθορίζουν την πορεία της εκπαίδευσης και, τελικά, την ποιότητα του αποτελέσματος. Οι κύριοι μοχλοί ρύθμισης της εκπαίδευσης μπορούν να χωριστούν σε τέσσερις κατηγορίες: Τις απλές ετικέτες που υποδηλώνουν κατηγορικές ρυθμίσεις, πάνω στις οποίες εφαρμόζονται εντολές ελέγχου if, τις απλές αριθμητικές παραμέτρους που εφαρμόζονται ως έχουν, τις εκπαιδευόμενες τιμές, οι οποίες προσαρμόζονται ανά μοντέλο και τα σύνθετα αντικείμενα. Στον πίνακα 4.3 παρατίθεται το εύρος παραμετροποίησης που έχει υλοποιηθεί.

Πίνακας 4.3: Διαθέσιμη Παραμετροποίηση Μοντέλων

Κατηγορία	Παράμετροι
Ετικέτες	Τύπος Αλγόριθμου, Εξισορρόπηση Κλάσεων
Αριθμητικές	Ορίζοντας Στόχου, Αναλογία Δοκιμαστικών Στιγμιότυπων, Αριθμός Τυχαίας Κατάστασης, Στόχος Ακρίβειας, Στόχος Ανάκλησης
Εκπαιδευόμενες	Υπερπαράμετροι (διαφέρουν ανάλογα με το μοντέλο), Κατώφλια Κινδύνου
Σύνθετες	Σύνολο Δεδομένων, Βήματα Προετοιμασίας Δεδομένων (συγκεκριμένα σε κάθε μοντέλο)

4.8.2 Δημιουργία Συνόλου Εκπαίδευσης

Η παράμετρος του συνόλου δεδομένων αποτελεί αναφορά σε συγκεκριμένο, αποθηκευμένο αντικείμενο, το οποίο περιέχει όλα τα στοιχεία που χρειάζονται για την δημιουργία των δεδομένων εκπαίδευσης, ήτοι, το χρονικό διάστημα που καλύπτει, το σύνολο χαρακτηριστικών πρόβλεψης που περιλαμβάνει και το πλέγμα διαχωρισμού της περιοχής.

Η διαδικασία ξεκινάει με την ανάκτηση των μετρήσεων, από τα αρχειοθετημένα δεδομένα (*archive*) για την δεδομένη περίοδο, και προχωράει στην δημιουργία των στιγμιότυπων. Έπειτα, υπολογίζονται τα χαρακτηριστικά για κάθε στιγμιότυπο με την χρήση του υποσυστήματος προεπεξεργασίας 4.7.2. Τέλος, οι πυρκαγιές του αρχείου τοποθετούνται στα στιγμιότυπα με ένα απλό ερώτημα χωροχρονικής αντιστοίχισης. Το ολοκληρωμένο σύνολο δεδομένων αποθηκεύεται τοπικά σε μορφή CSV, ώστε να μην χρειάζεται να δημιουργηθεί εκ νέου, κάθε φορά που παράγεται και αξιολογείται ένα μοντέλο που χρησιμοποιεί το συγκεκριμένο σύνολο, εφόσον δεν έχει υποστεί αλλαγές.

Καθώς η δημιουργία συνόλου εκπαίδευσης διαχειρίζεται εκατομμύρια εγγραφές ταυτόχρονα, είναι από τις πλέον απαιτητικές διεργασίες που εκτελεί το σύστημα. Η χρήση της βιβλιοθήκης Polars που αναθέτει εγγραφές στην μνήμη βάσει των αναγκών (*lazy loading*) και διαθέτει βελτιστοποιημένες και διανυσματοποιημένες (*vectorized*) υπορουτίνες γραμμένες σε Rust, αύξησε σημαντικά τον όγκο των δεδομένων εκπαίδευσης που είναι ικανό να διαχειριστεί το σύστημα, σε σχέση με την βιβλιοθήκη Pandas [60].

4.8.3 Ροή Εκπαίδευσης

Οι αλγόριθμοι ML στο scikit-learn, αλλά και σε άλλες βιβλιοθήκες, όπως η XGBoost, παρέχουν γενικευμένες διεπαφές που επιτυγχάνουν μεγάλο επίπεδο τυποποίησης και διευκολύνουν σημαντικά την υλοποίηση της ενιαίας, παραμετροποιήσιμης ροής εκπαίδευσης που ακολουθείται. Επιπλέον, οι βιβλιοθήκες ML παρέχουν και δυνατότητες παράλληλης επεξεργασίας, που αξιοποιούνται για διαμοιρασμό του φόρτου εργασίας στους διαθέσιμους επεξεργαστές και, επομένως, ταχύτερη διεκπεραίωση της ρύθμισης και εκπαίδευσης των μοντέλων. Η διαδικασία της εκπαίδευσης έχει ως εξής:

1. Το σύνολο δεδομένων δημιουργείται ή ανακτάται από τοπικό αρχείο και φορτώνεται στην μνήμη.

2. Τα στιγμιότυπα με θετικό σήμα πυρκαγιάς το απλώνουν στα χρονικά πλησιέστερα στιγμιότυπα του ίδιου τομέα, σύμφωνα με την τιμή της παραμέτρου ορίζοντα στόχου, όπως περιγράφηκε στην ενότητα 3.7.1.
3. Το σύνολο δεδομένων χωρίζεται σε σύνολο εκπαίδευσης και σύνολο ελέγχου, βάσει της τιμής της παραμέτρου αναλογίας δοκιμαστικών στιγμιότυπων. Ο διαχωρισμός γίνεται χρονολογικά, με τα πιο πρόσφατα έτη να χρησιμοποιούνται για επικύρωση.
4. Καθορίζεται ο τρόπος χειρισμού της ανισορροπίας κλάσεων καθώς και ο σπόρος κατάστασης για τους αλγόριθμους με εγγενή τυχαιότητα.
5. Δημιουργείται η ροή βημάτων προετοιμασίας των δεδομένων για τον αλγόριθμο. Οι διάφορες ροές βημάτων που μπορούν να χρησιμοποιήσουν τα μοντέλα βρίσκονται αποθηκευμένες στην ΒΔ και δηλώνονται ως παράμετρος στο κάθε μοντέλο. Τα βήματα είναι πλήρως προσαρμόσιμα και έχουν ήδη υλοποιηθεί μετατροπές όπως η κανονικοποίηση και η συμπλήρωση κενών τιμών.
6. Εφόσον δεν έχει ήδη ολοκληρωθεί για το συγκεκριμένο μοντέλο, αρχίζει η εξαιρετικά χρονοβόρα διαδικασία ρύθμισης υπερπαραμέτρων με την μέθοδο στρωματοποιημένου Cross Validation του scikit-learn που διαχωρίζει το σύνολο εκπαίδευσης σε ημερολογιακά έτη, όπως περιγράφηκε στην ενότητα 3.7.5. Ο χώρος αναζήτησης υπερπαραμέτρων είναι προκαθορισμένος για κάθε αλγόριθμο.
7. Ορίζεται ο τελικός αλγόριθμος ML με τις ρυθμισμένες υπερπαραμέτρους, τις υπόλοιπες παραμέτρους μοντέλου και τα βήματα προεπεξεργασίας.
8. Με την ίδια εφαρμογή της τεχνικής CV γίνεται η ρύθμιση των κατωφλίων κινδύνου, όπως περιγράφηκε στην ενότητα 3.7.6 για τον τελικό αλγόριθμο.
9. Ξεκινάει η εκπαίδευση του μοντέλου με τον τελικό αλγόριθμο πάνω σε ολόκληρο το σύνολο εκπαίδευσης.

4.8.4 Αξιολόγηση Μοντέλου

Η αξιολόγηση των εκπαιδευμένων μοντέλων πραγματοποιείται πάνω στο σύνολο δοκιμής, με τα στιγμιότυπα του οποίου δεν είχαν καμία επαφή, σε κανένα στάδιο της ρύθμισης και της εκπαίδευσης. Υπολογίζονται οι μετρικές απόδοσης που αναφέρθηκαν στην ενότητα 3.7.2, τα αποτελέσματα του δείκτη κινδύνου, καθώς και συμπληρωματικά στοιχεία, όπως τα βάρη των χαρακτηριστικών. Οι υπολογισμοί γίνονται κυρίως με την χρήση έτοιμων μεθόδων που παρέχει η scikit-learn.

4.8.5 Αποθήκευση Μοντέλου

Το μοντέλο φέρει, σε όλη την διάρκεια της εκπαίδευσης, ένδειξη που υποδηλώνει ότι δεν είναι έτοιμο. Μόλις τελειώσει και η αξιολόγηση, τα αποτελέσματα της αξιολόγησης αποθηκεύονται σε ξεχωριστό πίνακα καταγραφής αποτελεσμάτων και η ένδειξη του μοντέλου αλλάζει, ώστε να δείξει πως είναι έτοιμο, ενώ ενημερώνεται και η ώρα ολοκλήρωσης, οι υπερπαραμέτροι που χρησιμοποιεί και τα υπολογισμένα κατώφλια για τα επίπεδα κινδύνου. Εάν το μοντέλο είχε ήδη εκπαιδευτεί στο

παρελθόν και έχουν πραγματοποιηθεί αλλαγές στις ρυθμίσεις ή τα δεδομένα εκπαίδευσης, τότε καταχωρείται ως νέα έκδοση.

Το ωφέλιμο αποτέλεσμα της εκπαίδευσης, που είναι σε θέση να παράγει εκτιμήσεις, αποθηκεύεται σε καθορισμένη θέση, ώστε να είναι προσβάσιμο από την διαδικασία περιοδικής παραγωγής προβλέψεων. Η αποθήκευση και η φόρτωση γίνεται σε μορφή παγωμένου αντικειμένου `pkl` με την βοήθεια του εργαλείου `joblib` της Python.

4.9 Υποσύστημα Πρόβλεψης

Το υποσύστημα προβλέψεων αποτελείται από μία κλάση που δέχεται στιγμιότυπα με χαρακτηριστικά πρόβλεψης και ένα μοντέλο και, με χρήση απλών και περιεκτικών μεθόδων, παράγει τις πιθανότητες εκδήλωσης πυρκαγιάς. Αυτό γίνεται σε δύο απλά βήματα: Πρώτα, φορτώνεται το μοντέλο από την θέση στην οποία αναμένεται να βρίσκεται, βάση του τύπου και της έκδοσης του. Έπειτα, το μοντέλο εφαρμόζει τα δικά του βήματα προετοιμασίας των δεδομένων και αναλύει τα στιγμιότυπα που δόθηκαν, ώστε να εξάγει μία πιθανότητα για κάθε στιγμιότυπο. Το σύνολο των πιθανοτήτων των στιγμιοτύπων επιστρέφει ως αποτέλεσμα και καταχωρείται στο σύστημα αποθήκευσης με ερώτημα τύπου `Upsert` ανά στιγμιότυπο.

4.10 Εκτέλεση Εργασιών

4.10.1 Διαχείριση Ουράς Εργασιών

Οι εργασίες που γίνονται από τα επιμέρους υποσυστήματα των στρωμάτων Δεδομένων και ML υλοποιήθηκαν ως θεματικές κλάσεις που καλούνται από υποπρογράμματα Python (*scripts*) και, για αυτόν το λόγο, έπρεπε να οριστούν συγκεκριμένα πρωτόκολλα εκτέλεσης τους. Σε αυτό το έργο ενσωματώθηκε πρωτόκολλο ουράς εργασιών που καθορίζει τις προϋποθέσεις και την ροή εκτέλεσης των εργασιών. Αυτό υλοποιήθηκε με το Celery [61], μία καθιερωμένη και δοκιμασμένη βιβλιοθήκη της Python, η οποία διαθέτει εργαλεία για κατανεμημένη, παράλληλη και ασύγχρονη εκτέλεση εργασιών.

4.10.2 Παραγωγοί Εργασιών

Το σύστημα έχει τους εξής παραγωγούς εργασιών: Ο πρώτος είναι η τακτική περιοδική διαδικασία ανάκτησης δεδομένων και παραγωγής προβλέψεων. Σε αυτήν την περίπτωση, το σύστημα εκμεταλλεύεται τα εργαλεία χρονοδρομολόγησης του Celery, που επιτρέπουν τον προγραμματισμό περιοδικών εργασιών στο ρυθμό c που ακολουθεί το σύστημα. Χρονοδρομολόγηση αξιοποιεί και ο δεύτερος παραγωγός, ο οποίος είναι οι διαδικασίες που ανανεώνουν τα γεωχωρικά στρώματα, ελέγχοντας περιοδικά την "φρεσκάδα" κάθε στρώματος βάσει του καθορισμένου ρυθμού ανανέωσης που έχει οριστεί για το καθένα.

Σήματα για εργασίες μπορούν να προκύψουν και από τις ενέργειες ενός χρήστη. Συγκεκριμένα, η διαδικασία εκπαίδευσης νέου μοντέλου από τον χειριστή του συστήματος ενδέχεται να είναι χρονοβόρα, και, άρα, είναι σκόπιμο να μεταφερθεί σε μία κρυφή διεργασία η οποία δεν θα παρακωλύει τις άλλες ενέργειες του χειριστή. Όταν τελειώσει η εκπαίδευση, η κατάσταση του μοντέλου θα ενημερωθεί και ο χειριστής θα ενημερωθεί για το αποτέλεσμα.

4.10.3 Ροή Εκτέλεσης Εργασιών

Η περιοδική ροή παραγωγής προβλέψεων που εκτελείται κάθε 30 λεπτά έχει ως εξής:

1. Γίνεται η συλλογή των μετρήσεων από όλους τους μετεωρολογικούς σταθμούς που ανήκουν στην περιοχή και το επόμενο βήμα προχωράει μόνο όταν τελειώσουν όλες οι εργασίες ανάκτησης (*chord*).
2. Οι τιμές καιρικών μεταβλητών που ανακτήθηκαν αποθηκεύονται και παραδίδονται στην επόμενη εργασία.
3. Δημιουργούνται τα χαρακτηριστικά πρόβλεψης για κάθε κελί της περιοχής από το υποσύστημα επεξεργασίας και παραδίδονται στην επόμενη εργασία.
4. Τα κελιά χωρίζονται σε παρτίδες συγκεκριμένου μεγέθους και δημιουργείται νέα εργασία για κάθε παρτίδα.
5. Γίνεται η εξαγωγή προβλέψεων με όλα τα ενεργά μοντέλα για την κάθε παρτίδα και τα αποτελέσματα αποθηκεύονται.

Ο διαχωρισμός σε εργασίες έχει σκοπιμότητα. Στο 1ο βήμα, κάθε ανάκτηση μετρήσεων από σταθμό είναι ξεχωριστή εργασία με διακριτό αποτέλεσμα εκτέλεσης, ώστε οι διακοπές λειτουργίας να είναι εύκολα εντοπίσιμες και να καταγράφονται. Στο 4ο βήμα, η παρτιδοποίηση (*batching*) των προβλέψεων μειώνει τον αντίκτυπο που έχει στον χρόνο εκτέλεσης η επιβάρυνση (*overhead*), που προκύπτει από επαναλαμβανόμενες κλήσεις μεθόδων. Ταυτόχρονα, ο διαχωρισμός αυτός εξασφαλίζει την μελλοντική δυνατότητα κλιμάκωσης του συστήματος μέσω του διαμοιρασμού των εργασιών πρόβλεψης σε κατανομημένους εργάτες, ιδιαίτερα σημαντικό για πλέγματα με υψηλή ανάλυση που ενδέχεται να εμπεριέχουν δεκάδες χιλιάδες τομείς.

4.10.4 Διαχείριση Σφαλμάτων

Η ροή εργασιών μπορεί να συναντήσει δύο είδη σφαλμάτων, τα παροδικά σφάλματα και τα σφάλματα εκτέλεσης. Τα παροδικά σφάλματα προκύπτουν κυρίως από αστάθειες στην σύνδεση των υπηρεσιών με άλλες υπηρεσίες. Αυτά τα συμβάντα μετριάζονται με την χρήση μηχανισμού αυτόματης επανεκπομπής (*retry*) με καθυστέρηση, εντός των μεθόδων που είναι επιρρεπείς σε διακοπές, όπως η ανάκτηση δεδομένων με HTTP αιτήματα. Σε περίπτωση που η σύνδεση δεν επανέλθει εντός των προσπαθειών επανεκπομπής, η διακοπή λειτουργία καταχωρείται και οι χρήστες ενημερώνονται.

Κάποια σφάλματα εκτέλεσης έχουν προβλεφθεί, και η διαχείρισή τους γίνεται ανά περίπτωση με δομές *try-catch*, που στοχοποιούν συγκεκριμένους τύπους σφαλμάτων. Ένα σφάλμα εκτέλεσης που δεν έχει προβλεφθεί, αποτελεί σοβαρό συμβάν, διότι υποδηλώνει αβλεψία στην σύνταξη του κώδικα ή στην ρύθμιση του συστήματος, και καταγράφεται για την ενημέρωση του διαχειριστή και την αποσφαλμάτωση.

4.11 Διαδικτυακή Διεπαφή

4.11.1 Διαδικτυακές Τεχνολογίες

Η διεπαφή της εφαρμογής είναι το τμήμα του συστήματος που το συνδέει με τον έξω κόσμο και το μετατρέπει, από ερευνητικό έργο, σε ένα πρακτικό εργαλείο για την μελέτη και ενημέρωση σχετικά με τον κίνδυνο δασικών πυρκαγιών. Η διεπαφή του Aristaeos, λαμβάνει την μορφή ενός διαδικτυακού REST API, γραμμένο με την βιβλιοθήκη FastAPI [62] που εκτελείται σε διακομιστή διαδικτυακής εφαρμογής unicorn. Αυτά τα εργαλεία, συμπορεύονται ομαλά με τις υπόλοιπες υπηρεσίες Python. Επίσης, παρέχουν ασύγχρονες δυνατότητες για εύκολη μελλοντική κλιμάκωση των υπηρεσιών του συστήματος για πολλούς ταυτόχρονους χρήστες, ευκολία πρωτοτύπησης και ισχυρή ακεραιότητα δεδομένων μέσω καλά ορισμένων μοντέλων αιτήσεων και τύπων μεταβλητών.

4.11.2 Υλοποίηση REST API

Το REST API αποτελείται από μία συλλογή πόρων, ή αλλιώς τερματικών σημείων (*endpoints*), ο κάθε ένας από τους οποίους ταυτοποιείται από μία καλά ορισμένη μοναδική διεύθυνση (*URL*). Αυτοί οι πόροι επεκτείνουν τις μεθόδους του διαδικτυακού πρωτοκόλλου HTTP, που σημαίνει πως είναι προσβάσιμοι σε εξωτερικούς χρήστες μέσω αιτημάτων HTTP, και αποστέλλουν το αποτέλεσμα της εκτέλεσης στον χρήστη. Οι εξωτερικοί χρήστες μπορεί να είναι αυτοί που χειρίζονται την εφαρμογή από περιηγητή ιστού, αναλυτές που επιθυμούν πρόσβαση σε δεδομένα του συστήματος ή ακόμα και αυτοματοποιημένες διεργασίες από εφαρμογές τρίτων που παρακολουθούν την εξέλιξη των προβλέψεων.

Για την επεξεργασία των αιτημάτων, οι πόροι διαβιβάζουν ή μετατρέπουν στοιχεία που είναι καταχωρημένα στο σύστημα αποθήκευσης, καλώντας CRUD υπορουτίνες της διεπαφής ΒΔ. Έτσι, υπάρχει η δυνατότητα διανομής των αποτελεσμάτων της εκτέλεσης του ETL Pipeline, αλλά και η δυνατότητα παροχής μοχλών παραμετροποίησης της εκτέλεσης του σε πιστοποιημένους χρήστες.

Εντός της μελέτης υλοποιήθηκαν οι πόροι που διανέμουν τις εκτιμήσεις του συστήματος πρόβλεψης. Η διανομή γίνεται με τρεις διαφορετικές μορφές: Σε μορφή εγγραφών ανά κελί για ανάλυση και σύγκριση αποτελεσμάτων, σε μορφή GeoJSON για αξιοποίηση από εφαρμογές GIS, και σε μορφή διανυσματικών πλακιδίων χάρτη (*Mapbox Vector Tiles MVT*), για ομαλή προβολή από εργαλεία οπτικοποίησης στους περιηγητές ιστού.

4.12 Διεπαφή Χρήστη

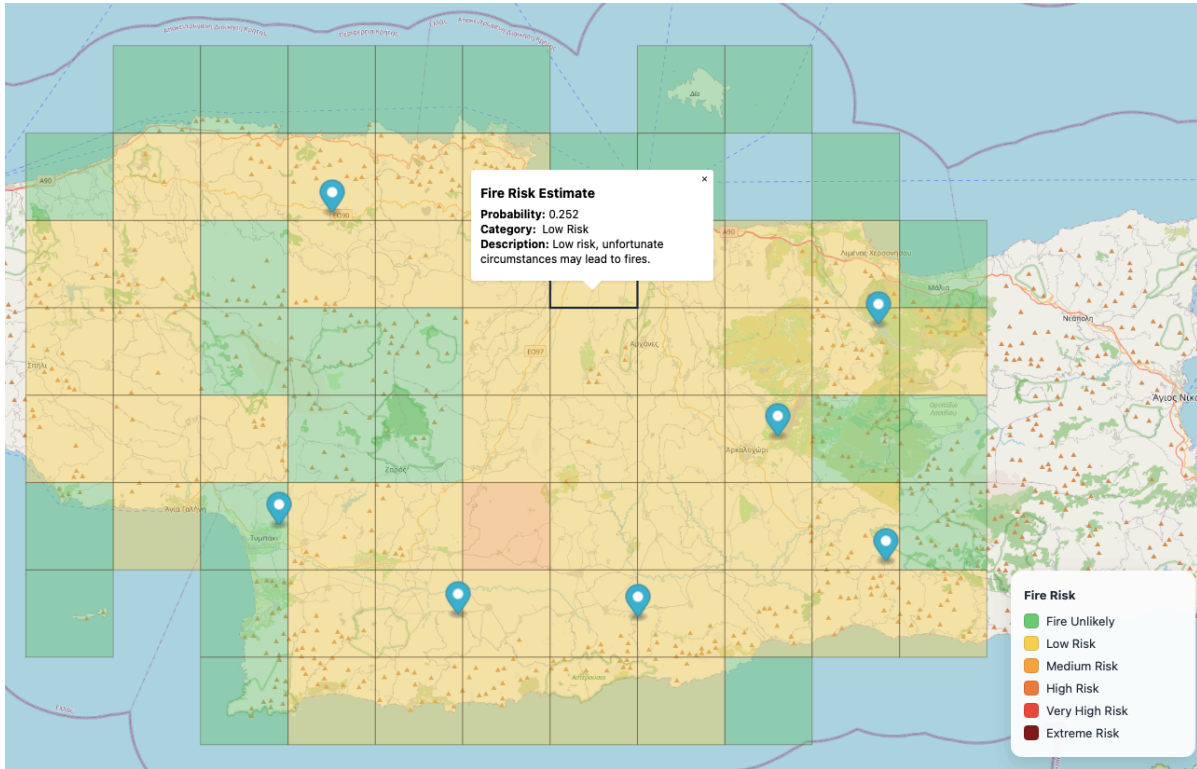
Η διεπαφή χρήστη (*User Interface - UI*) είναι απαραίτητη για τον διαμοιρασμό των αποτελεσμάτων στους ενδιαφερόμενους και, ταυτόχρονα, μπορεί να διευκολύνει την διαχείριση του συστήματος. Στο πρωτότυπο σύστημα, υλοποιείται μόνο μία απλή παρουσίαση των προβλέψεων, η οποία παρέχεται με την μορφή HTML σελίδας που σερβίρεται από έναν HTTP διακομιστή NGINX.

4.12.1 Οπτικοποίηση Κινδύνου

Η διαδικτυακή διεπαφή παρέχει ανά πάσα στιγμή τις πιο πρόσφατες πιθανότητες για τους τομείς των περιοχών αρμοδιότητάς της. Επιπλέον, για να παρουσιαστεί ορθά το εκτιμώμενο επίπεδο κινδύνου, χρειάζονται τα κατώφλια του μοντέλου που εκτέλεσε την πρόβλεψη που υπολογίζονται

στο βήμα 8 της εκπαίδευσης 4.8.3, καθώς και οι ετικέτες των επιπέδων που είναι αποθηκευμένες σε πίνακα στην ΒΔ. Αυτές οι τιμές ανακτώνται με αιτήματα HTTP προς το REST API που γίνονται από κώδικα JavaScript.

Η οπτικοποίηση των αποτελεσμάτων γίνεται με γεωχωρική παρουσίαση των στρωμάτων των τομέων και των σταθμών, ως διανυσματικά γραφικά, επί υπόβαθρου κατατοπιστικού αναφορικά με την πραγματική τοποθεσία των στοιχείων, σε διαδραστικό χάρτη 4.4. Η συμβολολογία που ακολουθείται περιλαμβάνει απλές πινέζες (*pins*) για τους σταθμούς και ημιδιαφανή χρωματισμό των τομέων με κατηγορική χρωματική κωδικοποίηση, κατά την οποία τα υψηλά επίπεδα κινδύνου βρίσκονται κοντά στο κόκκινο και τα χαμηλά στο πράσινο.



Σχήμα 4.4: Οπτικοποίηση Εκτίμησης Κινδύνου με Γεωχωρική Ανάλυση 10χμ. **Υπόβαθρο:**© OpenStreetMap contributors, available under the Open Database License.

Ο αριθμός των διανυσμάτων N_v που παράγει ο διαχωρισμός μίας περιοχής επιφάνειας a με ομοιόμορφη ανάλυση r υπολογίζεται με την σχέση 4.1. Αυτό σημαίνει πως για μία μεγάλη περιοχή με υψηλή ανάλυση διαχωρισμού που φαίνεται σε πλήρη προβολή χωρίς μεγέθυνση (*zoom*), δηλαδή με μικρή τιμή z , προκύπτουν δεκάδες χιλιάδες διανύσματα που θα πρέπει να ανακτηθούν, να διαβαστούν και να ζωγραφιστούν, μία κατάσταση που ενδέχεται να επιβαρύνει δυσανάλογα το πρόγραμμα περιήγησης.

$$N_v = \frac{a}{r} \quad (4.1)$$

4.12.2 Τεχνολογίες Γεωχωρικής Απεικόνισης

Η επίδραση του αριθμού των διανυσμάτων σε ένα στρώμα μπορεί να μετριάσει με την κατάλληλη οργάνωση της παρουσίασης και ρύθμιση του αριθμού και της μορφής δεδομένων. Η μορφή MVT είναι ένας βελτιστοποιημένος τρόπος αναπαράστασης διανυσματικών γεωχωρικών στρώσεων που επιτρέπει γρήγορη διανομή των διανυσματικών στοιχείων που απαιτούνται για την οπτικοποίηση του στρώματος πάνω σε συγκεκριμένες τιμές συντεταγμένων και μεγέθυνσης x , y και z . Η PostGIS υποστηρίζει την μετατροπή των δεδομένων σε MVT και επιτρέπει την ανάπτυξη προσαρμοσμένης λογικής επιπέδων μεγέθυνσης για ακόμα μεγαλύτερη βελτιστοποίηση της απόδοσης σε εκτενή σύνολα δεδομένων.

Η μορφή GeoJSON, που θυμίζει λεξικό Python με γεωμετρίες, είναι απλούστερη στην υλοποίηση, αλλά κάθε αντικείμενο θα πρέπει να ερμηνευτεί στο πρόγραμμα περιήγησης για την ύπαρξη και ανάγνωση των γεωμετριών, γεγονός που συνεπάγεται υψηλή υπολογιστική επιβάρυνση.

Η JavaScript βιβλιοθήκη MapLibreGL [63] επιλέχθηκε για την υλοποίηση της παρουσίασης των προβλέψεων, καθώς εξειδικεύεται στην δημιουργία διανυσματικών γραφικών σε διαδραστικούς χάρτες για φυλλομετρητές ιστού με χρήση MVT δεδομένων. Πέραν της έμφυτης ενσωμάτωσης της μορφής MVT, η βιβλιοθήκη αυτή εφαρμόζει αυτόματα τεχνικές αύξησης της απόδοσης, όπως βελτιστοποίηση της χρήσης κρυφής μνήμης για τμήματα των χαρτών και επιλεκτική ανάκτηση δεδομένων. Αυτά τα χαρακτηριστικά θα διασφαλίσουν ομαλή εμπειρία χρήστη στα μεγέθη δεδομένων που αναμένεται να διαχειρίζεται το σύστημα.

4.13 Υποσύστημα Καταγραφής

Οι ανάγκες του χειρισμού σφαλμάτων, κατά την διάρκεια της εκτέλεσης των υπορουτίνων του συστήματος, καθώς και η ανάγκη της παρακολούθησης της ποιότητας των μετρήσεων, επιτάσσουν την ανάπτυξη ενός στιβαρού και ευέλικτου συστήματος καταγραφής και επιτήρησης. Σε υψηλό επίπεδο, το σύστημα καταγραφής είναι ένας μηχανισμός που αιχμαλωτίζει συμβάντα από έναν παραγωγό (*producer*), και τα μεταφέρει σε έναν ή περισσότερους αποδέκτες (*sinks*). Η κατηγοριοποίηση αυτών των συμβάντων καθιστά την καταγραφή τους ωφέλιμη για την επιτήρηση της λειτουργίας και την ταχύτερη και αποτελεσματικότερη επιδιόρθωση των βλαβών.

Ακολουθώντας αυτές τις αρχές, το υποσύστημα καταγραφής του συστήματος υλοποιήθηκε ως πακέτο Python, αντικείμενα του οποίου μπορούν να αρχικοποιηθούν κατά την εκτέλεση οποιουδήποτε πιθανού παραγωγού συμβάντων, ώστε να κληθεί η υπορουτίνα καταγραφής σε οποιοδήποτε σημείο κρίνεται σκόπιμο, κυρίως εντός δομών *if-else* και *try-except*. Ο προγραμματιστής έχει την ευχέρεια να ορίσει δικούς του κωδικούς συμβάντων, με ονόματα, περιγραφές και επίπεδο κρισιμότητας, και να τα διαβιβάσει στην υπορουτίνα καταγραφής ως παράμετρο, μαζί με το υπόλοιπο πλαίσιο εκτέλεσης του προγράμματος την στιγμή του συμβάντος και την ώρα καταγραφής. Με αυτόν τον τρόπο επιτυγχάνεται υψηλό επίπεδο πληροφόρησης.

Σε αρχικό στάδιο, η επιτήρηση με το υποσύστημα καταγραφής έχει ενσωματωθεί στην περιοδική ροή εργασιών, για την εύρεση των σφαλμάτων που δεν έχουν προβλεφθεί, καθώς και την καταγραφή των παροδικών διακοπών λειτουργίας 4.10.4. Τα συμβάντα αποστέλλονται σε δύο αποδέκτες: Το αρχείο καταγραφής (*logs*) του προγράμματος και την ΒΔ, από όπου μπορούν να φτάσουν στον ενδιαφερόμενο. Οι διαχειριστές πρέπει να ειδοποιηθούν για κρίσιμα σφάλματα λειτουργίας που χρήζουν επισκευής, ενώ οι χρήστες πρέπει να ενημερώνονται για διακοπές στις μετρήσεις, που ενδεχομένως να υποβαθμίσουν την αξιοπιστία των προγνωστικών.

4.14 Εγκατάσταση και Ανάπτυξη Συστήματος

4.14.1 Υπηρεσίες σε Κοντέινερ

Για να εξυπηρετηθεί η απαίτηση της εύκολης και γρήγορης ανάπτυξης του συστήματος στα μεταβλητά περιβάλλοντα των ενδιαφερόμενων φορέων, ήταν αναγκαία η υλοποίηση μίας αναπαράξιμης και τυποποιημένης διαδικασίας εγκατάστασης και ανάπτυξης των υπηρεσιών του συστήματος. Η φιλοξενία υπηρεσιών σε περιβάλλοντα προσομοίωσης συνηθίζεται για την διασφάλιση της σταθερότητας κατά την ανάπτυξη και διαμοιρασμό του λογισμικού. Σε αυτόν τον τομέα, η ανάπτυξη με κοντέινερ (*Containers*) στο Docker [64] έχει σε μεγάλο βαθμό αντικαταστήσει τις ψηφιακές μηχανές (*Virtual Machines - VMs*) και έχει υιοθετηθεί στο παρόν έργο.

Οι λόγοι που προτιμήθηκε το Docker για την ανάπτυξη είναι οι εξής: Τα κοντέινερ είναι αυτόνομα, ελαφρύτερα και πιο ευέλικτα σε σχέση με VM, ενώ παρέχουν αντίστοιχη απομόνωση. Επίσης, το Docker παρέχει την δυνατότητα σύνταξης ρουτινών δημιουργίας και ρύθμισης περιβάλλοντος για απρόσκοπτη και προβλέψιμη εγκατάσταση και, επιπλέον, επιτρέπει την χρήση έτοιμων προτύπων. Αξιοποιώντας τις δυνατότητες αυτές κατά την ανάπτυξη του συστήματος, ελαττώθηκε σημαντικά ο όγκος του τυποποιημένου κώδικα που χρειάστηκε να συνταχθεί.

4.14.2 Οργάνωση Ανάπτυξης με Docker

Η ανάπτυξη και εγκατάσταση των υπηρεσιών του συστήματος έχει οργανωθεί σε τρία επίπεδα, σύμφωνα με την τυπική προσέγγιση ανάπτυξης με Docker.

Επίπεδο Υπηρεσιών

Το σύστημα χωρίζεται υπηρεσίες που έχουν διαφορετικές απαιτήσεις και ανάγκες και, για αυτόν τον λόγο, εκτελούνται αυτόνομα και ανεξάρτητα. Η περιγραφή αυτών των υπηρεσιών γίνεται με την σύνταξη `docker-compose` σε ένα `yaml` έγγραφο, όπου διατυπώνονται οι οδηγίες δημιουργίας της υπηρεσίας. Στις οδηγίες περιλαμβάνονται τα αρχεία πηγαίου κώδικα της υπηρεσίας, οι πόροι του υπολογιστικού συστήματος που της διατίθενται, ο τρόπος επικοινωνίας με άλλες υπηρεσίες, οι ρυθμίσεις λειτουργίας και οι κρυφές παράμετροι. Εδώ ορίζονται και οι τόμοι του δίσκου όπου γίνεται η μόνιμη αποθήκευση των δεδομένων της ΒΔ.

Οι ανεξάρτητες υπηρεσίες που απαιτούνται για την λειτουργία του πρωτότυπου συστήματος είναι η ΒΔ, η υπηρεσία ETL που φιλοξενεί και την εκπαίδευση των μοντέλων, ο διαμεσολαβητής μηνυμάτων για την εκτέλεση των εργασιών, η διαδικτυακή διεπαφή και, τέλος, η διεπαφή χρήστη.

Επίπεδο Κοντέινερ

Οι οδηγίες χτισίματος κάθε κοντέινερ αποτυπώνονται σε ένα αρχείο `Dockerfile` εντός του πηγαίου κώδικα της υπηρεσίας. Τα βήματα ξεκινάνε με την χρήση έτοιμου προτύπου και ακολουθούνται από εντολές Docker που εγκαθιστούν τα απαραίτητα εργαλεία του λειτουργικού συστήματος, κάποια από τα οποία είναι κοινά για όλες τις υπηρεσίες. Έπειτα αντιγράφονται τα αρχεία που χρησιμοποιεί η υπηρεσία εντός του κοντέινερ, όπου βρίσκεται και η λίστα με απαιτούμενες βιβλιοθήκες για την εκτέλεση του κώδικα, οι οποίες εγκαθίστανται. Τέλος, καλούνται τα βήματα εκτέλεσης της εφαρμογής.

Επίπεδο Εκτέλεσης

Η διαδικασία εκτέλεσης παραμένει όσο το δυνατόν πιο λιτή γίνεται· ιδανικά, είναι απλά μία εντολή που ξεκινάει την λειτουργία της υπηρεσίας. Σε κάποιες περιπτώσεις, όμως, είναι απαραίτητη η παραμετροποίηση της εντολής εκτέλεσης ή η αρχικοποίηση κάποιων τμημάτων της εφαρμογής με μία υπορουτίνα SQL ή shell script.

4.14.3 Παραμετροποίηση Περιβάλλοντος Υπηρεσιών

Καθώς οι ρυθμίσεις που καθορίζουν την λειτουργία των υπηρεσιών ενδέχεται να χρησιμοποιούνται σε πολλά τμήματα του κώδικα, ακόμα και σε ξεχωριστές υπηρεσίες, ήταν σκόπιμο να καταγράφονται σε αρχεία ρυθμίσεων περιβάλλοντος (*env files*), από τα οποία τις εισάγουν οι υπηρεσίες που τις χρησιμοποιούν.

Κάποιες από τις μεταβλητές περιβάλλοντος, όπως ο κωδικός διαχειριστή της ΒΔ και τα κλειδιά πρόσβασης σε εξωτερικά API, είναι νευραλγικές για την ασφάλεια του συστήματος, και η διαρροή τους επιφέρει σημαντικές επιπτώσεις. Αυτές οι μεταβλητές δεν κοινοποιούνται και είναι ευθύνη του διαχειριστή που εγκαθιστά το σύστημα για πρώτη φορά να τις τοποθετήσει σε καθορισμένο σημείο των αρχείων όπου είναι προσβάσιμες στο επίπεδο των υπηρεσιών ως μυστικά (*secrets*), για να μεταφερθούν εντός των κοντέινερ.

4.14.4 Διασύνδεση Υπηρεσιών

Οι υπηρεσίες είναι τυπικά ανεξάρτητες, ωστόσο, η επικοινωνία μεταξύ τους είναι απαραίτητη για την εκτέλεση συγκεκριμένων λειτουργιών· κατά την καταχώρηση και την ανάκτηση στοιχείων από την ΒΔ από την ETL και το API ή την λήψη σήματος από το API για εκτέλεση εργασίας, για το οποίο πρέπει να ειδοποιηθεί η ETL μέσω του διαμεσολαβητή μηνυμάτων. Για αυτόν τον σκοπό, δημιουργείται ένα μικρό εικονικό δίκτυο στο επίπεδο των υπηρεσιών, στο οποίο οι υπηρεσίες ανοίγουν συγκεκριμένες πόρτες (*port binding*) ώστε να είναι δυνατό να καλούνται οι λειτουργίες τους από τις υπόλοιπες υπηρεσίες του συστήματος, χωρίς όμως αυτό να τις εκθέτει στο διαδίκτυο.

Το API και το UI είναι εξωστρεφείς υπηρεσίες, καθώς, για να φέρουν εις πέρας το έργο τους, απαιτείται η σύνδεση τους στο διαδίκτυο. Το άνοιγμα μίας υπηρεσίας στο διαδίκτυο δεν είναι καθόλου απλή υπόθεση και απαιτεί πολύ προσεκτικό σχεδιασμό και εξειδικευμένες ενέργειες για την περιφρούρηση της ασφάλειας των δεδομένων, των πόρων και της διαθεσιμότητας των υπηρεσιών και, για αυτούς τους λόγους, είναι εκτός του πλαισίου του πρωτότυπου συστήματος.

4.14.5 Βήματα Εγκατάσταση Συστήματος

Η δομή της ανάπτυξης που υλοποιήθηκε με το Docker αυτοματοποιεί το μεγαλύτερο μέρος των εργασιών εγκατάστασης και, έτσι, επιτυγχάνεται η φορητότητα του συστήματος. Εντούτοις, κατά την πρώτη εγκατάσταση, ο διαχειριστής του συστήματος θα πρέπει να εισάγει κάποια στοιχεία στα σωστά τμήματα του πηγαίου κώδικα. Θα πρέπει οι μυστικές μεταβλητές περιβάλλοντος να οριστούν εντός του φακέλου με τα μυστικά και να συμπληρωθούν ορθά οι υπόλοιπες μεταβλητές περιβάλλοντος στα αρχεία ρυθμίσεων. Όποτε αλλάζουν ή προστίθενται μεταβλητές περιβάλλοντος, θα πρέπει τα κοντέινερ που κάνουν χρήση τους να χτίζονται εκ νέου.

Επιπλέον, καλό είναι να αποτυπωθεί η αρχική κατάσταση του συστήματος στα αρχεία shell και SQL που εκτελούνται κατά την αρχικοποίηση της ΒΔ, καθώς και στα αρχεία CSV, geotiff και geojson, από τα οποία αντλούνται τα αρχικά δεδομένα, έτσι ώστε το σύστημα να ξεκινήσει την λειτουργία του με πλήρες αρχείο και έτοιμες ρυθμίσεις για άμεση παραγωγή.

4.15 Σενάρια Εκτέλεσης

Ο διαχειριστής ξεκινάει την ρύθμιση του Aristaeos. Ορίζει την περιοχή ευθύνης, κατά προτίμηση μέσω μία γεωχωρικής διεπαφής χρήστη. Ορίζει τις μετρήσιμες μετεωρολογικές μεταβλητές των σταθμών του δικτύου που χρησιμοποιεί, καθώς και τους μετεωρολογικούς σταθμούς του δικτύου με τις συντεταγμένες τους και τις διευθύνσεις ανάκτησης δεδομένων. Τέλος, ορίζει τα γεωχωρικά στρώματα που χρησιμοποιούνται και τα ανεβάζει μέσω διεπαφής, ή δίνει τις διαδικτυακές τοποθεσίες τους και τον ρυθμό ανανέωσης. Σε αυτό το σημείο το σύστημα ξεκινάει να συλλέγει μετρήσεις.

Για την παραγωγή προγνωστικών, θα πρέπει αρχικά να αντιστοιχιστούν, μέσω φόρμας η της ΒΔ, οι καιρικές μεταβλητές με καιρικά χαρακτηριστικά και τα γεωχωρικά στρώματα με γεωχωρικά χαρακτηριστικά, ρυθμίζοντας τις σχέσεις σύμφωνα με την παράγραφο για την μηχανική χαρακτηριστικών της ενότητας 4.7.2. Έπειτα, επιλέγει έναν διαχωρισμό πλέγματος και δημιουργούνται οι τομείς διαχωρισμού της περιοχής. Δεδομένης της ύπαρξης ιστορικού αρχείου, ο διαχειριστής μπορεί να επιλέξει ένα σύνολο χαρακτηριστικών και ένα χρονικό διάστημα για την δημιουργία συνόλου δεδομένων. Τέλος, δημιουργεί μοντέλο με ένα από τα σύνολα δεδομένων και τις παραμέτρους που επιθυμεί και θα ειδοποιηθεί για την απόδοσή του, όταν είναι έτοιμο. Εάν κρίνει πως οι μετρικές είναι ικανοποιητικές, το ενεργοποιεί για την περιοχή που μελετάει και το σύστημα ξεκινάει άμεσα να παράγει προβλέψεις με το νέο μοντέλο.

Ο χρήστης που επιθυμεί να ενημερωθεί για τον κίνδυνο πυρκαγιάς, ανοίγει την ιστοσελίδα σε φυλλομετρητή ιστού, περιηγείται μέσω του διαδραστικού χάρτη στην περιοχή που τον ενδιαφέρει και μελετάει τον χάρτη κινδύνου.

4.16 Έλεγχος και Δοκιμές

4.16.1 Τοπικές Δοκιμές

Κατά την πρωτοτύπηση του έργου, το μεγαλύτερο μέρος των δοκιμών (*testing*) διενεργήθηκε σε τοπικό περιβάλλον, ήτοι, στο μηχάνημα που χρησιμοποιήθηκε για την ανάπτυξη του λογισμικού.

Έλεγχος Λειτουργικότητας

Οι δοκιμές λειτουργικότητας του κώδικα ήταν δομημένες με απλό τρόπο και πραγματοποιήθηκαν με την παρακολούθηση μεταβλητών και σφαλμάτων εκτέλεσης. Τα σενάρια δοκιμών νέων λειτουργιών ξεκινούσαν με απομονωμένη εκτέλεση των προστιθέμενων τμημάτων κώδικα για την επιβεβαίωση της λειτουργικότητάς τους. Εφόσον αυτό έγινε με επιτυχία, ακολουθούσε εκτέλεση του συνόλου της ροής εργασιών, συμπεριλαμβανμένης της προσθήκης, σε δοκιμαστικά δεδομένα και δεδομένα παραγωγής, με στόχο τον εντοπισμό απρόσμενων επιδράσεων. Το τελικό στάδιο δοκιμών ήταν η προσθήκη του συνόλου των αλλαγών στην εκτέλεση της περιοδικής διαδικασίας

συλλογής και ανάλυσης δεδομένων και παραγωγής προβλέψεων για ένα διάστημα. Εφόσον δεν παρουσιαστούν προβλήματα στο διάστημα ζωντανής δοκιμής, οι αλλαγές προχωρούσαν στο αποθετήριο του έργου στο git.

Για τις δοκιμές των υπηρεσιών εντός κοντέινερ, αξιοποιήθηκε το εργαλείο αυτόματης ανανέωσης (*watch*) που παρέχει το Docker. Με την κατάλληλη ρύθμιση στο επίπεδο των υπηρεσιών, επιτεύχθηκε άμεση ανανέωση των αρχείων εντός των κοντέινερ με την αποθήκευση αλλαγών του πηγαίου κώδικα, γεγονός που βοήθησε στην απρόσκοπτη δοκιμή της λειτουργικότητας των υπηρεσιών στο σύνολό τους και της μεταξύ τους επικοινωνίας. Οι αλλαγές στο σχήμα της βάσης, στις μεταβλητές περιβάλλοντος ή στις απαιτούμενες βιβλιοθήκες και εργαλεία πρέπει είτε να δοκιμαστούν με εκ νέου χτίσιμο του κοντέινερ, είτε να γίνουν χειροκίνητα.

Δοκιμές Παραμετροποίησης

Καθώς οι παράμετροι που επηρεάζουν την εκτέλεση των εργασιών βρίσκονται στην ΒΔ, η δοκιμή διαφορετικών τιμών έγινε με την εισαγωγή νέων αντικειμένων ή την αλλαγή των τιμών των υπάρχουσών εγγραφών και την εκτέλεση της ροής εργασιών που αφορά τις συγκεκριμένες παραμέτρους.

4.16.2 Δοκιμές Ανάπτυξης Συστήματος

Οι δοκιμές ανάπτυξης του συστήματος έγιναν σε εξωτερικό πάροχο υπηρεσιών (*hosting service*). Για την σύνδεση στο απομακρυσμένο μηχάνημα χρησιμοποιήθηκε το πρωτόκολλο SSH, η αντιγραφή και ανανέωση του πηγαίου κώδικα έγινε μέσω git, τα κρυφά αρχεία και τα μεγάλα αρχεία δεδομένων μεταφέρθηκαν με το εργαλείο SCP. Αυτό το σύνολο εργαλείων επέτρεψε το συγχρονισμό του απομακρυσμένου μηχανήματος με τον πηγαίο κώδικα, την άμεση αρχικοποίηση και χτίσιμο των υπηρεσιών του συστήματος και τον συνεχή έλεγχο του μέσω των εγγραφών του υποσυστήματος καταγραφής στα logs και την ΒΔ.

4.17 Περιορισμοί και Αδυναμίες Υλοποίησης

Μολονότι η μελέτη καλύπτει το σύνολο της προσχεδιασμένης ροής εργασιών, από την συλλογή δεδομένων μέχρι την παρουσίαση των προβλέψεων, το υλοποιημένο σύστημα παραμένει ένα ερευνητικό πρωτότυπο. Τα αδύναμα σημεία που εντοπίζονται, είναι μεν περιοριστικά για τα θετικά πορίσματα πρακτικότητας που μπορούν να συναχθούν, αλλά παρέχουν κατεύθυνση για τις μελλοντικές εργασίες αναβάθμισης.

4.17.1 Περιορισμοί Συστήματος

Το τεχνικό επίπεδο της υλοποίησης ενισχύει, μεν, σημαντικά την ευελιξία και ευχρηστία του συστήματος, αλλά, οι κύριοι παράγοντες που κρίνουν την αξία του ως πρακτικό εργαλείο είναι ανεξάρτητοι από αυτό: η επιχειρησιακή επιτυχία του κρίνεται από την χρησιμότητα των παραγόμενων προβλέψεων, δηλαδή την αξιοπιστία τους. Η ορθή μεθοδολογία, ο όγκος των διαθέσιμων δεδομένων και η ποιότητά τους είναι κρίσιμοι συντελεστές, οι οποίοι, απαιτούν μία οργανωμένη υλοποίηση για να αξιοποιηθούν στο έπακρο, δεν αναπληρώνονται όμως, από αυτή.

Η υποτυπώδης υλοποίηση του API και της διεπαφής χρήστη, σημαίνει πως το μεγαλύτερο τμήμα των εντολών και ρυθμίσεων στο σύστημα γίνονται σε χαμηλό επίπεδο. Ένας ερευνητής που ασχολείται με το έργο στην τρέχουσα κατάσταση, θα πρέπει να είναι εξοικειωμένος με συστήματα διαχείρισης ΒΔ, την γραμμή εντολών και την γλώσσα Python.

Η ανεπάρκεια των μοχλών χειρισμού υψηλού επιπέδου, σε συνδυασμό με την έλλειψη συστήματος

αυθεντικοποίησης και την απουσία μεθοδικών δοκιμών υπό υψηλό φόρτο, διασφαλίζουν πως το σύστημα απέχει πολύ ακόμα από την ένταξη στην παραγωγή.

4.17.2 Συμβιβασμοί στην Σχεδίαση

Οι περιορισμοί χρόνου, μέσων και διαθέσιμων δεδομένων, ώθησαν την λήψη αποφάσεων σχεδίασης, οι οποίες αποσκοπούσαν στην ελάττωση της πολυπλοκότητας, την μειωμένη χρήση υπολογιστικών πόρων και την συντομότερη ανάπτυξη του πρωτότυπου.

Κατά τον αρχικό σχεδιασμό, τα επεξεργασμένα χαρακτηριστικά εισόδου, καθώς και οι προβλέψεις, επρόκειτο να αποθηκεύονται μόνιμα. Εντούτοις, για εξοικονόμηση αποθηκευτικού χώρου, και δεδομένου ότι είναι δυνατή επαναδημιουργία τους από τα πρωτογενή δεδομένα ανά πάσα στιγμή, δεν κρίθηκε σκόπιμη η μόνιμη καταχώρηση τους.

Καθώς η πρωτότυπη μοντελοποίηση αξιοποίησε ελάχιστες γεωχωρικές μεταβλητές, η ανάλυση διαχωρισμού της περιοχής μελέτης παρέμεινε μικρή. Συνεπώς, δεν είχε και νόημα η υλοποίηση αναπτυγμένης μεθόδου παρεμβολής των καιρικών συνθηκών· η εύρεση κατάλληλης μεθόδου παρεμβολής συνθηκών για την πρόβλεψη πυρκαγιών απαιτεί και αξίζει επιπλέον έρευνα.

Ανοιχτό θέμα είναι και η μέθοδος εξαγωγής κατωφλιών απόφασης, η υλοποίηση της οποίας στο πρωτότυπο σύστημα ήταν υποτυπώδης.

Η χρηστικότητα της τελικής εφαρμογής για το κοινό εξακολουθεί να είναι περιορισμένη, εφόσον δεν έχει εξασφαλιστεί η αξιοπιστία των προβλέψεων. Για αυτό, αποφασίστηκε πως δεν είναι συνετό να προχωρήσουν οι εργασίες υλοποίησης διεπαφής και εμπειρίας χρήστη (UI/UX), στην παρούσα φάση.

4.18 Μελλοντική Υλοποίηση

Οι επιθυμητές μελλοντικές εργασίες αφορούν την διευθέτηση συγκεκριμένων κενών στην λειτουργικότητα, καθώς και την άμβλυνση των ζητημάτων που τέθηκαν στην ενότητα 4.17. Τα επίπεδα λειτουργικότητας των υποσυστημάτων δίνονται επιγραμματικά στον πίνακα 4.4. Πλήρως λειτουργικά θεωρούνται τα συστήματα που δεν χρήζουν άμεσων βελτιώσεων. Λειτουργικά θεωρούνται αυτά που καλύπτουν το μεγαλύτερο μέρος των αρμοδιοτήτων τους, αλλά επιδέχονται κάποιων προσθηκών που θα βοηθήσουν σημαντικά το σύστημα. Πιλοτικά είναι αυτά που έχουν υποτυπώδη υλοποίηση και απαιτούν αρκετές εργασίες ακόμα για να λάβουν την τελική τους μορφή.

4.18.1 Υπολειπόμενη Λειτουργικότητα

Επεκτάσεις Ροής ETL

Το κομμάτι της διαχείρισης γεωχωρικών στρώσεων που απομένει να προγραμματιστεί, αφορά την αποθήκευση και εξαγωγή χαρακτηριστικών από τις διανυσματικές στρώσεις, καθώς και την αυτόματη ανανέωση των ταχέως μεταβαλλόμενων στρώσεων. Αναφορικά με τις μεθόδους διαχωρισμού, πρέπει να συμπεριληφθεί και ο διαχωρισμός σε εξάγωνα, που θεωρείται πιο αποτελεσματική δομή για την γεωχωρική ανάλυση.

Η κλάση που διαχειρίζεται την ανάθεση μετεωρολογικών τιμών στα κελιά θα πρέπει, κατόπιν έρευνας, να εμπλουτιστεί με πιο προηγμένες μεθόδους παρεμβολής. Αυτές οι μέθοδοι, θα πρέπει να

Πίνακας 4.4: Στάδια Υλοποίησης Υποσυστημάτων

Υποσύστημα	Βαθμός Ολοκλήρωσης	Μελλοντικές Εργασίες
Βάση Δεδομένων	Λειτουργικό	Δημιουργία ρουτίνας backup, Δημιουργία ρουτίνας migration, Υποστήριξη διανυσματικών στρώσεων, δημιουργία υπορουτίνας διαχωρισμού σε εξάγωνα
Υποσύστημα Εκπαίδευσης	Λειτουργικό	Βελτιστοποίηση διαχείρισης πόρων, Προσθήκη νέων αλγόριθμων, Προσθήκη νέων παραμέτρων εκπαίδευσης, Δημιουργία ρουτίνας επανεκπαίδευσης
Υποσύστημα Συλλογής Δεδομένων	Λειτουργικό	Δημιουργία ρουτίνας ανανέωσης γεωχωρικών στρώσεων, Προσθήκη νέων πηγών δεδομένων
Υποσύστημα Επεξεργασίας	Λειτουργικό	Επέκταση παρεμβολής καιρικών συνθηκών, Υποστήριξη παραγόντων ιστορικού πυρκαγιών, Προσθήκη νέων πυρικών παραγόντων
Υποσύστημα Πρόβλεψης	Πλήρως Λειτουργικό	-
Υποσύστημα Καταγραφής	Πλήρως Λειτουργικό	-
Διαδικτυακή Διεπαφή	Πιλοτικό	Σχεδίαση και εφαρμογή συστήματος ταυτοποίησης, Προστασία από εξωτερικές επιθέσεις, Βελτίωση διαθεσιμότητας, Υλοποίηση πόρων διανομής συμβάντων, Υλοποίηση λειτουργιών διαχειριστή
Διεπαφή Χρήστη	Πιλοτικό	Σχεδίαση και δημιουργία φορμών για λειτουργίες διαχειριστή, Δημιουργία μηχανισμού ειδοποιήσεων, Βελτίωση εμφάνισης και εμπειρίας χρήστη

βελτιώσουν την ακρίβεια της ανάθεσης καιρικών συνθηκών στους τομείς, ειδικά όταν η πυκνότητα των σταθμών δεν επαρκεί για την ικανοποιητική κάλυψη της μορφολογίας του εδάφους.

Επεκτάσεις Εργαλείων ML

Στο υποσύστημα ML, το εύρος της παραμετροποίησης είναι άμεσα διαθέσιμο για διεύρυνση, κυρίως ως προς τον κύριο αλγόριθμο και τα βήματα προετοιμασίας των δεδομένων. Κάποιοι αλγόριθμοι ML που μπορούν να εξεταστούν στο μέλλον είναι εναλλακτικές τεχνικές Gradient Boosting παραπλήσιες με τον XGB, η παραλλαγή του RF Extremely Random Trees και τα SVM, που αναφέρονται στην βιβλιογραφία.

Σε επόμενο στάδιο, μπορούν να υλοποιηθούν ρουτίνες για την εκπαίδευση και χρήση αλγορίθμων DL. Οι αλγόριθμοι DL που λειτουργούν σε χρονικά παράθυρα, όπως ο RNN, η σε χωρικούς τομείς, όπως οι αλγόριθμοι CNN, ταιριάζουν στην χωροχρονική φύση του προβλήματος. Απαιτούν, όμως, διαφορετικά πρωτόκολλα δημιουργίας συνόλου δεδομένων και προετοιμασίας των δεδομένων που θα πρέπει να ερευνηθούν και να υλοποιηθούν.

Ήδη έχουν γίνει ενέργειες για την άμβλυνση του προβλήματος της κλίμακας των μεγάλων συνόλων δεδομένων, όπως η χρήση βελτιστοποιημένων μεθόδων και βιβλιοθηκών. Για πολύ μεγάλα αρχεία με υψηλή ανάλυση, μπορούν να μελετηθούν και να εφαρμοστούν τεχνικές αποδεδκατισμού (*downsampling*) των στιγμιοτύπων με μικρή επίδραση στα αποτελέσματα της εκπαίδευσης.

Για την διατήρηση της απόδοσης του συστήματος στις μεταβαλλόμενες συνθήκες, θα πρέπει τελικά να υλοποιηθούν και μέθοδοι επανεκπαίδευσης των μοντέλων πάνω στα πιο πρόσφατα δεδομένα.

Τέλος, θα πρέπει να προγραμματιστεί η ρουτίνα παραλαβής σήματος εργασίας, ώστε η εκκίνηση εκπαίδευσης να είναι διαθέσιμη σε εργαλεία διαχείρισης υψηλού επιπέδου.

Επεκτάσεις Διαδικτυακού API

Προτού το σύστημα προταθεί σε φορείς, θα πρέπει να συνοδεύεται από μία ολοκληρωμένη εργαλειοθήκη διαχείρισης σε υψηλό επίπεδο, η οποία προτείνεται να αναπτυχθεί στο API. Αυτά τα εργαλεία συμπεριλαμβάνουν πόρους για εύκολη πρόσβαση σε ενέργειες, όπως αυτές που περιγράφονται στα σενάρια εκτέλεσης 4.15, σε συνδυασμό με ένα σύστημα αυθεντικοποίησης, που θα διασφαλίζει την εξουσιοδοτημένη πρόσβαση σε ρουτίνες παραμετροποίησης.

Επιπροσθέτως, θα πρέπει να δημιουργηθεί και μία σειρά από πόρους που θα διανέμουν τα πρόσφατα συμβάντα που αιχμαλωτίζει το υποσύστημα καταγραφής. Ο στόχος με αυτό είναι η ενημέρωση των χρηστών για μη αξιόπιστες προβλέψεις, καθώς και των διαχειριστών για τυχόν δυσλειτουργίες. Όταν η εφαρμογή βρεθεί σε ώριμο στάδιο, μπορούν να δημιουργηθούν και κανάλια ενημέρωσης των χρηστών για πολύ υψηλά επίπεδα κινδύνου.

Επεκτάσεις Διεπαφής Χρήστη

Η ανάπτυξη της διεπαφής χρήστη ακολουθεί στενά την ανάπτυξη του API, καθώς βασίζεται στις δυνατότητες που παρέχονται από αυτό: Για την αποτελεσματική διαχείριση του συστήματος απαιτείται μία σειρά από φόρμες (*forms*) για την εισαγωγή και την επεξεργασία οντοτήτων του συστήματος. Επιπλέον, για την αξιοποίηση των πόρων ενημέρωσης, συνιστάται η χρήση συστήματος ειδοποιήσεων (*notifications*). Για την τυποποίηση, την οργάνωση και την ευκολία διαχείρισης

του σημαντικού αριθμού των φορμών, των γραφικών στοιχείων και των ρουτινών ανάκτησης και αποστολής, προτείνεται η χρήση ολοκληρωμένου πλαισίου διαχείρισης JavaScript, πχ. React.

4.18.2 Μελέτη Συστήματος Παραγωγής

Η ένταξη του Aristaeos σε παραγωγικές διαδικασίες απαιτεί ξεχωριστή μελέτη αναφορικά με τις απαραίτητες εργασίες. Πέρα από την εφαρμογή συστήματος αυθεντικοποίησης, αναμένεται να χρειαστούν εργασίες στους εξής τομείς: Ορισμός τυποποιημένης ροής εφαρμογής αλλαγών στο git και εφαρμογής αλλαγών στην ΒΔ (*migrations*). Εύρεση και τακτοποίηση των κενών ασφαλείας, ακολουθούμενη από ελεγχόμενο άνοιγμα επιλεγμένων τμημάτων του συστήματος στο διαδίκτυο. Ορισμός ρουτίνας δημιουργίας αντιγράφων ασφαλείας (*backups*). Στήσιμο υποδομών εξισορρόπησης φόρτου από κλιμακούμενο αριθμό ταυτόχρονων χρηστών. Δημιουργία υπηρεσίας ανάθεσης πόρων κατανεμημένης επεξεργασίας για εργασίες ML. Επέκταση του υποσυστήματος καταγραφής και δημιουργία οργανωτικού πίνακα (*dashboard*) διαχειριστή.

4.19 Σύνοψη Κεφαλαίου

Η ανάλυση της αρχιτεκτονικής και της τεχνικής υλοποίησης του Aristaeos, παρέχει κατεύθυνση σε όποιον θέλει να το αξιοποιήσει και να το επεκτείνει. Αν και το κλειδί της επιτυχίας του προτεινόμενου συστήματος είναι η ορθή μεθοδολογία, η δομημένη υλοποίηση είναι το όχημα που θα οδηγήσει στην πρακτική εφαρμογή του.

Κεφάλαιο 5

Αποτελέσματα

5.1 Επισκόπηση Κεφαλαίου

Σε αυτό το κεφάλαιο θα παρουσιαστούν επιγραμματικά τα αποτελέσματα των δοκιμών του συστήματος. Οι δοκιμές καλύπτουν την απόδοση και την χρήση πόρων από τις εργασίες, την ανάλυση των δεδομένων εκπαίδευσης, καθώς και τα πειράματα απόδοσης των μοντέλων. Θα δοθεί έμφαση στις μετρικές αξιολόγησης των αποτελεσμάτων και την ανάλυση των περιπτώσεων που το σύστημα αποτυγχάνει.

5.2 Μετρικές Ροής Εργασιών

Η δοκιμή ανάπτυξης του συστήματος σε εξωτερικό πάροχο hosting διήρκησε από την 23-11-2025 μέχρι και την 11-05-2026, σύνολο 169 ημέρες. Η δοκιμή προσομοίωσε την τυπική ροή εργασιών πρόβλεψης, για 8 εικονικούς σταθμούς. Κατά το διάστημα αυτό, οι μετρήσεις συλλέχθηκαν ανελλιπώς από το εξωτερικό API, πλην από ένα κενό μισής ώρας για το διάστημα 25-02-2026 23:00-23:30, λόγω εκτεταμένης διακοπής της σύνδεσης.

Κατά την λειτουργία, καταχωρήθηκαν 649.144 εγγραφές καιρικών συνθηκών, με την κάθε εγγραφή να μετράει μία από 11 καιρικές μεταβλητές. Από τις εγγραφές προκύπτουν 64.896 στιγμιότυπα με σταθμό και χρονικό διάστημα. Το σύνολο των εγγραφών έχει χωριστεί σε 25 υποπίνακες από την TimescaleDB. Ο χώρος που καταλαμβάνουν στον δίσκο είναι 146, 40 MB.

Η ροή εργασιών περιλάμβανε ανάκτηση από τους σταθμούς, αποθήκευση των μετρήσεων, επεξεργασία και εξαγωγή προβλέψεων με μοντέλο RF για κελιά ανάλυσης 10 χμ. (77 τομείς). Ο χρόνος εκτέλεσης του συνόλου της ροής στο απομακρυσμένο μηχάνημα κυμάνθηκε στα 15s – 50s, ενώ στο μηχάνημα ανάπτυξης ανέρχεται σε 4s. Για την ίδια ροή με διαχωρισμό 1χμ (7.800 τομείς), χωρισμένους σε παρτίδες των 500, ο συνολικός χρόνος εκτέλεσης στο μηχάνημα ανάπτυξης ήταν 6s.

5.3 Αποτίμηση Διαδικασίας Εκπαίδευσης

5.3.1 Περιγραφή Συνόλου Εκπαίδευσης

Το σύνολο εκπαίδευσης αφορά μετρήσεις 5 ετών από 7 μετεωρολογικούς σταθμούς, κατόπιν αφαίρεσης του ορεινού σταθμού ΑΝΩΓΕΙΑ. Για την εκπαίδευση, αξιοποιήθηκαν τα 17 χαρακτηριστικά πρόβλεψης του πίνακα 5.1.

Πίνακας 5.1: Χαρακτηριστικά Πρόβλεψης

Κατηγορία	Χαρακτηριστικά
Απλά Καιρικά	Θερμοκρασία Αέρα, Ατμοσφαιρική Υγρασία, Ηλιακή Ακτινοβολία, Υετός Ημίωρου, Ατμοσφαιρική Πίεση
Αθροιστικά Καιρικά	Υετός Ημέρας, Υετός Εβδομάδος, Υετός Μήνα, Υετός Δίμηνου, Υετός Τρίμηνου, Υετός Εξάμηνου
Φάση	Φάση Ημέρας ημω, Φάση Ημέρας συνω, Φάση Εβδομάδος ημω, Φάση Εβδομάδος συνω, Φάση Έτους ημω, Φάση Έτους συνω,
Γεωχωρικά	Μέσο Υψόμετρο

Το αποτύπωμα του συνόλου των στιγμιότυπων στο σύστημα, μετά το φιλτράρισμα και την αντιστοίχιση με τομείς από πλέγματα διαχωρισμού, φαίνεται στον πίνακα 5.2. Η εκτίμηση χώρου στην μνήμη τυχαίας προσπέλασης αφορά μορφή Polars DataFrame.

Πίνακας 5.2: Αποτύπωμα Συνόλου Δεδομένων για Διαφορετικούς Τρόπους Διαχωρισμού

Μέγεθος	Πλέγμα Voronoi	Πλέγμα 10χμ	Πλέγμα 5χμ
Αριθμός Τομέων	7	77	336
Αριθμός Στιγμιότυπων	608.882	6.699.791	23.080.086
Χρόνος Δημιουργίας	26,04s	31,01s	42,94s
Χώρος στην Μνήμη	102,41 MB	1.126,83 MB	3.796,87 MB

5.3.2 Μετρικές Εκπαίδευσης

Η εκπαίδευση έγινε στο μηχάνημα ανάπτυξης λογισμικού: Κεντρική Μονάδα Επεξεργασίας 10 πυρήνων μέγιστης συχνότητας 4.5 GHz και 24 GB μνήμης τυχαίας προσπέλασης. Για τους RF και XGB χρησιμοποιήθηκαν όλοι οι πυρήνες ενώ ο LR εκτελέστηκε χωρίς παραλληλοποίηση. Τα αποτελέσματα χρόνου εκτέλεσης, σε δευτερόλεπτα, των βημάτων εκπαίδευσης επιλεγμένων μοντέλων με διαχωρισμό Voronoi φαίνονται στον πίνακα 5.3.

Πίνακας 5.3: Χρόνος Διεκπεραίωσης (s) Βημάτων Εκπαίδευσης

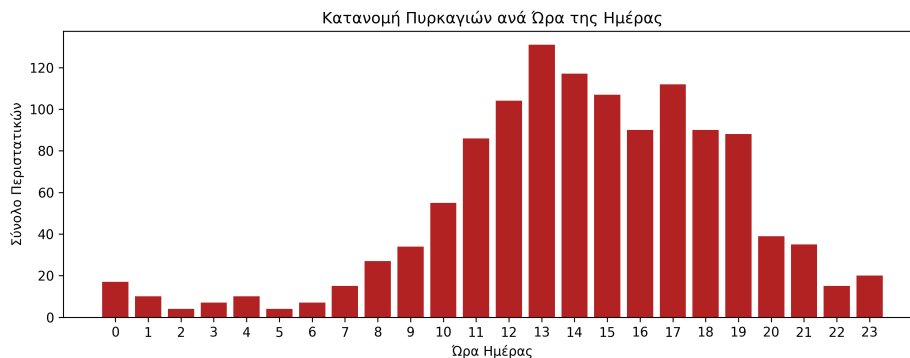
Μοντέλο	Υπολογισμός Κατωφλίων	Ρύθμιση Υπερπαραμέτρων	Εκπαίδευση
LR με PF	129,50	923,00	45,09
RF	126,03	6.108,89	42,88
XGB	4,25	4.999,89	1,32
XGB (5χμ)	108,74	-	31,93

Η υλοποίηση του RF και του LR, που παρέχεται από το scikit-learn, χρησιμοποιεί Pandas DataFrames και NumPy πίνακες εσωτερικά. Αυτό έχει ως αποτέλεσμα το σύνολο δεδομένων να καταλαμβάνει πολλαπλάσια μνήμη και να τερματίζει απότομα, όταν οι τομείς είναι πολλοί. Η υλοποίηση του XGB που χρησιμοποιήθηκε μπορεί να υποστηρίξει σύνολα τομέων υψηλής ανάλυσης πολύ πιο αποτελεσματικά.

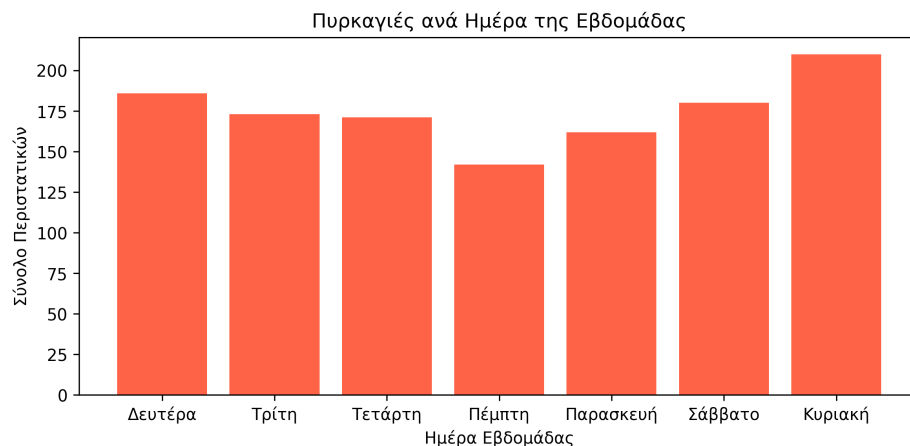
5.4 Στατιστικά Εκδήλωσης Πυρκαγιών

5.4.1 Περιγραφικά Στατιστικά Συμβάντων

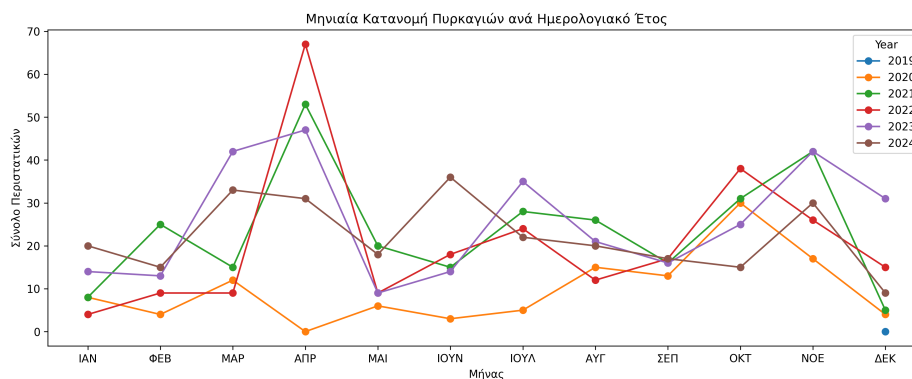
Για την εκπαίδευση χρησιμοποιήθηκαν τα συμβάντα πυρκαγιών εντός της περιοχής μελέτης και του χρονικού διαστήματος που καλύπτει το σύνολο εκπαίδευσης. Κατόπιν της αφαίρεσης των συμβάντων χωρίς καμμένη έκταση και σε χωματερές, απομένουν 1.224 περιστατικά πυρκαγιών.



Σχήμα 5.1: Κατανομή Αριθμού Πυρκαγιών ανά Ώρα της Ημέρας (Περιοχή Μελέτης 2020 - 2025)



Σχήμα 5.2: Κατανομή Αριθμού Πυρκαγιών ανά Ημέρα της Εβδομάδος (Περιοχή Μελέτης 2020 - 2025)



Σχήμα 5.3: Κατανομή Αριθμού Πυρκαγιών ανά Μήνα και Ημερολογιακό Έτος (Περιοχή Μελέτης 2020 - 2025)

Η χρονική κατανομή αυτών των συμβάντων μέσα στην ημέρα, όπως φαίνεται στο σχήμα 5.1 ακολουθεί το γενικότερο μοτίβο των πυρκαγιών στην Κρήτη (Σχ. 3.7), με το 60.8% των περιστατικών να λαμβάνουν χώρα κατά τις μεσημεριανές εργάσιμες ώρες 11-17. Η κατανομή των συμβάντων ανά τις μέρες της εβδομάδας είναι σχετικά ομοιόμορφη, με τις Πέμπτες να έχουν ελαφρώς μειωμένη εκδήλωση περιστατικών και τις Κυριακές ελαφρώς αυξημένη (Σχ. 5.2), σε συμφωνία με τα γενικότερα μοτίβα (Σχ. 3.8). Η κατανομή των πυρκαγιών εντός του έτους στην Π.Ε Ηρακλείου παρουσιάζει διαφορές σε σχέση με την γενική κατανομή των συμβάντων 25ετίας στην Κρήτη (Σχ. 3.3): Στην περιοχή μελέτης την τελευταία 5ετία, εκδηλώνεται αυξημένος αριθμός συμβάντων σε συγκεκριμένους μήνες όπως ο Απρίλιος, για τα έτη μετά το 2020, και ο Νοέμβριος· αυτό παρουσιάζεται στο γράφημα 5.3. Όμως, όσον αφορά την καμένη έκταση, οι πυρκαγιές της θερινής περιόδου, και συγκεκριμένα του Ιουλίου, διατηρούν σταθερά την πρωτιά, με μεγάλη διαφορά.

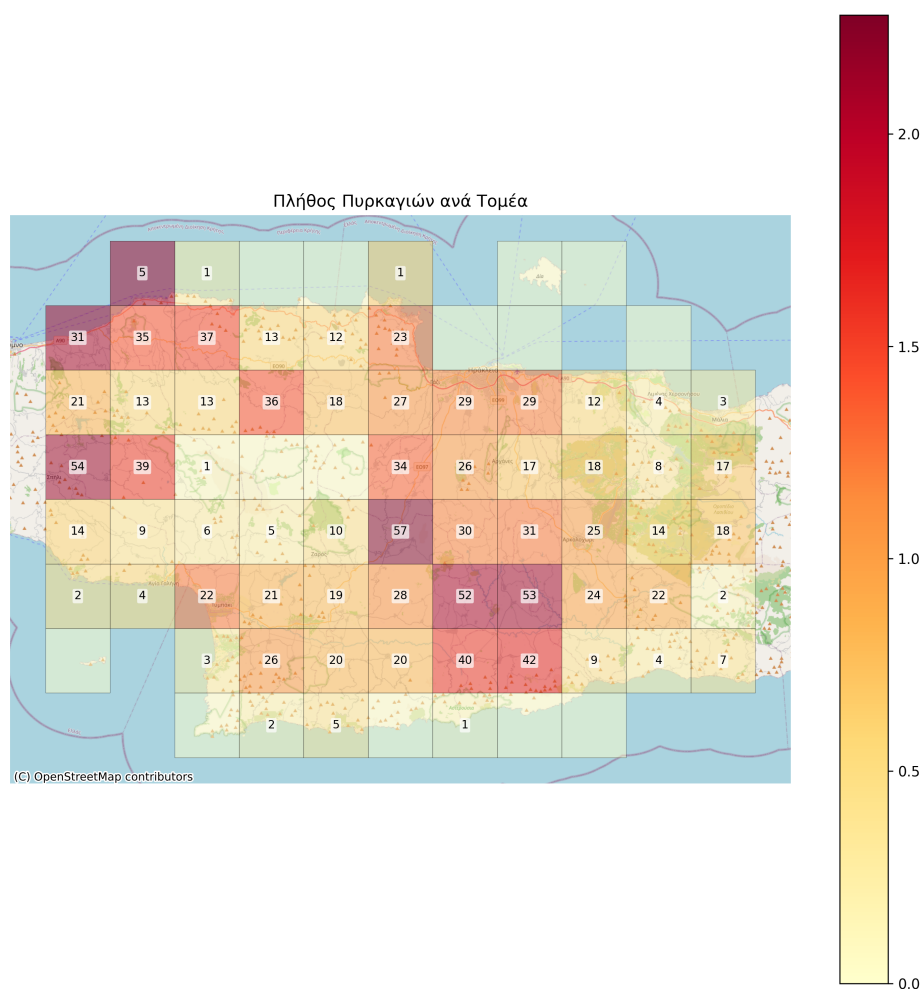
Η χωρική κατανομή του αριθμού πυρκαγιών στην περιοχή μελέτης παρουσιάζει ανομοιομορφίες. Συγκεκριμένοι τομείς εμφανίζουν σημαντικά αυξημένο αριθμό περιστατικών, όπως φαίνεται στην πυκνότητα πυρκαγιών ανά τομέα (Σχ. 5.4). Η μεγαλύτερη συγκέντρωση απαντάται στο ανατολικό τμήμα της πεδιάδας της Μεσσαράς (γύρω από τα χωριά Πύργος και Αγ. Βαρβάρα), στο ανατολικό άκρο της κοιλάδας του Αμαρίου (κοντά στο χωριό Σπήλι), και στο ανατολικό άκρο της κοιλάδας του Μυλοποτάμου.

Υψηλό ενδιαφέρον παρουσιάζει και η ζώνη γύρω από το Ηράκλειο· η πυκνότητα εκδήλωσης πυρκαγιών στις ελαιοκομικές και φρυγανικές εκτάσεις του WUI φαίνεται αυξημένη, εάν ληφθεί υπ' όψιν η μεγάλη αστική και βιομηχανική έκταση των τομέων της πόλης. Γενικότερα, η ενδοχώρα του Ηρακλείου έχει διάσπαρτες "θερμές ζώνες" με σχετικά συχνή εκδήλωση πυρκαγιών.

Την χαμηλότερη εκδήλωση πυρκαγιών παρουσιάζουν, αναμενόμενα, περιοχές με ελάχιστη βλάστηση, όπως το όρος Ψηλορείτης, η οροσειρά Δίκτη, καθώς και οι νότιες ακτές του νησιού στα Αστερούσια όρη. Στις λίγες και διάσπαρτες εκτάσεις πυκνής φυσικής βλάστησης, όπως το Δάσος Ρούβα, η πυκνότητα εκδήλωσης πυρκαγιών είναι περιορισμένη.

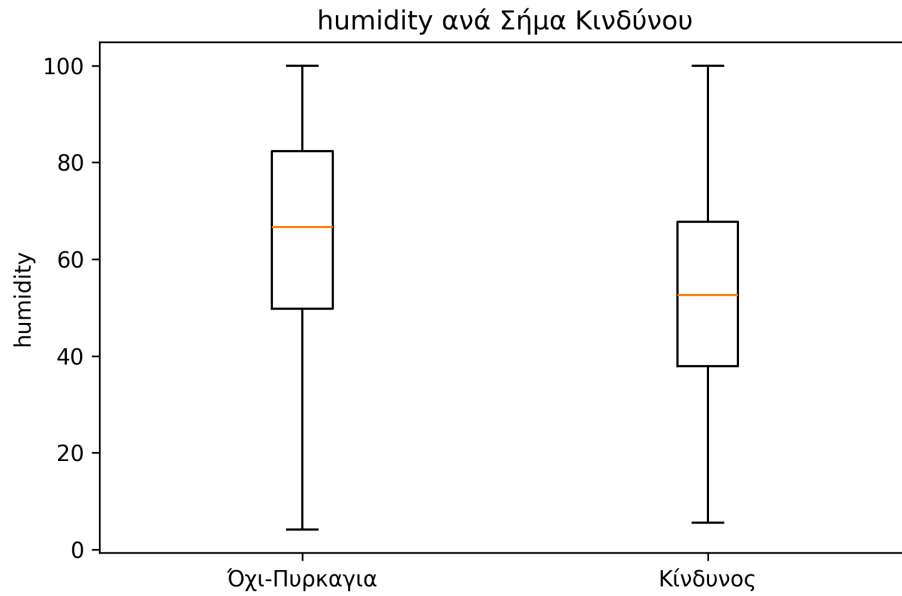
5.4.2 Στατιστικά Επικίνδυνων Συνθηκών

Εξετάζοντας τις συνθήκες που επικρατούσαν 3 ώρες πριν την εκδήλωση πυρκαγιάς διαμορφώνεται καλύτερη εικόνα των επικίνδυνων καταστάσεων που οδηγούν σε πυρκαγιές. Η κατανομή των επικίνδυνων καταστάσεων αναφορικά με την θερμοκρασία (Σχ. 5.5), δείχνει μία θετική σχέση



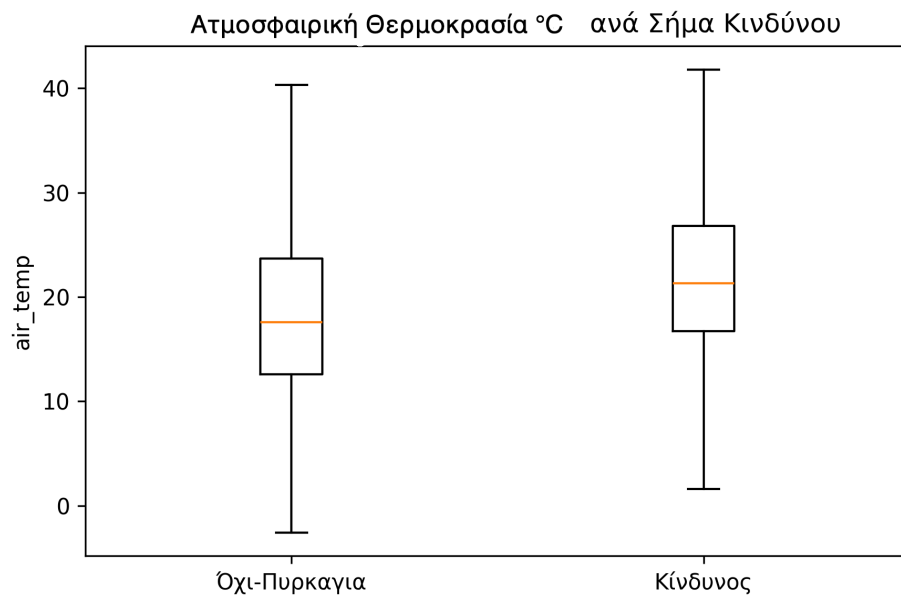
Σχήμα 5.4: Κατανομή Αριθμού Πυρκαγιών ανά Τομέα (Περιοχή Μελέτης 2020 - 2025) **Υπόβαθρο:**© OpenStreetMap contributors, available under the Open Database License.

υψηλής θερμοκρασίας και επικίνδυνων συνθηκών. Αντίστοιχα, η υγρασία (Σχ. 5.6) και ο πρόσφατος υετός δείχνουν αρνητική σχέση. Επιπλέον πληροφορία μπορεί να αντληθεί εξετάζοντας αυτές τις μεταβλητές σε σχέση με την εποχή του χρόνου (Σχ. 5.7 και 5.8)· οι ακραίες, για την εποχή, συνθήκες ζέστης και ξηρασίας φαίνεται να σχετίζονται με αυξημένο κίνδυνο κατά τους μήνες της άνοιξης και ειδικά του καλοκαιριού. Εάν, πάλι, εξεταστούν μαζί η θερμοκρασία με την υγρασία (Σχ. 5.9), παρατηρείται μία "επικίνδυνη ζώνη" συνθηκών που χαρακτηρίζεται από πολύ χαμηλή υγρασία και υψηλές θερμοκρασίες. Αυτές οι συνθήκες έχουν σημαντικές πιθανότητες να οδηγήσουν σε πυρκαγιά, εντός των ωρών που τις ακολουθούν. Παρόμοια κατάσταση επικρατεί και με τις βροχοπτώσεις του προηγούμενου διαστήματος σε συνάρτηση με την υγρασία. Εντοπίζονται συγκεκριμένες καταστάσεις πρόσφατης βροχόπτωσης, πχ του προηγούμενου μήνα (Σχ. 5.10) ή των 2 προηγούμενων μηνών (Σχ. 5.11) που, σε συνδυασμό με χαμηλή υγρασία, ανεβάζουν σημαντικά τον κίνδυνο πυρκαγιάς. Κάποιες μεταβλητές είναι εξαρτημένες από άλλους παράγοντες και είναι πιο χρήσιμο να εξεταστούν σε συνδυασμό με αυτούς. Η ηλιακή ακτινοβολία μεταβάλλεται μέσα στην ημέρα· τα υψηλότερα επίπεδα ακτινοβολίας φαίνεται να έχουν μία θετική συσχέτιση με την πρόκληση πυρκαγιάς, αλλά αυτό δεν φαίνεται να ισχύει για τις μέρες με εξαιρετικά έντονη ηλιοφάνεια.

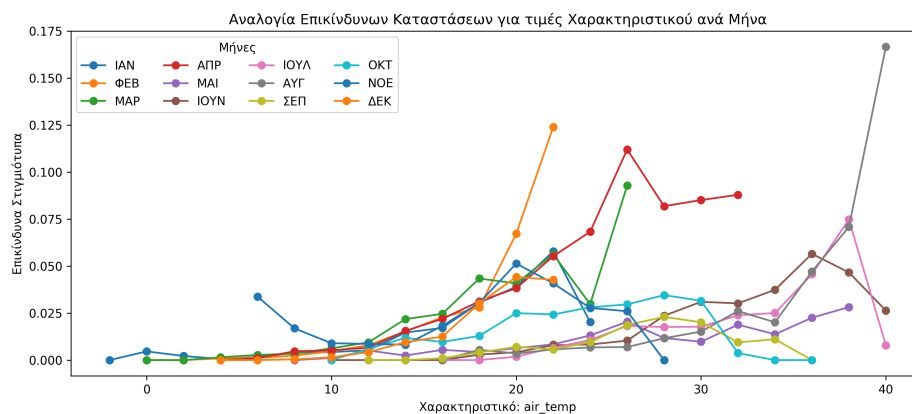


Σχήμα 5.5: Κατανομή Τιμών Θερμοκρασίας (Περιοχή Μελέτης 2020 - 2025)

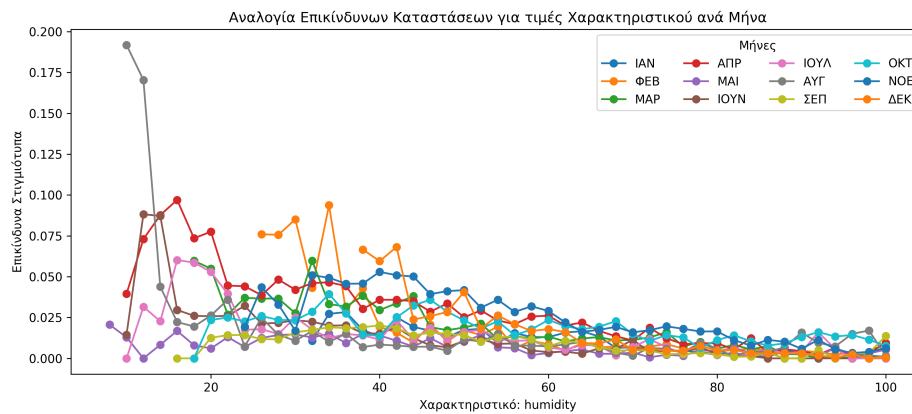
νεια (Σχ. 5.12). Η ατμοσφαιρική πίεση μειώνεται με το υψόμετρο, άρα είναι φυσικό να μελετηθεί σε συνδυασμό με αυτό· παρουσιάζονται αποκλίσεις για συγκεκριμένα επίπεδα πίεσης (Σχ. 5.13), αν και η σχέση της ατμοσφαιρικής πίεσης με τα επικίνδυνα διαστήματα δεν είναι τόσο ισχυρή.



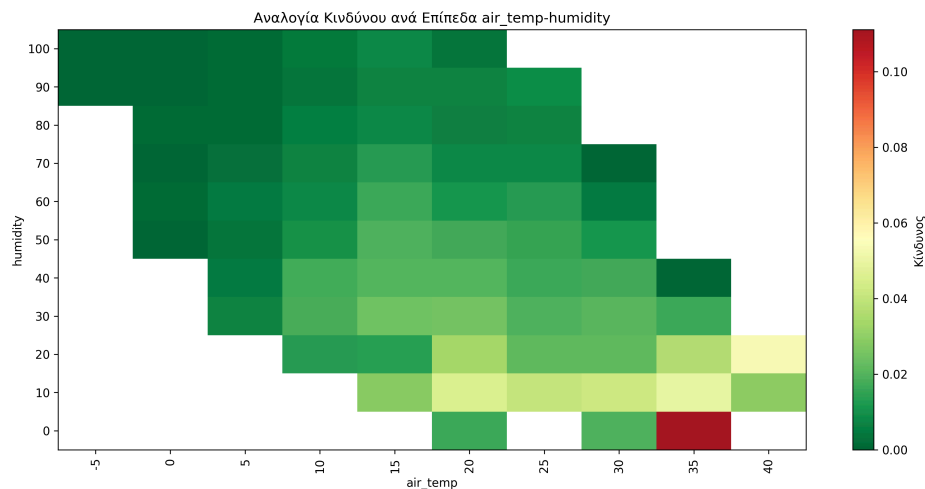
Σχήμα 5.6: Κατανομή Τιμών Υγρασίας (Περιοχή Μελέτης 2020 - 2025)



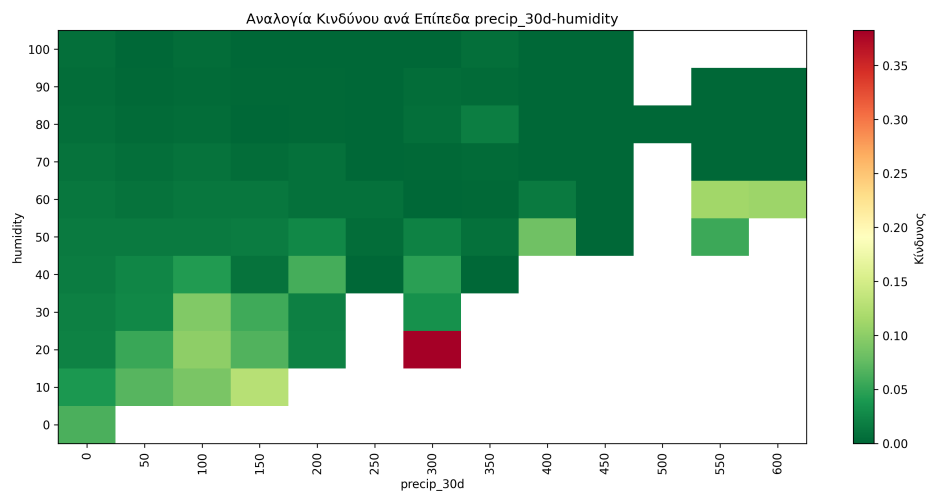
Σχήμα 5.7: Αναλογία Επικίνδυνων Στιγμιωτύπων ανά Επίπεδο Θερμοκρασίας και Μήνα (Περιοχή Μελέτης 2020 - 2025)



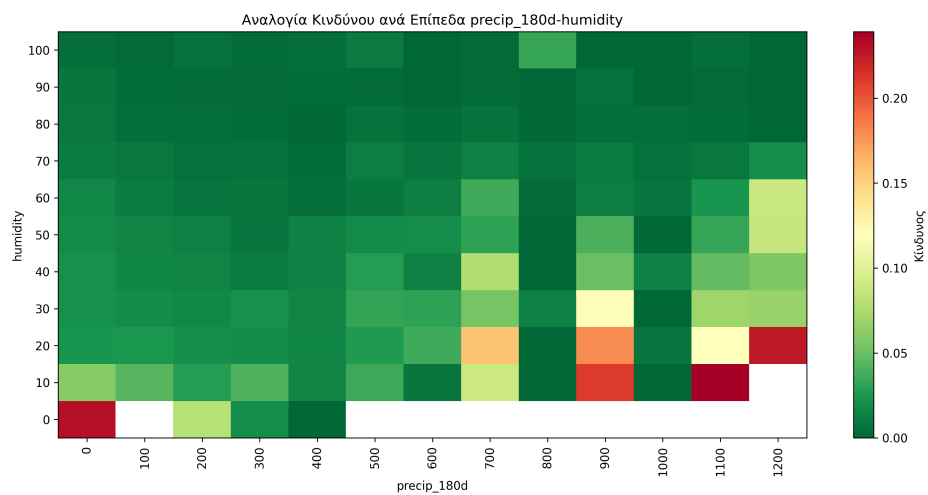
Σχήμα 5.8: Αναλογία Επικινδυνών Στιγμιοτύπων ανά Επίπεδο Υγρασίας και Μήνα (Περιοχή Μελέτης 2020 - 2025)



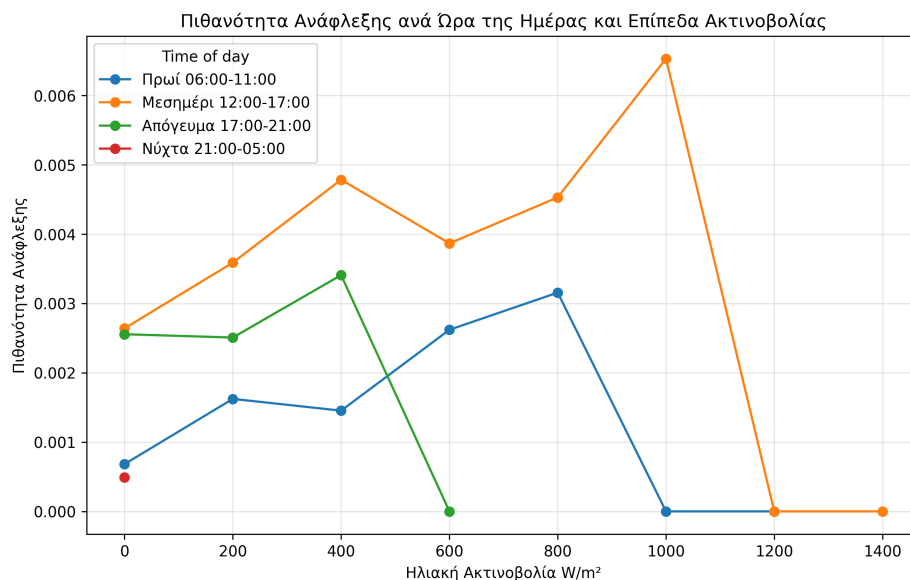
Σχήμα 5.9: Αναλογία Επικινδυνών Στιγμιοτύπων ανά Συνδυασμό Επίπεδου Υγρασίας και Θερμοκρασίας (Περιοχή Μελέτης 2020 - 2025)



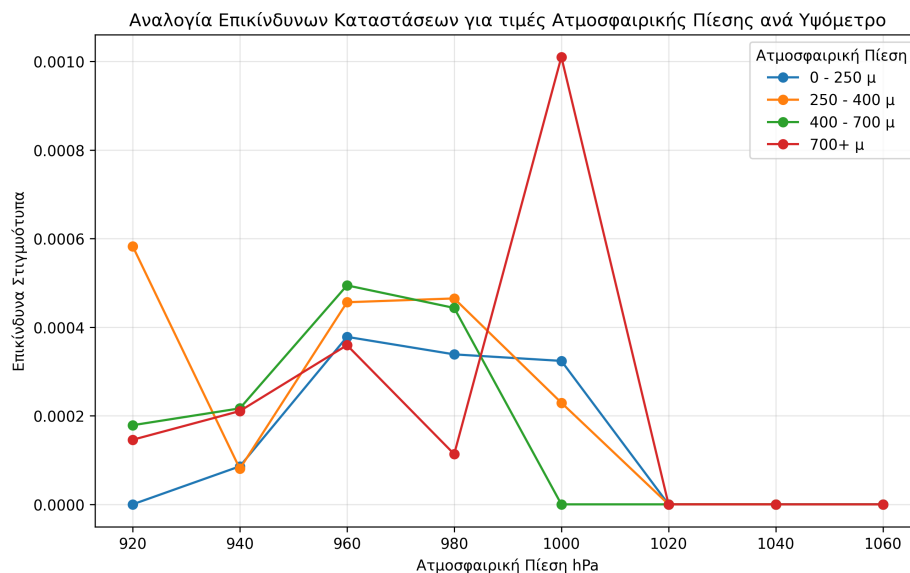
Σχήμα 5.10: Αναλογία Επικίνδυνων Στιγμιотύπων ανά Συνδυασμό Επιπέδου Υγρασίας και Βροχόπτωσης Μήνα (Περιοχή Μελέτης 2020 - 2025)



Σχήμα 5.11: Αναλογία Επικίνδυνων Στιγμιотύπων ανά Συνδυασμό Επιπέδου Υγρασίας και Βροχόπτωσης Εξαμήνου (Περιοχή Μελέτης 2020 - 2025)



Σχήμα 5.12: Αναλογία Επικίνδυνων Στιγμιότυπων ανά Επίπεδο Ηλιακής Ακτινοβολίας και Περίο-
δου της Ημέρας (Περιοχή Μελέτης 2020 - 2025)



Σχήμα 5.13: Αναλογία Επικίνδυνων Στιγμιότυπων ανά Επίπεδο Ατμοσφαιρικής Πίεσης και Υψό-
μετρο (Περιοχή Μελέτης 2020 - 2025)

5.5 Αποτελέσματα Μοντέλων

5.5.1 Αποτελέσματα Συγκριτικών Δοκιμών Μοντέλων

Τα αποτελέσματα των συγκριτικών δοκιμών αφορούν μοντέλα που εκπαιδεύτηκαν στα δεδομένα των ετών 2020-2023, σύνολο 487.105 στιγμιότυπα, με στόχο τον εντοπισμό κινδύνου πυρκαγιάς σε ορίζοντα 3 ωρών. Οι μετρικές υπολογίστηκαν βάσει των προβλέψεων για το σύνολο δοκιμών που αποτελείται από όλα τα στιγμιότυπα του έτους 2024, συνολικά 121.777, με επικράτηση θετικής κλάσης 1, 39%. Το F2Score υπολογίστηκε στο καλύτερο κατώφλι. Η σύγκριση των μοντέλων στο σύνολο εκπαίδευσης παρατίθεται στον πίνακα 5.4. Για τους αλγόριθμους με έμφυτους τυχαίους παράγοντες, η διαδικασία επαναλήφθηκε για 20 διαφορετικές τυχαίες καταστάσεις και καταγράφηκε η μέση τιμή της μετρικής, με τους συντελεστές διακύμανσης των μετρικών να καταγράφονται στον πίνακα 5.5. Οι αλγόριθμοι με τεχνικές εξισορρόπησης κλάσεων έχουν αισθητά μεγαλύτερη διακύμανση στα αποτελέσματα.

Πίνακας 5.4: Σύγκριση Μετρικών Απόδοσης Εξεταζόμενων Μοντέλων

Μοντέλο	log loss	PR-AUC	F2Score
Απλός Κανόνας (Υγρ < 30% & Θερμ. > 30 °C)	N/A*	0,0136	0,058
LR με Κανονικοποίηση	0,599	0,036	0,135
LR με Μετατροπή Δύναμης	0,616	0,032	0,130
LR με PF	0,588	0,052	0,150
RF	0,063	0,044	0,144
RF με Ισορροπία Κλάσεων	0,077	0,036	0,125
RF με Υποδειγμάτοληψία	0,077	0,037	0,131
XGB	0,063	0,048	0,152
XGB με Ισορροπία Κλάσεων	0,063	0,051	0,154

*Ο απλός κανόνας παράγει δυαδικές προβλέψεις, άρα δεν υφίσταται log loss για αυτόν.

Πίνακας 5.5: Συντελεστές Διακύμανσης Μετρικών μη-Ντετερμινιστικών Αλγορίθμων

Μοντέλο	log loss	PR-AUC	F2Score
RF	0,068	2,744	1,218
RF με Ισορροπία Κλάσεων	1,816	1,120	2,363
RF με Υποδειγμάτοληψία	1,224	0,967	3,159
XGB	0,059	1,382	0,853
XGB με Ισορροπία Κλάσεων	10,670	13,129	16,688

Ο LR με την χρήση PF παρέχει την μεγαλύτερη σταθερότητα στα αποτελέσματα των εκτιμήσεων, με κριτήριο την PR-AUC. Εντούτοις, όπως και οι άλλες εκδοχές του LR, παρουσιάζει μεγάλη λογαριθμική απώλεια και, επομένως, χάσμα μεταξύ των πιθανοτήτων και του στόχου. Οι αλγόριθμοι XGB φαίνεται να παρέχουν συνολικά την καλύτερη απόδοση στο συγκεκριμένο πρόβλημα. Ο XGB με ρυθμίσεις για ισορροπία κλάσεων και τις υπερπαραμέτρους που δίνονται στον πίνακα 5.6 δίνει ελαφρώς καλύτερες μετρικές σε σχέση με τον σκέτο XGB.

Πίνακας 5.6: Βελτιστοποιημένες Παράμετροι Επιλεγμένου Αλγορίθμου XGB, Βιβλιοθήκη Python XGBoost

Παράμετρος	Τιμή
gamma	0.00626
max_depth	3
reg_alpha	0.0566
subsample	0.771
reg_lambda	0.0657
n_estimators	344
learning_rate	0.0146
max_delta_step	1
colsample_bytree	0.723
min_child_weight	11
scale_pos_weight	1

Ο πίνακας σύγχυσης για το καλύτερο κατώφλι απόφασης του XGB με ισορροπία κλάσεων είναι ο 5.7. Υπολογίζοντας την ανύψωση του PR-AUC για την επικράτηση της θετικής κλάσης σε αυτό το μοντέλο έχουμε $Lift_{PR-AUC} = 3,66$, για σύγκριση, η ανύψωση του απλού κανόνα είναι 0,98, λίγο χειρότερο από τυχαία επιλογή.

Πίνακας 5.7: Πίνακας Σύγχυσης για κατηγοριοποίηση επικίνδυνων συνθηκών με XGB

Αλήθεια	Πρόβλεψη	
	Κίνδυνος	Όχι Κίνδυνος
Κίνδυνος	849	754
Όχι Κίνδυνος	20.103	100.071

5.5.2 Στατιστική Σημασία Χαρακτηριστικών

Κατά την τεχνική μεταθέσεων (Permutations), εκτελούνται δοκιμές αφαίρεσης χαρακτηριστικών πάνω στα δεδομένα ελέγχου. Με αυτόν τον τρόπο, προκύπτει μία εκτίμηση για την την επίδραση του κάθε χαρακτηριστικού στα αποτελέσματα ενός εκπαιδευμένου μοντέλου.

Επιπλέον, οι περισσότεροι αλγόριθμοι ML έχουν έναν τρόπο να εκφράζουν την βαρύτητα των χαρακτηριστικών πρόβλεψης. Οι συντελεστές σημαντικότητας του μοντέλου XGB, με κριτήριο σημαντικότητας το κέρδος πληροφορίας, παραθέτονται στον πίνακα 5.8. Αξίζει να σημειωθεί πως τα χαρακτηριστικά που επικαλύπτονται από άλλα έχουν χαμηλό κέρδος πληροφορίας. Αυτός είναι ο λόγος που ο νετός ημιώρου έχει τόσο χαμηλό δείκτη· όλη η πληροφορία που μπορεί να προσφέρει, καλύπτεται ήδη από τον νετό ημέρας και την υγρασία.

Πίνακας 5.8: Κέρδος Πληροφορίας Χαρακτηριστικών Αλγόριθμου XGB

Χαρακτηριστικό	Επίδραση Μεταθέσεων	Κέρδος Πληροφορίας
Ατμοσφαιρική Πίεση	0,00205	0,047
Φάση Ημέρας συνω	0,00171	0,033
Μέσο Υψόμετρο	0,00151	0,062
Ηλιακή Ακτινοβολία	0,00110	0,296
Ατμοσφαιρική Υγρασία	0,00109	0,053
Θερμοκρασία Αέρα	0,00052	0,050
Υετός Ημέρας	0,00045	0,044
Φάση Ημέρας ημω	0,00044	0,133
Φάση Έτους ημω	0,00043	0,041
Φάση Έτους συνω	0,00021	0,040
Υετός 3μήνου	0,00003	0,024
Φάση Εβδομάδας συνω	0,00003	0,013
Υετός Μήνα	0,00001	0,040
Υετός Εβδομάδας	0,00001	0,016
Υετος Ημιώρου	0,00000	0,001
Φάση Εβδομάδας ημω	-0,00001	0,016
Υετός 6μήνου	-0,00011	0,051

5.5.3 Αποτελέσματα Δοκιμών Παραμετροποίησης

Επιλεγμένες αλλαγές δοκιμάστηκαν στον XGB, τα αποτελέσματα παρατίθενται στον πίνακα 5.9. Η επικράτηση της θετικής κλάσης διαφέρει σε κάποιες περιπτώσεις.

Πίνακας 5.9: Δοκιμές XGB με Διαταραχές

Αλλαγή	log loss	PR-AUC	F2Score	Ανύψωση
Καμία Αλλαγή	0,062	0,051	0,154	3,66
Κανονικοποίηση	0,063	0,053	0,151	3,81
Μετατροπή Δύναμης	0,062	0,054	0,151	3,88
RF	0,062	0,054	0,158	3,88
Χωρίς Ορίζοντα Στόχου	0,012	0,013	0,062	7,22
Ορίζοντας 90λ	0,040	0,032	0,103	4,21
Ορίζοντας 360λ	0,088	0,073	0,207	3,70
Χωρίς Φάση Ημέρας	0,063	0,049	0,152	3,51
Χωρίς Φάση Εβδομάδος	0,063	0,053	0,154	3,81
Χωρίς Φάση Έτους	0,063	0,053	0,153	3,81
Χωρίς Τιμές Φάσης	0,063	0,051	0,149	3,66
Χωρίς Υετό Μηνών	0,062	0,051	0,156	3,66
Χωρίς Αθροιστικό. Υετό	0,063	0,048	0,150	3,43
Διαχωρισμός 10χμ.	0,009	0,005	0,026	3,84
Διαχωρισμός 5χμ.	0,003	0,001	0,007	2,7

Επιπροσθέτως, το μοντέλο δοκιμάστηκε σε συγκεκριμένες περιόδους του έτους· για το σύνολο

εκπαίδευσης και δοκιμής κρατήθηκαν μόνο τα στιγμιότυπα των συγκεκριμένων εποχών, τα αποτελέσματα βρίσκονται στον πίνακα 5.9.

Πίνακας 5.10: Δοκιμές XGB ανά Περίοδο του Έτους

Περίοδος	log loss	PR-AUC	F2Score	Ανύψωση
Μόνο Θερινοί Μήνες	0,082	0,036	0,127	2,11
Μόνο Άνοιξη-Φθινόπωρο	0,064	0,063	0,155	4,50
Μόνο Χειμώνας	0,044	0,030	0,108	3,69
Θερ. Χωρίς Φάσεις	0,082	0,036	0,126	2,15
Α-Φ Χωρίς Φάσεις	0,065	0,052	0,155	3,77
Χειμ. Χωρίς Φάσεις & Υετό Μηνών	0,043	0,063	0,128	7,68

5.5.4 Ανάλυση Προβλέψεων

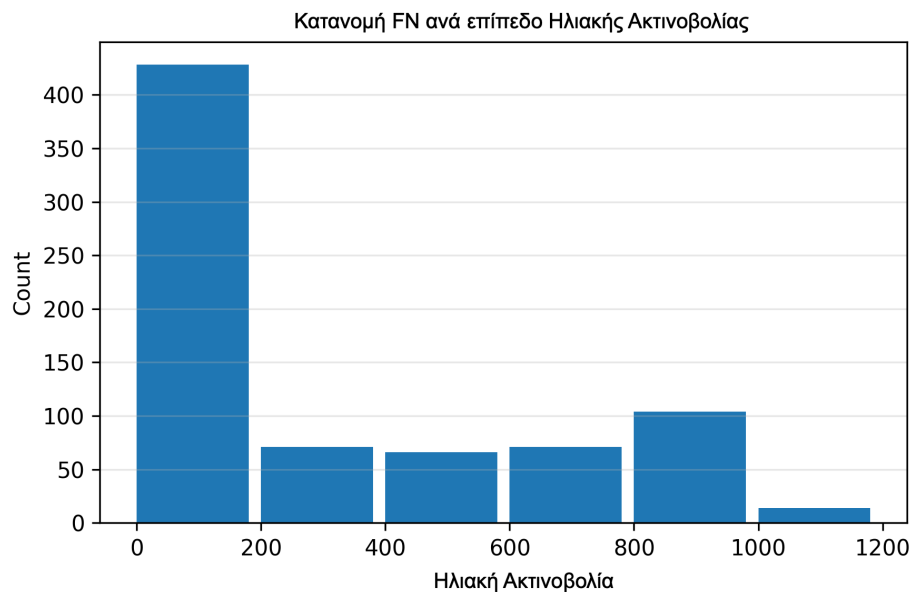
Το επιλεγμένο μοντέλο (XGB με ισορροπία κλάσεων) δεν πέτυχε αρκετά υψηλή ακρίβεια, επομένως, τα κατώφλια απόφασης για τα επίπεδα κινδύνου ορίστηκαν με FScore. Τα αποτελέσματα για εκτίμηση κινδύνου στο σύνολο ελέγχου, φαίνονται στον πίνακα 5.11. Τα Θετικά είναι το σύνολο των προβλέψεων που έδιναν κίνδυνο για προσεχή πυρκαγιά, τα TP είναι οι πυρκαγιές που ανιχνεύτηκαν επιτυχώς, και τα FN είναι αυτές που δεν προβλέφθηκαν. Η ανάκληση δηλώνει το ποσοστό ανίχνευσης και η ακρίβεια υποδεικνύει τους λάθος συναγεμμούς.

Πίνακας 5.11: Αποτελέσματα Επιπέδων Κινδύνου με XGB

Επίπεδο Κινδύνου	Τιμή b	Θετικά	TP	FN	Ανάκληση	Ακρίβεια
Ακραίο	0,5	9.784	456	1.144	0,285	0,046
Πολύ Υψηλό	1	14.550	613	990	0,382	0,042
Υψηλό	2	20.952	849	754	0,530	0,040
Μέτριο	4	33.360	1.136	467	0,708	0,034
Χαμηλό	8	52.304	1.359	244	0,850	0,025

Εξετάζοντας τα εσφαλμένα αρνητικά σε σχέση με τα πιο σημαντικά χαρακτηριστικά μπορούν να γίνουν κάποιες παρατηρήσεις:

1. 50 - 95% των FN, από το ακραίο μέχρι και το χαμηλό επίπεδο κινδύνου αντιστοιχούν σε στιγμιότυπα με πολύ χαμηλή ηλιοφάνεια $< 200W/m^2$ (Σχ. 5.14).
2. 50 - 80% των FN, από το ακραίο μέχρι και το χαμηλό επίπεδο κινδύνου αντιστοιχούν σε στιγμιότυπα με μεσαία και χαμηλή θερμοκρασία $< 20C$ (Σχ. 5.15).
3. 85% των FN, εκδηλώθηκαν σε υγρασία $> 40\%$ (Σχ. 5.16).
4. Ο Απρίλιος και ο Οκτώβριος είναι και οι μήνες με τα λιγότερα FN, συγκριτικά με τον αριθμό των περιστατικών που σημειώνουν. Ο Ιούλιος και οι χειμερινοί μήνες, έχουν τα περισσότερα FN συγκριτικά με τον αριθμό των περιστατικών που σημειώνουν.

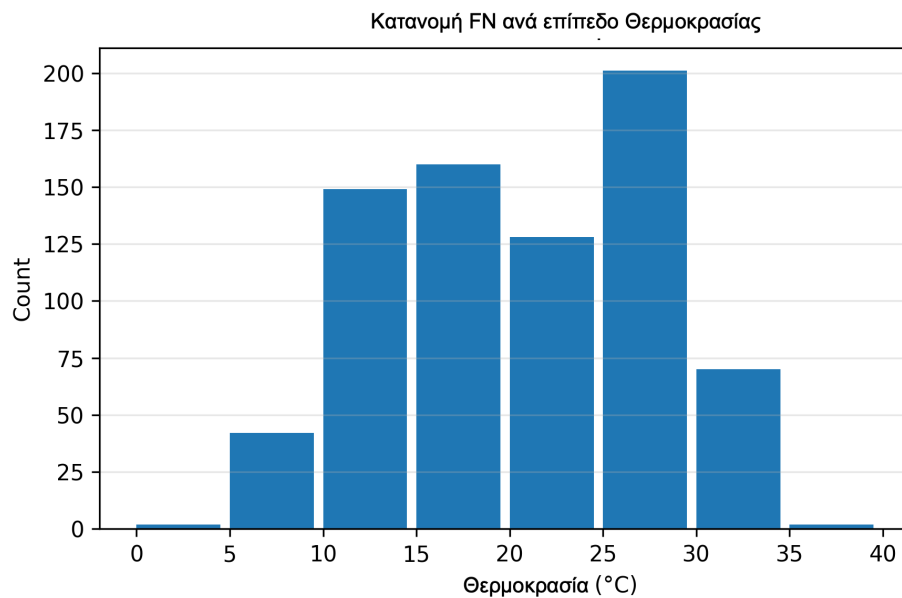


Σχήμα 5.14: Κατανομή Εσφαλμένων Αρνητικών Προβλέψεων ανά επίπεδο Ηλιακής Ακτινοβολίας

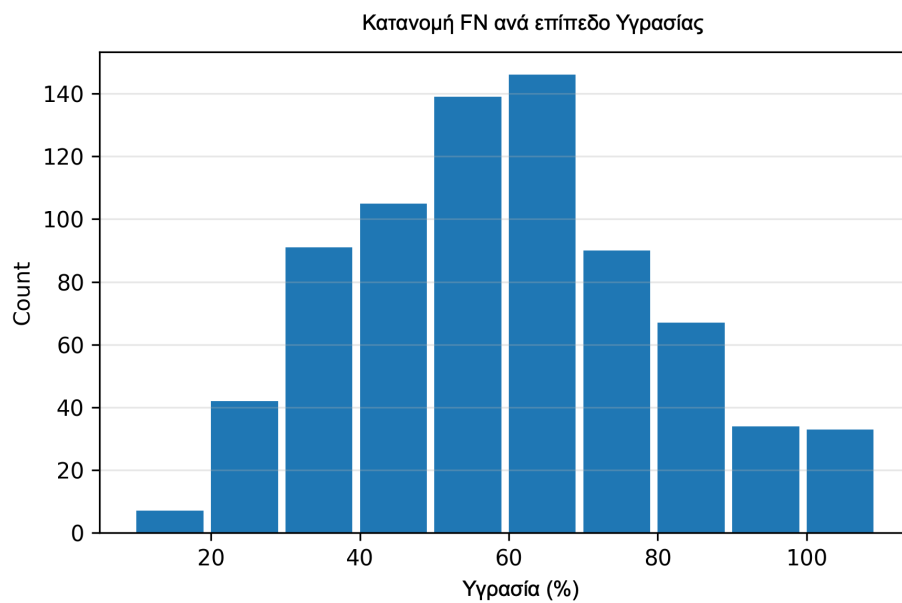
Τα επίπεδα κινδύνου συμπεριφέρονται διαφορετικά εάν εξετάσουμε μόνο τους θερινούς μήνες, που έχουν 30.470 στιγμιότυπα. Όπως φαίνεται στον πίνακα 5.12, η ακρίβεια είναι πολύ πιο χαμηλή, με αποτέλεσμα να πέφτει η μετρική PR-AUC, αλλά η ανάκληση είναι σημαντικά βελτιωμένη. Από την άλλη, οι δοκιμές με χρήση αποκλειστικά στιγμιότυπων άνοιξης-φθινοπώρου είχε περίπου 15% καλύτερη απόδοση, σε σχέση με τα μοντέλα που εκπαιδεύτηκαν και αξιολογήθηκαν στα στιγμιότυπα όλου του έτους.

Πίνακας 5.12: Αποτελέσματα Επιπέδων Κινδύνου με XGB (Θερινοί Μήνες)

Επίπεδο Κινδύνου	Τιμή b	Θετικά	TP	FN	Ανάκληση	Ακρίβεια
Ακραίο	0,5	14.275	413	108	0,801	0,028
Μέτριο	4	16.446	442	73	0,858	0,026
Χαμηλό	8	29.603	515	0	1	0,017



Σχήμα 5.15: Κατανομή Εσφαλμένων Αρνητικών Προβλέψεων ανά επίπεδο Θερμοκρασίας



Σχήμα 5.16: Κατανομή Εσφαλμένων Αρνητικών Προβλέψεων ανά επίπεδο Υγρασίας

Κεφάλαιο 6

Συζήτηση

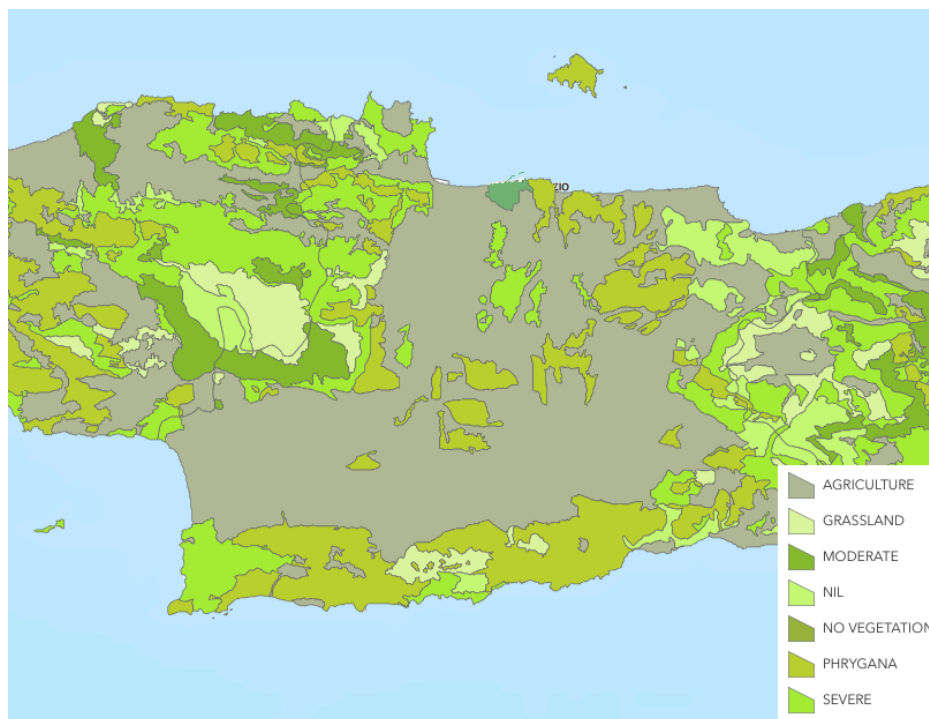
6.1 Ερμηνεία Δεδομένων

6.1.1 Ερμηνεία Χωρικής και Χρονικής Κατανομής

Η εκδήλωση των πυρκαγιών φαίνεται να ακολουθεί στενά την διεπαφή των καλλιεργήσιμων εκτάσεων με την άγρια βλάστηση (Σχ.6.1)· σε όλες τις ζώνες με πολύ υψηλό δείκτη εκδήλωσης πυρκαγιών, λαμβάνει χώρα καλλιέργεια ελαιόδεντρων και αμπελώνων σε κοντινή απόσταση από φρύγανα. Αυτό το χαρακτηριστικό, όμως, δεν συνεπάγεται από μόνο του υψηλή συχνότητα πυρκαγιών· η λεκάνη του Αποσελέμη φιλοξενεί πυκνούς ελαιώνες και μεγάλη συγκέντρωση βλάστησης, αλλά απολαμβάνει συγκριτικά υψηλότερες βροχοπτώσεις και υγρασία. Ο σταθμός ΠΟΤΑΜΙΕΣ της περιοχής, κατέγραψε συνολικά κοντά στα 3.300 χιλιοστά υετού την τελευταία 5ετία, ενώ οι σταθμοί της Νότιας Κρήτης, όπως αυτός του Πύργου, κατέγραψαν κατά μέσο όρο περίπου 1800 χιλιοστά.

Οι ώρες που εκδηλώνονται οι πυρκαγιές εντός της ημέρας έχουν απλή εξήγηση: Συμπίπτουν στενά με τις ώρες δραστηριοτήτων των ανθρώπων, οι οποίες συνήθως γίνονται με το φως του ηλίου, είτε πρόκειται για ψυχαγωγία, είτε για εργασίες στην ύπαιθρο. Η κατανομή των συμβάντων μέσα στο έτος, από την άλλη, αρχικά φαίνεται παράξενη, ειδικά ο αυξημένος αριθμός περιστατικών μικρής κλίμακας τον Νοέμβριο και τον Απρίλιο. Η αιτία αυτής της αλλαγής είναι ξεκάθαρη, εάν ληφθεί υπ' όψιν το διάστημα της αντιπυρικής περιόδου, η οποία στην Ελλάδα ορίζεται από τη 1 Μαΐου μέχρι και τις 31 Οκτώβρη [65][8]· οι κανονισμοί για αυτό αυστηροποιήθηκαν το 2021[66]. Κατά συνέπεια, όλες οι εργασίες που ενδέχεται να προκαλέσουν ανάφλεξη και πυρκαγιά παύουν και, επομένως, οι αγρότες και οι μελισσοκόμοι σπεύδουν να τις διεκπεραιώσουν αμέσως πριν ή μετά την αντιπυρική περίοδο. Άλλοι παράγοντες μπορεί να είναι ο καθαρισμός των οικοπέδων για το καλοκαίρι και η Μεγάλη Εβδομάδα.

Η χωροχρονική κατανομή των πυρκαγιών συνθέτει μία λεπτομερή εικόνα του φαινομένου στην περιοχή μελέτης: Η έντονη αγροτική δραστηριότητα στην ενδοχώρα συμπεριλαμβάνει εποχικές εργασίες, όπως το κάψιμο κλαδιών και η χρήση μελισσοκομικού καπνού, οι οποίες προκαλούν ένα μεγάλο μέρος των αναφλέξεων. Η πιθανότητα να τεθούν εκτός ελέγχου οι φωτιές αυξάνεται όταν αυτές οι δραστηριότητες πραγματοποιούνται σε κοντινή απόσταση από τα εύφλεκτα φρύγανα της νησιωτικής υπαίθρου. Το νομοθετικό πλαίσιο που θεσπίστηκε για την πρόληψη των πυρκαγιών, απο αυτήν την άποψη, είναι επιτυχημένο: Ένας μεγάλος όγκος των περιστατικών μετατοπίστηκε σε περιόδους όπου τα οικοσυστήματα είναι λιγότερο ευπαθή στην διάδοση της φωτιάς, με αποτέλεσμα, να προκύπτει ένας μεγάλος αριθμός μικρών σε έκταση περιστατικών σε αυτές τις περιόδους.



Σχήμα 6.1: Χρήσεις Γής, Βλάστηση **Υπόβαθρο:**© © 2017-2026 Αποκεντρωμένη Διοίκηση Κρήτης, Διεύθυνση Πληροφορικής και Επικοινωνιών.

Γενικότερα, η περιοχή μελέτης δεν κατέγραψε μεγάλες σε έκταση πυρκαγιές τα τελευταία χρόνια, γεγονός που θα μπορούσε να αποδοθεί στην διατήρηση των παραδοσιακών κτηνοτροφικών δραστηριοτήτων και την συνεχή ενασχόληση με την ύπαιθρο. Αυτά τα χαρακτηριστικά, έχουν προταθεί σε προηγούμενη έρευνα του Elhag κα. στα φρυγανικά οικοσυστήματα της Κρήτης ως ρυθμιστές της βλάστησης και του κινδύνου πυρκαγιών [18].

6.1.2 Ερμηνεία Καιρικών Συνθηκών

Οι καιρικές συνθήκες φαίνεται πως κατευθύνουν όχι απλά την συχνότητα, αλλά και την ένταση των πυρκαγιών. Τα αποτελέσματα υποδεικνύουν πως η υγρασία στην ατμόσφαιρα και η πρόσφατη βροχόπτωση είναι καλές μετρικές προσέγγισης του FMC. Η χαμηλή ατμοσφαιρική υγρασία και οι υψηλές θερμοκρασίες αποτελούν σημαντικό παράγοντα κινδύνου και είναι κύριοι δείκτες της περιόδου πυρκαγιών, όπου η συχνότητα και, κυρίως η έκταση των περιστατικών είναι αυξημένες. Αυτοί οι μηχανισμοί είναι πιο έντονοι στην μέση και το τέλος του καλοκαιριού, οπότε και η υγρασία στα οικοσυστήματα βρίσκεται στα πιο χαμηλά επίπεδα, λόγω παρατεταμένης ξηρασίας και εξάτμισης.

Η επίδραση άλλων καιρικών μεταβλητών είναι λιγότερο ξεκάθαρη: Τα αποτελέσματα των αλγόριθμων δείχνουν πως η ηλιακή ακτινοβολία έχει μεγάλη στατιστική σημασία. Δεν είναι, όμως, εύκολα αντιληπτό κατά πόσο αυτό οφείλεται στην επίδραση που έχει στην ευφλεκτότητα, και κατά πόσο είναι προϊόν της επιρροής της ηλιοφάνειας στην ανθρώπινη δραστηριότητα. Θα πρέπει να συνυπολογιστεί ότι η ηλιακή ακτινοβολία είναι χαμηλή κατά την διάρκεια της νύχτας, του χειμώνα και όταν βρέχει. Παρατηρώντας την σχέση της ηλιακής ακτινοβολίας, σε συνδυασμό την περίοδο



Σχήμα 6.2: Photo: Freideriki Dionysopoulou / Wikimedia Commons, Burning olive branches after olive harvesting, CC BY-SA 3.0.

της ημέρας, με την πιθανότητα ανάφλεξης, μπορούμε να πούμε ότι το μεγαλύτερο μέρος του αντίκτυπου αυτής της μεταβλητής είναι στην ανθρώπινη δραστηριότητα. Ο λόγος για αυτό είναι πως κατά τις ώρες με ακραία ηλιοφάνεια (πάντα βρίσκονται μεταξύ Μαΐου και Ιουλίου), δεν υπάρχουν σχεδόν καθόλου αναφλέξεις· είναι πολύ πιθανό τέτοιες μέρες, αν και δεν είναι πολλές σε αριθμό, να έχουν αποτρεπτικό, ή και απαγορευτικό χαρακτήρα για υπαίθριες δραστηριότητες.

Η ατμοσφαιρική πίεση φαίνεται, βάσει της κατανομής των συμβάντων, αλλά και της στατιστικής σημασίας της για τους αλγόριθμους ML, να έχει επίδραση στην εκδήλωση πυρκαγιών. Γενικά, η χαμηλή ατμοσφαιρική πίεση μπορεί να αποτελέσει ένδειξη κεραυνών και πιθανής ανάφλεξης. Αυτό βέβαια έχει περιορισμένη επίδραση, εφόσον η συντριπτική πλειοψηφία των πυρκαγιών έχει ανθρωπογενή αίτια. Μία άλλη εξήγηση θα ήταν πως η διαφορά πίεσης προκαλεί ανέμους, άρα, εάν η ένδειξη του βαρόμετρου παρεκκλίνει από το αναμενόμενο, υπάρχουν και οι ευνοϊκές συνθήκες για να τεθεί μία φωτιά εκτός ελέγχου. Σε αυτήν την ερμηνεία, η ατμοσφαιρική πίεση δρα ως υποκατάστατο της ταχύτητας ανέμου.

Η χρήση του αθροιστικού υετού έχει ανάμεικτα αποτελέσματα. Η εκδήλωση υετού μέσα στην μέρα είναι σημαντική ένδειξη για χαμηλό κίνδυνο, αλλά έχει και σημαντική επικάλυψη με την υγρασία του αέρα, λόγω της εξάτμισης των βρόχινων υδάτων. Η ωφέλιμη πληροφορία που παρέχουν οι αθροιστικοί υετοί σε πιο μακροπρόθεσμα διαστήματα είναι δύσκολο να εκτιμηθούν. Ο υετός των τελευταίων 6 μηνών και του τελευταίου μήνα θεωρούνται στατιστικά σημαντικοί από τους αλγόριθμους, αλλά ο λόγος δεν είναι αυτονόητος. Είναι πιθανόν η υπόθεση που έγινε ως προς την ιδιότητα αυτών των τιμών να προβλέπουν τον όγκο και την ευφλεκτότητα της βλάστησης να είναι ορθή. Μπορεί όμως οι αλγόριθμοι απλά να αποτυπώνουν τα μοτίβα του κάθε σταθμού και να προβλέπουν για αυτά. Το γεγονός, ότι ο υετός εξαμήνου έχει υψηλό κέρδος πληροφορίας αλλά αρνητική επίδραση μετατοπίσεων, είναι ένδειξη πως το μοντέλο υπερπροσαρμόζεται στο συγκεκριμένο χαρακτηριστικό.

6.2 Ανασκόπηση Πειραμάτων Μοντελοποίησης

6.2.1 Καταλληλότητα Αλγόριθμων

Ο αλγόριθμος LR δεν είναι κατάλληλος για μη γραμμικά προβλήματα, εξ' ου και η κακή απόδοσή του στα πειράματα. Ο λόγος που η χρήση RF ανεβάζει τόσο πολύ τις μετρικές, είναι η ανάλυση σε βαθύτερο επίπεδο που επιτυγχάνεται σε τα ζεύγη χαρακτηριστικών πρόβλεψης. Έτσι, είναι δυνατόν να αναπληρωθεί η γραμμικότητα του LR, σε κάποιο βαθμό.

Οι αλγόριθμοι που δημιουργούν δέντρα αποφάσεων χειρίζονται πολλά "μονοπάτια απόφασης" που μπορούν να αποτυπώσουν μη γραμμικά μοτίβα [52]. Ανάμεσα στον RF και τον XGB, ο XGB έχει κάποια χαρακτηριστικά που ενδέχεται να τον καθιστούν πιο κατάλληλο για αυτό το πρόβλημα: Κατά την εκπαίδευση, τα δέντρα αποφάσεων δημιουργούνται με την σειρά και το κάθε δέντρο χρησιμοποιεί τα προηγούμενα για να προσαρμόσει τα επόμενα [51]. Επιπλέον, η μέθοδος εκπαίδευσης του παρέχει πολύ περισσότερες επιλογές παραμετροποίησης. Ο XGB, γενικότερα, είναι πιο προηγμένος σε σχέση με τον RF, γεγονός που τον κάνει πιο αποδοτικό, αλλά και πιο επιρρεπή σε overfitting λόγω λάθος ρύθμισης [52].

Από τους δοκιμασμένους αλγόριθμους, πιο κατάλληλος κρίνεται ο XGB, αλλά περισσότερα πειράματα σε διαφορετικές καταστάσεις είναι απαραίτητα.

6.2.2 Επίδραση Παραμέτρων

Γενικότερα, καμία μεμονωμένη αλλαγή δεν επέφερε μεγάλη βελτίωση στην απόδοση. Το μοντέλο χωρίς ορίζοντα πρόβλεψης φαίνεται να προβλέπει πιο στοχευμένα, έχοντας διπλάσια ανύψωση σε σχέση με το επιλεγμένο μοντέλο. Εντούτοις, οι χαμηλότερες μετρικές στην ακρίβεια και την ανάκληση, καθώς και η επιχειρησιακή δυσκολία που παρουσιάζει σε σύγκριση με μία προσέγγιση ορίζοντα κινδύνου, αφήνουν ερωτήματα ως προς το εάν είναι πιο πρακτικό. Ο τρόπος με τον οποίο θα ενσωματωθεί η συγκεντρωτική επίδραση των συνθηκών των προηγούμενων ωρών παραμένει ανοιχτός σε διερεύνηση.

Επίσης, το υποσύστημα εκπαίδευσης δεν φαίνεται να είναι σε θέση ακόμη να χειριστεί πλέγματα μεγαλύτερης ανάλυσης με επιτυχία, εν μέρη, λόγω της μεγάλης ανισορροπίας κλάσεων. Η χρήση PF φαίνεται να βελτιώνει ελαφρώς την απόδοση - το κόστος βέβαια είναι υψηλό, καθώς ο αριθμός των χαρακτηριστικών τετραγωνίζεται. Ο XGB έχει καλή δυνατότητα προσαρμογής σε μη γραμμικές σχέσεις, όμως, το ευνοϊκό αποτέλεσμα της δοκιμής με PF, υποδεικνύει πως υπάρχουν κάποιες σχέσεις που είναι πολύ σημαντικό να βρίσκονται μαζί, όπως υγρασία x θερμοκρασία.

6.2.3 Ερμηνεία Προβλέψεων Μοντέλων

Από τα αποτελέσματα της ενότητας 5.5.4, μπορούν να εξαχθούν αρκετά χρήσιμα πορίσματα για την χρησιμότητα του μοντέλου και, τελικά, την δυνητική επιτυχία του συστήματος. Εκ πρώτης όψεως, το σύστημα αποτυγχάνει στο κριτήριο αποτελεσματικότητας: το μοντέλο αδυνατεί να πετύχει υψηλά ποσοστά ανάκλησης χωρίς να αναθέσει λανθασμένα ένδειξη πυρκαγιάς σε ένα υπερβολικά μεγάλο τμήμα των υπόλοιπων στιγμιотύπων. Έτσι, προκαλούνται πολλά FP σήματα και, επομένως, λάθος συναγερμοί. Αυτό μεν ισχύει, αξίζει, όμως, να εξεταστούν με λεπτομέρεια οι προβλέψεις, ώστε να εξηγηθούν τα αίτια.

Η κατανομή των FN δεν είναι ομοιόμορφη, αλλά, έχει τάσεις προς συγκεκριμένες τιμές για συ-

γκεκριμένα χαρακτηριστικά. Όπως αναφέρθηκε, η ηλιοφάνεια έχει πολυδιάστατο ρόλο. Το γεγονός ότι οι φωτιές με χαμηλή ηλιοφάνεια είναι δύσκολο να εκτιμηθούν, κατά πάσα πιθανότητα, οφείλεται στο ότι εκδηλώνονται νυχτερινές ώρες, ή κατά την διάρκεια έντονης συννεφιάς. Οι περιπτώσεις των νυχτερινών πυρκαγιών είναι ιδιάζουσες, καθώς, απαιτούνται ενέργειες πέραν των συνηθισμένων και αναμενόμενων καθημερινών εργασιών για να προκληθεί ανάφλεξη ή να προκύψει κάποια βλάβη ή να πέσει κάποιος κεραυνός ή κάποιο άλλο σπάνιο σενάριο.

Συγκρίνοντας την κατανομή των θερμοκρασιών και της υγρασίας των FN, φαίνεται πως ο αριθμός λάθος προβλέψεων για ακραίες τιμές θερμοκρασίας και υγρασίας είναι πολύ μικρός. Αυτό σημαίνει πως το μοντέλο ήδη αντιλαμβάνεται ορθά τις επικίνδυνες καιρικές συνθήκες και τις συνδέει επιτυχώς με την πιθανότητα πυρκαγιάς. Το πρόβλημα παρουσιάζεται με τις πυρκαγιές που εκδηλώνονται υπό συνθήκες που δεν ταιριάζουν στα συνηθισμένα μοτίβα (*outliers*), δηλαδή την πρόκληση πυρκαγιάς ενώ η υγρασία είναι υψηλή. Οι ενδιάμεσες τιμές θερμοκρασίας και υγρασίας φαίνονται επίφοβες ως προς το πότε οδηγούν σε πυρκαγιά. Είναι πιθανό πως, σε αυτό το σημείο, χρειάζεται υψηλότερη ανάλυση μετρήσεων, καθώς για κάποια στιγμιότυπα συντρέχει μεγάλη απόσταση από τον πλησιέστερο σταθμό, με αποτέλεσμα, οι μετρήσεις να μην ισχύουν. Εναλλακτικά, να ενσωματωθούν κάποιοι επιπλέον παράγοντες, που θα αποσαφηνίζουν τον κίνδυνο της πυρκαγιάς, στις περιπτώσεις που η θερμοκρασία, η υγρασία, και τα άλλα χαρακτηριστικά δεν αρκούν. Στην χειρότερη περίπτωση, το υπόλοιπο σύστημα εκδήλωσης πυρκαγιάς είναι εντελώς στοχαστικό, άρα, οι μη-ακραίες μετρήσεις θα παραμένουν ασαφείς.

Εποχικότητα

Τα συμπεράσματα της ενότητας 6.1.1 ενισχύονται εξετάζοντας τα FN, ως προς τους μήνες του έτους. Το μοντέλο φαίνεται να προσπαθεί να ισορροπήσει 2 ή 3 διαφορετικές εκδοχές του προβλήματος με αντιθετικά μοτίβα και έχοντας ανεπαρκή στοιχεία. Ο χωρισμός του προβλήματος σε υποπροβλήματα, βάσει των διακριτών περιόδων πυρκαγιών, φαίνεται να βοηθάει στην καλύτερη αποτύπωση των εποχικών μοτίβων. Το αρχικό μοντέλο προσαρμοζόταν καλύτερα στα μοτίβα αγροτικών εργασιών, μιας και αποτελούν την πλειοψηφία των πυρκαγιών, με αποτέλεσμα να έχει πολύ καλύτερη απόδοση για τον Απρίλιο και τον Οκτώβριο. Τα αποτελέσματα του μοντέλου στους θερινούς μήνες έχουν και αυτά ενδιαφέρον: Η ένδειξη των μισών στιγμιότυπων του συνόλου δεδομένων ως επικίνδυνα φαίνεται υπερβολική. Ωστόσο, οι συνθήκες στην καρδιά της αντιπυρικής περιόδου μπορούν να γίνουν εξαιρετικά επικίνδυνες για μεγάλο διάστημα.

Η γενικότερη επιτυχία των προβλέψεων παρουσιάζει έντονη εποχική διακύμανση, με το μοντέλο εκτίμησης ρίσκου για άνοιξη και φθινόπωρο να είναι περίπου 2 φορές καλύτερο, βάσει ανύψωσης, από το αντίστοιχο για το καλοκαίρι, με μικρότερη λογαριθμική απώλεια και καλύτερη μέση ακρίβεια. Αυτές οι σχέσεις ισχύουν ακόμα και όταν εξετάζεται ξεχωριστά η άνοιξη και το φθινόπωρο, επομένως ο αριθμός των στιγμιότυπων εκπαίδευσης δεν είναι η κύρια αιτία της διαφοράς.

Αναλύοντας τα αποτελέσματα σε σχέση με τα χαρακτηριστικά πρόβλεψης, είναι προφανές πως η επίδραση τους στην ικανότητα εκτίμησης διαφέρει σημαντικά ανάλογα με την περίοδο του έτους. Τα χαρακτηριστικά φάσης αυξάνουν σημαντικά την απόδοση του μοντέλου στις ενδιάμεσες εποχές, αλλά έχουν ελαφρώς αρνητική επιρροή στον χειμώνα και το καλοκαίρι. Αυτό δείχνει ξεκάθαρα πως η προσέγγιση της ανθρώπινης δραστηριότητας καθ' αυτόν τον τρόπο είναι επιτυχημένη κατά τις περιόδους των επικίνδυνων αγροτικών δραστηριοτήτων, αλλά δεν είναι κατάλληλη για την προσομοίωση των καλοκαιρινών και χειμερινών συνθηκών. Η ερμηνεία για αυτό μπορεί να είναι πως τα αίτια ανάφλεξης αυτές τις εποχές είναι ανεξάρτητα από τις ώρες εργασίας, όπως νυχτερινοί εμπρη-

σμοί, τουρισμός, ηλεκτρολογικές βλάβες, πυροτεχνήματα κλπ., και θα προσεγγίζονταν καλύτερα με μετρικές όπως η απόσταση από δρόμους και οι χρήσεις γης.

Αυτές οι διακυμάνσεις στην εποχική απόδοση συμβάλλουν στην αποκρυστάλλωση των εποχικών μοτίβων πυρκαγιών στην περιοχή μελέτης: Οι πυρκαγιές την άνοιξη και το φθινόπωρο είναι μεγάλες σε αριθμό και ακολουθούν στενά τις ώρες εργασίας στην ύπαιθρο, οπότε και διενεργούνται εργασίες με φωτιά. Κατά το θέρος κυριαρχούν οι απρόσμενες αιτίες πυροδότησης, λόγω της ευφλεκτότητας της καύσιμης ύλης που αναφλέγεται χωρίς δυσκολία. Η ισχυρότερη επιρροή των μετεωρολογικών συνθηκών, συμπεριλαμβανομένου του ανέμου, κατά την αντιπυρική περίοδο, είναι ένας πιθανός λόγος που το σύστημα δυσκολεύτηκε περισσότερο. Η ανύψωση του χειμερινού μοντέλου χωρίς χαρακτηριστικά φάσης και υετό μηνών ήταν συγκριτικά η καλύτερη. Φαίνεται πως οι καταστάσεις που οδηγούν σε χειμερινές πυρκαγιές εντοπίζονται εύκολα από τις καιρικές συνθήκες, κυρίως οι ιδιαίτερα ζεστές για την εποχή ημέρες. Ο χειμώνας δεν θα πρέπει να υποτιμάται ως πιθανή περίοδος πυρκαγιών, όπως φάνηκε στις πυρκαγιές της Καλιφόρνια το 2025 [67].

6.2.4 Αποτίμηση Αποτελεσμάτων

Η ακρίβεια, παραμένει ένα σημαντικό αγκάθι για τις προβλέψεις, για αυτό και τα αποτελέσματα απέχουν πολύ από το να είναι πρακτικά αξιοποιήσιμα. Ο μεγάλος αριθμός λάθους συναγερμών που καταγράφει το σύστημα στην προσπάθεια να προβλέψει την εκδήλωση των πυρκαγιών είναι γεγονός. Προτάσσονται δύο εξηγήσεις για την αιτιολόγηση αυτής της παθογένειας: Πρώτον, το σύστημα δεν έχει πλήρη εικόνα των συνθηκών. Αυτό γίνεται εύκολα κατανοητό, συγκρίνοντας απλά τους πυρικούς παράγοντες που περιγράφονται στις ενότητες 2.3 2.5 2.6 με τα χαρακτηριστικά πρόβλεψης του συστήματος που αναγράφονται στον πίνακα 5.1. Δεύτερον, υπάρχουν κάποιες μεταβλητές που είναι πρακτικά απρόσιτες σε κάθε υπολογιστικό σύστημα. Μπορεί, κάλλιστα, ένα σημαντικό ποσοστό των "λάθους συναγερμών" να είχαν οδηγήσει σε πυρκαγιές, εάν οι σκέψεις και οι πράξεις ενός ατόμου ήταν διαφορετικές εκείνη την στιγμή.

Η πιο ρεαλιστική λύση για ένα σύστημα σαν και αυτό είναι να γίνουν απόπειρες προσέγγισης των επικίνδυνων καταστάσεων· επειδή οι προσεγγίσεις δεν είναι αρκετά περιγραφικές, τα αποτελέσματα υστερούν. Η χαμηλή ακρίβεια κρατάει το σύστημα σε ένα καθαρά πιλοτικό στάδιο σε αυτήν την φάση, διότι οι εκτιμήσεις του δεν μπορούν σε καμία περίπτωση να χρησιμοποιηθούν αξιόπιστα σε πρακτικό επίπεδο. Παρόλα αυτά, η τιμή 3,66 για την ανύψωση υποδεικνύει ότι το μοντέλο είναι σημαντικά καλύτερο από έναν τυχαίο κατηγοριοποιητή (τιμή 1) στο να εντοπίζει τον κίνδυνο.

Οι θετικές ενδείξεις επαρκούν για να αποδείξουν πως η μεθοδολογία μοντελοποίησης μπορεί να απομονώσει συγκεκριμένα μοτίβα της εκδήλωσης πυρκαγιών και να παράγει χρήσιμα συμπεράσματα, αλλά και κάποιο αποτέλεσμα. Πλέον, το ερώτημα που διερευνά το έργο, δεν είναι ΕΑΝ μπορεί να γίνει, αλλά ΠΩΣ, ποια δεδομένα και τεχνικές πρέπει να εξευρεθούν, ώστε το αποτέλεσμα να ικανοποιεί.

6.3 Αποτίμηση Υλοποίησης

Η υλοποίηση θεωρείται επιτυχημένη, καθώς καλύπτει με επιτυχία όλα τα λειτουργικά μέρη της διαδικασίας πρόβλεψης.

Το πρωτότυπο σύστημα ακολούθησε στενά της αρχές σχεδίασης που τέθηκαν στην ενότητα 3.9: Η φορητότητά του επέτρεψε τις απρόσκοπτες δοκιμές σε ανεξάρτητη υποδομή, κατά τις οποίες

επέδειξε εξαιρετική σταθερότητα. Οι πόροι αξιοποιήθηκαν επιτυχώς και έγιναν οι κατάλληλες βελτιστοποιήσεις για την κάλυψη των επιχειρησιακών απαιτήσεων απόκρισης. Τέλος, δόθηκαν μοχλοί ρύθμισης για το μεγαλύτερο μέρος της λειτουργικότητας που επέτρεψε την δημιουργία και δοκιμή σημαντικού αριθμού μοντέλων.

Οι διαπαφές της εφαρμογής με τον εξωτερικό κόσμο έφεραν εις πέρας την διανομή των δεδομένων με τρόπο που τα καθιστά αξιοποιήσιμα από το κοινό.

6.4 Σύγκριση με Άλλες Μελέτες

Τα αποτελέσματα της μελέτης αντικατοπτρίζουν σε μεγάλο βαθμό προηγούμενες έρευνες για την μοντελοποίηση πυρκαγιών στην Ελλάδα· οι μετεωρολογικές συνθήκες, η βλάστηση, το έδαφος και η ανθρώπινη δραστηριότητα καθοδηγούν την εκδήλωση των αγροτοδοσικών πυρκαγιών. Η γεωχωρική ανάλυση στην Χίο των Adaktylou κα. [15] αξιοποιεί πολύ μεγαλύτερη γκάμα γεωχωρικών παραγόντων και αναφέρει επιτυχία στον εντοπισμό ιστορικών πυρκαγιών, χωρίς όμως να παρέχει συγκεκριμένες μετρικές. Οι Simeonidis κα. [68] χρησιμοποίησαν τον XGBoost για ανάλυση επικινδυνότητας με 83% επιτυχία στον ελληνικό χώρο, χρησιμοποιώντας γεωχωρικά χαρακτηριστικά και το ίδιο αρχείο του Πυροσβεστικού Σώματος. Αυτές οι μελέτες αποσκοπούν στον διαχωρισμό του εδάφους σε ζώνες επικινδυνότητας, αλλά δεν κάνουν προβλέψεις για τις τρέχουσες συνθήκες.

Η έρευνα των Apostolakis κα. [14] για την καθημερινή πρόβλεψη πυρκαγιάς με βαθιά μάθηση επικεντρώνεται στην εξηγησιμότητα των αποτελεσμάτων και την τεχνική ανάθεσης επιπέδου κινδύνου και όχι στην επίτευξη συγκεκριμένων μετρικών απόδοσης για τις προβλέψεις. Οι Gholami κα. [13] ακολουθούν μεθοδολογία παρόμοια με αυτήν που ακολουθήθηκε για αυτήν την έρευνα και κάνουν εκτίμηση κινδύνου στην Ινδία. Για πλέγμα κελιών 500μ και χρονικό διαχωρισμό, κατέγραψαν PR-AUC για LR, RF και XGB $0,0045 \pm 0,0025$ και $0,01 \pm 0,001$ για τις περιοχές έρευνας. Δεδομένης της πολύ υψηλής χωρικής ανάλυσης, το Lift αυτών των αποτελεσμάτων ενδέχεται να είναι υψηλότερο από αυτό των μοντέλων του Aristaeos.

Οι μετρικές αξιολόγησης των αποτελεσμάτων σε κάποιες μελέτες δεν είναι αρκετά κατατοπιστικές ως προς την πρακτική απόδοση των μοντέλων. Οι Moghim κα. [45] παραθέτουν μόνο τιμές ROC-AUC και Ευστοχίας (*Accuracy*), χωρίς να αναφέρουν το πλήθος των θετικών στιγμιότυπων ή το FScore. Το ROC-AUC δείχνει πόσο καλό είναι ένα μοντέλο στο να ξεχωρίζει τα επικίνδυνα στιγμιότυπα από αυτά που δεν είναι επικίνδυνα και είναι ανεξάρτητο από επικράτηση θετικής κλάσης και κατώφλι απόφασης. Ενδεικτικά, το επιλεγμένο μοντέλο του συστήματος καταγράφει 0,714. Σε επιχειρησιακό επίπεδο, αυτό που έχει σημασία για την εκτίμηση κινδύνου είναι η ανίχνευση της θετικής κλάσης και οι λάθος συναγερμοί, δηλαδή η ισορροπία μεταξύ ακρίβειας και ανάκλησης.

Η υλοποίηση ενός συστήματος, όπως αυτό που πραγματεύεται η μελέτη, εμπεριέχει συγκεκριμένα βήματα, όπως η συλλογή, η επεξεργασία και η γεωχωρική παρουσίαση των προβλέψεων, που είναι κοινά και απαραίτητα. Το έργο FireWeather.eu [17] αποτελεί ένα καλό παράδειγμα ολοκληρωμένου συστήματος εκτίμησης με χρήση των δεδομένων του δικτύου meteo [49]. Οι πυρικοί παράγοντες που αναπαριστώνται από τους ημερήσιους χάρτες του περιλαμβάνουν την συσσώρευση νεκρής βλάστησης, το FMC, την ταχύτητα εξάπλωσης, το FWI και την δυσκολία ελέγχου. Επίσης, παρέχει πυρομετεωρολογική πληροφόρηση για 3 δήμους τις Ελλάδος.

Το Aristaeos καινοτόμησε με την διενέργεια ζωντανής εκτίμησης κινδύνου πυρκαγιάς για περιοχές της Ελλάδος με ξεκάθαρη επιχειρησιακή κατεύθυνση. Η πλειοψηφία των μελετών και έργων που αφορούν τον ελληνικό χώρο κάνουν στατική χωρική ανάλυση επικινδυνότητας, ενώ μία μειο-

ψηφία επικεντρώνεται στις ημερήσιες εκτιμήσεις βάσει μετεωρολογικών προγνώσεων. Οι ζωντανές πυρικές εκτιμήσεις συγκαταλέγονται στην έρευνα αιχμής στην Ελλάδα και διεθνώς.

Κεφάλαιο 7

Συμπεράσματα

7.1 Αντίκτυπος της Μελέτης

Η λεπτομερής ανάλυση των δεδομένων που διενεργήθηκε, απέφερε ωφέλιμα συμπεράσματα για τα χαρακτηριστικά των πυρκαγιών στην περιοχή μελέτης, στα οποία περιλαμβάνεται και ο αντίκτυπος νομοθετικών ενεργειών του κράτους. Επιβεβαιώθηκε η ύπαρξη κοινών μοτίβων με άλλες Μεσογειακές περιοχές και, ταυτοχρόνως, τονίστηκαν οι τοπικές ιδιαιτερότητες και ο τρόπος που επηρεάζουν την εκδήλωση του φαινομένου.

Από τα πορίσματα της έρευνας, αυτό που ξεχωρίζει περισσότερο είναι πως η υλοποίηση ενός ολοκληρωμένου συστήματος εκτίμησης κινδύνου για μία περιοχή μελέτης ήταν επιτυχημένη. Η υποδομή και η πρακτική τεχνογνωσία που προέκυψαν μπορούν να γίνουν καταλύτες της συνεχόμενης έρευνα πάνω στο φαινόμενο των πυρκαγιών και την τυποποίηση της ανάλυσης διάφορων γεωγραφικών περιοχών.

Αναφορικά με την απόδοση και την αξιοπιστία των παραγόμενων προβλέψεων, εκτιμάται πως μπορεί να υπάρξει σταδιακή βελτίωση: Αφενός, με την εύρεση των κατάλληλων πρωτογενών δεδομένων και χαρακτηριστικών πρόβλεψης, αφετέρου, με την εφαρμογή πιο εξεζητημένων μεθόδων και πιο κατάλληλων παραμέτρων.

Με την κατάλληλη ρύθμιση, τις προσθήκες που ειπώθηκαν στην ενότητα 4.18, και τις αντίστοιχες πηγές δεδομένων, το σύστημα θα μπορούσε να τεθεί σε ζωντανή δοκιμαστική λειτουργία για οποιαδήποτε περιοχή. Ένα τέτοιο εργαλείο, επιτρέπει στις αρμόδιες αρχές να κατευθύνουν τα μέσα καταπολέμησης σε περιοχές υψηλού κινδύνου και να εντοπίσουν παθογένειες στον σχεδιασμό αντιμετώπισης των συμβάντων.

Το έργο κινήθηκε με ξεκάθαρους, μολονότι φιλόδοξους, επιχειρησιακούς στόχους αναφορικά με την επιτυχή πρόβλεψη των πυρκαγιών σε συνδυασμό με περιορισμό των λάθους συναγερμών. Η εφαρμογή αυτού του προτύπου στην υπερσύγχρονη προσέγγιση για ζωντανές εκτιμήσεις που ακολουθεί το υλοποιημένο σύστημα είναι αξιοσημείωτη. Μέσω σταδιακών βελτιώσεων της μεθοδολογίας και της υλοποίησης, το Aristaeos μπορεί να γίνει ένα μέσο άμεσης ενημέρωσης των κατοίκων περί του κινδύνου, ώστε να φροντίζουν την προσωπική τους ασφάλεια και να κινούνται με αυξημένη προσοχή.

7.2 Περιορισμοί

Το εγχείρημα αυτό έπρεπε να κινηθεί εντός ενός πλαισίου που συνεπάγεται κάποιους περιορισμούς· ενίοτε ήταν αίτιες για επινόηση εναλλακτικών μεθόδων, τις περισσότερες φορές, όμως, έθεσαν όρια στον σχεδιασμό και εμπόδια στην υλοποίηση. Το χρονοδιάγραμμα, στο οποίο έπρεπε να κινηθεί η εκπόνηση της μελέτης ήταν συγκεκριμένο, γεγονός που οδήγησε σε συμβιβασμούς που ευνόησαν την απλότητα, σε σχέση με εξεζητημένες και περίπλοκες προσεγγίσεις. Οι χρηματικοί πόροι και ο εξοπλισμός που ήταν διαθέσιμα στο έργο ήταν, επίσης, περιορισμένα.

Το εύρος των αξιοποιούμενων δεδομένων υπέφερε το μεγαλύτερο πλήγμα από τα περιθώρια που τέθηκαν στην μελέτη, καθώς δεν ήταν δυνατή η εύρεση, η αγορά, ή η υλοποίηση των απαραίτητων μεθόδων για την χρήση συγκεκριμένων δεδομένων. Αυτό είχε ως αποτέλεσμα κάποιοι πολύ σημαντικοί παράγοντες στην εκδήλωση των πυρκαγιών, όπως η πυκνότητα και η υγρασία της βλάστησης να προσεγγιστούν χονδρικά, διότι είναι ταχέα μεταβαλλόμενοι και δύσκολοι στην μέτρηση [12]. Άλλοι πυρικοί παράγοντες έμειναν εντελώς εκτός διερεύνησης, όπως ο άνεμος, λόγω των ελαττωματικών μετρήσεων, ή η καμένη έκταση και το οδικό δίκτυο, λόγω της περιπλοκότητας εφαρμογής τους εντός του χρονικού πλαισίου.

Επιθυμητός θα ήταν και μεγαλύτερος έλεγχος πάνω στους σταθμούς μετρήσεων: Δυστυχώς δεν ήταν δυνατή η εύρεση τρεχουσών μετρήσεων για την περιοχή μελέτης ώστε να γίνει ολοκληρωμένος έλεγχος της ζωντανής απόδοσης. Επιπροσθέτως, για τα μετεωρολογικά δεδομένα που χρησιμοποιήθηκαν, είναι αμφίβολο κατά πόσο η αραιή πυκνότητα των σταθμών μπόρεσε να αποτυπώσει τις συνθήκες στην ετερογενή περιοχή μελέτης. Παρομοίως, η χωρική λεπτομέρεια των συνθηκών παρέμεινε χαμηλή, λόγω περιορισμένων μεθόδων διαχείρισης πολύ μεγάλων συνόλων και περιορισμένης διαθέσιμης μνήμης.

Η θεώρηση της επάρκειας των διαθέσιμων δεδομένων είναι μία εργασία που θα ακολουθεί το σύστημα σε κάθε μελλοντική έρευνα και πιθανή εφαρμογή του.

Απόρροια των περιορισμών είναι και το περιορισμένο βάθος της ανάλυσης που εκτελείται από το Aristaeos. Τα μοντέλα είναι σε θέση να εντοπίσουν στατιστικές συσχετίσεις ανάμεσα στις πυρκαγιές και τους παράγοντες εκδήλωσης, δεν εντοπίζουν, όμως, τα ακριβή αίτια. Επίσης, δεν λαμβάνεται υπ' όψιν καθόλου ο πιθανός αντίκτυπος και η έκταση που ενδέχεται να λάβει μία πυρκαγιά, στοιχεία που θα καθιστούσαν τις εκτιμήσεις εξαιρετικά χρήσιμες στους μηχανισμούς απόκρισης.

7.3 Μελλοντική Έρευνα

Οι μελλοντικές έρευνες πάνω στο θέμα θα πρέπει να επικεντρωθούν κυρίως στην αξιοπιστία των προβλέψεων και την ανάπτυξη καλά τεκμηριωμένων μεθόδων μοντελοποίησης των συνθηκών. Χτίζοντας πάνω στην υπάρχουσα υποδομή, είναι δυνατόν να ενσωματωθεί και να δοκιμαστεί μία μεγάλη ποικιλία παραγόντων εκδήλωσης πυρκαγιών, ώστε να αποδοθούν πολύ πιο αποτελεσματικά οι διεργασίες που οδηγούν σε ανάφλεξη.

Η έλλειψη του ανέμου (Παρ. 3.10.1) από τα χαρακτηριστικά πρόβλεψης ήταν ιδιαίτερα ατυχής. Όπως αναφέρθηκε στην παράγραφο 2.3.2, και γενικότερα στην συντριπτική πλειοψηφία των ερευνών, η ύπαρξη ανέμου είναι πολύ ισχυρή ένδειξη πυρικής επικινδυνότητας. Η έλλειψη μεγάλου αριθμού καταγραφών κάποιας σημαντικής καιρικής μεταβλητής είναι κάτι που μπορεί να συμβεί με οποιονδήποτε επίγειο σταθμό. Εάν το πλήθος των χαμένων μετρήσεων είναι αρκετά μεγάλο, η απλή συμπλήρωση δεν είναι επαρκής και η μεταβλητή πρέπει να απορριφθεί. Για την βελτίωση

των αποτελεσμάτων της έρευνας, αλλά και μεγαλύτερη ανοχή του συστήματος στις αδυναμίες των επίγειων σταθμών, προτείνεται η έρευνα στις μεθόδους επανανάλυσης μετεωρολογικών δεδομένων. Ενδεικτικά, η ενσωμάτωση του ERA5 του Copernicus [69] ενδέχεται να αναπληρώνει για τις κενές μετρήσεις μίας καιρικής μεταβλητής πολύ πιο αποτελεσματικά από την απλή συμπλήρωση διάμεσου.

Το σύστημα είναι σε θέση να υποστηρίξει την δοκιμή επιπλέον αλγόριθμων, να επεκταθεί με επιπλέον παραμετροποιήσεις της εκπαίδευσης και να εμβαθύνει η μέθοδος ανάθεσης επιπέδου κινδύνου. Αυτά είναι οργανικά μονοπάτια για μελλοντική έρευνα που μπορούν να διερευνηθούν σχεδόν άμεσα. Με κάποιες προσθήκες, μπορούν να εξεταστούν ακόμη και εντελώς διαφορετικοί τρόποι λειτουργίας του συστήματος: Για παράδειγμα, ένα μοντέλο λειτουργίας σε ξεχωριστές περιόδους με εποχικό χαρακτήρα, το οποίο αποσκοπεί στην πρόληψη των εξαιρετικά επικίνδυνων και εκτεταμένων πυρκαγιών κατά την αντιπυρική περίοδο, και την αποτελεσματικότερη ανάθεση μέσων κατάσβεσης κατά τις αγροτικές περιόδους· μία τέτοια προσέγγιση ταιριάζει στην εποχικότητα στα διακριτά εποχικά μοτίβα που παρατηρήθηκαν στην περιοχή μελέτης. Ενδιαφέρον έχει και η εκδοχή ενός συστήματος που αξιοποιεί έναν συνδυασμό μοντέλων συσχέτισης και κατηγοριοποίησης των πυρκαγιών και των καιρικών μοτίβων, για βαθύτερη ανάλυση και πιο παραγωγικά συμπεράσματα. Σε αυτό το πλαίσιο, μπορεί να ερευνηθεί και η ενσωμάτωση αλγορίθμων DL.

Η χωρική ανάλυση των συνθηκών αποτέλεσε ένα κώλυμα στην μοντελοποίηση· η μεγαλύτερη λεπτομέρεια στους χάρτες κινδύνου είναι επιθυμητή από τους χρήστες και μπορεί να παρέχει μεγαλύτερη ακρίβεια στις συνθήκες επί του πεδίου και, επομένως, τις προβλέψεις. Για αυτό, κρίνεται σκόπιμη η διερεύνηση μεθόδων που αποδίδουν καλά σε υψηλές χωρικές αναλύσεις, καθώς και τεχνικών ελέγχου του όγκου των στιγμιοτύπων. Την αύξηση της χωρικής ανάλυσης υποστηρίζει η μέθοδος μετεωρολογικής παρεμβολής. Καθώς δεν μπορεί να υπάρξει σταθμός σε κάθε μικρή υποδιαίρεση της περιοχής, είναι απαραίτητη η ανάπτυξη κατάλληλων τεχνικών υπολογισμού των καιρικών συνθηκών ανά τομέα, βάσει των μετρήσεων που καταγράφουν οι προσκείμενοι σταθμοί, της τοπολογίας και των χαρακτηριστικών του εδάφους.

7.4 Επίλογος

Αναλύοντας τις πυρκαγιές, αποκαλύπτονται οι φανεροί και οι λιγότερο φανεροί μηχανισμοί που διέπουν τα φαινόμενα γύρω μας. Μπορούμε να εντοπίσουμε κάποιους από αυτούς, εξετάζοντας το ιστορικό των περιστατικών και των συνθηκών. Το επόμενο βήμα, είναι να αποτυπωθούν σε μία διαδικασία, που θα επιτρέπει την πρόβλεψη των πυρκαγιών, ώστε να έχουμε έτσι την ευκαιρία να ελαχιστοποιήσουμε τις καταστροφές.

Η μετάφραση της επιστημονικής έρευνας σε υπολογιστικές ρουτίνες ενός πρακτικού πληροφοριακού συστήματος ήταν το κυρίαρχο μοτίβο της μελέτης. Αυτό είναι ένα δύσκολο έργο που απαιτεί ορθή επιστημονική έρευνα, προσεκτική σχεδίαση και τεχνική αριστεία, και δεν θα έχει ολοκληρωθεί πλήρως με το πέρας αυτής της εργασίας. Η μελέτη, ωστόσο, μας αφήνει με χρήσιμες παρατηρήσεις για τις πυρκαγιές στην Π.Ε Ηρακλείου και ωφέλιμη μεθοδολογία για τα συστήματα εκτίμησης κινδύνου. Συνάμα, η υλοποίηση απέφερε ένα πρωτότυπο σύστημα ζωντανής πρόγνωσης, που δύναται να βοηθήσει την μελλοντική έρευνα.

Τα αποτελέσματα δείχνουν πως χρειάζεται περαιτέρω διερεύνηση πάνω στην πρακτική μοντελοποίηση πυρκαγιών, ειδικά όσον αφορά τις ζωντανές εκτιμήσεις και την έγκαιρη ενημέρωση. Η βελτίωση των μετρικών απόδοσης, η αντιπροσωπευτική μετάφραση των προβλέψεων σε επίπεδα

κινδύνου και η φορητότητα των μοντέλων είναι κάποια μόνο από τα ανοιχτά ζητήματα.

Βιβλιογραφία

- [1] C. Eberle και O. Higuera Roa, *Technical Report: Mediterranean wildfires*, Αύγ. 2022. doi: 10.53324/VCEB1752
- [2] N. Koutsias, M. Arianoutsou, A. Kallimanis, G. Mallinis, J. Halley και P. Dimopoulos, "Where did the fires burn in Peloponnisos, Greece the summer of 2007? Evidence for a synergy of fuel and weather", *Agricultural and Forest Meteorology*, τόμ. 156, σσ. 41–53, Απρ. 2012. doi: 10.1016/j.agrformet.2011.12.006
- [3] N. Efthimiou, E. Psomiadis και P. Panagos, "Fire severity and soil erosion susceptibility mapping using multi-temporal Earth Observation data: The case of Mati fatal wildfire in Eastern Attica, Greece", *CATENA*, τόμ. 187, σ. 104320, Νοέ. 2019. doi: 10.1016/j.catena.2019.104320
- [4] A. Y. Troumbis, C. M. Gaganis και H. Sideropoulos, "Probabilistic Wildfire Risk Assessment and Modernization Transitions: The Case of Greece", *Fire*, τόμ. 6, αρθμ. 4, 2023, issn: 2571-6255. doi: 10.3390/fire6040158 διεύθν.: <https://www.mdpi.com/2571-6255/6/4/158>
- [5] A. Ganteaume, A. Camia, M. Jappiot, J. San-Miguel-Ayanz, M. Long-Fournel και C. Lampin, "A Review of the Main Driving Factors of Forest Fire Ignition Over Europe", *Environmental management*, τόμ. 51, Οκτ. 2012. doi: 10.1007/s00267-012-9961-z
- [6] N. Koutsias κ.ά., "On the relationships between forest fires and weather conditions in Greece from long-term national observations (1894–2010)", *International Journal of Wildland Fire*, τόμ. 22, σσ. 493–507, Δεκ. 2012. doi: 10.1071/WF12003
- [7] M. Calviño-Cancela, M. L. Chas-Amil, E. D. García-Martínez και J. Touza, "Wildfire risk associated with different vegetation types within and outside wildland-urban interfaces", *Forest Ecology and Management*, τόμ. 372, σσ. 1–9, 2016, issn: 0378-1127. doi: <https://doi.org/10.1016/j.foreco.2016.04.002> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S0378112716301529>
- [8] Ministry of Climate Crisis and Civil Protection. "Ξεκινά αύριο 1η Μαΐου η αντιπυρική περίοδος -- Έναρξη έκδοσης του Ημερήσιου Χάρτη Πρόβλεψης Κινδύνου Πυρκαγιάς". States that the Greek fire-protection period starts on 1 May and lasts until 31 October., Υπουργείο Κλιματικής Κρίσης και Πολιτικής Προστασίας, επίσκεψη 25 Μάι. 2026. διεύθν.: <https://civilprotection.gov.gr/deltia-tupou/xekina-ayrio-li-maiou-i-antipyriki-periodos-enarxi-ekdosis-toy-imerisiou-harti>

- [9] A. Dimitrakopoulos, M. Vlahou, C. Anagnostopoulou και I. Mitsopoulos, "Impact of drought on wildland fires in Greece: Implications of climatic change?", *Climatic Change*, τόμ. 109, σσ. 331–347, Δεκ. 2011. doi: 10.1007/s10584-011-0026-8
- [10] M. Arianoutsou, G. Athanasakis, D. Kazanis και A. Christopoulou, "Attica: A Hot Spot for Forest Fires in Greece", *Fire*, τόμ. 7, αρθμ. 12, 2024, issn: 2571-6255. doi: 10.3390/fire7120467 διεύθυν.: <https://www.mdpi.com/2571-6255/7/12/467>
- [11] I. Zacharakis και V. A. Tsihrintzis, "Integrated wildfire danger models and factors: A review", *Science of The Total Environment*, τόμ. 899, σ. 165 704, 2023, issn: 0048-9697. doi: <https://doi.org/10.1016/j.scitotenv.2023.165704> διεύθυν.: <https://www.sciencedirect.com/science/article/pii/S0048969723043279>
- [12] Z. Xu κ.ά. "Deep Learning for Wildfire Risk Prediction: Integrating Remote Sensing and Environmental Data". arXiv: 2405.01607v5 [cs.LG], επίσκεψη 20 Μάι. 2026. διεύθυν.: <https://arxiv.org/abs/2405.01607>
- [13] S. Gholami, N. Kodandapani, J. Wang και J. Lavista Ferres, *Where there's Smoke, there's Fire: Wildfire Risk Predictive Modeling via Historical Climate Data*, Ιούλ. 2021.
- [14] A. Apostolakis, S. Girtsou, K. Alexis, G. Giannopoulos, N. S. Bartsotas και C. Kontoes, "Enhancing Daily Wildfire Risk Prediction Application Through Interpretable Machine Learning Results", στο *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, 2024, σσ. 2066–2071. doi: 10.1109/IGARSS53475.2024.10641551
- [15] N. Adaktylou, D. Stratoulas και R. Landenberger, "Wildfire Risk Assessment Based on Geospatial Open Data: Application on Chios, Greece", *ISPRS International Journal of Geo-Information*, τόμ. 9, αρθμ. 9, 2020, issn: 2220-9964. doi: 10.3390/ijgi9090516 διεύθυν.: <https://www.mdpi.com/2220-9964/9/9/516>
- [16] K. Lagouvardos, V. Kotroni, T. M. Giannaros και S. Dafis, "Meteorological Conditions Conducive to the Rapid Spread of the Deadly Wildfire in Eastern Attica, Greece", *Bulletin of the American Meteorological Society*, τόμ. 100, αρθμ. 11, σσ. 2137–2145, 2019. doi: 10.1175/BAMS-D-18-0231.1 διεύθυν.: <https://journals.ametsoc.org/view/journals/bams/100/11/bams-d-18-0231.1.xml>
- [17] FLAME Pyrometeorological Team. "FireWeather.eu", επίσκεψη 21 Μάι. 2026. διεύθυν.: <https://fireweather.eu/>
- [18] M. Elhag, N. Yilmaz και C. A. Coman Dumitrache, "Post-fire fuel and vegetation dynamics in an ungrazed phryganic community of Crete, Greece", *Applied Ecology and Environmental Research*, τόμ. 16, σσ. 3289–3303, Ιαν. 2018. doi: 10.15666/aeer/1603_32893303
- [19] Αποκεντρωμένη Διοίκηση Κρήτης. "Γεωπύλη Αποκεντρωμένης Διοίκησης Κρήτης: Gallery of Maps", επίσκεψη 21 Μάι. 2026. διεύθυν.: <https://geoportal.apdkritis.gov.gr/gis/home/gallery.html?sortField=title&sortOrder=asc&focus=maps>
- [20] M. Tzouveleki, M. Hatzaki, V. Antoniou, A. Karali και C. Giannakopoulos, "Modelling Future Forest Fire Risk for the Tourism Sector of Crete", *Environmental Sciences Proceedings*, τόμ. 26, αρθμ. 1, 2023, issn: 2673-4931. doi: 10.3390/environsciproc2023026150 διεύθυν.: <https://www.mdpi.com/2673-4931/26/1/150>

- [21] S. Archibald και G. P. Hempson, "Competing consumers: contrasting the patterns and impacts of fire and mammalian herbivory in Africa", *Philosophical Transactions of the Royal Society B: Biological Sciences*, τόμ. 371, αρθμ. 1703, σ. 20150309, Σεπτ. 2016, issn: 0962-8436. doi: 10.1098/rstb.2015.0309 eprint: <https://royalsocietypublishing.org/rstb/article-pdf/doi/10.1098/rstb.2015.0309/112831/rstb.2015.0309.pdf>. διεύθν.: <https://doi.org/10.1098/rstb.2015.0309>
- [22] L. Kobziar κ.ά., "Principles of fire ecology", *Fire Ecology*, τόμ. 20, Απρ. 2024. doi: 10.1186/s42408-024-00272-0
- [23] I. Michael, "Application of Thermodynamic Methods to the Study of Plant Biomass and Its Components—A Review", *Applied Biosciences*, τόμ. 3, αρθμ. 4, σσ. 577–616, 2024, issn: 2813-0464. doi: 10.3390/applbiosci3040036 διεύθν.: <https://www.mdpi.com/2813-0464/3/4/36>
- [24] J. Ruffault, T. Curt, N. K. Martin-StPaul, V. Moron και R. M. Trigo, "Extreme wildfire events are linked to global-change-type droughts in the northern Mediterranean", *Natural Hazards and Earth System Sciences*, τόμ. 18, αρθμ. 3, σσ. 847–856, 2018. doi: 10.5194/nhess-18-847-2018 διεύθν.: <https://nhess.copernicus.org/articles/18/847/2018/>
- [25] A. Karali, M. Hatzaki, C. Giannakopoulos, A. Roussos, G. Xanthopoulos και V. Tenentes, "Sensitivity and evaluation of current fire risk and future projections due to climate change: The case study of Greece", *Natural Hazards and Earth System Sciences Discussions*, τόμ. 1, σσ. 4777–4800, Σεπτ. 2013. doi: 10.5194/nhessd-1-4777-2013
- [26] J. P. Argañaraz, M. A. Landi, C. M. Scavuzzo και L. M. Bellis, "Determining fuel moisture thresholds to assess wildfire hazard: A contribution to an operational early warning system", *PLOS ONE*, τόμ. 13, αρθμ. 10, σσ. 1–19, Οκτ. 2018. doi: 10.1371/journal.pone.0204889 διεύθν.: <https://doi.org/10.1371/journal.pone.0204889>
- [27] C. Tziaila, D. Veresoglou, D. Papakosta και A. Mamolos, "Changes in soil characteristics and plant species composition along a moisture gradient in a Mediterranean pasture", *Journal of Environmental Management*, τόμ. 80, αρθμ. 1, σσ. 90–98, 2006, issn: 0301-4797. doi: <https://doi.org/10.1016/j.jenvman.2005.08.017> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S0301479705002744>
- [28] A. Dimitrakopoulos, "A statistical classification of Mediterranean species based on their flammability components", *International Journal of Wildland Fire*, τόμ. 10, σσ. 113–118, Ιούλ. 2001. doi: 10.1071/WF01004
- [29] J. Madrigal κ.ά., "Do cone age and heating mode determine the opening of serotinous cones during wildfires? A new bench scale approach applied to *Pinus halepensis* Mill.", *Science of The Total Environment*, τόμ. 763, σ. 144222, 2021, issn: 0048-9697. doi: <https://doi.org/10.1016/j.scitotenv.2020.144222> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S0048969720377536>
- [30] E. Kostopoulou και G. Stavridis, "Wildfire Risk Assessment Using the Fire Weather Index (FWI) in Greece", *Climate*, τόμ. 13, αρθμ. 6, 2025, issn: 2225-1154. doi: 10.3390/cli13060109 διεύθν.: <https://www.mdpi.com/2225-1154/13/6/109>

- [31] S. Liodakis, D. Bakirtzis και A. Dimitrakopoulos, "Ignition characteristics of forest species in relation to thermal analysis data", *Thermochimica Acta*, τόμ. 390, αρθμ. 1, σσ. 83–91, 2002, issn: 0040-6031. doi: [https://doi.org/10.1016/S0040-6031\(02\)00077-1](https://doi.org/10.1016/S0040-6031(02)00077-1) διεύθν.: <https://www.sciencedirect.com/science/article/pii/S0040603102000771>
- [32] F. Tedim κ.ά., "Forest Fire Causes and Motivations in the Southern and South-Eastern Europe through Experts' Perception and Applications to Current Policies", *Forests*, τόμ. 13, αρθμ. 4, 2022, issn: 1999-4907. doi: 10.3390/f13040562 διεύθν.: <https://www.mdpi.com/1999-4907/13/4/562>
- [33] J. S. Pereira και M. M. Chaves, "Plant Responses to Drought Under Climate Change in Mediterranean-Type Ecosystems", στο *Global Change and Mediterranean-Type Ecosystems*, J. M. Moreno και W. C. Oechel, επιμελητές. New York, NY: Springer New York, 1995, σσ. 140–160, isbn: 978-1-4612-4186-7. doi: 10.1007/978-1-4612-4186-7_7 διεύθν.: https://doi.org/10.1007/978-1-4612-4186-7_7
- [34] I. Aguado, E. Chuvieco, R. Borén και H. Nieto, "Estimation of dead fuel moisture content from meteorological data in Mediterranean areas. Applications in fire danger assessment", *International Journal of Wildland Fire*, τόμ. 16, αρθμ. 4, σσ. 390–397, Αύγ. 2007, issn: 1049-8001. doi: 10.1071/WF06136 eprint: <https://connectsci.au/wf/article-pdf/16/4/390/159571/wf06136.pdf>. διεύθν.: <https://doi.org/10.1071/WF06136>
- [35] A. Rovithakis κ.ά., "Future climate change impact on wildfire danger over the Mediterranean: The case of Greece", *Environmental Research Letters*, τόμ. 17, σ. 045 022, Απρ. 2022. doi: 10.1088/1748-9326/ac5f94
- [36] C. Ricotta, S. Bajocco, D. Guglietta και M. Conedera, "Assessing the Influence of Roads on Fire Ignition: Does Land Cover Matter?", *Fire*, τόμ. 1, αρθμ. 2, 2018, issn: 2571-6255. doi: 10.3390/fire1020024 διεύθν.: <https://www.mdpi.com/2571-6255/1/2/24>
- [37] S. Gonzalez και L. Ghermandi, "How to define the wildland-urban interface? Methods and limitations: towards a unified protocol", *Frontiers in Environmental Science*, τόμ. 11, σ. 1284631, Ιαν. 2024. doi: 10.3389/fenvs.2023.1284631
- [38] A. Colantoni κ.ά., "Sustainable Land Management, Wildfire Risk and the Role of Grazing in Mediterranean Urban-Rural Interfaces: A Regional Approach from Greece", *Land*, τόμ. 9, αρθμ. 1, 2020, issn: 2073-445X. doi: 10.3390/land9010021 διεύθν.: <https://www.mdpi.com/2073-445X/9/1/21>
- [39] European Environment Agency, *Climate-related economic losses | Greece | Europe's environment 2025*, <https://www.eea.europa.eu/en/europe-environment-2025/countries/greece/climate-related-economic-losses>, Published 29 September 2025, modified 23 September 2025, accessed 20 May 2026, 2025.
- [40] C. Papalexatou και M. Matsaganis, *In focus -- Economic losses related to weather and climate in EU countries*, <https://www.eliamep.gr/en/in-focus-oi-oikonomikes-zimies-pou-schetizontai-me-ton-kairo-kai-to-klima-stis-chores-tis-ee/>, Accessed 20 May 2026, Αύγ. 2023.

- [41] K. Papagiannaki, T. Giannaros, S. Lykoudis, V. Kotroni και K. Lagouvardos, "Weather-related thresholds for wildfire danger in a Mediterranean region: The case of Greece", *Agricultural and Forest Meteorology*, τόμ. 291, σ. 108 076, Σεπτ. 2020. doi: 10.1016/j.agrformet.2020.108076
- [42] G. Papavasileiou και T. M. Giannaros, "Synoptic-scale drivers of fire weather in Greece", *Science of The Total Environment*, τόμ. 925, σ. 171 715, 2024, issn: 0048-9697. doi: <https://doi.org/10.1016/j.scitotenv.2024.171715> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S0048969724018576>
- [43] A. Economou και J. C. Karkazis, "Fires in Greece: Causes, Consequences and Measures for the Ecosystems Protection", *Journal of Management Sciences and Regional Development*, αρθμ. 7, σσ. 141–154, 2011. επίσκεψη 19 Μάι. 2026. διεύθν.: https://www.stt.aegean.gr/geopolab/_private/MANAGEMENT%20SCIENCES%20AND%20REGIONAL%20DEVELOPMENT%207%20-%20FIRES%20IN%20GREECE.%20CAUSES%20CONSEQUENCES%20AND%20MEASURES%20FOR%20ECOSYSTEMS%20PROTECTION%20-%20Economou.doc
- [44] M. Mishra κ.ά., "Spatial analysis and machine learning prediction of forest fire susceptibility: a comprehensive approach for effective management and mitigation", *The Science of The Total Environment*, σ. 171 713, Μαρ. 2024. doi: 10.1016/j.scitotenv.2024.171713
- [45] S. Moghim και M. Mehrabi, "Wildfire assessment using machine learning algorithms in different regions", *Fire Ecology*, τόμ. 20, Δεκ. 2024. doi: 10.1186/s42408-024-00335-2
- [46] Πυροσβεστικό Σώμα Ελλάδος. "Σύνολα Δεδομένων". Open datasets of the Hellenic Fire Service, επίσκεψη 22 Μάι. 2026. διεύθν.: <https://www.fireservice.gr/el/synola-dedomenon>
- [47] Αποκεντρωμένη Διοίκηση Κρήτης. "Κατάλογος Ανοιχτών Δεδομένων: Μετεωρολογικοί -- Υδρολογικοί Σταθμοί". Open data catalogue; publisher: Διεύθυνση Υδάτων, επίσκεψη 22 Μάι. 2026. διεύθν.: <https://data.apdkritis.gov.gr/search?f%5B0%5D=topic%3A56>
- [48] OpenWeather. "Weather API". Weather data API documentation, επίσκεψη 22 Μάι. 2026. διεύθν.: <https://openweathermap.org/api>
- [49] Meteo.gr. "Χάρτης θερμοκρασιών Ελλάδας". Online temperature map; National Observatory of Athens / IERSD, επίσκεψη 22 Μάι. 2026. διεύθν.: https://www.meteo.gr/greece_heat_Map.cfm
- [50] Copernicus Data Space Ecosystem. "Copernicus DEM: Global and European Digital Elevation Model". Digital Surface Model; Copernicus Contributing Missions, επίσκεψη 22 Μάι. 2026. διεύθν.: <https://dataspace.copernicus.eu/explore-data/data-collections/copernicus-contributing-missions/collections-description/COP-DEM>

- [51] T. Chen και C. Guestrin, "XGBoost: A Scalable Tree Boosting System", στο *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, 2016, σσ. 785–794. doi: 10.1145/2939672.2939785
- [52] S. Fatima, A. Hussain, S. Amir, S. H. Ahmed και S. Aslam, "XGBoost and Random Forest Algorithms: An in Depth Analysis", *Pakistan Journal of Scientific Research*, τόμ. 3, σσ. 26–31, Οκτ. 2023. doi: 10.57041/pjosr.v3i1.946
- [53] scikit-learn developers. "Tuning the decision threshold for class prediction". scikit-learn User Guide, επίσκεψη 22 Μάι. 2026. διεύθν.: https://scikit-learn.org/stable/modules/classification_threshold.html
- [54] "Advantages of the relational approach", στο *The Relational Model for Database Management: Version 2*. USA: Addison-Wesley Longman Publishing Co., Inc., 1990, isbn: 0201141922.
- [55] PostgreSQL Global Development Group. "PostgreSQL", επίσκεψη 22 Μάι. 2026. διεύθν.: <https://www.postgresql.org/>
- [56] PostGIS Project Steering Committee. "PostGIS: Spatial and Geographic Objects for PostgreSQL", επίσκεψη 22 Μάι. 2026. διεύθν.: <https://postgis.net/>
- [57] Tiger Data. "TimescaleDB Documentation". PostgreSQL extension for time-series data, επίσκεψη 24 Μάι. 2026. διεύθν.: <https://www.tigerdata.com/docs>
- [58] Polars. "Polars User Guide". DataFrame library for manipulating structured data, επίσκεψη 24 Μάι. 2026. διεύθν.: <https://docs.pola.rs/>
- [59] F. Pedregosa κ.ά., "Scikit-learn: Machine Learning in Python", *Journal of Machine Learning Research*, τόμ. 12, σσ. 2825–2830, 2011. διεύθν.: <https://jmlr.org/papers/v12/pedregosa11a.html>
- [60] W. McKinney, "Data Structures for Statistical Computing in Python", στο *Proceedings of the 9th Python in Science Conference*, 2010, σσ. 56–61. doi: 10.25080/Majora-92bf1922-00a
- [61] Celery Project. "Celery: Distributed Task Queue", επίσκεψη 22 Μάι. 2026. διεύθν.: <https://docs.celeryq.dev/>
- [62] FastAPI. "FastAPI Documentation". Python web framework for building APIs based on standard type hints, επίσκεψη 24 Μάι. 2026. διεύθν.: <https://fastapi.tiangolo.com/>
- [63] MapLibre. "MapLibre GL JS Documentation". TypeScript/WebGL library for interactive browser maps, επίσκεψη 24 Μάι. 2026. διεύθν.: <https://www.maplibre.org/maplibre-gl-js/docs/>
- [64] Docker Inc. "Docker Documentation". Documentation for containerizing applications, επίσκεψη 24 Μάι. 2026. διεύθν.: <https://docs.docker.com/>

- [65] *Πυρ. Διατ. 9/2000: Κανονισμός ρύθμισης μέτρων για την πρόληψη και αντιμετώπιση πυρκαγιών σε δασικές και αγροτικές εκτάσεις*, Useful for documenting the pre-2021 regulatory baseline; repealed by Fire Regulation 9/2021., 30 Νοέ. 2000. επίσκεψη 25 Μάι. 2026. διεύθν.: <https://www.elinyae.gr/ethniki-nomothesia/pyrdiat-92000-fek-1459b-30112000>
- [66] *Υ.Α. 15005 οικ. Φ.700.9/2021: Έγκριση της υπ' αρ. 9/2021 Πυροσβεστικής Διάταξης με θέμα: Καθορισμός μέτρων για την πρόληψη και αποφυγή εκδήλωσης πυρκαγιών σε δασικές, αγροτικές εκτάσεις και οικοπεδικούς χώρους*, 13 Μάι. 2021. επίσκεψη 25 Μάι. 2026. διεύθν.: https://www.fireservice.gr/documents/20184/1592806/fek_b_1923_2021.pdf/842ab3bf-759c-47c5-a4ec-ff024e0a2620
- [67] CAL FIRE, *2025 Fire Season Incident Archive*, Accessed: 2026-05-29, California Department of Forestry και Fire Protection, 2025. διεύθν.: <https://www.fire.ca.gov/incidents/2025>
- [68] P. Symeonidis, T. Vafeiadis, D. Ioannidis και D. Tzovaras, ``Wildfire Susceptibility Mapping in Greece Using Ensemble Machine Learning'', *Earth*, τόμ. 6, αρθμ. 3, 2025, issn: 2673-4834. doi: 10.3390/earth6030075 διεύθν.: <https://www.mdpi.com/2673-4834/6/3/75>
- [69] H. Hersbach κ.ά., *ERA5 hourly data on single levels from 1940 to present*, Copernicus Climate Change Service Climate Data Store, 2023. doi: 10.24381/cds.adbb2d47 επίσκεψη 28 Μάι. 2026. διεύθν.: <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels>