



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Χαρτογράφηση καμένων εκτάσεων απο διαχρονικά δορυφορικά
δεδομένα Sentinel-2 στο περιβάλλον του Google Earth Engine**

Ψυχογιός Ιωάννης

Αθήνα, Ιούλιος 2021



NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF RURAL AND SURVEYING ENGINEERING
DEPARTMENT OF TOPOGRAPHY
LABORATORY OF REMOTE SENSING

DIPLOMA THESIS

**Burned area mapping from multitemporal Sentinel-2 satellite data
in Google Earth Engine**

Psichogios Ioannis

Athens, July 2021



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Χαρτογράφηση καμένων εκτάσεων απο διαχρονικά δορυφορικά
δεδομένα Sentinel-2 στο περιβάλλον του Google Earth Engine**

Ψυχογιός Ιωάννης

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή στις

Κ. Καραντζαλος

Β.Τσιχριτζής

Μ. Παπαδοπούλου

.....

.....

.....

Αν. Καθηγητής ΕΜΠ

Καθηγητής ΕΜΠ

Καθηγήτρια ΕΜΠ

Επιβλέπον



RSLab
Remote Sensing Laboratory
National Technical University of Athens
✓ Sensing ✓ Analytics ✓ Monitoring



Copyright © All rights reserved I. Ψυχογιός, 2021 Με επιφύλαξη παντός δικαιώματος

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ' ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικό ή ερευνητικής φύσης, υπό την προϋπόθεση να αναγράφεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς το συγγραφέα.

Ευχαριστίες

Η εργασία αυτή σηματοδοτεί την ολοκλήρωση των σπουδών μου στο Ε.Μ.Π. , και συνεπώς θα ήθελα να ευχαριστήσω όλους όσους συνέβαλαν στην περάτωση της. Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέπων καθηγητή κύριο Κωνσταντίνο Καράντζαλο για τη βοήθεια και την καθοδήγηση που μου προσέφερε. Στην συνέχεια οφείλω επίσης να ευχαριστήσω τον υποψήφιο διδάκτορα Ζαχαρία Κανδυλάκη, τον Δρ. Πολυχρόνη Κολοκούση αλλά και τον μεταπτυχιακό φοιτητή Αλέξανδρο Φαλάγγα για τη βοήθεια που μου προσέφεραν τόσο στο τεχνικό κομμάτι με τους κώδικες προγραμματισμού, όσο και για τις συμβουλές τους και την κρίση τους στο κομμάτι που αφορά την αξιολόγηση των αποτελεσμάτων. Τέλος οφείλω ένα μεγάλο ευχαριστώ στην οικογένεια μου για την πνευματική και υλική συμπαράσταση που μου προσέφερε καθόλη τη διάρκεια της φοίτησης μου.

Περίληψη

Στόχος της παρούσας διπλωματικής εργασίας είναι η δημιουργία ενός αλγορίθμου για χαρτογράφηση καμένων εκτάσεων χρησιμοποιώντας καινοτόμες μεθόδους αλγορίθμων που μας προσφέρουν οι σύγχρονες τεχνολογίες. Συγκεκριμένα, ο αλγόριθμος προγραμματίστηκε στη διαδικτυακή εφαρμογή Google Earth Engine (GEE), μια εφαρμογή η οποία έχει πολλές δυνατότητες καθώς προσφέρει στο χρήστη άμεση, εύκολη και γρήγορη πρόσβαση σε πληθώρα μεγάλων δορυφορικών δεδομένων μέσω χρονοσειρών δορυφορικών εικόνων οι οποίες είναι αποθηκευμένες σε βάσεις δεδομένων.

Για τις ανάγκες της παρούσας μελέτης χρησιμοποιήθηκαν ατμοσφαιρικά διορθωμένα δορυφορικά δεδομένα Sentinel-2. Ο δορυφόρος Sentinel-2 εξυπηρετεί με πολύ ικανοποιητικό βαθμό τις ανάγκες της μελέτης αυτής λόγω της υψηλής του χωρικής του ανάλυσης (10 m), αλλά και λόγω της συχνότητας παρατήρησης της γήινης επιφάνειας, καθώς προσφέρει δεδομένα ανά 5 η 6 ημέρες για κάθε τμήμα της φυσικής γήινης επιφάνειας .

Η χαρτογράφηση επιτεύχθηκε με χρήση επιβλεπόμενης ταξινόμησης. Οι αλγόριθμοι μηχανικής μάθησης που βασίζονται σε δένδρα απόφασης για ταξινόμηση κρίνονται ως σαφώς πιο αποτελεσματικοί από τους συμβατικούς αλγορίθμους (K-means, minimum distance, κοκ), και για αυτό ο αλγόριθμος που χρησιμοποιήσαμε είναι τα τυχαία δάση. Αναγκαία προϋπόθεση για υψηλή ακρίβεια και ικανοποιητικό αποτέλεσμα είναι η συλλογή των κατάλληλων δεδομένων εκπαίδευσης και ελέγχου. Για το σκοπό αυτό, δημιουργήθηκαν ψευδοεικόνες στον Ελλαδικό χώρο στις περιοχές Κύθηρα , Κινέτα, Μάτι, Γραμματικό και Κοντοδεσπότι. Συγκεκριμένα, ορίστηκαν στο περιβάλλον του GEE πολύγωνα ενδιαφέροντος για τις παραπάνω περιοχές και οι τιμές φωτεινότητας που τους αντιστοιχούν στα εικονοστοιχεία δεν πάρθηκαν από μια ημερομηνία, αλλά από ένα μεγάλο χρονικό εύρος.

Για την πραγματοποίηση των ταξινομήσεων επιλέχθηκαν 6 και 7 κατηγορίες αντίστοιχα για Ελλάδα και Αυστραλία, ικανές να εκφράσουν το σύνολο της μεταβλητότητας των περιοχών μελέτης στο επιθυμητό επίπεδο ανάλυσης, δηλαδή στον εντοπισμό των καμένων εκτάσεων. Συγκεκριμένα, για Ελλάδα οι κατηγορίες αυτές ήταν καμένα, νερό, βλάστηση, αστικός ιστός, καλλιέργειες και γυμνό έδαφος. Για Αυστραλία ήταν έρημος, νερό, καλλιέργειες, καμένα, αστικός ιστός, δάσος 1 και δάσος 2. Τα δεδομένα εκπαίδευσης και ελέγχου ψηφιοποιήθηκαν από το φυσικό έγχρωμο σύνθετο RGB 4-3-2, με εξαίρεση τα καμένα όπου έγινε χρήση του έγχρωμου συνθέτου RGB 12-8-2, στις 4 περιοχές που προαναφέρθηκαν. Πραγματοποιήθηκαν διαφορετικοί συνδυασμοί των παραπάνω περιοχών ως δεδομένα εκπαίδευσης και ελέγχου. Παράλληλα δημιουργήθηκαν διαγράμματα για να ελεγχθεί η φασματική συνεκτικότητα των πολυγώνων που ψηφιοποιήθηκαν για τις κατηγορίες αυτές. Επίσης οι καμένες εκτάσεις εμβαδομετρήθηκαν και έγινε σύγκριση των τιμών με τα δεδομένα που έδωσε η πυροσβεστική για Ελλάδα, και οι τοπικές υπηρεσίες για Αυστραλία αντίστοιχα.

Τέλος, έγινε προσπάθεια χαρτογράφησης των καμένων εκτάσεων στην Νέα Νότια Ουαλία της Αυστραλίας, όσο αφορά το μεγάλο κύμα δασικών πυρκαγιών που έπληξε τη χώρα τα έτη 2019 και 2020. Για τις ανάγκες της παρούσας διπλωματικής εργασίας πραγματοποιήθηκε μόνο ποιοτικός έλεγχος με φωτοερμηνεία σε ορισμένες περιοχές για να εξαχθούν συμπεράσματα όσο αφορά την αποδοτικότητα της διαδικασίας.

Λέξεις κλειδιά: Ταξινόμηση, Μηχανική μάθηση, Google Earth Engine, Sentinel-2, Ελλάδα, Αυστραλία .

Abstract

At the current thesis, there was developed an algorithm suitable for burned area mapping, using innovative techniques the technology enables us to use. Specifically, the algorithm was programmed in the web application Google Earth Engine(GEE), a web application with many capabilities since it enables us to have direct, easy and fast attendance in a large database with big satellite data via time series of satellite images.

In the current research, there were used atmospheric corrected satellite data from the satellite Sentinel – 2 . This satellite was considered as suitable for the needs of the current research, due to its high spatial resolution (10 m) , as well as for the frequency it detects the earth surface, estimated each 5 to 6 days.

The mapping was achieved through supervised classification. Machine learning algorithms based in decision trees for classification are being considered as way more effective and suitable than standard algorithms (K-means, minimum distance etc), and considering that we used the random forest classification algorithm. Collecting the right training and validation data is essential for high accuracy and a satisfying result. For that purpose, there were created pseudo-images at Greece, at the regions Kithira, Kineta, Mati, Grammatiko and Kontodespoti. Specifically, there were created at the Google Earth Engine environment regions of interest as for the referred regions , and there was set up and temporal range at which there were detected all the Sentinel – 2 images that overlapped the polygons of the regions of interest. The final spectral value of each pixel is the mean of the all the pixels of the satellite images that overlapped.

There were assigned 6 categories for the classification, due to the fact that this number is considered as suitable for detecting burned areas from the rest categories, and for a successful classification result overall. Both training and validation data were being taken from the color composite RGB 4-3-2, aside from the burned data which were taken from the composite RGB 12- 8-2, for each of the four regions that were referred before. The region that got classified is the Kontodespoti in Eboea, with all 6 combinations of the training and validation data, after we created diagrams to make sure that each of the polygons of the training and validation data has a close spectral behavior to the mean of all of them. The best result of all 6 was given by the combination of training at Kineta and Grammatiko and validation at Mati and Kithira. This came to a conclusion after photo – interpretation as well comparing the estimated value in the Google Earth Engine to the one the fire service gave.

Finally, we tried to classify the New South Wales of Australia, due to the large scale bushfires that took place in the region at the years 2019 and 2020. The procedure was similar to Greece, even though 7 categories were used in the classification. It was impossible to classify New South Wales with the training data for Greece, due to the fact that different climate conditions occur in this area. Validation data were not used due to the fact that it is impossible to create a confusion matrix for such a large area of interest, as it is beyond the Google Earth Engine limitations. Aside from that, it became clear that the result is viable both with photo – interpretation as well as with comparison of the estimated value with the value the local services gave, in hectares. Maps were created for all the above.

Key words: Classification, Machine learning, Google Earth Engine, Sentinel – 2, Greece, Australia

Πίνακας περιεχομένων

Περίληψη.....	
Abstract	
1. Εισαγωγή.....	1
1.1 Αντικείμενο και στόχοι διπλωματικής.....	2
1.2 Η διαδικτυακή πλατφόρμα Google Earth Engine	3
2. Θεωρητικό υπόβαθρο και σχετική βιβλιογραφία	5
2.1 Τηλεπισκόπηση και γεωχωρικά δεδομένα.....	5
2.1.1 Χαρακτηριστικά τηλεπισκοπικών δεδομένων	6
2.1.2 Ατμοσφαιρικές και ραδιομετρικές διορθώσεις	7
2.2 Μηχανική μάθηση	12
2.2.1 Δένδρα απόφασης	13
2.3 Το Ευρωπαϊκό Πρόγραμμα COPERNICUS.....	18
2.3.1 Ο δορυφόρος Sentinel - 2	18
2.4 Πυρκαγιές.....	20
3. Μεθοδολογία και υλοποίηση.....	21
3.1 Πυρκαγιές στις επιμέρους περιοχές μελέτης της Ελλάδας.....	21
3.2 Νέα Νότια Ουαλία και κύμα δασικών πυρκαγιών	23
3.3 Δημιουργία ψευδοεικόνων	26
3.3.1 Επαναπροβολή εικονοστοιχείων	27
3.4 Δεδομένα εκπαίδευσης και ελέγχου	28
3.4.1 Ελλάδα.....	29
3.4.2 Αυστραλία	34
3.5 Πειράματα και αξιολόγηση τους	39
3.5.1 Πίνακες σύγχυσης και επιμέρους ακρίβειες	39
4. Χάρτες και σχολιασμός	43
5. Συμπεράσματα - Προτάσεις	48

Παράρτημα εικόνων

Εικόνα 1.1 Το περιβάλλον του Google Earth Engine	3
Εικόνα 2.1 . Δένδρο απόφασης	14
Εικόνα 2.2 . Δείκτης gini	15
Εικόνα 2.3 Παράδειγμα ταξινόμησης τυχαίου εικονοστοιχείου άγνωστης τάξης με τον αγόριθμο τυχαίων δασών.....	16
Εικόνα 2.4. Κανάλια δορυφόρου Sentinel – 2	19
Εικόνα 3.1 . Περιοχές μελέτης στον Ελλαδικό χώρο	22
Εικόνα 3.2 Καμένη έκταση στην Εύβοια	23
Εικόνα 3.3 Η Νέα Νότια Ουαλία (Πολιτικός χάρτης)	24
Εικόνα 3.4 . Πυρκαγιές από πρωτότυπο χάρτη	25
Εικόνα 3.5 .Οι πυρκαγιές όπως φαίνονται από δορυφορική εικόνα	25
Εικόνα 3.6 Φυσικό έγχρωμο σύνθετο ψευδοεικόνας στο Γραμματικό	26
Εικόνα 3.7. Πρίν την προβολή	27
Εικόνα 3.8. Μετά την προβολή	27
Εικόνα 3.9 . Πολύγωνα λήψης για καμένα από τα Κύθηρα	29
Εικόνα 3.10 . Πολύγωνα λήψης από φυσικό έγχρωμο σύνθετο στην Κινέτα	29
Εικόνα 3.11. Φυσικό έγχρωμο σύνθετο (RGB: 432) για την περιοχή της Νέας Νότιας Ουαλίας	36
Εικόνα 3.12 . Λήψη δεδομένων για καμένα στη Νέα Νότια Ουαλία απο ψευδέχρωμο σύνθετο RGB 12-8-2	38
Εικόνα 3.13. Έγχρωμο σύνθετο RGB 12-8-2 για Κοντοδεσπότη Εύβοιας.....	43
Εικόνα 3.14. Αποτέλεσμα ταξινόμησης περιοχής Κοντοδεσπότη	44
Εικόνα 3.15. Καμένα περιοχής Κοντοδεσπότη	45
Εικόνα 3.16. Εικονοστοιχεία ακρογραμμής	45
Εικόνα 3.17. Καμένα περιοχής Νέας Νότιας Ουαλίας	46
Εικόνα 3.18. Χάρτης απο πρωτότυπες πηγές Νέας Νότιας Ουαλίας	46

Παράρτημα πινάκων

Πίνακας 3.1. Πολύγωνα ταξινόμησης	30
Πίνακας 3.2. Πίνακας σύγκρισης για πείραμα 2	41
Πίνακας 3.3. Πίνακας σύγκρισης για πείραμα 6	41
Πίνακας 3.3. Πίνακας σύγκρισης για Νέα Νότια Ουαλία	42

Παράρτημα διαγραμμάτων

Διάγραμμα 3.1 . Σύνολο παρατηρήσεων για καμένα από τα Κινέτα	32
Διάγραμμα 3.2. Μέσος όρος και τυπική απόκλιση για καμένα από τα Κινέτα	32
Διάγραμμα 3.3 . Σύνολο παρατηρήσεων για έδαφος από τα Κινέτα	32
Διάγραμμα 3.4. Μέσος όρος και τυπική απόκλιση για έδαφος από τα Κινέτα	32
Διάγραμμα 3.5 . Σύνολο παρατηρήσεων για τα καμένα από το Γραμματικό	33
Διάγραμμα 3.6. Μέσος όρος και τυπική απόκλιση για καμένα από Το Γραμματικό	33
Διάγραμμα 3.7 . Σύνολο παρατηρήσεων για έδαφος από το Γραμματικό	33
Διάγραμμα 3.8. Μέσος όρος και τυπική απόκλιση για έδαφος από το Γραμματικό	33
Διάγραμμα 3.9. Σύνολο παρατηρήσεων για καλλιέργειες από τα Κινέτα	34
Διάγραμμα 3.10 Μέση τιμή και τυπική απόκλιση για καλλιέργειες από τα Κινέτα	34
Διάγραμμα 3.11. Σύνολο παρατηρήσεων για δάσος1 από την Αυστραλία	37
Διάγραμμα 3.12 .Μέση τιμή και τυπική απόκλιση για δάσος1 από την Αυστραλία	37
Διάγραμμα 3.13. Σύνολο παρατηρήσεων για καμένα από την Αυστραλία	37
Διάγραμμα 3.14 Μέση τιμή και τυπική απόκλιση για καμένα από την Αυστραλία	37

1. Εισαγωγή

Η Τηλεπισκόπηση είναι από τον ορισμό της μια τεχνική παρατήρησης, επεξεργασίας και εξαγωγής συμπερασμάτων δεδομένων από απόσταση. Η σύγχρονη τεχνολογία μας παρέχει άμεση πρόσβαση σε πληθώρα δορυφορικών δεδομένων, τα οποία και μπορούν να χρησιμοποιηθούν για πολλούς σκοπούς και εφαρμογές. Οι σύγχρονοι τηλεπισκοπικοί δέκτες βρίσκονται σε συνεχή τροχιά πάνω από τη γη και ανιχνεύουν την ηλιακή ακτινοβολία που ανακλάται από τη γήινη επιφάνεια σε μεγάλο εύρος του ηλεκτρομαγνητικού φάσματος, προσφέροντας τη δυνατότητα στο χρήστη να δει και χρησιμοποιήσει εικόνες πέρα των περιορισμών του φυσικού έγχρωμου σύνθετου, αυτού δηλαδή που μπορεί να αντιληφθεί ο ανθρώπινος οφθαλμός. Έτσι μπορεί ο χρήστης να δημιουργήσει έγχρωμα σύνθετα με τηλεπισκοπικά κανάλια από το υπέρυθρο και το υπεριώδες, και συνεπώς να δει χαρακτηριστικά που ειδάλλως θα ήταν δυσδιάκριτα έτσι όπως τα αντιλαμβάνεται ο ανθρώπινος οφθαλμός.

Η χωρική ανάλυση είναι ιδιαίτερα υψηλή, με τα δεδομένα που παρέχονται από τις σύγχρονες τεχνολογίες, και αυτό καθιστά τις επιμέρους εφαρμογές Τηλεπισκόπησης πιο ακριβείς και αποδοτικές εν γένει. Για τις ανάγκες της παρούσας μελέτης χρησιμοποιήθηκαν δορυφορικά δεδομένα Sentinel-2, τα κανάλια του οποίου έχουν 10, 20 και 60 μέτρα χωρική ανάλυση. Με χρήση προγραμματιστικών τεχνικών η ανάλυση όλων των καναλιών του δορυφόρου Sentinel-2 βελτιώθηκε στα 10 μέτρα.

Οι τηλεπισκοπικές ταξινομήσεις είναι ιδιαίτερα χρήσιμες για τη δημιουργία εικόνων με καλύψεις γης. Υπάρχουν δυο ειδών ταξινομήσεις, η επιβλεπόμενη όπου ο χρήστης ορίζει τις κατηγορίες που επιθυμεί και την εκπαιδεύει με δεδομένα για κάθε μια από αυτές και η μη επιβλεπόμενη, όπου και ο χρήστης αναγνωρίζει εκ των υστέρων τις ταξινομημένες κατηγορίες από ηλεκτρονικές πηγές χαρτογραφικά υπόβαθρα κοκ. Έχει διαπιστωθεί ότι η επιβλεπόμενη ταξινόμηση δίνει καλύτερα αποτελέσματα, τόσο από τη διεθνή βιβλιογραφία όσο και από ερευνητικές ομάδες.

1.1 Αντικείμενο και στόχοι διπλωματικής

Στην παρούσα διπλωματική εργασία, σκοπός ήταν η δημιουργία ενός συνόλου αλγορίθμων για εντοπισμό και χαρτογράφηση καμένων εκτάσεων από τηλεπισκοπικά δορυφορικά δεδομένα Sentinel-2.

Κρίθηκε αναγκαία η εφαρμογή επιβλεπόμενης ταξινόμησης με λήψη δεδομένων εκπαίδευσης και ελέγχου σε περιοχές διαφορετικές από αυτή που θα ταξινομηθεί. Για τον εντοπισμό καμένων εκτάσεων στο σύνολο του ελλαδικού χώρου κρίθηκε αναγκαία η λήψη δεδομένων εκπαίδευσης και ελέγχου από πολλαπλές περιοχές.

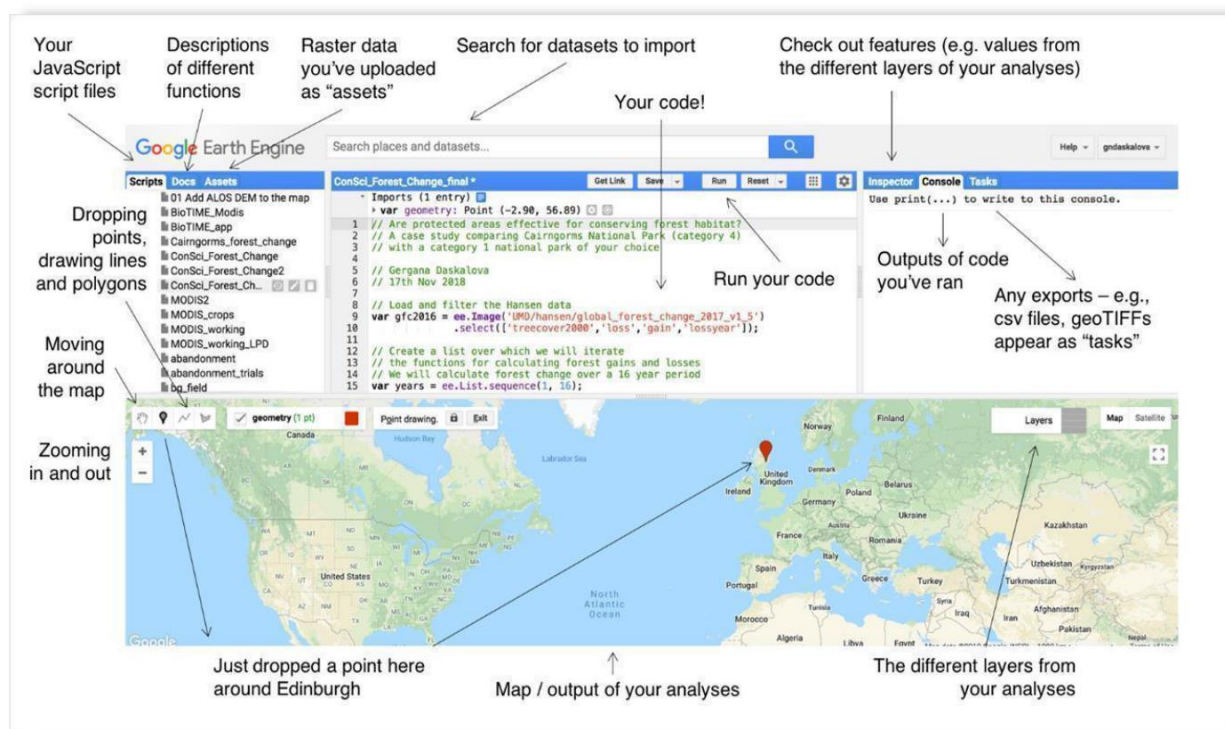
Οι νέες τεχνολογίες προσφέρουν σε μηχανικούς, ερευνητές και επιστήμονες πληθώρα δορυφορικών δεδομένων, τα οποία είναι εύκολα προσβάσιμα, έχουν υψηλή χωρική ανάλυση και υψηλούς ρυθμούς παρατήρησης της φυσικής γήινης επιφάνειας, και προσφέρουν στο χρήστη δυνατότητες για πληθώρα εφαρμογών. Η μηχανική μάθηση προσφέρει επίσης αποτελεσματικές μεθόδους για παλινδρόμηση, ταξινόμηση, ομαδοποίηση κι άλλες πολλές εφαρμογές.

Στόχος της διπλωματικής εργασίας είναι η χαρτογράφηση καμένων εκτάσεων με μέθοδο ταξινόμησης σε κατηγορίες κάλυψης γης, και ο έλεγχος του καλύτερου συνδυασμού γι' αυτό με ποσοτικό και ποιοτικό έλεγχο σε πειράματα που εφαρμόστηκαν, με την υψηλή χωρική ανάλυση (10 m) και το υψηλό εύρος παρατήρησης που προσφέρουν οι δορυφόροι Sentinel-2.

Τέλος, επιμέρους στόχοι είναι η ποσοτική αξιολόγηση της διαδικασίας και του αποτελέσματος με σύγκριση αριθμητικών τιμών, αυτές που έδωσαν οι ταξινομήσεις με τις αντίστοιχες άλλων υπηρεσιών, για να μπορέσουν να εξαχθούν ακριβέστερα συμπεράσματα όσο αφορά την ακρίβεια της διαδικασίας.

1.2 Η διαδικτυακή πλατφόρμα Google Earth Engine

Το Google Earth Engine είναι μια διαδικτυακή εφαρμογή προωθημένου προγραμματισμού, όπου ο χρήστης έχει πρόσβαση σε μια μεγάλη βάση δεδομένων με δορυφορικά δεδομένα, τόσο σε εικόνες όσο και σε δεδομένα ραντάρ (SAR). Μπορεί ο χρήστης να χρησιμοποιήσει δεδομένα από πολλούς δορυφόρους, όπως Sentinel-2, Landsat-8, Landsat-5, Landsat-2, Landsat-1 κ.ο.κ. Για τις ανάγκες της παρούσας εργασίας αυτής χρησιμοποιήθηκαν ατμοσφαιρικά διορθωμένα δορυφορικά δεδομένα Sentinel-2 από τη χρονοσειρά S2_SR. Στην εικόνα 1.1 παρουσιάζεται ενδεικτικά το περιβάλλον του Google Earth Engine, και αναλύονται οι δυνατότητες και τα βασικά εργαλεία του.



Εικόνα 1.1. Το περιβάλλον του Google Earth Engine

Στο Google Earth Engine ο χρήστης έχει τη δυνατότητα να προγραμματίσει είτε σε Python είτε σε JavaScript. Στην εργασία αυτή χρησιμοποιήθηκε κώδικας γραμμένος σε JavaScript κι αυτό είναι προτιμότερο γενικά διότι και το στο εγχειρίδιο χρήσης τα κομμάτια κώδικα είναι σε αυτή τη γλώσσα

προγραμματισμού. Όπως φαίνεται και στην εικόνα 1.1 , στο περιβάλλον της εφαρμογής υπάρχει ένας παγκόσμιος χάρτης σε μερκατορική προβολή όπου ο χρήστης μπορεί να κάνει την ανάλυση του. Υπάρχει η επιλογή για τον χρήστη να ορίσει μια γεωμετρία αναφοράς για την περιοχή μελέτης είτε με σημείο, περίπτωση στην οποία θα μπορεί να έχει ένα εικονοστοιχείο, είτε με τετράγωνο είτε και με ακανόνιστο πολύγωνο. Η επιλογή για ορισμό γεωμετρίας βρίσκεται κάτω αριστερά στο geometry. Η μετακίνηση στο χάρτη γίνεται με χρήση του κέρσορα. Πάνω αριστερά στην καρτέλα Scripts παραθέτονται τα διάφορα επιμέρους στοιχεία , όπως ο αλγόριθμος τυχαία δάση που χρησιμοποιήθηκε σε αυτή τη μελέτη. Στην καρτέλα Docs υπάρχει περιγραφή για τις δυνατότητες των εργαλείων αυτών και στην καρτέλα Assets ο χρήστης μπορεί να ανεβάσει αρχεία εικόνων σε .GEO TIFF format αλλά και shapefiles που μπορούν να χρησιμοποιηθούν ως γεωμετρίες στον κώδικα.

Πάνω στη μηχανή αναζήτησης δίνεται η δυνατότητα στο χρήστη να αναζητήσει δεδομένα όπως χρονοσειρές δορυφορικών εικόνων, περιοχές στο χάρτη και όχι μόνο. Στον code editor δημιουργείται το script, δεξιά στο Console εμφανίζονται όσα έχει αποφασίσει η προγραμματιστής να εμφανίσει με την εντολή 'print'. Αυτό έχει ιδιαίτερη σημασία καθώς πολλές φορές είναι σημαντικό να φαίνονται πράγματα όπως το προβολικό σύστημα, τα metadata, το πλήθος των εικόνων που επικαλύπτουν ένα πολύγωνο κτλ. Ιδιαίτερα σημαντική είναι και η καρτέλα Tasks, όπου μπορεί να γίνει εξαγωγή στον Google Drive, στα Assets κτλ δεδομένων από το περιβάλλον του Google Earth Engine. Η πρακτική εφαρμογή της καρτέλας αυτής είναι πως έτσι μπορεί να εξάχθει ένας ταξινομημένος χάρτης ή ένα .csv αρχείο για να γίνει επεξεργασία κάπου αλλού, όπως και έγινε και σε αυτή την εργασία. Επίσης καλό είναι να αναφερθεί πως το Google Earth Engine έχει ορισμένους περιορισμούς, οπότε αν η χρήστης επιθυμεί να κάνει έναν υπολογισμό που διαρκεί πολλή ώρα μπορεί να κάνει ταυτόχρονα export την υπολογισμένη τιμή σε .csv, ειδάλλως θα εμφανιστεί error και δε θα γίνει ο υπολογισμός λόγω του περιορισμού στο computation time.

Τέλος, κάτω δεξιά εμφανίζονται τα επιμέρους θεματικά επίπεδα που έχουν προκύψει ως αποτέλεσμα από την εκτέλεση του κώδικα. Ο χρήστης έχει τη δυνατότητα να τα επεξεργαστεί όπως αυτός επιθυμεί και από εκεί, ξεχωριστά από τον κώδικα. Μπορεί να γίνει ενίσχυση του ιστογράμματος, αλλαγή στο έγχρωμο σύνθετο κοκ. Ο κώδικας αποθηκεύεται σε link και ανοίγει κάθε φορά που εφαρμοστεί αυτό στον web browser.

Συμπερασματικά, το Google Earth Engine είναι μια ιδιαίτερα εύχρηστη και φιλική προς τον χρήστη εφαρμογή, με πολλές δυνατότητες, εύκολη και γρήγορη πρόσβαση σε μεγάλο όγκο δεδομένων και πολλή ποικιλομορφία. Οι επιμέρους περιορισμοί αποτελούν τροχοπέδη για το χρήστη σε πολλές περιπτώσεις, άλλες φορές όμως μπορούν να παρακαμφθούν με τεχνάσματα.

2. Θεωρητικό υπόβαθρο και σχετική βιβλιογραφία

Στο κεφάλαιο αυτό αναλύεται το θεωρητικό υπόβαθρο της εργασίας. Συγκεκριμένα, εξηγείται το τι είναι η Τηλεπισκόπηση, τι χαρακτηριστικά, ιδιότητες και περιορισμούς έχει. Αναλύεται επίσης το κομμάτι της μηχανικής μάθησης που βασίζεται σε δένδρα απόφασης και το πώς λειτουργεί ο αλγόριθμος τυχαία δάση που εφαρμόστηκε. Τέλος παρουσιάζεται ο δορυφόρος Sentinel-2, τα χαρακτηριστικά του και οι δυνατότητες που προσφέρει.

2.1 Τηλεπισκόπηση και γεωχωρικά δεδομένα

Ο όρος Τηλεπισκόπηση μπορεί να γενικευθεί ως οποιαδήποτε ενέργεια καταγραφής από απόσταση και ανάλυσης των εικόνων που προκύπτουν (ακόμη και των ανθρώπινων ματιών). Στη σημερινή εποχή λόγω της πολυφασματικότητας έχει καθιερωθεί ο όρος για την καταγραφή και την ανάλυση δορυφορικών εικόνων. Η έννοια της λέξης «δορυφόρος» εστιάζεται κυρίως στους τεχνητούς δορυφόρους που είναι τεχνητά συστήματα που πετούν γύρω από την γη ή γενικότερα από κάποιο ουράνιο σώμα και έχουν κατασκευαστεί για ποικίλες χρήσεις, οι κυριότερες των οποίων είναι οι τηλεπικοινωνίες και η παρατήρηση μέσω δορυφορικών εικόνων.

Οι δορυφόροι έχουν γενικά μια καθορισμένη εκ των προτέρων τροχιά. Η τροχιά αυτή υπακούει στους νόμους του Κέπλερ, που είναι οι εξής:

- Η τροχιά ενός δορυφόρου βρίσκεται σε ένα επίπεδο. Για την ακρίβεια, ο δορυφόρος διαγράφει μια έλλειψη της οποίας μια εστία (από τις δύο) κατέχει η γη.
- Ο δορυφόρος διαγράφει ίσες επιφάνειες σε ίσους χρόνους.

Οι δορυφόροι ανιχνεύουν και καταγράφουν την ανακλώμενη ηλιακή και την θερμική ακτινοβολία στα ορατά και στα υπέρυθρα μήκη κύματος του ηλεκτρομαγνητικού φάσματος. Είναι εφοδιασμένοι με έναν πολυφασματικό σαρωτή ο οποίος καταγράφει το ηλεκτρομαγνητικό σήμα που προέρχεται από την ίδια περιοχή της γήινης επιφάνειας, την ίδια χρονική στιγμή αλλά σε διάφορα μήκη κύματος, παρέχοντας με αυτόν τον τρόπο τα διαφορετικά φασματικά κανάλια μιας πολυφασματικής δορυφορικής εικόνας (SEOS 2011).

Οι δορυφόροι αυτοί περιλαμβάνουν διάφορων τύπων όργανα τα κυριότερα των οποίων καταγράφουν στα ορατό, εγγύς υπέρυθρο, θερμικό υπέρυθρο, και τμήματα της περιοχής των μικροκυμάτων του Η/Μ φάσματος. Αυτά είναι:

Επιταχυνσιόμετρο: καταγράφει την γραμμική επιτάχυνση σε μία ή περισσότερες διαστάσεις καθώς και την γωνιακή επιτάχυνση.

Ραδιόμετρο: μετρά ποσοτικά την ένταση της Η/Μ ακτινοβολίας σε ορισμένα τμήματα του Η/Μ φάσματος. Υπάρχουν ραδιόμετρα για το ορατό, το υπέρυθρο ή τα μικροκύματα.

Απεικονιστικό Ραδιόμετρο: έχει την ιδιότητα να σαρώνει και να δημιουργεί έναν διδιάστατο πίνακα εικονοστοιχείων από τον οποίο δημιουργείται και η δορυφορική εικόνα ενός φασματικού καναλιού. Η σάρωση πραγματοποιείται ηλεκτρονικά ή μηχανικά.

Φασματόμετρο: είναι ένα όργανο που ανιχνεύει, μετρά και αναλύει το φασματικό περιεχόμενο της προσπίπτουσας Η/Μ ακτινοβολίας. Τα συνήθη απεικονιστικά φασματόμετρα χρησιμοποιούν σχάρες ή πρίσματα για τον διασπορά της ακτινοβολίας στον φασματικό διαχωρισμό.

Τα παραπάνω αφορούν τους παθητικούς δορυφόρους, όπως και το Sentinel – 2 . Με χρήση πολυφασματικών εικόνων οι οποίες προκύπτουν από πολυφασματικούς σαρωτές μπορεί να γίνει ψηφιακή επεξεργασία και να προκύψουν έγχρωμα σύνθετα, σαν αυτά που δημιουργεί μια φωτογραφική μηχανή, αλλά σε μεγάλο εύρος του ηλεκτρομαγνητικού φάσματος.

2.1.1 Χαρακτηριστικά τηλεπισκοπικών δεδομένων

Η ποιότητα των δορυφορικών εικόνων εξαρτάται από τα χαρακτηριστικά των φασματόμετρων που διαθέτουν οι αντίστοιχοι δορυφόροι. Τέσσερις είναι οι βασικές ιδιότητες που χαρακτηρίζουν μια δορυφορική εικόνα οι οποίες συνθέτουν με ένα όνομα την διακριτική ικανότητα του αισθητήρα κάποιου δορυφόρου :

- Η χωρική διακριτική ικανότητα αναφέρεται στο ελάχιστο μέγεθος της γήινης επιφάνειας που μπορεί να διακρίνει ο συγκεκριμένος αισθητήρας, δηλαδή στην γήινη έκταση που αντιπροσωπεύει ένα εικονοστοιχείο της εικόνας του συγκεκριμένου αισθητήρα (του συγκεκριμένου δορυφόρου). Εικονοστοιχείο ονομάζουμε το στοιχειώδες κομμάτι εικόνας, συνήθως στην οθόνη του υπολογιστή, (αλλά και στις ψηφιακές εικόνες γενικότερα) και που αντιπροσωπεύει ένα συγκεκριμένο μέγεθος της γήινης επιφάνειας.
- Η φασματική διακριτική ικανότητα αναφέρεται στον αριθμό των μηκών κύματος που μπορεί να καταγράψει την επιφάνεια της γης ο συγκεκριμένος αισθητήρας συνεπώς ουσιαστικά αναφέρεται στον αριθμό των φασματικών καναλιών από τα οποία αποτελείται η συγκεκριμένη δορυφορική εικόνα.
- Η ραδιομετρική διακριτική ικανότητα αναφέρεται στον αριθμό των διαφορετικών διαβαθμίσεων της έντασης της καταγεγραμμένης ακτινοβολίας που μπορεί να καταγράψει ο συγκεκριμένος αισθητήρας (του συγκεκριμένου δορυφόρου). Ουσιαστικά πρόκειται για τον αριθμό των bits που περιέχει η τιμή του εικονοστοιχείου στο δυαδικό σύστημα και που αντιστοιχούν σε συγκεκριμένο ακέραιο αριθμό στο δεκαδικό σύστημα. Π.χ. οι εικόνες του δορυφόρου Sentinel – 2 είναι καταγεγραμμένες σε οκτώ (16) bits = 2 bytes δηλαδή κάθε εικονοστοιχείο μιας εικόνας του δορυφόρου Sentinel – 2 παίρνει 65535 ακέραιες τιμές (από το 0 έως και το 65535) και άρα μπορεί κατ' αντιστοιχία να φωτιστεί με 65536 διαφορετικές εντάσεις κάποιου χρώματος (ή της ασπρόμαυρης κλίμακας).
- Η χρονική διακριτική ικανότητα είναι χρόνος που διέρχεται ο δορυφόρος από την ίδια περιοχή της γήινης επιφάνειας.

2.1.2 Ατμοσφαιρικές και ραδιομετρικές διορθώσεις

Τα ψηφιακά δεδομένα που καταγράφονται από αισθητήρες είτε αερομεταφερόμενους είτε από δορυφόρους ενδέχεται να εμπεριέχουν σφάλματα τόσο στην γεωμετρία τους όσο και στις τιμές φωτεινότητας (ραδιομετρικές τιμές των εικονοστοιχείων τους (pixels)). Στη δεύτερη περίπτωση οι αλλοιώσεις των τιμών των pixels αναφέρονται ως ραδιομετρικά σφάλματα και μπορεί να οφείλονται τόσο στους αισθητήρες καταγραφής όσο και στις επιδράσεις της ατμόσφαιρας. Δύο είναι οι βασικές κατηγορίες ραδιομετρικών σφαλμάτων: (α) Τα εσωτερικά σφάλματα, που προκύπτουν από το ίδιο το σύστημα καταγραφής. Πρόκειται για συστηματικά σφάλματα των αισθητήρων, τα οποία είναι δυνατόν εφόσον εντοπιστούν να διορθωθούν, όπως για παράδειγμα το πρόβλημα απορρύθμισης ανιχνευτών και (β) τα εξωτερικά σφάλματα, που προέρχονται από διαφορετικούς εξωγενείς παράγοντες, όπως η ατμόσφαιρα, το υψόμετρο, οι κλίσεις ή ο προσανατολισμός του εδάφους. Ορισμένα εξωτερικά σφάλματα μπορούν να διορθωθούν με τη βοήθεια επιτόπιων μετρήσεων καθώς και με την εφαρμογή ατμοσφαιρικών μοντέλων. Η διόρθωση των ραδιομετρικών σφαλμάτων αποτελεί μια πολύπλοκη αλλά αναγκαία διαδικασία προ-επεξεργασίας των δορυφορικών δεδομένων, για την βελτίωση της ποιότητας των εικόνων και την ασφαλή εξαγωγή ποιοτικών και ποσοτικών συμπερασμάτων. Στη χρονοσειρά που χρησιμοποιήθηκε στο Google Earth Engine οι εικόνες είναι προ- διορθωμένες τόσο ατμοσφαιρικά όσο και ραδιομετρικά .

Ραδιομετρικά σφάλματα και διορθώσεις

Τα ψηφιακά δεδομένα που καταγράφονται από αισθητήρες ή αερομεταφερομένους ή δορυφόρων είναι δυνατό να εμπεριέχουν σφάλματα τόσο στη γεωμετρία τους όσο και στις τιμές ραδιομετρικές τιμές των εικονοστοιχείων τους. Στη δεύτερη περίπτωση οι αλλοιώσεις των τιμών των pixels αναφέρονται ως ραδιομετρικά σφάλματα και μπορεί να οφείλονται τόσο στους αισθητήρες καταγραφής όσο και στις επιδράσεις της ατμόσφαιρας. Γενικά, δύο είναι οι βασικές κατηγορίες ραδιομετρικών σφαλμάτων (Jensen 1996):

- Τα εσωτερικά σφάλματα, που προκύπτουν από το ίδιο το σύστημα καταγραφής. Πρόκειται για συστηματικά σφάλματα των αισθητήρων, τα οποία αν εντοπιστούν, μπορούν να διορθωθούν. Χαρακτηριστικό παράδειγμα είναι το πρόβλημα απορρύθμισης των ανιχνευτών.
- Τα εξωτερικά σφάλματα, που προκύπτουν από διαφορετικούς εξωτερικούς παράγοντες, όπως η ατμόσφαιρα, το υψόμετρο, οι κλίσεις, ο προσανατολισμός του εδάφους. Ορισμένα εξωτερικά σφάλματα μπορούν να διορθωθούν με τη βοήθεια επιτόπιων μετρήσεων καθώς και με την εφαρμογή ατμοσφαιρικών μοντέλων.

Επίσης, όταν τα ραδιομετρικά σφάλματα των εικόνων εξετάζονται υπό το πρίσμα εντοπισμού διαχρονικών αλλαγών, μπορούν να ταξινομηθούν στις ακόλουθες κατηγορίες (Paolini et al. 2006) σύμφωνα με τις:

- Διαφορές στις ραδιομετρικές αποκρίσεις μεταξύ των δορυφόρων.
- Αλλαγές στην βαθμονόμηση των αισθητήρων των δορυφόρων.
- Διαφορές στη γωνία παρατήρησης και στη φωτεινότητα.
- Ατμοσφαιρικές μεταβολές και επιδράσεις του ανάγλυφου.

Η συγκεκριμένη κατηγοριοποίηση αφορά την επεξεργασία δύο ή περισσότερων πολυφασματικών εικόνων, σε διαφορετικές ημερομηνίες, που χρησιμοποιούνται για την παρατήρηση διαχρονικών αλλαγών στη γήινη επιφάνεια. Οι ραδιομετρικές διαφοροποιήσεις των εικόνων εξαιτίας των ανωτέρω σφαλμάτων μπορεί να οδηγήσει σε εσφαλμένα συμπεράσματα της ποιοτικής και της ποσοτικής εκτίμησης των μεταβολών των καλύψεων γης.

Από τα παραπάνω γίνεται φανερό ότι η διόρθωση των ραδιομετρικών σφαλμάτων αποτελεί μια πολύπλοκη αλλά αναγκαία διαδικασία προεπεξεργασίας των δεδομένων για την βελτίωση της ποιότητας των εικόνων και την ασφαλή εξαγωγή ποιοτικών και ποσοτικών συμπερασμάτων. Η αναγκαιότητα η μη των ραδιομετρικών διορθώσεων εξαρτάται τόσο από τη φύση των δεδομένων όσο και από το είδος των πληροφοριών που πρέπει να εξαχθούν.

Ο στόχος των ραδιομετρικών διορθώσεων των εικόνων είναι η αφαίρεση ή η μείωση των ελαττωματικών λειτουργιών (θόρυβος, κλπ) των αισθητήρων καθώς και των ατμοσφαιρικών και τοπογραφικών επιδράσεων, ώστε να καταγράφονται στην δορυφορική εικόνα οι πραγματικές ιδιότητες της γήινης επιφάνειας. Υπάρχουν διαφορετικές προσεγγίσεις σχετικά με το ερώτημα εφαρμογής ραδιομετρικών βαθμονομήσεων και ατμοσφαιρικών διορθώσεων σε δεδομένα Τηλεπισκόπησης για την εξαγωγή γεωφυσικών πληροφοριών, όπως (Cracknell et al. 1993):

- Να αγνοηθούν πλήρως οι ατμοσφαιρικές επιδράσεις.
- Να γίνει βαθμονόμηση των εικόνων με χρήση επιτόπιων μετρήσεων.
- Να γίνει χρήση ατμοσφαιρικών μοντέλων με παραμέτρους που προέρχονται από ιστορικά στοιχεία.
- Να γίνει χρήση ατμοσφαιρικών μοντέλων με παραμέτρους που προέρχονται από σύγχρονα μετεωρολογικά δεδομένα. Η πολυπλοκότητα της ατμοσφαιρικής διόρθωσης εγείρει το ερώτημα σε ποιες περιπτώσεις είναι αναγκαία και σε ποιες όχι. Γενικά δεν υπάρχει ανάγκη ραδιομετρικής διόρθωσης όταν γίνεται ανάλυση και ταξινόμηση της εικόνας σε μια ημερομηνία. Αυτό ισχύει υπό την προϋπόθεση τα δειγματοληπτικά δεδομένα να εξάγονται από την ίδια την εικόνα και όχι από άλλη με διαφορετική ημερομηνία λήψης. Αντίθετα, η ραδιομετρική διόρθωση επιβάλλεται σε περιπτώσεις όπως:

- Σύγκριση εικόνων για εντοπισμό αλλαγών, με διαφορετικές ημερομηνίες λήψης (διαχρονική ανάλυση).
- Σύγκριση δορυφορικών δεδομένων ανάκλασης της γήινης επιφάνειας από διαφορετικούς δέκτες με παρόμοια φασματικά χαρακτηριστικά (Landsat TM κανάλι 3 με SPOT κανάλι 2).
- Σύγκριση και επαλήθευση των τιμών ανάκλασης της γήινης επιφάνειας από δορυφορικά δεδομένα με τις αντίστοιχες επιτόπιες μετρήσεις ανάκλασης στο έδαφος.
- Εξαγωγή ποσοτικών πληροφοριών (τιμές βιομάζας – συγκεντρώσεις ουσιών στο νερό κ.α.). Η επιλογή της μεθόδου διόρθωσης εξαρτάται:
 - από το πρόβλημα που πρέπει να επιλυθεί,
 - από το είδος των δορυφορικών δεδομένων,
 - από τις ατμοσφαιρικές πληροφορίες και τέλος
 - από τη διαθεσιμότητα σε εξειδικευμένο προσωπικό.

Γενικά, οι ραδιομετρικές διορθώσεις ταξινομούνται σε τρία βήματα:

- Διορθώσεις για την ελαχιστοποίηση των σφαλμάτων των αισθητήρων – θέσης δορυφόρου.
- Διορθώσεις για την ελαχιστοποίηση των επιδράσεων της ατμόσφαιρας (ατμοσφαιρικές διορθώσεις).
- Τοπογραφικές διορθώσεις. Ένα απλοποιημένο διάγραμμα των ραδιομετρικών διορθώσεων απεικονίζεται στην εικόνα που ακολουθεί.

Ατμοσφαιρικά σφάλματα και διορθώσεις

Οι δορυφορικές εικόνες στο ορατό και στο υπέρυθρο μήκος κύματος του ΗΜΦ επηρεάζονται από τα σωματίδια και τα αέρια της ατμόσφαιρας μέσω των διεργασιών της σκέδασης και της απορρόφησης. Ένα καλό μοντέλο διόρθωσης πρέπει να μπορεί να προσομοιάζει τις διεργασίες που υφίσταται η ΗΜΑ στην ατμόσφαιρα και να διορθώνει τις επιπτώσεις στις εικόνες όπως η «θολούρα», η αλλαγή στις ραδιομετρικές τιμές κλπ. Αρκετές μέθοδοι ατμοσφαιρικών διορθώσεων έχουν αναπτυχθεί και παρουσιαστεί στη διεθνή βιβλιογραφία. Γενικά ταξινομούνται σε 2 κατηγορίες (Lu et al. 2002):

- Απόλυτες Ατμοσφαιρικές διορθώσεις.
- Σχετικές Ατμοσφαιρικές διορθώσεις.

Οι απόλυτες διορθώσεις των δορυφορικών εικόνων της συγκεκριμένης κατηγορίας οδηγούν σε τιμές ανάκλασης και θερμοκρασίας της επιφάνειας της γης. Βάσει των χαρακτηριστικών και της πολυπλοκότητάς τους υποδιαιρούνται σε 2 υποκατηγορίες:

A) Μοντέλα βασιζόμενα στα χαρακτηριστικά της εικόνας (image-based atmospheric corrections). Η συγκεκριμένη κατηγορία ατμοσφαιρικών διορθώσεων βασίζεται στην εξέταση καλύψεων γης που έχουν γνωστές τιμές ανάκλασης στις πολυφασματικές εικόνες. Οι γνωστές τιμές ανάκλασης των αντικειμένων στην επιφάνεια της γης και η σύγκρισή τους με τις αντίστοιχες τιμές των pixels των εικόνων μπορεί να συμβάλει στην εκτίμηση των ατμοσφαιρικών επιδράσεων που υφίστανται. Όμως στην πράξη δεν είναι πάντα εφικτό να υπάρχουν μετρήσεις της ανάκλασης αντικειμένων που εμφανίζονται στις εικόνες. Γι' αυτό έχουν αναπτυχθεί μέθοδοι όπως το μοντέλο φαινόμενης ανάκλασης (apparent reflectance model) σύμφωνα με το οποίο μετατρέπεται η TOA ανάκλαση (at-satellite reflectance) σε ανάκλαση στην επιφάνεια (surface reflectance). Σ' αυτή την κατηγορία ανήκουν μοντέλα όπως:

- Η Αφαίρεση του Μελανού Σώματος (Dark Object Subtraction – DOS) .
- Πίνακας συνδιακύμανσης εκτιμώμενων παραμέτρων (covariance matrix) .
- Παλινδρόμηση (Regression analysis) Το μοντέλο DOS .

Β) Αριθμητικά Μοντέλα. Τα αριθμητικά μοντέλα είναι πιο πολύπλοκα αλλά παρέχουν και μεγαλύτερη ακρίβεια στη μετατροπή των ψηφιακών τιμών (ή των τιμών TOA) των εικόνων σε τιμές ανάκλασης στην επιφάνεια της γης. Περιγράφουν με μαθηματικά μοντέλα τη συμπεριφορά της ακτινοβολίας κατά την είσοδό της στην ατμόσφαιρα και την αλληλεπίδρασή της με τα μόρια και τα σωματίδια του αέρα (μετάδοση, σκέδαση, απορρόφηση). Τα αριθμητικά μοντέλα έχουν βασικά πλεονεκτήματα όπως η ακρίβεια και το μεγάλο εύρος εφαρμογής τους. Όμως, έχουν και σημαντικά μειονεκτήματα όπως η ανάγκη για συλλογή λεπτομερών ατμοσφαιρικών πληροφοριών σχετικά με την ατμοσφαιρική υγρασία και τις συγκεντρώσεις των σωματιδίων του αέρα. Τέτοιου είδους δεδομένα είναι δύσκολο να αποκτηθούν με μεγάλη λεπτομέρεια. Επιπλέον, οι ατμοσφαιρικές συνθήκες διαφέρουν ανάλογα με το ύψος από την επιφάνεια της γης. Η συλλογή τέτοιων δεδομένων, σε αρκετές περιπτώσεις, είναι αδύνατη με ίσως μοναδική δυνατότητα να μπορούν να υλοποιηθούν μερικές σημειακές μετρήσεις που αφορούν το σύνολο μιας δορυφορικής πολυφασματικής εικόνας (π.χ. Landsat TM). Σ' αυτό το πλαίσιο έχει αναπτυχθεί ένα σύνολο μοντέλων υπό το όνομα «Διόρθωση βάση του μοντέλου διάδοσης της ακτινοβολίας» (Radiative Transfer Model based correction). Τα μοντέλα αυτά υπολογίζουν την αλληλεπίδραση της ακτινοβολίας με τα σωματίδια της ατμόσφαιρας. Για την εφαρμογή τους απαιτείται ύπαρξη πληροφοριών όπως συγκεντρώσεις αερολυμάτων/σωματιδίων καθώς και η ατμοσφαιρική υγρασία. Σε περίπτωση απουσίας τέτοιων πληροφοριών χρησιμοποιούν τυποποιημένα ατμοσφαιρικά προφίλ που καλύπτουν μεγάλο εύρος ατμοσφαιρικών χαρακτηριστικών. Μεταξύ των πιο ευρέως χρησιμοποιούμενων αριθμητικών μοντέλων είναι τα ακόλουθα:

- LOWTRAN (LOW resolution atmospheric TRANsmittance).
- MODTRAN (MODerate resolution atmospheric TRANsmittance).
- ATCOR (Atmospheric and Topographic Correction model).
- FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes).
- 6S (Second Simulation of the Satellite Signal in the Solar Spectrum).

2.2 Μηχανική μάθηση

Η μηχανική μάθηση είναι ένα υποπεδίο της επιστήμης των υπολογιστών που αναπτύχθηκε από τη μελέτη της αναγνώρισης προτύπων και της υπολογιστικής θεωρίας μάθησης στην τεχνητή νοημοσύνη. Η μηχανική μάθηση θα μπορούσε να οριστεί ως ένα πεδίο μελέτης που δίνει την ικανότητα στους υπολογιστές να μαθαίνουν χωρίς να έχουν ρητά προγραμματιστεί (Arthur Samuel , 1959). Εν γένει είναι η μελέτη και η κατασκευή αλγορίθμων που εκπαιδεύονται με ένα μεγάλο όγκο δεδομένων και κάνουν εκτιμήσεις και προβλέψεις σχετικά με αυτά.

Στο πεδίο της ανάλυσης δεδομένων, η μηχανική μάθηση είναι μια μέθοδος που χρησιμοποιείται για την επινόηση πολύπλοκων μοντέλων και αλγορίθμων που οδηγούν στην πρόβλεψη. Τα αναλυτικά μοντέλα επιτρέπουν στους ερευνητές , τους επιστήμονες δεδομένων, στους μηχανικούς και τους αναλυτές να παράγουν αξιόπιστες αποφάσεις και αποτελέσματα και να αναδείξουν αλληλοσυσχετίσεις μέσω της μάθησης από ιστορικές σχέσεις και τάσεις στα δεδομένα.

Ένας ακόμα καλός και κατατοπιστικός ορισμός της μηχανικής μάθησης είναι ο παρακάτω :

«Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία E ως προς κάποια κλάση εργασιών T και μέτρο απόδοσης P, αν η απόδοσή του σε εργασίες από το T, όπως μετρείται από το P, βελτιώνεται μέσω της εμπειρίας E.» Mitchell , 1997

Εν γένει, ο τομέας της Μηχανικής Μάθησης αναπτύσσει τρεις τρόπους μάθησης, ανάλογους με τους τρόπους με τους οποίους μαθαίνει ο άνθρωπος: επιβλεπόμενη μάθηση, μη επιβλεπόμενη μάθηση και ενισχυτική μάθηση.

Πιο αναλυτικά:

- Επιβλεπόμενη Μάθηση (Supervised Learning) είναι η διαδικασία όπου ο αλγόριθμος κατασκευάζει μια συνάρτηση που απεικονίζει δεδομένες εισόδους (σύνολο εκπαίδευσης) σε γνωστές επιθυμητές εξόδους, με απώτερο στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο. Χρησιμοποιείται σε προβλήματα:
 - Ταξινόμησης (Classification)
 - Πρόγνωσης (Prediction)
 - Διερμηνείας (Interpretation)
- Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning), όπου ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων υπό μορφή παρατηρήσεων χωρίς να γνωρίζει τις επιθυμητές εξόδους. Χρησιμοποιείται σε προβλήματα:
 - Ανάλυσης Συσχετισμών (Association Analysis)
 - Ομαδοποίησης (Clustering)
- Ενισχυτική Μάθηση (Reinforcement Learning), όπου ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών μέσα από άμεση αλληλεπίδραση με το περιβάλλον. Χρησιμοποιείται κυρίως σε προβλήματα Σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ και η βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους.

Για κάθε πρόβλημα προς επίλυση στο χώρο της Μηχανικής Μάθησης υπάρχει ένας κατάλληλος τρόπος μάθησης και για κάθε τρόπο μάθησης υπάρχει τουλάχιστον ένας κατάλληλος αλγόριθμος που μπορεί να χρησιμοποιηθεί. Για τις ανάγκες της εργασίας αυτής χρησιμοποιήθηκαν επιβλεπόμενη μάθηση για πρόβλημα ταξινόμησης.

2.2.1 Δένδρα απόφασης

Τα Δένδρα Απόφασης-ΔΑ (Decision Trees) είναι ο γνωστότερος αλγόριθμος επιβλεπόμενης Επαγωγικής Μάθησης και έχει εφαρμοστεί με επιτυχία σε πολλούς τομείς όπου απαιτείται ταξινόμηση: ενδεικτικά, στην αναγνώριση προσώπων σε εικόνες, στην ιατρική για διάγνωση περιστατικών, για προβλέψεις απαραίτητες στη διαφήμιση, για προώθηση προϊόντων και, γενικότερα, για εξόρυξη γνώσης. Ο αλγόριθμος ΔΑ οδηγεί στη δημιουργία μιας δενδροειδούς μορφής που τα φύλλα της αποτελούν κατηγορίες ταξινόμησης (classes). Η δενδροειδής αυτή μορφή μπορεί να αναγνωστεί και ως ένα σύνολο κανόνων που καλούνται κανόνες ταξινόμησης (classification rules).

Βασικές προϋποθέσεις για τη λειτουργία ενός αλγόριθμου επαγωγικής μάθησης είναι:

- καθορισμός ενός συνόλου χαρακτηριστικών (features set - FS) ως των προϋποθέσεων του επιδιωκόμενου προς εξαγωγή κανόνα ταξινόμησης:

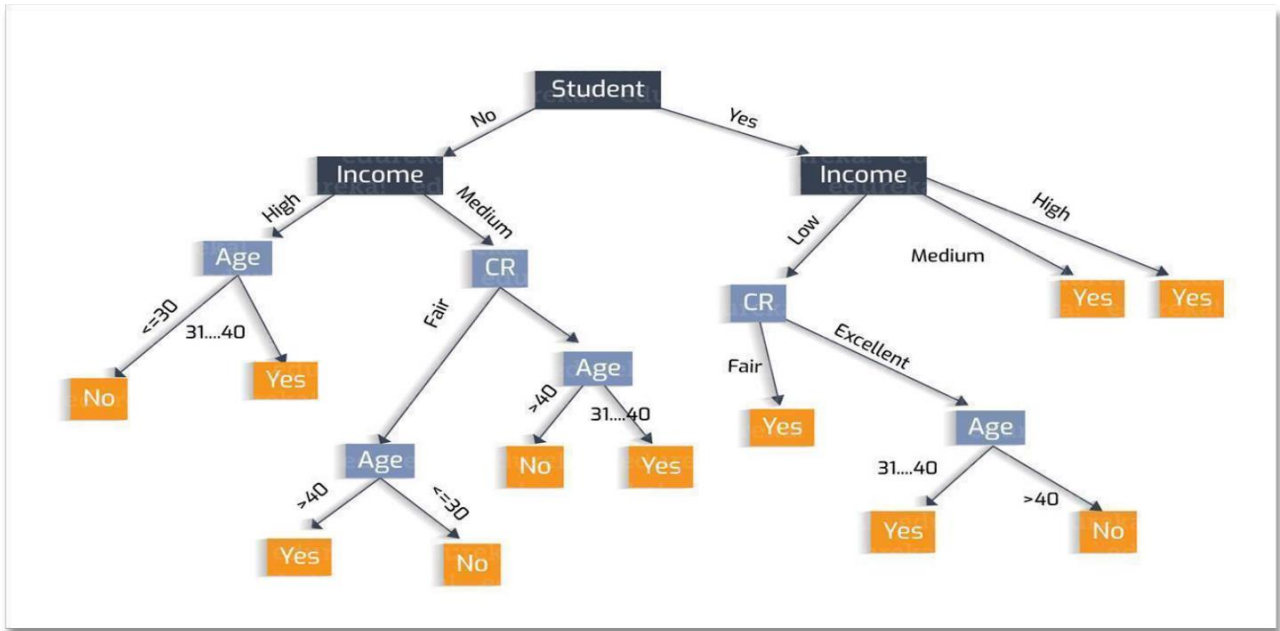
$$FS = \{F_1, F_2, F_3, \dots, F_{|FS|}\}$$

- ύπαρξη προκαθορισμένων διακριτών κατηγοριών ταξινόμησης (classes – C) ως στόχου του διαχωρισμού τον οποίο θα επιδιώξει ο αλγόριθμος και, στη συνέχεια, ως συμπερασμάτων (conclusions) των κανόνων στους οποίους θα οδηγήσει η αναγνώριση της δενδροειδούς μορφής που θα αναπτύξει ο αλγόριθμος:

$$C = \{c_1, c_2, c_3, \dots, c_{|c|}\}$$

- ύπαρξη επαρκούς αριθμού δειγμάτων που θα προκύψουν από παρατηρήσεις και θα χρησιμοποιηθούν για τη δημιουργία του εκπαιδευτικού συνόλου (training set - TS).

Παρακάτω παρουσιάζεται ενδεικτικά σε μια εικόνα ένα δέντρο απόφασης στην εικόνα 2.1 .



Εικόνα 2.1 . Δένδρο απόφασης

Ο αλγόριθμος τυχαία δάση

Ο ταξινομητής Random Forest (RF) , είναι ένας από τους πιο γνωστούς ομοιογενείς, συνδυαστικούς ταξινομητές, ο οποίος χρησιμοποιεί πολλά, ασυσχέτιστα μεταξύ τους δέντρα αποφάσεων για να πραγματοποιήσει μία εκτίμηση (Belgiu M. and Dragut L. (2016)). Η βασική ιδέα πίσω από το μοντέλο Random Forest, είναι η μείωση της συσχέτισης μεταξύ των ταξινομητών που το απαρτίζουν και του φαινομένου overfitting που προκαλεί την προσαρμογή του μοντέλου στον θόρυβο που υπάρχει στα δεδομένα.

Κατασκευή του αλγόριθμου

Η ανεξάρτητη ανάπτυξη των δέντρων που απαρτίζουν το συνδυαστικό μοντέλο, βασίζεται στην επιλογή τυχαίων υποσυνόλων δεδομένων εκπαίδευσης (Bagging), με αντικατάσταση. Αυτό σημαίνει ότι μέρος του συνόλου των δειγμάτων που χρησιμοποιούνται για την εκπαίδευση ενός δέντρου, μπορεί να επιλεγεί περισσότερες από μία φορές για την εκπαίδευση και άλλων δέντρων μέσα στο σύμπλεγμα, ενώ άλλα δείγματα ενδέχεται να μην επιλεγούν καθόλου. Σε κάθε περίπτωση, τα υποσύνολα εκπαίδευσης που δημιουργούνται είναι διαφορετικά μεταξύ τους και κατά συνέπεια στατιστικά ανεξάρτητα (Rokach L. (2010)) Η δεύτερη πηγή τυχαιότητας στο μοντέλο Random Forest, αφορά την τυχαία επιλογή των μεταβλητών (καναλιών), βάση των οποίων θα τεθούν τα φασματικά όρια διαχωρισμού σε κάθε εσωτερικό κόμβο (internal splitting node). Το πλήθος των καναλιών που χρησιμοποιούνται σε κάθε δέντρο απόφασης, καθορίζεται από το σύνολο των μεταβλητών (V) που

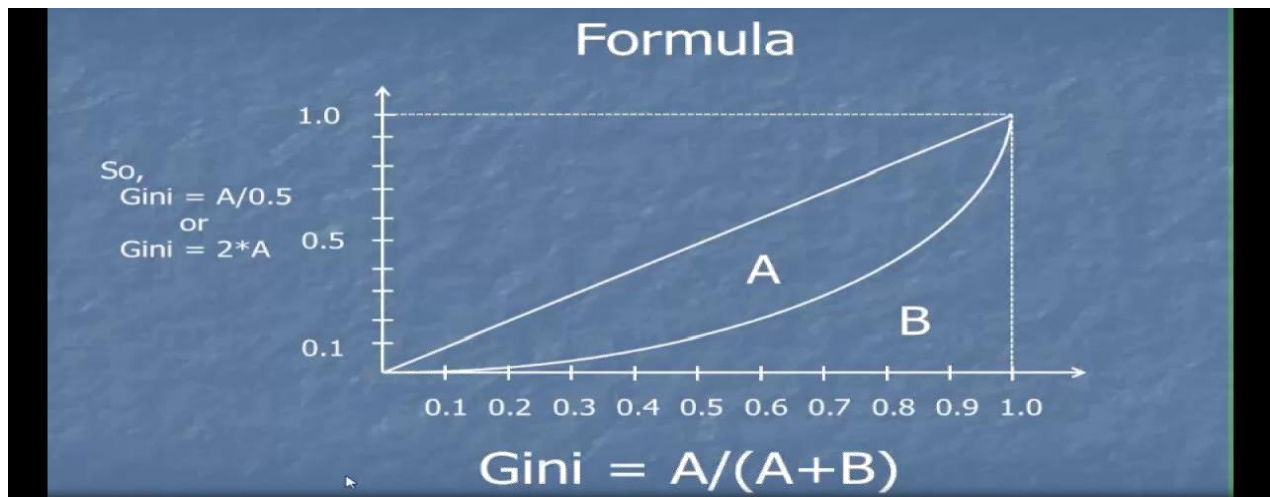
χρησιμοποιούνται στην ταξινόμηση και συνηθίζεται να είναι $mtry = \text{sqrt}(V)$. Η επιλογή ενός υποσυνόλου εκπαίδευσης, επιδρά θετικά στη μείωση του υπολογιστικού χρόνου ενώ η επιλογή μερικώς ή εξ' ολοκλήρου διαφορετικού υποσυνόλου δεδομένων εκπαίδευσης για κάθε δέντρο, βοηθά στη δημιουργία διαφορετικών εκτιμητών, ικανών να προσαρμοστούν σε διαφορετικά είδη δεδομένων, για να βελτιωθεί έτσι η απόδοση του αλγορίθμου (Criminisi et al., 2011). Ακόμα, σύμφωνα με τους Hastie T. et al, (2009), η επιλογή κάθε φορά $mtry$ τυχαίων μεταβλητών σε κάθε εσωτερικό κόμβο διαχωρισμού (internal splitting node), βοηθά επίσης στη μείωση της συσχέτισης (ρ) μεταξύ κάθε πιθανού ζεύγους δέντρων και κατ' επέκταση μείωση της μεγάλης αστάθειας (ύπαρξη 31 τυχαίων σφαλμάτων) που παρουσιάζουν τα δενδροειδές μοντέλα από τα οποία απαρτίζεται το δάσος.

$$\bar{\sigma}^2 = \rho * \sigma^2 + (((1-\rho)/B) * \sigma^2$$

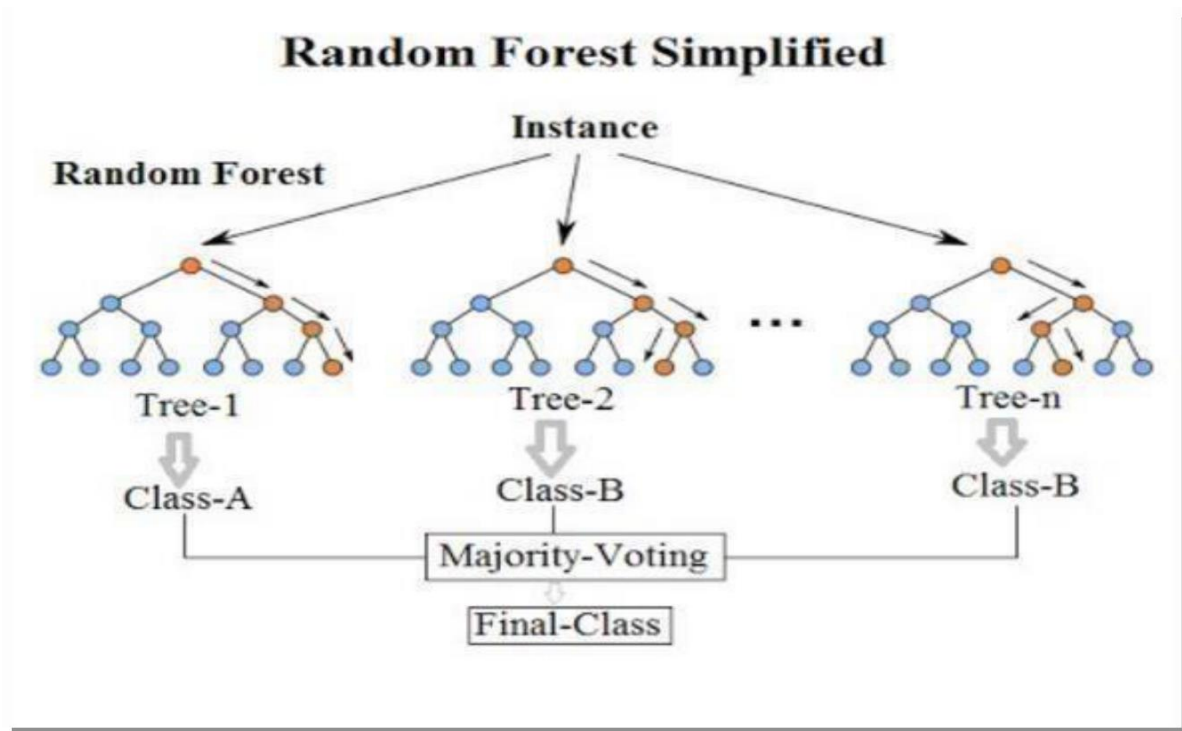
Εξίσωση μέσης διακύμανσης του αλγορίθμου τυχαία δάση

Στην παραπάνω εξίσωση βλέπουμε τη μέση διακύμανση που υπάρχει στο δάσος ($\bar{\sigma}^2$) και μπορούμε πολύ εύκολα να συμπεράνουμε ότι αν διατηρήσουμε όλους τους όρους σταθερούς και αυξήσουμε το πλήθος των δέντρων (B) που απαρτίζουν το δάσος, τότε ο δεύτερος όρος τείνει στο 0 και μας μένει ο 1ος όρος. Η μείωση της συσχέτισης (ρ) μεταξύ κάθε πιθανού ζεύγους δέντρων, επιτυγχάνεται μέσω της επιλογής τυχαίου πλήθους m μεταβλητών για κάθε εσωτερικό κόμβο διαχωρισμού.

Αφού επιλεγούν τα κανάλια και τα υποσύνολα των δεδομένων εκπαίδευσης που θα χρησιμοποιηθούν για την ανάπτυξη του κάθε δέντρου, σε κάθε εσωτερικό κόμβο επιλέγεται το σημείο διαχωρισμού μεταξύ των υποψήφιων φασματικών κλάσεων. Τα κριτήρια που θα καθορίσουν τη διαδρομή αυτού του εικονοστοιχείου μέχρι έναν τερματικό κόμβο (φύλλο), καθορίζονται από τα όρια διαχωρισμού. Αυτά προσδιορίζονται σε κάθε εσωτερικό κόμβο μέσα από μια επαναληπτική διαδικασία, κατά την οποία διάφορα όρια εξετάζονται και αξιολογούνται μέσω του υπολογισμού του δείκτη gini (εικόνα 2.2). Επιλέγεται το όριο με την χαμηλότερη τιμή του δείκτη, αφού με αυτό τον τρόπο επιτυγχάνεται η υπολογιστική ταχύτητα του αλγορίθμου, χωρίς να σπαταλιέται χρόνος που δε θα οδηγήσει το συνηθισμένο σε μία απόφαση.



Εικόνα 2.2 . Δείκτης gini



Εικόνα 2.3 Παράδειγμα ταξινόμησης τυχαίου εικονοστοιχείου άγνωστης τάξης με τον αλγόριθμο τυχαίων δασών

Εκτός από τα δεδομένα εκπαίδευσης και τον αριθμό των μεταβλητών που επιλέγεται τυχαία για την εκπαίδευση του κάθε δέντρου, ακόμα ένας παράγοντας πρέπει να καθορισθεί για την κατασκευή του αλγορίθμου και αυτός είναι ο αριθμός των δέντρων απόφασης που απαρτίζουν το δάσος. Θεωρητικά, όσο το πλήθος των δέντρων αυξάνεται, τόσο πιο ομαλά είναι τα όρια διαχωρισμού και κατά συνέπεια καλύτερη η απόδοση του ταξινομητή. Εμπειρικά, όπως αναφέρουν οι Belgiu M. and Dragut L. (2016) σε έρευνες των Ghosh et al., (2014) και Kulkarni and Sinha, (2012) έχουν δείξει ότι η ακρίβεια της ταξινόμησης όσο αφορά την παράμετρο του πλήθους των δέντρων, δεν είναι τόσο ευαίσθητη όσο στην παράμετρο Unknown Class Pixel. Παράδειγμα ταξινόμησης άγνωστης τάξης εικονοστοιχείου με τον αλγόριθμο Τυχαίων δασών (Criminisi et al. (2011)) 33 του πλήθους των μεταβλητών που θα καθορίσουν τα όρια διαχωρισμού σε κάθε εσωτερικό κόμβο. Στη μελέτη αυτή η παράμετρος τίθεται στα 100.

Ευαισθησία του ταξινομητή στα δείγματα εκπαίδευσης

Οι δειγματοληπτικές περιοχές που χρησιμοποιούνται για την εκπαίδευση των επιβλεπόμενων ταξινομητών, πρέπει να ικανοποιούν συγκεκριμένες απαιτήσεις. Μεταξύ αυτών, πρέπει να είναι αντιπροσωπευτικά του φαινομένου που απεικονίζουν, να είναι φασματικά πλήρη και αξιόπιστα. Επιπλέον, πρέπει να συντελούν στη δημιουργία φασματικά ομοιογενών κλάσεων, οι οποίες μεταξύ τους να είναι ανομοιογενής.

Εντούτοις, κάτι που παραλείφθηκε, είναι η κατάσταση ισορροπίας που υπάρχει μεταξύ των κλάσεων που απαρτίζουν το σύνολο δεδομένων. Αυστηρά μιλώντας, σχεδόν όλα τα σύνολα δεδομένων που υπάρχουν στον πραγματικό κόσμο είναι ανισόρροπα. Σε συνδυασμό με το γεγονός ότι η συλλογή τους είναι μια επίπονη, χρονοβόρα και δαπανηρή διαδικασία, οδηγούμαστε πολλές φορές στη δημιουργία ενός συνόλου δεδομένων εκπαίδευσης το οποίο αποτελείται από ομάδες που δεν αντιπροσωπεύονται εξίσου στο σύνολο των δεδομένων εκπαίδευσης που τις απαρτίζουν.

Η πρόκληση προέρχεται από το γεγονός ότι τα σπανιότερα εμφανιζόμενα δείγματα που αφορούν μια τυχαία κλάση A, συνήθως “επικαλύπτονται” από τα δείγματα της τυχαίας κλάσης B (κλάση πλειοψηφίας) και είναι κατά συνέπεια πολύ πιο δύσκολο να ταυτοποιηθούν κατά τη διαδικασία της ταξινόμησης. Οι αλγόριθμοι ταξινόμησης, συνήθως στοχεύουν στην επίτευξη υψηλής ολικής ακρίβειας, γεγονός που δημιουργεί μια εγγενή μεροληψία υπέρ των κλάσεων πλειοψηφίας, αφού οι κλάσεις μειονότητας επηρεάζουν λιγότερο την ακρίβεια της ταξινόμησης (Yuan B. And Liu W. (2012)). Έτσι οι κλάσεις με τα μικρότερα πλήθη εικονοστοιχείων, καταλήγουν με υψηλότερα ποσοστά σφάλματος. Επομένως η ανάλυση της ευαισθησίας του ταξινομητή ως προς την κατανομή των δειγμάτων εκπαίδευσης είναι απαραίτητη.

Στη βιβλιογραφία προτείνονται διάφορες τεχνικές για την τροποποίηση των κατανομών των κλάσεων. Μεταξύ αυτών η μείωση του πλήθους των παρατηρήσεων για τις κλάσεις που υπερέχουν αριθμητικά τις κλάσεις μειονότητας (Under-sampling method), η οποία είναι και η βέλτιστη στην περίπτωση που το προς ταξινόμηση σύνολο δεδομένων είναι τεράστιο (π.χ. δορυφορικές εικόνες) και η μείωση του πλήθους των δεδομένων εκπαίδευσης θα βοηθήσει στην υπολογιστική απόδοση του αλγόριθμου και αποφυγή προβλημάτων αποθήκευσης των δεδομένων. Παρ’ όλα αυτά, η τυχαία απόρριψη πληροφορίας από τις κλάσεις, επηρεάζει την φασματική πληρότητα στην περιγραφή κάθε θεματικής κατηγορίας (Yap B.W. et al (2013)).

Ακόμα, η μέθοδος oversampling, κατά την οποία προστίθενται επιπλέον δεδομένα εκπαίδευσης για όλες τις κλάσεις μέχρις ότου να έχουν ίσο πλήθος παρατηρήσεων με την κλάση πλειοψηφίας. Το πλεονέκτημα αυτής της μεθόδου, είναι η αποφυγή απώλειας σημαντικής πληροφορίας από τις κλάσεις, ενώ στα μειονεκτήματα το γεγονός ότι η επιπλέον πληροφορία θα προστεθεί στις κλάσεις είτε μέσω δημιουργίας επαναλήψεων των ήδη υπάρχοντων δεδομένων, κάτι που δεν θα ωφελήσει ιδιαίτερα την εκπαίδευση του αλγόριθμου, είτε μέσω συλλογής περισσότερων δεδομένων που μπορεί να μην υπάρχουν (Yap B.W. et al (2013)).

2.3 Το Ευρωπαϊκό Πρόγραμμα COPERNICUS

Με πρωτοβουλία της Επιτροπής Ευρωπαϊκών Κοινοτήτων (European Commission) και σε συνεργασία με τον Ευρωπαϊκό Οργανισμό Διαστήματος δημιουργήθηκε το Ευρωπαϊκό επιστημονικό Πρόγραμμα «Κοπέρνικος» το οποίο ουσιαστικά αποτελεί την συνέχεια του προγράμματος GMES (Global Monitoring for Environment and Security) δηλαδή του Ευρωπαϊκού Προγράμματος για την «Παγκόσμια Παρακολούθηση του Περιβάλλοντος και της Ασφάλειας».

Έχοντας ως βασικό σκοπό την μελέτη και παρακολούθηση του περιβάλλοντος της γης και πιο στοχευμένα της επιφάνειας της γης (ξηρά-θάλασσα), της ατμόσφαιρας, της αλλαγής του κλίματος, την διαχείριση έκτακτων καταστάσεων και την ασφάλεια το πρόγραμμα «Κοπέρνικος» έχει ένα ευρύ πεδίο εφαρμογών όπως παρακολούθηση κλιματικών αλλαγών, αειφόρο ανάπτυξη, προστασία του περιβάλλοντος διαχείριση αστικών περιοχών, τοπικό και περιφερειακό σχεδιασμό, γεωργία, δασολογία, ιχθυολογία, υγεία, μεταφορές, πολιτική προστασία και τουρισμό, κ.ά.

Από τεχνολογικής πλευράς, πρόκειται για ένα σύνολο συστημάτων που αποτελείται από διαφόρων τύπων συστήματα και όργανα συλλογής δεδομένων από διαφορετικές πηγές: δορυφόρους παρατήρησης της γήινης επιφάνειας, αερομεταφερόμενες πλατφόρμες εφοδιασμένες με ειδικούς αισθητήρες και καταγραφείς της γήινης επιφάνειας και τέλος σταθερούς αισθητήρες συλλογής και καταγραφής των διαφόρων παραμέτρων και μεταβλητών που αφορούν το γήινο περιβάλλον (Copernicus, Earth Observation Satellites 2015).

Αλλά υπάρχει και ένα δευτερεύον και εξίσου σημαντικό πλεονέκτημα του προγράμματος Κοπέρνικος. Χάρis στις πολλαπλές αποστολές και στην ποικιλία των δεδομένων που μπορούν να προμηθεύσουν οι δορυφόροι και οι αισθητήρες στα πλαίσια του προγράμματος, πολλές επιπλέον υπηρεσίες οι οποίες προσαρμόζονται στις εκάστοτε ιδιωτικές ή δημόσιες ανάγκες δημιουργούν νέα πλαίσια και θέσεις εργασίας σε εθνικά και πολυεθνικά επίπεδα. Πολλές μελέτες οικονομικού χαρακτήρα έχουν εστιαστεί και αποδείξει τις χρηστικές αυτές ιδιότητες των δορυφορικών δεδομένων του συγκεκριμένου (αλλά και άλλων) επιστημονικών προγραμμάτων.

2.3.1 Ο δορυφόρος Sentinel - 2

Ο δορυφόρος Sentinel-2 τέθηκε σε τροχιά στις 23/06/2015 μέσω ενός πυραύλου Vega από τον Ευρωπαϊκό Διαστημικό Σταθμό που βρίσκεται κοντά στο Κούρου στην Γαλλική Γουινέα. Μέσα στο έτος 2016 εκτοξεύθηκε ένας ακόμα δορυφόρος.

Έχει έναν υψηλής χωρικής διακριτικής ικανότητας πολυφασματικό σαρωτή με δέκα τρία φασματικά κανάλια, εύρος ανοίγματος σάρωσης 290 km, και συχνή επαναδιέλευση και καταγραφή της επιφάνειας της γης. Ουσιαστικά συνίσταται στην συνύπαρξη και στην συνέργεια δύο όμοιων δορυφόρων που καλύπτουν όλη την γήινη επιφάνεια. Κάθε ένας από τους δορυφόρους αυτούς ζυγίζει 1140 Kg.

Έχουν σχεδιαστεί για 7 έτη και προοπτικές για ακόμη μεγαλύτερη διάρκεια ύπαρξης. Πετάνε σε ύψος 786 km από την επιφάνεια της γης και έχουν γωνιακή απόκλιση 180^ο μεταξύ τους. Επειδή υπάρχει επικάλυψη των δορυφορικών δεδομένων τους με τον δορυφόρο Landsat 8, η NASA και η ESA σε συνεννόηση οδήγησαν τους δύο δορυφόρους σε επαναδιέλευση 3 ημερών (από τον Ισημερινό).

Οι δορυφόροι Sentinel-2 είναι προϊόντα συνεργασίας της Ευρωπαϊκής Επιτροπής Διαστήματος (European Space Agency – ESA), μιας κοινοπραξίας 60 εταιρειών με επικεφαλής την Airbus Defense and Space και υποστηρίζονται από το Εθνικό Κέντρο Διαστημικών Μελετών (Centre National d'Études Spatiales – CNES) της Γαλλίας και το Γερμανικό κέντρο Διαστημικής DLR, για την βελτιστοποίηση των καλύψεων της γήινης επιφάνειας.

Οι δορυφόροι αυτοί είναι κατασκευασμένοι για να παρέχουν δεδομένα τόσο για αγροτικές όσο και για δασικές περιοχές με σκοπό την διατροφική ασφάλεια του πλανήτη. Ειδικότερα από τα δεδομένα που παρέχουν προσδιορίζονται παράγοντες πολύ σημαντικοί για την ανθρώπινη διαβίωση και επιβίωση όπως δείκτες περιεκτικότητας νερού και χλωροφύλλης στις φυτοκοινωνίες.

Θεωρείται κατάλληλος για τις παρακάτω εφαρμογές:

- Διαχείριση των αλλαγών των καλύψεων – χρήσεων γης.
- Καταγραφή δασικών περιοχών και δασικών καταστροφών.
- Διαχείριση και καταγραφή ανάπτυξης φυτικών πληθυσμών.
- Καταγραφή της μόλυνσης παράκτιων και λιμναίων περιοχών.
- Καταγραφικές και συνεχείς απεικονίσεις φυσικών καταστροφών όπως των πλημμυρών, ηφαιστειακών εκρήξεων και καθιζήσεων-ολισθήσεων εδαφών.

Παρουσιάζονται παρακάτω τα 13 κανάλια του δορυφόρου (εικόνα 2.4.) , με τα μήκη κύματος που τους αντιστοιχούν. Αξίζει να αναφερθεί πως χρησιμοποιήθηκαν ως παράμετροι για το prediction στο random forest 8 κανάλια , τα B2 - B3 - B4 - B5 - B6 - B7 - B8 - B11 - B12 . Επίσης μέσα από τον κώδικα μειώσαμε σε όλα τα κανάλια την ανάλυση στα 10 μέτρα , αν και αυτό θα αναλυθεί στο κεφάλαιο 3.

Sentinel-2 Bands

Sentinel-2 Bands	Central Wavelength [micrometers]	Resolution [meters]
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

Εικόνα 2.4. Κανάλια δορυφόρου Sentinel – 2

2.4 Πυρκαγιές

Ένας από τους σημαντικότερους παράγοντες που επιδρούν αρνητικά στη διατήρηση της παραγωγικότητας των δασών είναι οι δασικές πυρκαγιές. Κανένας άλλος παράγοντας δεν προκαλεί τόσο γρήγορες και τόσο εκτεταμένες καταστροφές στο οικοσύστημα όσο οι φωτιές στα δάση. Στην πραγματικότητα, ωστόσο, παρόλο που φαντάζουν ένα εξαιρετικά καταστρεπτικό γεγονός, αποτελούν μέρος του κύκλου ζωής των δασών και σε ορισμένες περιπτώσεις μπορούν να θεωρηθούν και ωφέλιμες. Αυτό που φαίνεται σίγουρο είναι ότι οι πυρκαγιές, με τη δύναμη που διαθέτουν, προκαλούν ριζικές καταστροφές στα δάση.

Τα αίτια των δασικών πυρκαγιών και όλων των πυρκαγιών εν γένει οφείλονται στις παρακάτω περιπτώσεις:

- Φυσικά ή τυχαία αίτια
- Εμπρησμό από αμέλεια
- Εμπρησμό από πρόθεση
- Άγνωστα αίτια

Οι ζημιές από τις δασικές πυρκαγιές μπορούν να διακριθούν σε άμεσες και έμμεσες. Οι άμεσες είναι εκείνες που προκαλούνται από τη μερική ή ολική καταστροφή του ξυλώδους κεφαλαίου. Οι άμεσες είναι αυτές που προκαλούνται από μερική ή ολική καταστροφή του ξυλώδους κεφαλαίου και περιλαμβάνουν τους απώλειες των δασικών προϊόντων καθώς και της βλάστησης των βοσκότοπων. Επίσης, περιλαμβάνονται οι ζημιές που προκαλούνται σε γεωργικές καλλιέργειες, κτηνοτροφικές – μελισσοκομικές εγκαταστάσεις, κατοικίες, βιομηχανικές και βιοτεχνικές εγκαταστάσεις ή άλλες μονάδες, από τη μετάδοση σ' αυτές δασικών πυρκαγιών.

Οι έμμεσες ζημιές προκαλούνται στις έμμεσες ωφέλειες του δάσους και είναι:

- Οικολογικές
- Υδρολογικές
- Απώλειες θέσεων εργασίας
- Υπερβόσκηση
- Η αισθητική, τουριστική και πολιτιστική αξία του δάσους
- Οι υγειονομικές επιπτώσεις

3. Μεθοδολογία και υλοποίηση

Στο κεφάλαιο αυτό αναλύεται η μεθοδολογία για την ταξινόμηση που χρησιμοποιήθηκε, από τη λήψη των δεδομένων εκπαίδευσης και ελέγχου και τη δημιουργία των ψευδοεικόνων έως και τη δημιουργία του πίνακα σύγχυσης με τις επιμέρους ακρίβειες κάθε κατηγορίας της ταξινόμησης. Αναλύονται επίσης όλα τα πειράματα που έγιναν για να προσδιοριστεί αυτό που δίνει το καλύτερο αποτέλεσμα.

3.1 Πυρκαγιές στις επιμέρους περιοχές μελέτης της Ελλάδας

Ως προς την ετήσια συχνότητα των δασικών πυρκαγιών στη χώρα μας και μάλιστα των πιο καταστροφικών, αυτές συμβαίνουν, κατά μέσο όρο 80% περίπου, το χρονικό διάστημα από 1^η Ιουνίου έως και 31^η Οκτωβρίου, αν και δεν είναι σπάνιες πυρκαγιές σε ορισμένες περιοχές κατά τη διάρκεια του Νοεμβρίου η Μαΐου, ιδίως κατά τα ξηρά έτη. Κατά τη διάρκεια του 24ώρου, το 90% περίπου εκδηλώνεται από τις 08.00 έως τις 23.00 το βράδυ, με μέγιστο, στο διάστημα από 14.00 έως 17.00. Το φαινόμενο αυτό έχει σαφή σχέση με την ημερήσια πορεία διακύμανσης της σχετικής υγρασίας του αέρα και της θερμοκρασίας.

Για τις ανάγκες της μελέτης αυτής, τη λήψη των απαραίτητων δεδομένων και τις ταξινομήσεις δημιουργήθηκαν ψευδοεικόνες (κεφάλαιο 3.3) στις περιοχές Κύθηρα, Κινέτα, Μάτι, Γραμματικό, Κοντοδεσπότι, Νέα Νότιο Ουαλία. Όσο αφορά τις περιοχές της Ελλάδας, η πυροσβεστική έδωσε τις εκτιμήσεις που φαίνονται στον παρακάτω πίνακα για τις καμένες εκτάσεις σε στρέμματα που παρουσιάζονται στον Πίνακα 3.1, ενώ απεικονίζονται και σε χάρτη στην εικόνα

ΠΕΡΙΟΧΗ	ΚΑΜΜΕΝΗ ΕΚΤΑΣΗ	ΗΜΕΡΟΜΗΝΙΑ
ΚΥΘΗΡΑ	25568,6	13/8/2017
ΚΙΝΕΤΑ	54435	24/7/2018
ΜΑΤΙ	11160,7	24/7/2018
ΓΡΑΜΜΑΤΙΚΟ	25652,4	04/8/2017
ΚΟΝΤΟΔΕΣΠΟΤΙ	23849,1	13/8/2019

Πίνακας 3.1 Καμένες εκτάσεις

Τα Κύθηρα είναι ένα νησί που βρίσκεται στη Νότιο Ελλάδα, στα νότια της Πελοποννήσου και νοτιότερα της Ελλαφώνησου και του Κάβο Μάλια. Η Κινέτα είναι παραθαλάσσιος οικισμός της Δυτικής Αττικής που βρίσκεται κατά μήκος της Εθνικής οδού Αθηνών-Κορίνθου, 56 χιλιόμετρα δυτικά της Αθήνας. Το Μάτι είναι παραθεριστικός οικισμός στο δήμο Μαραθώνος της Ανατολικής Αττικής, στα διοικητικά όρια της Δημοτικής ενότητας Νέας Μάρκρης στο νοτιότερο σημείο των συνόρων με τη Ραφήνα. Το Γραμματικό είναι ημιορεινό χωριό της Βορειοανατολικής Αττικής, πλησίον του Μαραθώνα. Εκτείνεται στο βορειοανατολικό άκρο του Νομού, μεταξύ των ορεινών όγκων που απαρτίζουν τη μορφολογία της περιοχής. Το Κοντοδεσπότι είναι χωριό του Δήμου Διφύων-Μεσσαπίων της Περιφερειακής Ενότητας Εύβοιας. Παρακάτω, στην εικόνα 3.1 παρουσιάζονται οι περιοχές αυτές χοντρικά οριοθετημένες σε ένα χάρτη της Ελλάδας, και στην εικόνα 3.2 παρουσιάζεται ενδεικτικά μια καμένη έκταση.



Εικόνα 3.1 . Περιοχές μελέτης στον Ελλαδικό χώρο



Εικόνα 3.2 Καμένη έκταση στην Εύβοια

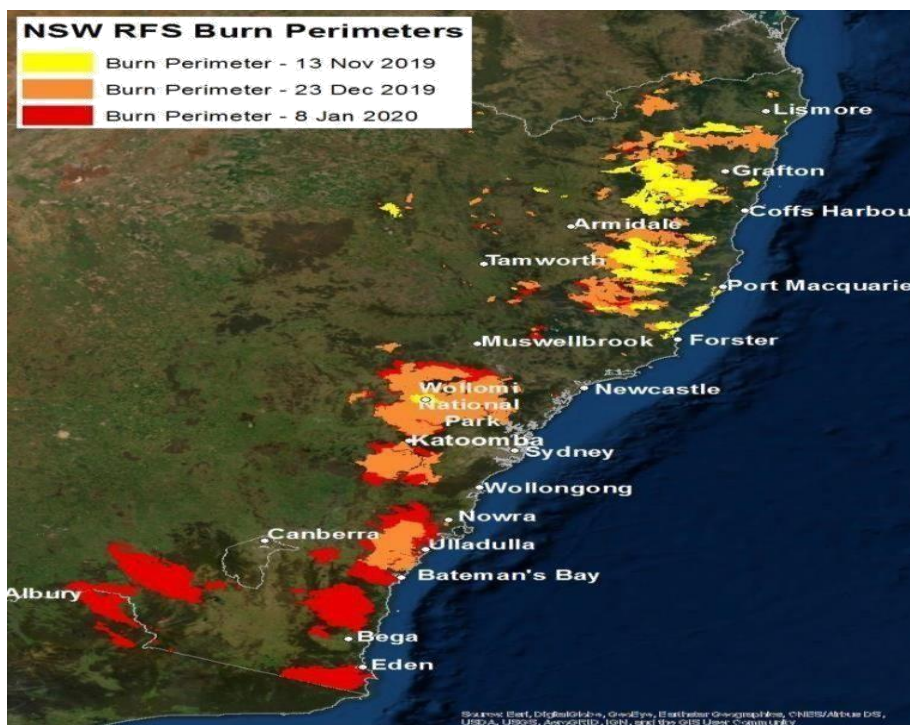
3.2 Νέα Νότια Ουαλία και κύμα δασικών πυρκαγιών

Ο αλγόριθμος που δημιουργήθηκε για χαρτογράφηση καμένων εκτάσεων με ταξινόμηση έχει εφαρμοστεί και στην περιοχή Νέα Νότια Ουαλία της Αυστραλίας(Εικόνα 3.3). Η περιοχή αυτή έχει έκταση 801.150 m² και κάηκαν κατά εκτίμηση 5,5 εκατομμύρια εκτάρια στο μεγάλο κύμα των δασικών πυρκαγιών που έπληξε τη χώρα τα έτη 2019 και 2020 . Είναι η πιο παλιά αποικία-πολιτεία της Κοινοπολιτείας της Αυστραλίας. Συνορεύει στο βορρά με το Κουίνσλαντ, στο νότο με τη Βικτώρια, στα δυτικά με τη Νότια Αυστραλία και εσωτερικά με το διαμέρισμα της Αυστραλιανής πρωτεύουσας και την περιοχή Τζέρβις Μπέι . Βρέχεται από τη θάλασσα της Τασμανίας και έχει ακτογραμμή μήκους 2.137 χλμ. Γεωφυσικά, η Νέα Νότια Ουαλία χωρίζεται σε τέσσερα μέρη. Την στενή παραλιακή λωρίδα, της οποίας το κλίμα κυμαίνεται από υποτροπικό στα σύνορα με το Κουϊνσλαντ έως δροσερό στο νότο, την ορεινή περιοχή των Ανατολικών Υψιπέδων, όπου βρίσκονται κορυφές που ξεπερνούν τα 1.000 μέτρα (όρος Κοσζιούσκο 2.229 μ.), τις «αγροτικές» πεδιάδες, που καταλαμβάνουν μεγάλο μέρος της αποικίας και τις Δυτικές πεδιάδες που καλύπτουν τα 2/3 της αποικίας.

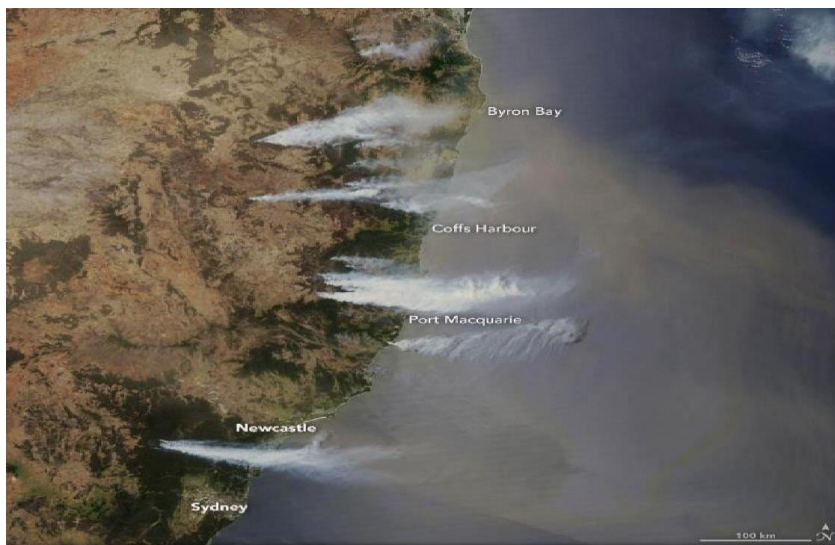


Εικόνα 3.3. Η Νέα Νότια Ουαλία (Πολιτικός χάρτης)

Η περίοδος 2019-2020 με το μεγάλο κύμα δασικών πυρκαγιών είναι η δυσμενέστερη που έχει υπάρξει στη Νέα Νότια Ουαλία. Θερμοκρασίες υψηλότερες από τις συνηθισμένες σε συνδυασμό με χαμηλά επίπεδα υγρασίας ακολουθούμενα από πολλά χρόνια ξηρασίας οδήγησαν σε καταστροφικές πυρκαγιές σε μεγάλο μέρος της έκτασης που καλύπτει η περιοχή. Σε μερικούς μήνες χάθηκαν 26 ανθρώπινες ζωές, καταστράφηκαν 2.448 σπίτια και κάηκαν σε σύνολο περίπου 5,5 εκατομμύρια εκτάρια. Το αντίκτυπο στους αγρότες, τις τοπικές επιχειρήσεις, την άγρια ζωή και την κοινωνία γενικότερα ήταν ανυπολόγιστο και η καταστροφή τεράστια. Ενδεικτικά, από τη 1^η Ιουλίου του 2019 έως και τη 31^η Μάρτη του 2020 συνέβησαν περίπου 11.400 δασικές πυρκαγιές στη Νέα Νότια Ουαλία. Περισσότερα από 1 δισεκατομμύριο ζώα πέθαναν στις δασικές πυρκαγιές, πολλά από αυτά σπάνια η υπο εξαφάνιση, με τον οριστικό χαμό ορισμένων από αυτών να θεωρείται δεδομένος. Η καταστροφή κοστολογείται σε 1,88 δισεκατομμύρια δολάρια Αυστραλίας (AUD). Στις παρακάτω εικόνες (3.4. και 3.5.) φαίνονται οι πυρκαγιές όπως τις κατέγραψαν οι τοπικές υπηρεσίες.



Εικόνα 3.4 . Πυρκαγιές από πρωτότυπο χάρτη



Εικόνα 3.5 .Οι πυρκαγιές όπως φαίνονται από δορυφορική εικόνα

3.3 Δημιουργία ψευδοεικόνων

Ενα σημαντικό πλεονέκτημα της δημιουργίας ψευδοεικόνων είναι πως είναι απαλλαγμένες απο νεφοκάλυψη συγκριτικά με πρωτογενή δεδομένα που αντιστοιχούν σε συγκεκριμένες ημερομηνίες. Για την παραγωγή τους ορίστηκαν περιοχές ενδιαφέροντος για τις περιοχές Κύθηρα, Γραμματικό, Κοντοδεσπότι, Μάτι και Κινέτα με χρήση και εφαρμογή στο χάρτη μιας γεωμετρίας (geometry). Η χρονοσειρά για την εύρεση και εισαγωγή δορυφορικών εικόνων είναι η 'COPERNICUS/S2_SR', με ατμοσφαιρικά διορθωμένα δορυφορικά δεδομένα Sentinel – 2 .

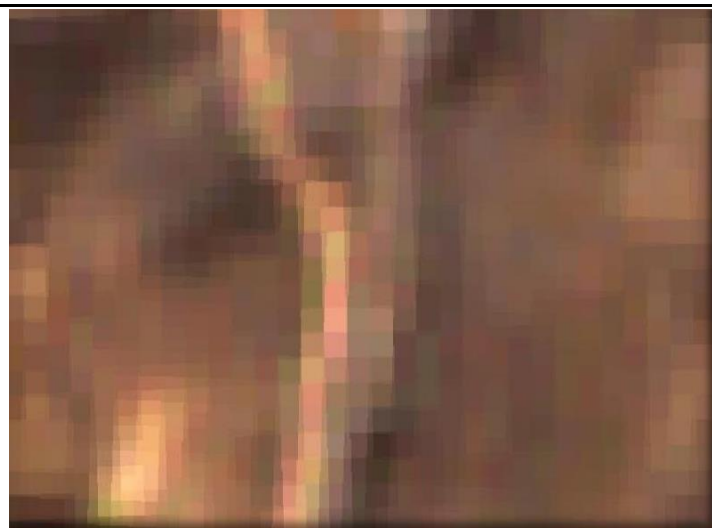
Αρχικά ορίστηκε ανώτατο όριο νεφοκάλυψης 1% στις επιμέρους εικόνες που επικαλύπτουν τα πολύγωνα σε ένα χρονικό εύρος περίπου 20 ημέρες μετά το πέρας της πυρκαγιάς, και ορίστηκε με clip μόνο το μέρος των εικόνων που βρίσκονται εντός των πολυγώνων. Ενδεικτικά τα πολύγωνα για Γραμματικό (Εικόνα 3.6.), Κοντοδεσπότι και Κύθηρα τα επικαλύπτουν 4,8 και 7 εικόνες αντίστοιχα. Στις περιοχές Κινέτα και Μάτι ορίστηκε ενιαίο geometry λόγω του ότι η πυρκαγιά έλαβε χώρα την ίδια ακριβώς ημέρα, και ορίστηκε ανώτατο όριο στα σύννεφα 10% για να υπάρξει πληθώρα εικόνων, 27 στο σύνολο και για τις δύο περιοχές. Το κάθε εικονοστοιχείο της τελικής ψευδοεικόνας έχει ψηφιακή τιμή το μέσο όρο των εικόνων που το επικαλύπτουν σε κάθε ένα του κανάλι. Σε κάθε κανάλι ορίστηκαν min & max 100 και 1600 αντίστοιχα με αποκοπή στα άκρα, δηλαδή τιμές φωτεινότητας μεγαλύτερες απο 1600 στο εύρος τιμών του Sentinel-2 πήραν την τιμή 1600 και τιμές μικρότερες απο 100 πήραν την τιμή 100. Στην Νέα Νότια Ουαλία η ψευδοεικόνα δημιουργήθηκε από 1436 εικόνες Sentinel – 2 σε χρονικό εύρος απο 10-01-2020 έως και 01-05-2020. Ξεκινά δηλαδή μετά το πέρας των πυρκαγιών σε όλη την περιφέρεια.



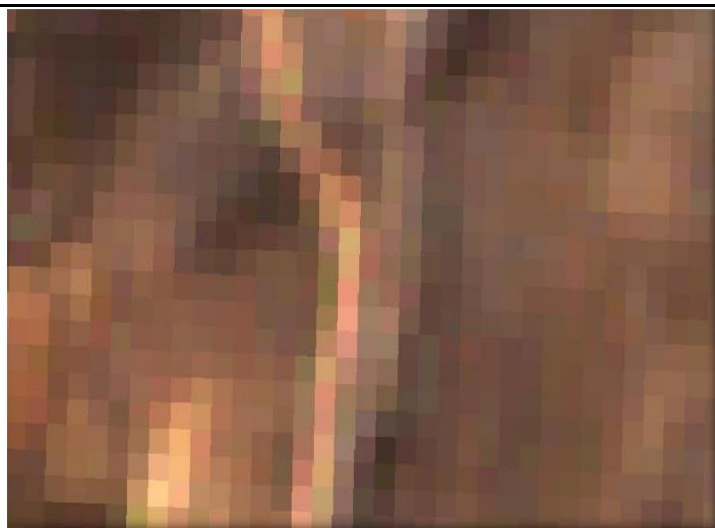
Εικόνα 3.6 Φυσικό έγχρωμο σύνθετο ψευδοεικόνας στο Γραμματικό

3.3.1 Επαναπροβολή εικονοστοιχείων

Για τη δημιουργία κάθε ψευδοεικόνας ενδέχεται να χρησιμοποιήθηκαν εικόνες οι οποίες αντιστοιχούν σε διαφορετικές τροχιές του δορυφόρου Sentinel-2. Οι αποκλίσεις στη γεωαναφορά, της τάξης μισού έως δύο εικονοστοιχεία που υπάρχουν ανάμεσα στις διαφορετικές τροχιές, έχουν ως συνέπεια τα σχήματα στις εικόνες να εμφανίζονται ελαφρώς παραμορφωμένα. Γι 'αυτό δημιουργήθηκε στον κώδικα το παρακάτω τέχνασμα: επιλέχθηκε το κανάλι 2 (B2) της εικόνας με τα λιγότερα σύννεφα από όλες αυτές που επικαλύπτουν το πολύγωνο, και επαναπροβλήθηκαν όλα τα εικονοστοιχεία της ψευδοεικόνας στην προβολή του B2 . Έτσι, η τελική εικόνα έχει εικονοστοιχεία σε κανονική ορθογώνια διάταξη, και όλα με ανάλυση 10 επί 10 μέτρα, όπως και το κανάλι B2. Για να γίνουν καλύτερα αντιληπτά τα παραπάνω παρουσιάζονται παρακάτω δύο εικόνες – αποσπάσματα από πριν και μετά την προβολή, για την ίδια περιοχή (Εικόνες 3.7. και 3.8.).



Εικόνα 3.7. Πρίν την προβολή



Εικόνα 3.8. Μετά την προβολή

Για να αποφευχθεί σύγχυση όσο αφορά τα ονόματα των κατηγοριών που χρησιμοποιήθηκαν στην ταξινόμηση, παρακάτω παρουσιάζεται ένας κατατοπιστικός πίνακας, με τα ονόματα έτσι όπως χρησιμοποιήθηκαν στους επιμέρους κώδικες και τις περιοχές που τους αντιστοιχούν.

3.4 Δεδομένα εκπαίδευσης και ελέγχου

Ο ταξινομητής είναι τα Τυχαία Δάση (Random Forest) και αναλύθηκε εκτενώς στο κεφάλαιο 2 του τεύχους. Χρειάζεται επαρκή δεδομένα εκπαίδευσης και ελέγχου για να εφαρμοστεί, όπως και παραμέτρους, στην προκειμένη περίπτωση τα κανάλια με τις φασματικές τιμές, για να κάνει το prediction. Χρησιμοποιήθηκαν 9 από τα 13 κανάλια του δορυφόρου, τα B2 – B3 – B4 – B5 – B6 – B7 – B8 – B11 – B12.

Όσο αφορά τα δεδομένα, πάρθηκαν 24 σετ δεδομένων για 6 κατηγορίες με πολύγωνα από 4 περιοχές με εφαρμογή Feature Collection απάνω στις ψευδοεικόνες. Οι περιοχές από τις οποίες πάρθηκαν τα πολύγωνα είναι τα Κύθηρα, το Μάτι, τα Κινέτα και το Γραμματικό, ενώ η ταξινόμηση έγινε στο Κοντοδεσπότι. Αντίστοιχα για την Αυστραλία πάρθηκαν όλα από την περιοχή της Νέας Νότιας Ουαλίας.

Κρίθηκε σκόπιμο να ληφθούν δεδομένα για 6 κατηγορίες στην Ελλάδα και για 7 στην Αυστραλία, και παρουσιάζονται παρακάτω.

Ελλάδα:

- Καμένα
- Νερό
- Αστικός ιστός
- Καλλιέργειες
- Βλάστηση
- Έδαφος

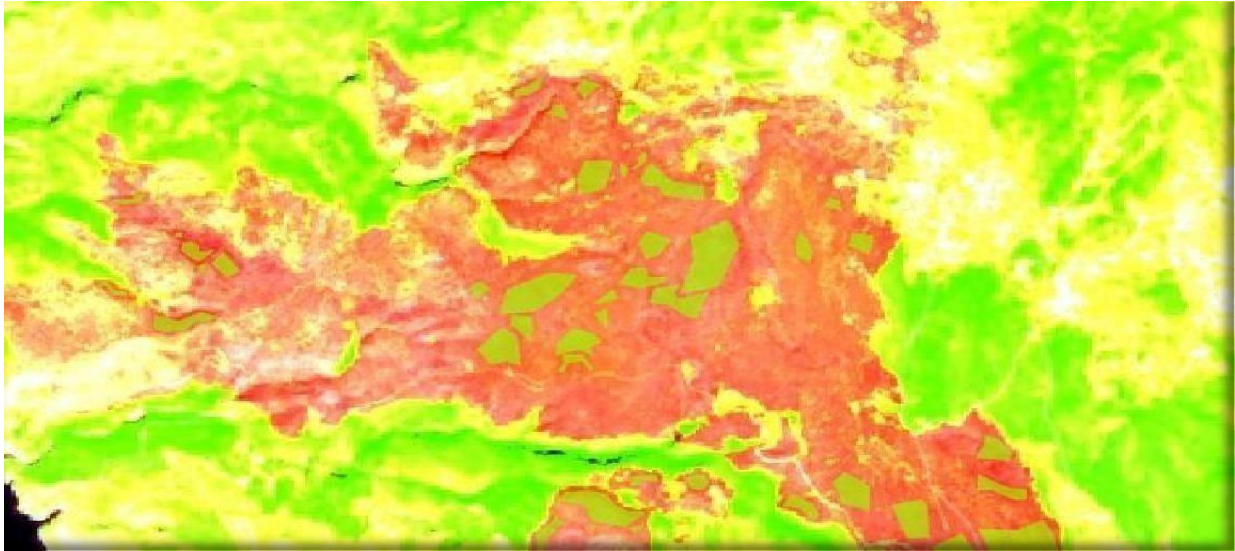
Αυστραλία:

- Έρημος
- Καμένα
- Αστικός ιστός
- Νερό
- Καλλιέργειες
- Δάσος 1
- Δάσος 2

Η κάθε μια κατηγορία αποτελείται από πολύγωνα που ψηφιοποιήθηκαν από τις ψευδοεικόνες. Στόχος για ένα ακριβές και ικανοποιητικό αποτέλεσμα είναι τα πολύγωνα αυτά να είναι όσο το δυνατότερο πιο αμιγή όσο αφορά τις κατηγορίες που τους αντιστοιχούν. Έτσι έγινε λεπτομερής ψηφιοποίηση στο φυσικό έγχρωμο σύνθετο RGB 4-3-2 , εκτός της κατηγορίας καμένα που είναι και η πιο σημαντική, η οποία και ψηφιοποιήθηκε στο RGB 12-8 -2 το οποίο εξυπηρετεί καλά την ανίχνευση των καμένων, καθώς η στάχτη φαίνεται σε έντονες αποχρώσεις του κόκκινου.

3.4.1 Ελλάδα

Δεδομένα λήφθηκαν για 6 κατηγορίες από 4 περιοχές, και παρακάτω παρουσιάζεται ενδεικτικά η λήψη των δεδομένων αυτών σε εικόνες (Εικόνες 3.9 και 3.10.) από το περιβάλλον του Google Earth Engine. Στο περιβάλλον τα πολύγωνα αυτά βρίσκονται στα geometry imports. Για τα καμένα, τα πολύγωνα από τα Κύθηρα ονομάζονται burned4, από τα Κινέτα burned1, από το Μάτι valburned1 και από το Γραμματικό valburned2. Αντίστοιχα και για τις υπόλοιπες κατηγορίες.



Εικόνα 3.9. Πολύγωνα λήψης για καμένα από τα Κύθηρα



Εικόνα 3.10 . Πολύγωνα λήψης από φυσικό έγχρωμο σύνθετο στην Κινέτα

Για να αποφευχθεί σύγχυση όσο αφορά τα ονόματα των κατηγοριών που χρησιμοποιήθηκαν στην ταξινόμηση, παρακάτω παρουσιάζεται ένας κατατοπιστικός πίνακας, με τα ονόματα έτσι όπως χρησιμοποιήθηκαν στους επιμέρους κώδικες και τις περιοχές που τους αντιστοιχούν.

burned4	KYΘHPA	burned	NEA NOTIA OYALIA
soil4	KYΘHPA	desert	NEA NOTIA OYALIA
water4	KYΘHPA	forest1	NEA NOTIA OYALIA
buildings4	KYΘHPA	water	NEA NOTIA OYALIA
vegetation4	KYΘHPA	forest2	NEA NOTIA OYALIA
kalliergeies4	KYΘHPA	urban	NEA NOTIA OYALIA
burned1	KINETA	crops	NEA NOTIA OYALIA
soil1	KINETA	valburned	NEA NOTIA OYALIA
water1	KINETA	valdesert	NEA NOTIA OYALIA
buildings1	KINETA	valforest1	NEA NOTIA OYALIA
vegetation1	KINETA	valwater	NEA NOTIA OYALIA
kalliergeies1	KINETA	valforest2	NEA NOTIA OYALIA
valburned1	MATI	valurban	NEA NOTIA OYALIA
valsoil1	MATI	valcrops	NEA NOTIA OYALIA
valwater1	MATI		
valbuildings1	MATI		
valvegetation1	MATI		
valkalliergeies1	MATI		
valburned2	ΓΡΑΜΜΑΤΙΚΟ		
valsoil2	ΓΡΑΜΜΑΤΙΚΟ		
valwater2	ΓΡΑΜΜΑΤΙΚΟ		
valbuildings2	ΓΡΑΜΜΑΤΙΚΟ		
valvegetation2	ΓΡΑΜΜΑΤΙΚΟ		
valkalliergeies2	ΓΡΑΜΜΑΤΙΚΟ		

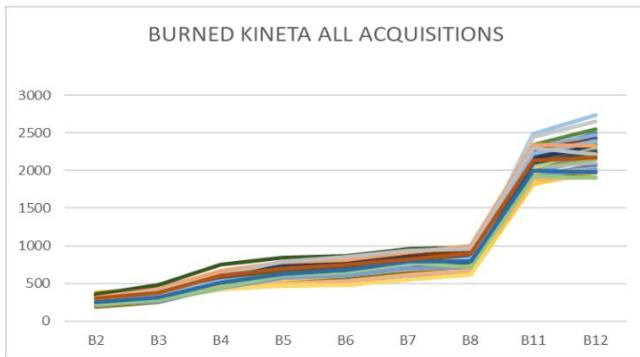
Πίνακας 3.1. Πολύγωνα ταξινόμησης

Στον παραπάνω πίνακα η κατηγορία burned αντιστοιχεί στα καμένα, η soil στο έδαφος, η water στο νερό, η buildings στα κτίρια, η vegetation στη βλάστηση και οι kalliergeies στις καλλιέργειες. Αυτά όσο αφορά στον Ελλαδικό χώρο. Για τη Νέα Νότια Ουαλία της Αυστραλίας, η κατηγορία burned αντιστοιχεί στα καμένα, η desert στην Έρημο , η κατηγορία forest1 σε ένα είδος δασικών εκτάσεων της βορειοδυτικής περιοχής της επικράτειας, η κατηγορία water στο νερό, η κατηγορία forest2 σε δασικές εκτάσεις οι οποίες βρίσκονται Νοτιοδυτικά στην περιοχή μελέτης, η κατηγορία urban στον αστικό ιστό και τα έργα υποδομής και τέλος η κατηγορία crops στις καλλιέργειες. Με μπλέ παρουσιάζονται στον πίνακα οι κατηγορίες που αντιστοιχούν στην εκπαίδευση και με πορτοκαλί αυτές που αντιστοιχούν στην αξιολόγηση.

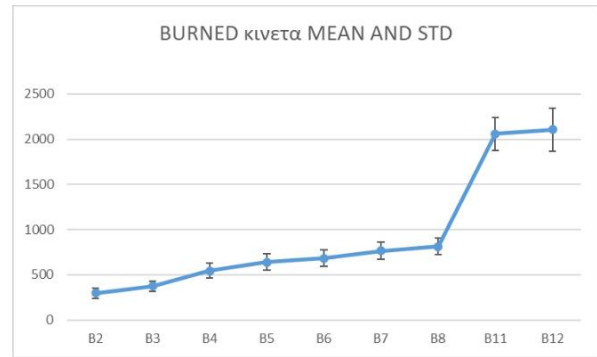
Διαγράμματα για δεδομένα από Ελλάδα και σχολιασμός

Σημαντικό είναι, από τον ορισμό του αλγορίθμου Random Forest, τα πολύγωνα από τα οποία αποτελείται το dataset να είναι όσο το δυνατόν πιο αμιγή όσο αφορά αυτά που απεικονίζουν, να μη διαφέρουν αισθητά όσο αφορά τη φασματική τους συμπεριφορά και να είναι επαρκή σε σύνολο. Εδώ αξίζει να αναφερθεί πως το Google Earth Engine έχει ορισμένους περιορισμούς στο computation time και συνεπώς στο σύνολο των δεδομένων που μπορούν να χρησιμοποιηθούν για την εκπαίδευση του μοντέλου και την πρόβλεψη. Σε κάθε περίπτωση τα δεδομένα για την εργασία αυτή είναι επαρκή, και λήφθηκαν με σχολαστικότητα και λεπτομέρεια για όλες τις κατηγορίες όσο αφορά το τμήμα της εργασίας που αντιστοιχεί στον Ελλαδικό χώρο.

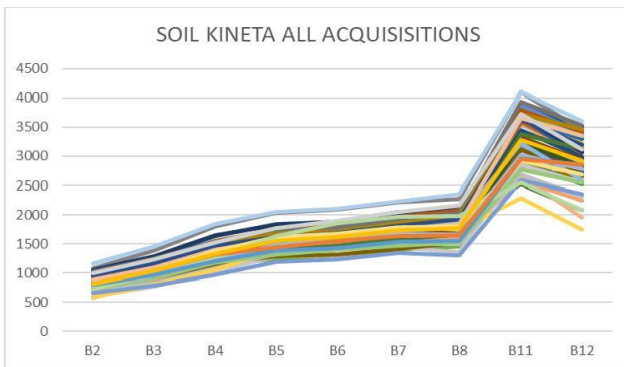
Παρακάτω παρουσιάζονται διαγράμματα μέσου όρου και τυπικής απόκλισης, για τις ανα κανάλι φασματικές τιμές των κατηγοριών καμένα και έδαφος, από τις περιοχές Κινέτα και Γραμματικό. Τα διαγράμματα δημιουργήθηκαν σε λογιστικό φύλλο excel , αφού προηγήθηκε εξαγωγή σε μορφή αρχείου .csv από το περιβάλλον του Google Earth Engine. Αυτό επιτεύχθηκε με κώδικα, στον οποίο ορίζεται ένα χρονικό εύρος μετά την ημερομηνία της πυρκαγιάς και γίνεται εξαγωγή σε μορφή αρχείου .csv ο μέσος όρος των ψηφιακών τιμών που για το κάθε πολύγωνο από όλες τις εικόνες που το επικαλύπτουν σε αυτό το χρονικό εύρος. Αυτό δεν αντιστοιχεί ακριβώς στις ψηφιακές τιμές των πολυγώνων της ψευδοεικόνας, αλλά είναι επαρκές και ικανοποιητικό για την εξαγωγή συμπερασμάτων για τη φασματική συνεκτικότητα των πολυγώνων από τα οποία αποτελείται το dataset.



Διάγραμμα 3.1 . Σύνολο παρατηρήσεων για καμένα από τα Κινέτα



Διάγραμμα 3.2. Μέσος όρος και τυπική απόκλιση για καμένα από τα Κινέτα



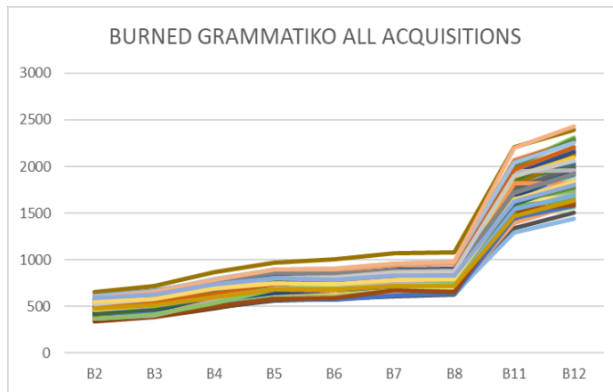
Διάγραμμα 3.3 . Σύνολο παρατηρήσεων για έδαφος από τα Κινέτα



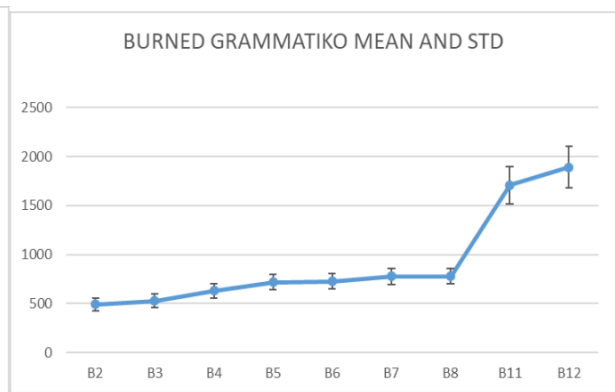
Διάγραμμα 3.4. Μέσος όρος και τυπική απόκλιση για έδαφος από τα Κινέτα

Τα αποτελέσματα είναι πολύ ικανοποιητικά, και γίνεται ξεκάθαρο από τα παραπάνω διαγράμματα 4.1 και 4.2 πως τα πολύγωνα δε διαφέρουν φασματικά σε μεγάλο βαθμό το ένα από το άλλο. Οι μεγαλύτερες τυπικές αποκλίσεις των ψηφιακών τιμών φωτεινότητας βρίσκονται στα κανάλια B11 και B12, κι αυτό οφείλεται στο γεγονός ότι τα καμένα ανακλούν πολύ σε αυτά τα μήκη κύματος (υπέρυθρο). Παρατηρούμε πως σε ό,τι αφορά την κατηγορία Έδαφος, ή φασματική συνεκτικότητα είναι σαφώς χαμηλότερη και έχουμε σημαντικά μεγαλύτερες τυπικές αποκλίσεις, ενώ οι τυπικές αποκλίσεις διογκώνονται στα κανάλια B11 και B12. Αυτό έγκειται στο γεγονός πως το έδαφος δεν έχει την ίδια φυσιογεωγραφική ομοιομορφία και φασματική ομοιομορφία με τα καμένα και συνεπώς τα επιμέρους πολύγωνα που το συνιστούν διαφέρουν σημαντικά το ένα με το άλλο.

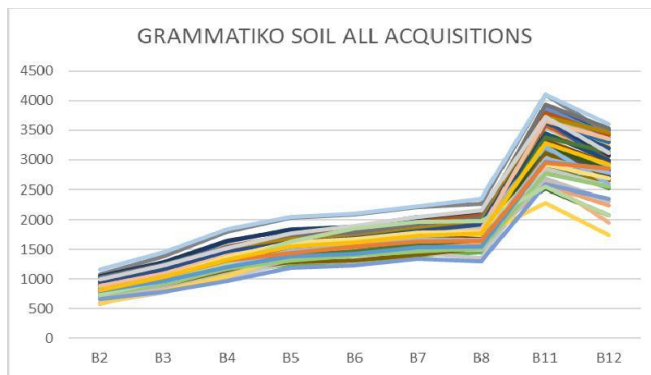
Αντίστοιχα παρουσιάζονται και τα διαγράμματα για τις κατηγορίες καμένα και έδαφος από το γραμματικό. Ο λόγος που επιλέχθηκαν να φαίνονται οι δύο αυτές κατηγορίες, από τις 2 αυτές περιοχές είναι πως από τις 2 αυτές περιοχές λήφθηκαν τα training data για το καλύτερο και αποτελεσματικότερο πείραμα, όπως επίσης και διότι έχουν παρόμοια φασματική συμπεριφορά και ενδεχομένως αυτό να προκαλέσει σύγχυση στην ταξινόμηση.



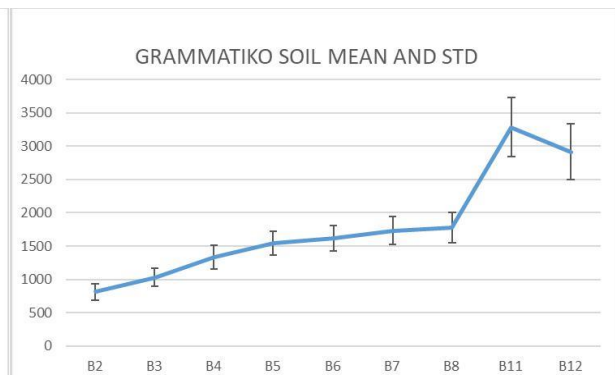
Διάγραμμα 3.5 . Σύνολο παρατηρήσεων για τα καμένα από το Γραμματικό



Διάγραμμα 3.6.. Μέσος όρος και τυπική απόκλιση για καμένα από Το Γραμματικό

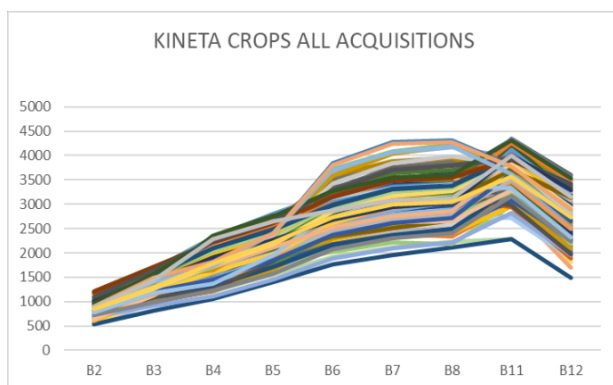


Διάγραμμα 3.7 . Σύνολο παρατηρήσεων για έδαφος από το Γραμματικό

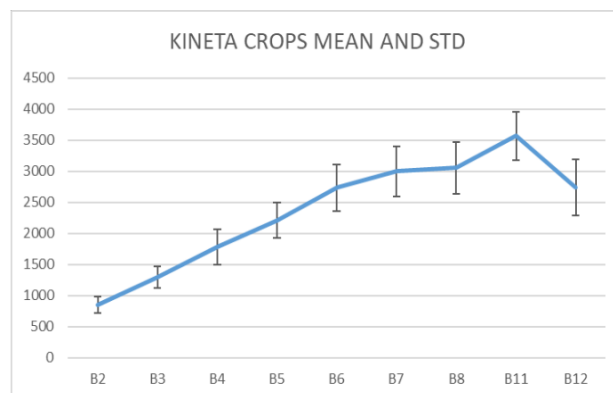


Διάγραμμα 3.8. Μέσος όρος και τυπική απόκλιση για έδαφος από το Γραμματικό

Παρατηρούμε πως όσο αφορά την κατηγορία καμένα (διαγράμματα 4.5 και 4.6), η φασματική συνεκτικότητα είναι πολύ ικανοποιητική και οι τυπικές αποκλίσεις μικρές, τόσο στα Κινέτα όσο και στο Γραμματικό. Αυτό έγκειται στο γεγονός ότι ψηφιοποιήθηκαν με πολλή ακρίβεια και σχολαστικότητα από ένα έγχρωμο σύνθετο (RGB 12-8-2), το οποίο επιτρέπει τον εύκολο εντοπισμό τους. Όσο αφορά το έδαφος, οι φασματική συνεκτικότητα είναι ασθενέστερη και οι αποκλίσεις μεγαλύτερες κι αυτό οφείλεται στο ότι, όπως προαναφέρθηκε, είναι μια φασματικά ετερογενής κατηγορία, καθώς εδάφη με διαφορετική σύσταση έχουν διαφορετική φασματική υπογραφή. Οι υπόλοιπες κατηγορίες που επιλέχθηκαν για την ταξινόμηση στον ελλαδικό χώρο παρουσιάζουν γενικά καλή διαχωριστικότητα σε σχέση με τα καμένα και το έδαφος παρότι είναι αρκετά γενικές και συμπεριλαμβάνουν αρκετές υποκατηγορίες. Παρακάτω παρουσιάζονται ενδεικτικά και τα διαγράμματα (4.9 και 4.10) για την κατηγορία καλλιέργειες στα Κινέτα, όπου τα επιμέρους διαγράμματα ψηφιακών τιμών των πολυγώνων διαφέρουν αισθητά και οι τυπικές αποκλίσεις είναι μεγάλες.



Διάγραμμα 3.9. Σύνολο παρατηρήσεων για καλλιέργειες από τα Κινέτα



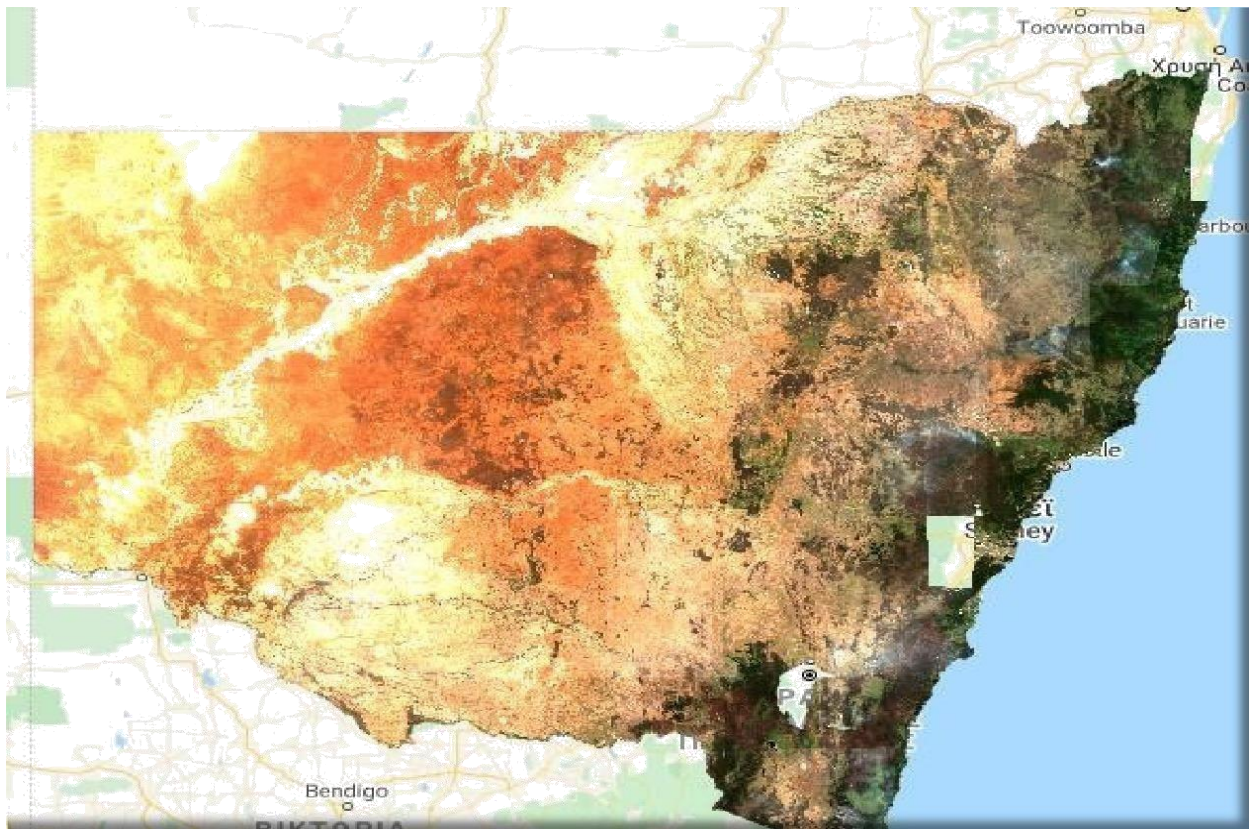
Διάγραμμα 3.10 Μέση τιμή και τυπική απόκλιση για καλλιέργειες από τα Κινέτα

3.4.2 Αυστραλία

Η ταξινόμηση μιας τόσο μεγάλης σε έκταση και τόσο διαφοροποιήσιμης και ιδιομορφής σε φυσιογεωγραφικά χαρακτηριστικά και μορφολογία περιοχής αποτελεί από μόνο του ένα δύσκολο, αν όχι αδύνατο εγχείρημα. Όπως έχει προαναφερθεί στο κεφάλαιο 3, η επικράτεια της Νέας Νότιας Ουαλίας καταλαμβάνει μια αχανή έκταση με περισσότερα από 800.000 m². Συνεπώς, οι καλύψεις γής είναι πολλές και διαφορετικές, και η ορθή, ακριβής και λεπτομερής ταξινόμηση της περιοχής πρακτικά αδύνατη. Επιπρόσθετα προβλήματα είναι το μεγάλο χρονικό εύρος που πρέπει να ληφθεί για να επικαλύπτεται με εικόνες όλη η επικράτεια της Νέας Νότιας Ουαλίας με εικόνες, αναλογιζόμενοι και το γεγονός πως αυτές θα έχουν ανώτατο όριο σε σύννεφα το 1% της επιφάνειάς τους, που καθιστά το χρονικό εύρος ακόμα μεγαλύτερο.

Η μεγαλύτερη τροχοπέδη του όλου εγχειρήματος όμως είναι το γεγονός ότι οι πυρκαγιές όχι μόνο είναι πολλές (χιλιάδες σε αριθμό), αλλά και το ότι συνέβησαν σε διαφορετικά χρονικά διαστήματα, με όλο το εύρος διάρκειας τους αθροιστικά για όλη τη Νέα Νότια Ουαλία να ξεκινά από τον Μάη του 2019 και να ολοκληρώνεται τον Ιανουάριο του 2020. Η φασματική συμπεριφορά των εικονοστοιχείων μετά το πέρας της πυρκαγιάς αλλάζει σταδιακά, και τείνει με το πέρας του χρόνου να διαφοροποιηθεί από καμένο και να έχει φασματική συμπεριφορά πιο κοντά σε μια άλλη κατηγορία. Λόγου χάρη, οι πυρκαγιές στο βορειοανατολικό μέρος της Νέας Νότιας Ουαλίας άρχισαν να μαίνονται τον Μάη του 2019 και ολοκληρώθηκαν τον Οκτώβρη του 2019, ενώ οι πυρκαγιές στο νοτιοανατολικό ολοκληρώθηκαν τον Ιανουάριο του 2020. Η ημερομηνία λήξης όλων των πυρκαγιών είναι στις 10/01/2020, όπως έδωσαν οι υπηρεσίες της Αυστραλίας, και συνεπώς αυτό είναι και το σημείο εκκίνησης για την εύρεση των δορυφορικών εικόνων και τη δημιουργία της ψευδοεικόνας της Νέας Νότιας Ουαλίας.

Με χρονικό εύρος από τις 10-01-2020 έως τις 01-05-2020 βρέθηκαν 1436 δορυφορικές εικόνες οι οποίες επικαλύπτουν τη Νέα Νότια Ουαλία και δημιουργούν την ψευδοεικόνα. Στην εικόνα 3.11 παρουσιάζεται μια ψευδοεικόνα σε φυσικό έγχρωμο σύνθετο print – screen από το περιβάλλον του Google Earth Engine της Νέας Νότιας Ουαλίας μετά την πυρκαγιά.



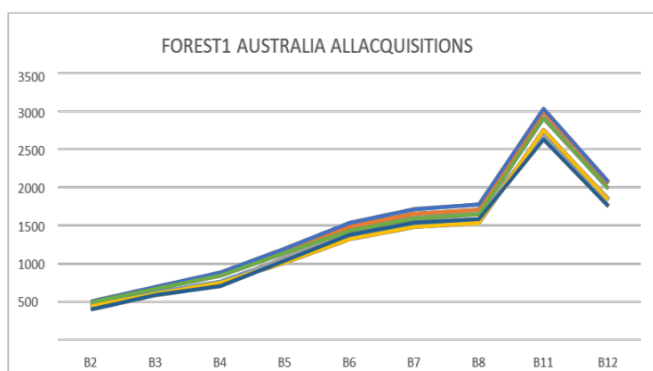
Εικόνα 3.11. Φυσικό έγχρωμο σύνθετο (RGB: 432) για την περιοχή της Νέας Νότιας Ουαλίας

Διαγράμματα για δεδομένα από Αυστραλία και σχολιασμός

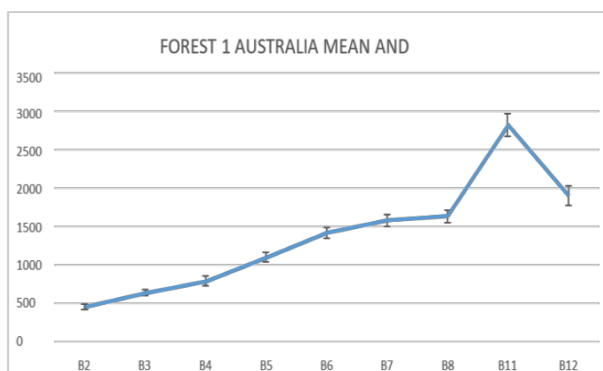
Από την περιοχή της Νέας Νότιας Ουαλίας κρίθηκε σκόπιμο να ληφθούν δεδομένα για 7 θεματικές κατηγορίες κάλυψης γής, καθώς λόγω των διαφορετικών κλιματολογικών συνθηκών είναι αδύνατο να ταξινομηθεί η περιοχή με δεδομένα από Ελλάδα. Οι 7 αυτές κατηγορίες είναι οι παρακάτω :

- Έρημος
- Νερό
- Καλλιέργειες
- Καμένα
- Αστικός ιστός
- Δάσος 1
- Δάσος 2

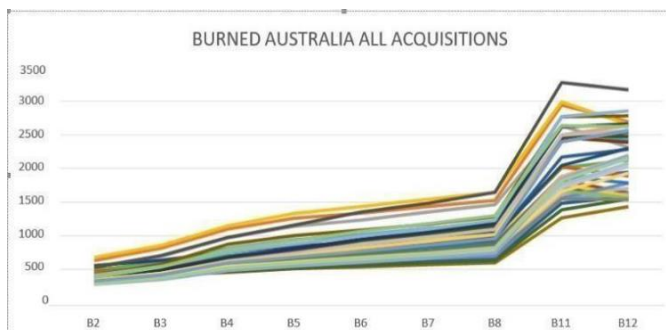
Παρακάτω παρουσιάζονται τα διαγράμματα για τις κατηγορίες καμένα και δάσος 1 της Αυστραλίας, και ακολουθεί σχολιασμός τους. Το χρονικό εύρος παρατηρήσεων σε αυτή την περίπτωση ορίστηκε να είναι από 10-01-2020 έως και 10-03-2020.



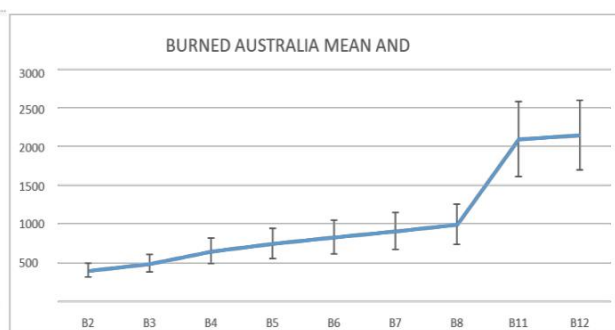
Διάγραμμα 3.11. Σύνολο παρατηρήσεων για δάσος1 από την Αυστραλία



Διάγραμμα 3.12. Μέση τιμή και τυπική απόκλιση για δάσος1 από την Αυστραλία



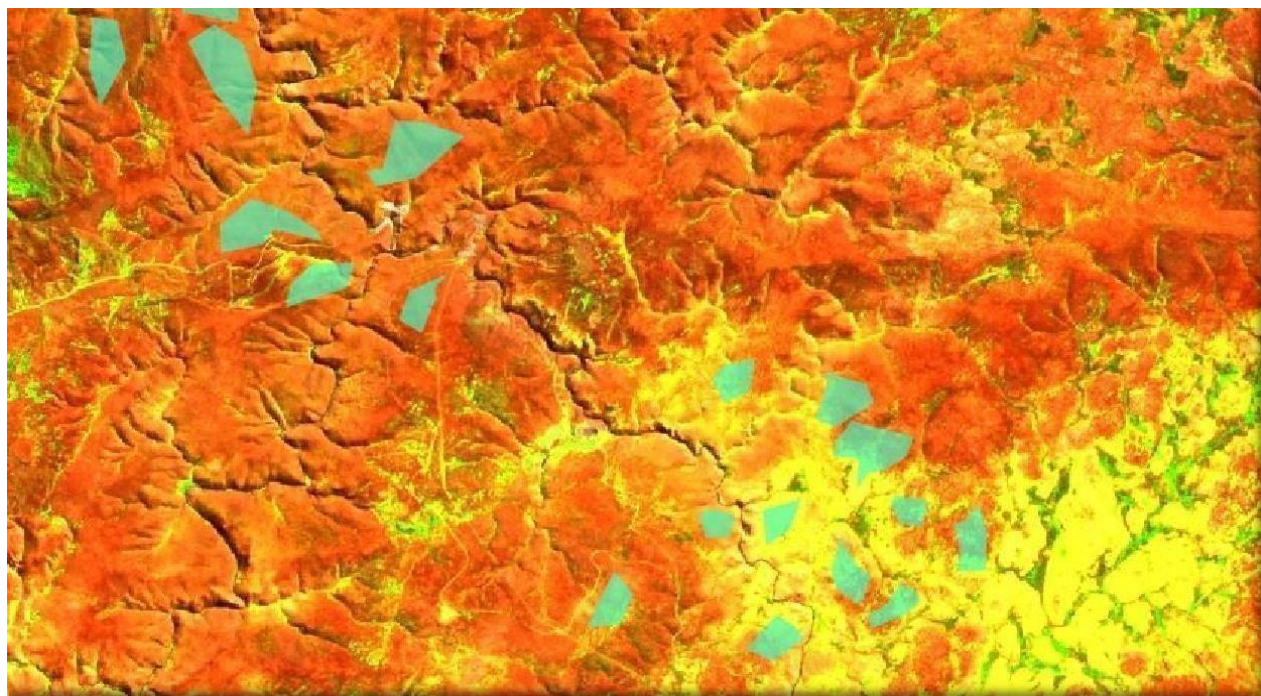
Διάγραμμα 3.13. Σύνολο παρατηρήσεων για καμένα από την Αυστραλία



Διάγραμμα 3.14 Μέση τιμή και τυπική απόκλιση για καμένα από την Αυστραλία

Γίνεται αντιληπτό πως σε ότι αφορά την κατηγορία δάσος 1 τα αποτελέσματα είναι πολύ ικανοποιητικά. Οι τυπικές αποκλίσεις είναι μικρές και οι ψηφιακές φασματικές τιμές των πολυγώνων είναι κοντά η μια στην άλλη. Για τις κατηγορίες δάσος2, καλλιέργειες, νερό και έρημος τα διαγράμματα είναι επίσης ικανοποιητικά. Όσο αφορά την κατηγορία αστικός ιστός αυτό δεν ισχύει, διότι ο αστικός ιστός σε μεγάλες πόλεις όπως το Sydney διαφέρει πολύ σε αποχρώσεις.

Η κατηγορία με ιδιαίτερη σημασία είναι τα καμένα όμως, διότι στόχος της μελέτης αυτής είναι η χαρτογράφηση καμένων εκτάσεων. Στο διάγραμμα 4.14 παρατηρείται πως η κατηγορία καμένα έχει μεγάλες τυπικές αποκλίσεις ανά κανάλι άρα και χαμηλή φασματική συνεκτικότητα στην περίπτωση της Αυστραλίας. Η παρατήρηση αυτή έρχεται σε αντίθεση με τη συμπεριφορά της αντίστοιχης κατηγορίας στον Ελλαδικό χώρο. Μια πιθανή εξήγηση, είναι πως αυτό οφείλεται στο μεγάλο πλήθος των διαφορετικών εστιών πυρκαγιών, αλλά κυρίως στη διαφορά τους στο διάστημα 8 μηνών. Συνέπεια αυτών, ήταν κάποιες απο τις καμένες εκτάσεις να έχουν ήδη αρχίσει να ανακάμπτουν όσον αφορά την υπάρχουσα βλάστηση, για το χρονικό διάστημα που πάρθηκαν οι ψευδοεικόνες. Στην εικόνα 3.12 παρουσιάζεται στιγμιότυπο απο το περιβάλλον του Google Earth Engine, στο οποίο διακρίνονται κάποια ψηφιοποιημένα πολύγωνα για την κατηγορία καμένα, επί έγχρωμου συνθέτου RGB 12-8-2.



Εικόνα 3.12 . Λήψη δεδομένων για καμένα στη Νέα Νότια Ουαλία απο ψευδέχρωμο σύνθετο RGB 12-8-2

3.5 Πειράματα και αξιολόγηση τους

Όπως προαναφέρθηκε στην ενότητα 1.2, δεδομένα λήφθηκαν με χρήση πολυγώνων – Feature Collections από τις περιοχές Κύθηρα – Γραμματικό – Μάτι – Κινέτα , και πραγματοποιήθηκε ταξινόμηση στην περιοχή Κοντοδεσπότι στην Εύβοια. Έπειτα απο τη λήψη των δεδομένων και τον έλεγχο της φασματικής τους συνεκτικότητας, πραγματοποιήθηκαν διαφορετικοί συνδυασμοί απο τις τέσσερις περιοχές ανα ζεύγη. Τα συνολικά πειράματα ανα συνδυασμό δυάδας δεδομένων εκπαίδευσης και ελέγχου ήταν 6, και παρουσιάζονται παρακάτω.

- 1) TRAINING: KINETA-MATI / VALIDATION: KYΘHPA – ΓPAMMATIKO -> 20.000
- 2) TRAINING: KINETA - ΓPAMMATIKO / VALIDATION: MATI – KYΘHPA ->23.000
- 3) TRAINING : KYΘHPA -KINETA / VALIDATION: MATI – ΓPAMMATIKO ->19.000
- 4) TRAINING : KYΘHPA - ΓPAMMATIKO / VALIDATION: MATI – KINETA -> 19.000
- 5) TRAINING : KYΘHPA - MATI / VALIDATION: ΓPAMMATIKO – KINETA -> 22.400
- 6) TRAINING : ΓPAMMATIKO - MATI / VALIDATION: KINETA – KYΘHPA -> 25.460

Η πυροσβεστική έδωσε εκτίμηση 23.800 όσο αφορά τα καμένα στρέμματα στο Κοντοδεσπότι, και συνεπώς την ακριβέστερη εκτίμηση τη δίνει το πείραμα 2. Αυτό επιβεβαιώνεται και από τον πίνακα σύγκυσης του πειράματος αυτού, και αυτό αναλύεται παρακάτω.

3.5.1 Πίνακες σύγκυσης και επιμέρους ακρίβειες

Ο πίνακας αυτός βασίζεται στα δεδομένα εκπαίδευσης σε σύγκριση με τα δεδομένα ελέγχου, και μας δείχνει το κατά πόσο έχουν ταξινομηθεί ορθά τα εικονοστοιχεία της κάθε κατηγορίας. Συγκεκριμένα, η κύρια διαγώνιος αντιπροσωπεύει τον αριθμό των εικονοστοιχείων που έχουν ταξινομηθεί ορθά, οι γραμμές τα σφάλματα παράλειψης, δηλαδή αυτά που εσφαλμένα έχουν ταξινομηθεί σε άλλη κατηγορία κι όχι σε αυτή στην οποία ανήκουν, και οι στήλες τα σφάλματα συμπερίληψης, αυτή δηλαδή τα οποία εσφαλμένα έχουν ταξινομηθεί στην κατηγορία αυτή.

Αντίστοιχα, στον πίνακα φαίνονται τα Users Accuracy και Precision Accuracy. Αυτά τα ποσοστά % εκφράζουν το τι μέρος 100% των εικονοστοιχείων έχει ταξινομηθεί ορθά, αναλογιζόμενοι τα σφάλματα συμπερίληψης και παράλειψης αντίστοιχα. Η συνολική ακρίβεια της ταξινόμησης (Overall Accuracy) προκύπτει ως ο μέσος όρος των επιμέρους ακριβειών. Αυτό που έχει σημασία όμως για τη μελέτη αυτή είναι τα Users Accuracy και Precision Accuracy που αντιστοιχούν στις καμένες εκτάσεις. Παρακάτω παρουσιάζεται ο καλύτερος πίνακας σύγκυσης από τα 6 πειράματα της Ελλάδας (πείραμα 2), και ο ασθενέστερος (πείραμα 6), με γνώμονα την ακρίβεια ταξινόμησης των καμένων εκτάσεων.

Στον πίνακα 3.1 παρατηρείται ότι υπάρχουν μικρά σφάλματα παράλειψης και συμπερίληψης (11 και 7) για τα καμένα με το έδαφος. Απο αυτό μπορούμε να αποφανθούμε πως για το πείραμα αυτό τα δεδομένα εκπαίδευσης και ελέγχου για την κατηγορία καμένα είναι πολύ ικανοποιητικά. Σημαντικά σφάλματα παράλειψης και συμπερίληψης παρατηρούνται για τις καλλιέργειες προς το έδαφος και για το έδαφος προς τις καλλιέργειες αντίστοιχα, κι αυτό ενδεχομένως να οφείλεται στο ότι ορισμένες καλλιέργειες που ψηφιοποιήθηκαν αποτελούνται απο γυμνό χώμα, όπως επίσης και στο ότι το έδαφος έχει πολλή αλληλουχία όσο αφορά τη φασματική του συμπεριφορά, και συνεπώς ορισμένα πολύγωνα που ψηφιοποιήθηκαν έχουν πανομοιότυπη συμπεριφορά με το έδαφος. Τέλος, σημαντικά σφάλματα παράλειψης και συμπερίληψης υφίστανται και ανάμεσα στις κατηγορίες έδαφος και νερό, κι αυτό οφείλεται στην ύπαρξη βραχώδης έκτασης που έχει φασματική ομοιότητα και με τις 2 αυτές κατηγορίες. Για το στόχο της παρούσας διπλωματικής όμως, που είναι η χαρτογράφηση των καμένων εκτάσεων, τα αποτελέσματα κρίνονται ως ικανοποιητικά για αυτό το πείραμα.

Αντίστοιχα, στον πίνακα για το δυσμενέστερο πείραμα που εξυπηρετεί το σκοπό της παρούσας διπλωματικής παρατηρούνται πανομοιότυπα σφάλματα παράλειψης και συμπερίληψης για τις κατηγορίες νερό, καλλιέργειες και έδαφος. Όσο αφορά την κατηγορία καμένα παρατηρείται σημαντικό σφάλμα συμπερίληψης με την κατηγορία νερό, και γι' αυτό το πείραμα αυτό κρίνεται ως ακατάλληλο. Το σφάλμα αυτό οφείλεται ενδεχομένως σε εσφαλμένα δεδομένα εκπαίδευσης και ελέγχου.

Αξίζει να σημειωθεί πως η συνολική ακρίβεια στο πείραμα 6 είναι σαφώς υψηλότερη από αυτή του πειράματος 2, αλλά οι επιμέρους ακρίβειες για UA και PA για τα καμένα είναι υψηλότερες, και αυτό καθιστά το πείραμα αυτό πιο ικανοποιητικό.

Τα αποτελέσματα είναι πολύ ικανοποιητικά και για την ταξινόμηση στη Νέα Νότια Ουαλία. Απ' ότι παρατηρείτε στον πίνακα σύγκυσης για Αυστραλία (Πίνακας 3.3), για την κατηγορία καμένα δεν υπάρχουν σημαντικά σφάλματα παράλειψης και συμπερίληψης, παρά λίγο με το δάσο1 με ένα σφάλμα παράλειψης 211 εικονοστοιχείων, κι ένα με τον αστικό ιστό με 184 εικονοστοιχεία. Επίσης, παρατηρείται και σημαντικό σφάλμα συμπερίληψης της κατηγορίας καλλιέργειες με την έρημο. Τα σφάλματα αυτά ενδεχομένως οφείλονται στο γεγονός ότι η Νέα Νότια Ουαλία περιέχει υψηλή ποικιλομορφία όσο αφορά την εδαφική της κάλυψη και συνεπώς ορισμένα εικονοστοιχεία υπάρχει υψηλή πιθανότητα να αντιστοιχούν σε περισσότερες απο κατηγορίες. Επίσης τα δεδομένα εκπαίδευσης και ελέγχου λήφθηκαν απο διαφορετικές περιοχές της τεράστιας αυτής επικράτειας και αυτό οδηγεί σε μια κάποια διαφοροποίηση της φασματικής τους συμπεριφοράς, με επακόλουθο τα σφάλματα που υπάρχουν στον πίνακα σύγκυσης.

		1	2	3	4	5	6		
		ΚΑΜΜΕΝΑ	ΚΑΛΛΙΕΡΓΕΙΕΣ	ΒΛΑΣΤΗΣΗ	ΕΔΑΦΟΣ	ΝΕΡΟ	ΚΤΙΡΙΑ	SUM	UA
1	ΚΑΜΜΕΝΑ	16268	0	0	7	0	0	16275	99,96%
2	ΚΑΛΛΙΕΡΓΕΙΕΣ	0	15795	0	952	33	999	17779	88,84%
3	ΒΛΑΣΤΗΣΗ	0	61	12916	5	0	0	12982	99,49%
4	ΕΔΑΦΟΣ	11	12453	0	8337	4376	0	25177	33,11%
5	ΝΕΡΟ	0	0	0	0	25657	0	25657	100,00%
6	ΚΤΙΡΙΑ	0	64	1	5	1	4258	4329	98,36%
	SUM	16279	28373	12917	9306	30067	5257	OA	81,44%
	PA	99,93%	55,67%	99,99%	89,59%	85,33%	81,00%		

Πίνακας 3.1 . Πίνακας σύγκρισης για πείραμα 2

		1	2	3	4	5	6		
		ΚΑΜΜΕΝΑ	ΚΑΛΛΙΕΡΓΕΙΕΣ	ΒΛΑΣΤΗΣΗ	ΕΔΑΦΟΣ	ΝΕΡΟ	ΚΤΙΡΙΑ	SUM	UA
1	ΚΑΜΜΕΝΑ	14886	0	0	2	67	0	14955	99,54%
2	ΚΑΛΛΙΕΡΓΕΙΕΣ	0	12606	12	172	12	1112	13914	90,60%
3	ΒΛΑΣΤΗΣΗ	0	0	29849	0	0	0	29849	100,00%
4	ΕΔΑΦΟΣ	8	7066	1	8073	3241	0	18389	43,90%
5	ΝΕΡΟ	2923	0	0	0	81909	0	84832	96,55%
6	ΚΤΙΡΙΑ	1	37	1	0	0	6040	6079	99,36%
	SUM	17818	19709	29863	8247	85229	7152	OA	91,28%
	PA	83,54%	63,96%	99,95%	97,89%	96,10%	84,45%		

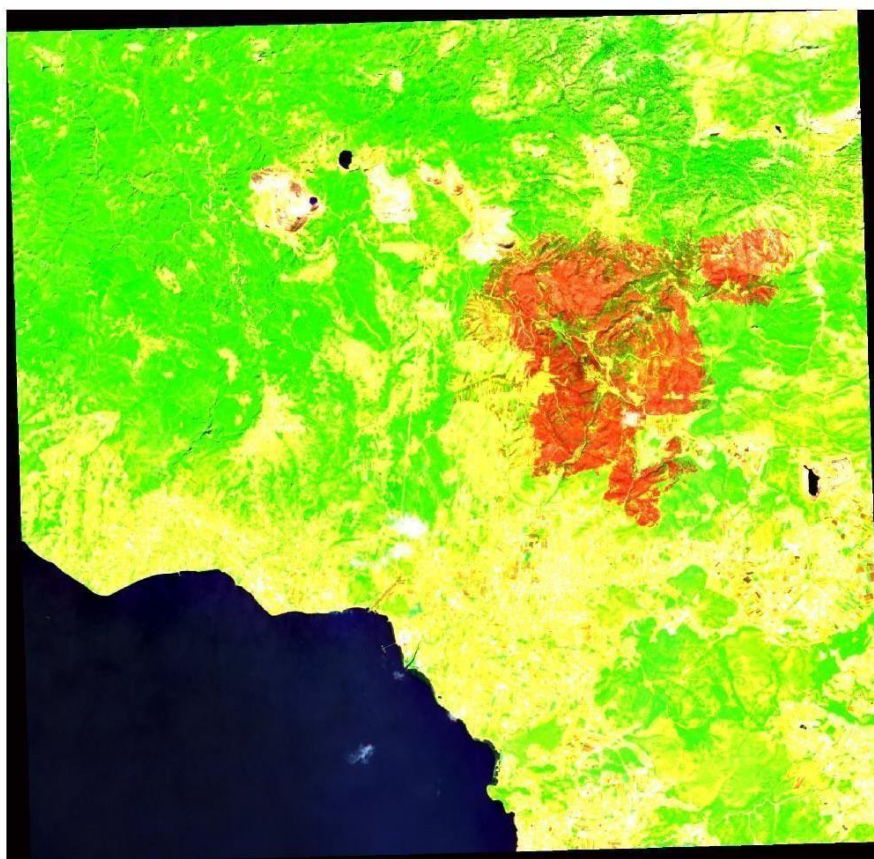
Πίνακας 3.2 . Πίνακας σύγκρισης για πείραμα 6

		1	2	3	4	5	6	7			
		ΚΑΜΕΝΑ	ΕΡΗΜΟΣ	ΔΑΣΟΣ1	ΔΑΣΟΣ2	ΑΣΤΙΚΟΣ ΙΣΤΟΣ	ΝΕΡΟ	ΚΑΛΛΙΕΡΓΕΙΕΣ	SUM		ΥΑ
1	ΚΑΜΕΝΑ	55211	0	211	41	184	12	0		55659	99,20%
2	ΕΡΗΜΟΣ	0	97974	0	0	0	0	0		97974	100,00%
3	ΔΑΣΟΣ1	0	0	24619	0	1	0	0		24620	100,00%
4	ΔΑΣΟΣ2	0	0	0	81815	0	0	0		81815	100,00%
5	ΑΣΤΙΚΟΣ ΙΣΤΟΣ	11	0	0	0	1312	0	0		1323	99,17%
6	ΝΕΡΟ	0	0	0	0	0	3061	0		3061	100,00%
7	ΚΑΛΛΙΕΡΓΕΙΕΣ	67	3015	0	0	18	0	90493		93593	96,69%
	SUM	55289	100989	24830	81856	1515	3073	90493	OA		0,99%
	ΥΑ	99,86%	97,01%	99,15%	99,95%	86,60%	99,61%	100,00%			

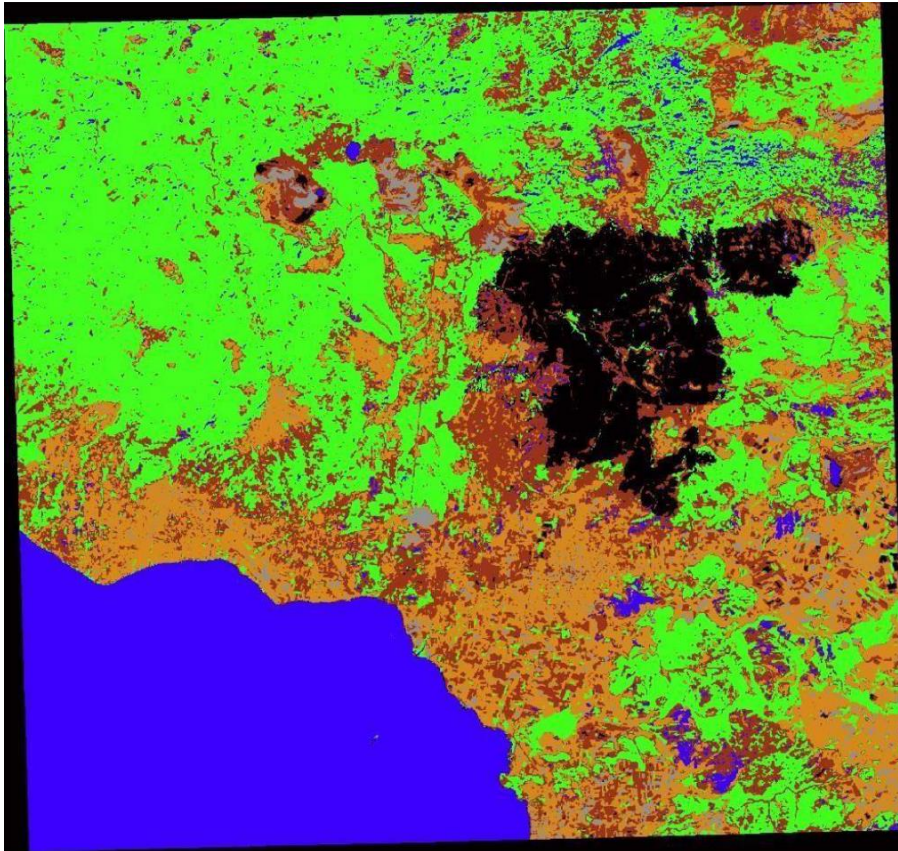
Πίνακας 3.3 Πίνακας σύγκρισης για Νέα Νότια Ουαλία

4. Χάρτες και σχολιασμός

Στο κεφάλαιο αυτό, παρουσιάζονται οι χάρτες που προέκυψαν στο περιβάλλον του Google Earth Engine. Έγιναν οι ταξινομημένες εικόνες για την Νέα Νότια Ουαλία (Εικόνες 3.16. και 3.17.) και το Κοντοδεσπότι (Εικόνες 3.13. 3.14, 3.15) αλλά και η ψευδοεικόνα για το Κοντοδεσπότι (Εικόνα 3.12) έγινε εξαγωγή σε μορφή αρχείου .TIF από το Google Earth Engine στον Google Drive. Ενδεικτικά, η εικόνα για το Κοντοδεσπότι έχει μέγεθος 580 kilobytes, η ψευδοεικόνα για το Κοντοδεσπότι έχει μέγεθος 91,1 Megabytes κι η ταξινομημένη Εικόνα της Νέας Νότια Ουαλίας 52 GigaBytes. Λόγω του μεγάλου μεγέθους της τελευταίας, η εξαγωγή της πραγματοποιήθηκε σε 20 επιμέρους τμήματα, δηλαδή διαφορετικά αρχεία. Στη συνέχεια αυτά τα αρχεία συνενώθηκαν με το εργαλείο merge στο QGIS.



Εικόνα 3.13. Έγχρωμο σύνθετο RGB 12-8-2 για Κοντοδεσπότι Εύβοιας

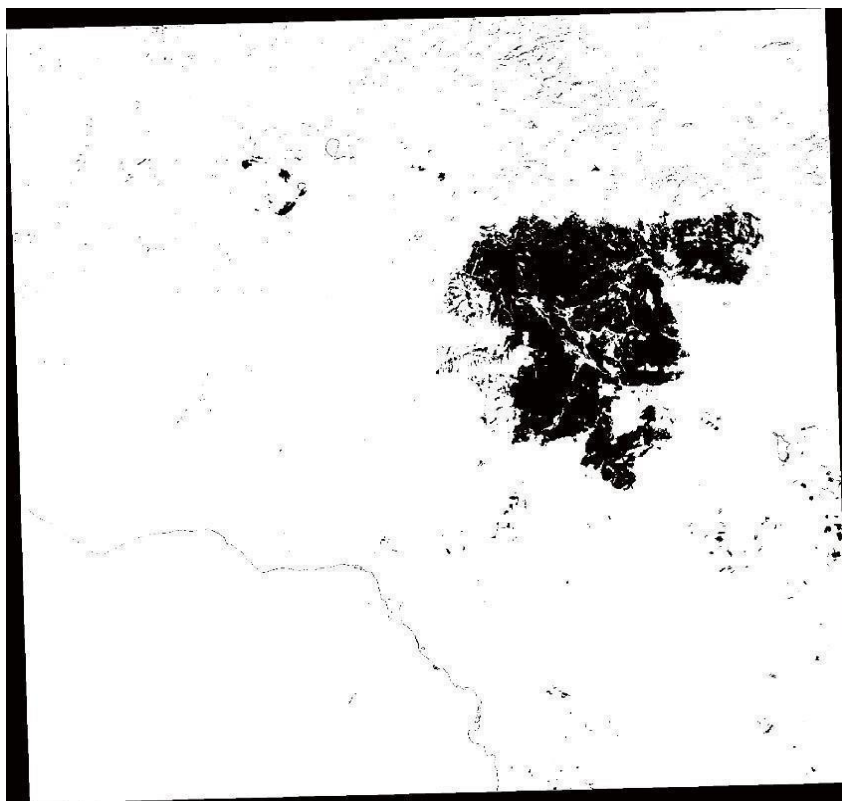


Εικόνα 3.14. Αποτέλεσμα ταξινόμησης περιοχής Κοντοδεσπότι

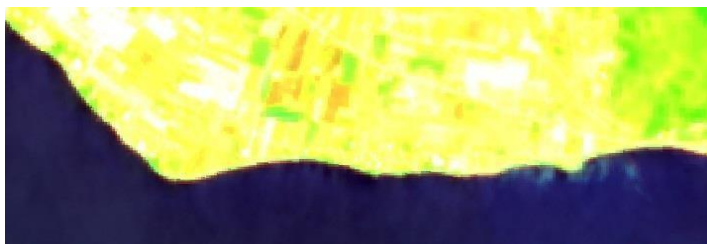
Όσο αφορά την ταξινόμηση στο Κοντοδεσπότι στον Ελλαδικό χώρο, το αποτέλεσμα φαίνεται αρκετά ικανοποιητικό από φωτοερμηνευτική άποψη. Οι καμένες εκτάσεις έχουν ταξινομηθεί ορθά, αλλά παράλληλα έχουν ταξινομηθεί και ως καμένα ορισμένα εικονοστοιχεία της παράκτιας γραμμής όπως και ορισμένα εικονοστοιχεία στις βραχώδεις εκτάσεις. Αυτό ενδεχομένως οφείλεται στην πανομοιότυπη φασματική συμπεριφορά που έχουν οι καμένες εκτάσεις με στην παραλία και τα βράχια και όσο καλά και αξιόπιστα δεδομένα να ληφθούν, δε γίνεται να εξαλειφθεί.

Επίσης είναι οφθαλμοφανές πως και πολλά χωράφια έχουν ταξινομηθεί ως καμένα. Αυτό συμβαίνει λόγω του ότι η ψευδοεικόνα δημιουργήθηκε στην καλοκαιρινή περίοδο, και αποτελεί συνήθης πρακτική για τους αγρότες να καίνε τα χωράφια τους την περίοδο αυτή. Συνεπώς το αποτέλεσμα κρίνεται ικανοποιητικό όσο αφορά τον Ελλαδικό χώρο, αλλά ορισμένες ατέλειες και τροχοπέδια είναι αδύνατο να εξαλειφθούν.

Τέλος, παρουσιάζεται ένας χάρτης με μονάχα τα εικονοστοιχεία ταξινομημένα ως καμένα για το Κοντοδεσπότη.

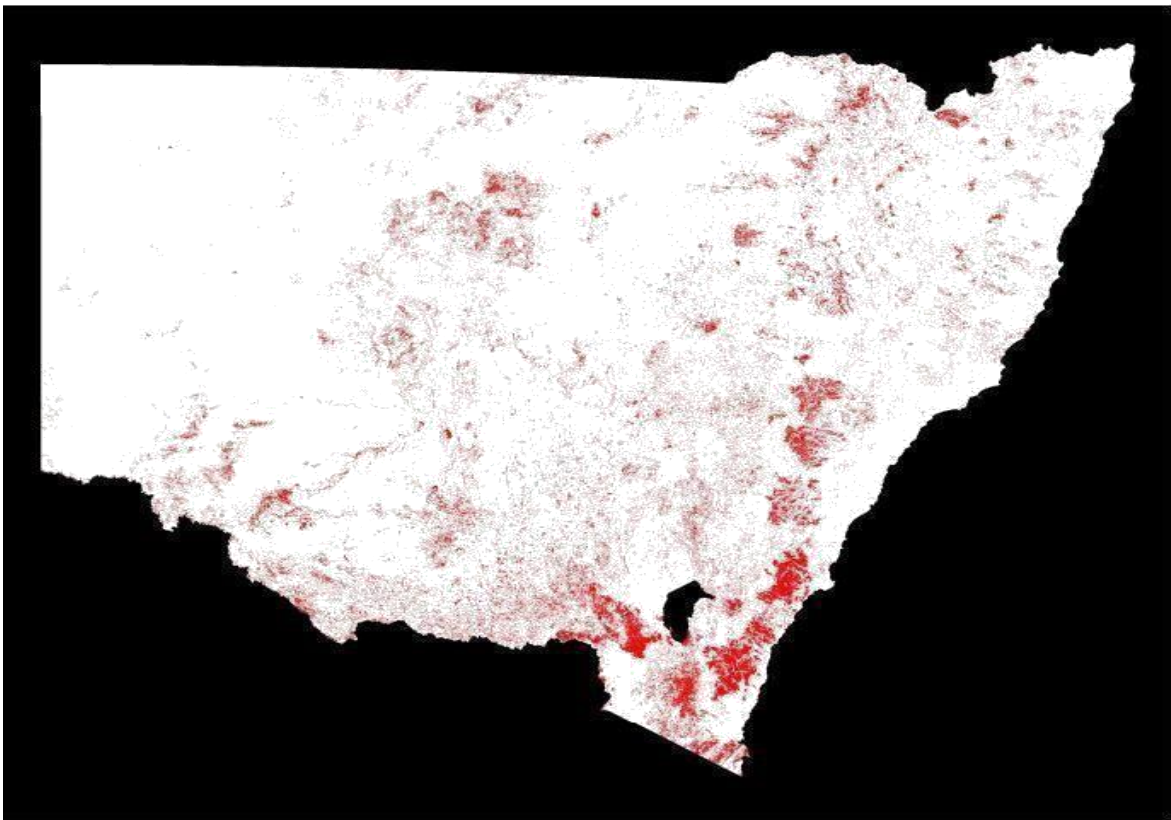


Εικόνα 3.15. Καμένα περιοχής Κοντοδεσπότη

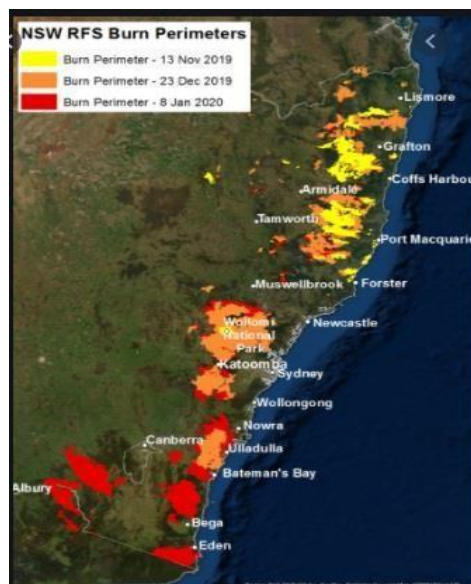


Εικόνα 3.16. Εικονοστοιχεία ακτογραμμής

Γίνεται αντιληπτό πως πολλά εικονοστοιχεία στην ακτογραμμή ταξινομήθηκαν εσφαλμένα ως καμένα. Αυτό οφείλεται στο γεγονός πως η άμμος στην ακτογραμμή έχει παρόμοια φασματική συμπεριφορά με τις καμένες εκτάσεις. Κάτι αντίστοιχο συμβαίνει και με τις ορεινές βραχώδεις περιοχές. Στο έγχρωμο σύνθετο RGB 12-8-2 τα εικονοστοιχεία των περιοχών αυτών παρουσιάζονται με κοκκινωπή απόχρωση, σαν τα καμένα, έχουν δηλαδή πανομοιότυπη φασματική συμπεριφορά. Έγινε προσπάθεια να ορισθούν στον αλγόριθμο ως ξεχωριστές κατηγορίες αλλά αυτό δημιούργησε σημαντικά σφάλματα στον πίνακα σύγχυσης και ως εκ τούτου πάρθηκε η απόφαση να παραμείνουν οι κατηγορίες της ταξινόμησης ως έχουν. Παρακάτω παρουσιάζεται μια εικόνα – απόσπασμα από την ακτογραμμή (Εικόνα 3.15.).



Εικόνα 3.17. Καμένα περιοχές Νέας Νότιας Ουαλίας



Εικόνα 3.18. Χάρτης απο πρωτότυπες πηγές Νέας Νότιας Ουαλίας

Όσο αφορά τη Νέα Νότια Ουαλία, είναι εύκολα αντιληπτό πως οι δασικές πυρκαγιές έπληξαν κυρίως το νοτιοανατολικό κομμάτι της περιφέρειας, με πολλές μεγάλες η και μικρότερες ταξινομημένες καμένες εκτάσεις να βρίσκονται διάσπαρτες σε σχεδόν όλη την περιφέρεια. Άλλωστε σε ολόκληρο το κύμα των δασικών πυρκαγιών το οποίο έπληξε την περιοχή τα από το Μάη του 2019 έως και τον Ιανουάριο του 2020, περίπου 11.400 ενεργές πυρκαγιές συνέβησαν. Ενδεχομένως όμως ορισμένες από αυτές τις εκτάσεις να είναι εσφαλμένα ταξινομημένες ως καμένες, λόγω της ανομοιομορφίας των δεδομένων εκπαίδευσης για τα καμένα, καθώς συμπεριλαμβάνουν και ημί – καμένα, για να συμπεριλάβουν στην ταξινόμηση και εικονοστοιχεία των οποίων η φασματική συμπεριφορά έχει αρχίσει να διαφοροποιείται από αυτή του καμένου μετά το πέρας της πυρκαγιάς.

Επίσης, η περιοχή αποτελείται από πολλές καλύψεις γης, και συνεπώς υπάρχει σημαντικό ενδεχόμενο να υπάρχουν πολλές περιοχές με φασματική συμπεριφορά παρόμοια με αυτή των δεδομένων εκπαίδευσης που λήφθηκαν για τα καμένα.

Σε κάθε περίπτωση, ο υπολογισμός στο Google Earth Engine των καμένων εκτάσεων έδωσε τιμή περίπου 5.700.000 εκτάρια ενώ οι εκτιμήσεις των τοπικών υπηρεσιών περίπου 5.500.000 εκτάρια. Αναλογίζοντας το ότι πολλά εικονοστοιχεία σε πολλές περιοχές έχουν χάσει τη φασματική συμπεριφορά του καμένου, είναι σαφές πως πολλές περιοχές έχουν εσφαλμένα ταξινομηθεί ως καμένες. Ακόμα, το χαρτογραφικό υπόβαθρο του Google Earth Engine είναι σε μερκατορική προβολή και συνεπώς η Αυστραλία, η οποία απέχει αρκετά από τον Ισημερινό έχει έκταση μεγαλύτερη απ' ότι στην πραγματικότητα. Αθροιστικά για μια τόσο μεγάλη σε έκταση καμένη περιοχή η μερκατορική προβολή δίνει μεγαλύτερο νούμερο καμένων εκτάσεων απ' ότι στην πραγματικότητα.

5. Συμπεράσματα - Προτάσεις

Ο αλγόριθμος που αναπτύχθηκε στα πλαίσια της εργασίας αυτής είναι ιδιαίτερα εύχρηστος αποτελεσματικός και αξιόπιστος, κι αυτό διότι δημιουργεί μια ψευδοεικόνα με πολύ λίγα έως και καθόλου σύννεφα και χρησιμοποιεί δεδομένα εκπαίδευσης και ελέγχου τα οποία είναι εύκολο να ληφθούν από ψευδοεικόνες με επίσης λίγα σύννεφα. Επιπρόσθετα , όλες οι εικόνες οι οποίες χρησιμοποιούνται για τη δημιουργία της τελικής ψευδοεικόνας είναι εύκολα προσβάσιμες, ατμοσφαιρικά και ραδιομετρικά διορθωμένες.

Η ακρίβεια του αποτελέσματος εξαρτάται από τα δεδομένα εκπαίδευσης και ελέγχου για την αξιολόγηση της , τα οποία ορίζονται από το χρήση, και η φασματική συνεκτικότητα και αξιοπιστία των οποίων μπορεί επίσης να ελεγχθεί με χρήση ενός λογιστικού φύλλου.

Εν προκειμένω, το αποτέλεσμα είναι πολύ ικανοποιητικό αλλά όχι αψεγάδιαστο όπως φαίνεται και στους χάρτες καθώς ατέλειες και εσφαλμένες ταξινομήσεις εικονοστοιχείων υπάρχουν και ενδεχομένως είναι ιδιαίτερα χρονοβόρο αν όχι αδύνατο να εξαλειφθούν πλήρως. Η μέθοδος που χρησιμοποιήθηκε όμως για χαρτογράφηση καμένων εκτάσεων είναι σαφώς πιο γρήγορη, οικονομική ενδεχομένως και αποτελεσματική και ακριβής από μετρήσεις πεδίου (γεωδαισία) η χρήση ορθοφωτοχαρτών (φωτογραμμετρία) στις περισσότερες των περιπτώσεων. Πολλές φορές, άλλωστε, αυτό που έχει σημασία δεν είναι η ακριβέστατη ποσοτική χαρτογράφηση και υπολογισμός της έκτασης που καταλαμβάνουν οι καμένες εκτάσεις, αλλά ποιοτικά να μπορεί να δεί ο χρήστης και ο πάσης φύσεως ενδιαφερόμενος το που συνέβη η πυρκαογιά.

Το σημαντικότερο στοιχείο του αλγορίθμου αυτού είναι η δυνατότητα χαρτογράφησης καμένων εκτάσεων σε οποιαδήποτε περιοχή επικρατούν οι ίδιες κλιματικές και γεωμορφολογικές συνθήκες με αυτές από τις οποίες πάρθηκαν τα δεδομένα εκπαίδευσης και ελέγχου. Αντίστοιχα, σε περιοχή αισθητά διαφορετική από αυτή της παρούσας ταξινόμησης απαιτεί τη λήψη νέων δεδομένων από το χρήστη με προσεκτική φωτοερμηνείας στην εκάστοτε περιοχή, αλλά με τη δομή του αλγορίθμου να παραμένει η ίδια.

Η εφαρμογή σε πολύ μικρή κλίμακα έχει ορισμένα αρνητικά τα οποία και έχουν αναλυθεί παραπάνω, αλλά μπορεί να φανεί ιδιαίτερα χρήσιμη σε ορισμένες περιπτώσεις για την εξαγωγή ποσοτικών αλλά και ποιοτικών συμπερασμάτων. Οι αλγόριθμοι μηχανικής μάθησης για ταξιμόηση, άλλωστε, προσφέρουν πολλές δυνατότητες και καινοτομίες για να ξεπεραστούν πολλοί περιορισμοί.

Προτάσεις για περαιτέρω διερεύνηση :

- Αλγόριθμος ενίσχυσης καμένων εκτάσεων (burned area growing algorithm), για εισαγωγή νέων εικονοστοιχείων στα καμένα .
- Προσπάθεια ταξινόμησης με άλλο αλγόριθμο μηχανικής μάθησης πέραν του random forest (SVM , CART, κοκ).
- Για εφαρμογή ταξινόμησης με μικρή κλίμακα, χωρισμός της περιοχής μελέτης σε υποπεριοχές και δημιουργία επιμέρους ψευδοεικόνων σε διαφορετικά χρονικά εύρη όσο το δυνατόν πιο κοντά στο τέλος των πυρκαγιών, για ακριβέστερη ταξινόμηση και καλύτερο αποτέλεσμα.
- Εφαρμογή αλγορίθμου για έλεγχο πληρότητας καύσης των εικονοστοιχείων (combustion completeness) , και μετα- ταξινόμηση των ήδη καμένων εικονοστοιχείων με κατηγορίες ανάλογα με το πόσο καμένα είναι.

Βιβλιογραφία :

Δημητρακόπουλος Κ. (2010), 'Χαρτογράφηση χρήσης/ κάλυψης γης με την αντικειμενοστραφή ταξινόμηση εικόνων SPOT', MSc thesis, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Θεσσαλονίκη

Παρχαρίδης Ι. (2015), "Αρχές Δορυφορικής Τηλεπισκόπησης - Θεωρία και Εφαρμογές", Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκονόμων, Εθνικό Μετσόβιο Πολυτεχνείο

Belgiu, M. and Draǵut L., (2016). "RandomForest in remote sensing: a review of applications and future directions". ISPRS J. Photogramm. Remote Sensing

Μερτίκας Π.Σ. (2006), "Τηλεπισκόπηση και Ψηφιακή Ανάλυση Εικόνας", ΙΩΝ, Περιστέρι, κεφ. Ταξινόμηση Εικόνας

Belgiu, M. and Draǵut L., (2016). "RandomForest in remote sensing: a review of applications and future directions". ISPRS J. Photogramm. Remote Sens.

Friedl M.A. and Brodley C.E. (1997), "Decision Tree Classification of Land Cover from Remotely Sensed Data", Remote Sensing of Environment

Tso B. and Mather P.M. (2009), Classification Methods for Remotely Sensed Data, Taylor & Francis Group, Decision Trees, Sound Parkway NW,, 183-220

Copernicus, 2015. Earth Observation Satellites, Διαθέσιμο στο: <http://www.copernicus.eu/main/satellites>

Digital Globe, Satellite Information. Διαθέσιμο από <https://www.digitalglobe.com/resources/satellite>

https://en.wikipedia.org/wiki/2019%E2%80%9320_Australian_bushfire_season#New_South_Wales

Βλαχάβας Ι κ.α., 2011, Παρουσίαση 'Τεχνητή Νοημοσύνη', Εκδόσεις Παν/μίου Μακεδονίας

ΚΩΔΙΚΕΣ ΠΟΥ ΕΦΑΡΜΟΣΤΗΚΑΝ ΓΙΑ ΤΗΝ ΕΡΓΑΣΙΑ(απαιτείται registration στο Google Earth Engine)

TESTING IN GREECE:

1) TRAINING: KINETA-MATI VALIDATION: KITHIRA - GRAMMATIKO

<https://code.earthengine.google.com/b8ee9dfabc1f12614a90b556d9a8cabcb>

2) TRAINING: KINETA-GRAMMATIKO - VALIDATION: MATI - KITHIRA

<https://code.earthengine.google.com/d73a1983fb7c9f3ffaaf457f1c3de35f>

3) TRAINING : KITHIRA -KINETA VALIDATION: MATI -GRAMMATIKO

<https://code.earthengine.google.com/fcae5b5683f2eb3a5d642ff593649ec4>

4) TRAINING : KITHIRA - GRAMMATIKO VALIDATION: MATI - KINETA

<https://code.earthengine.google.com/e8f626224680f7273b9d68f7d0646152>

5) TRAINING : KITHIRA - MATI VALIDATION: GRAMMATIKO - KINETA

<https://code.earthengine.google.com/3f9245e7237f61b050f15c143833aa59>

6) TRAINING : GRAMMATIKO - MATI VALIDATION: KINETA - KITHIRA

<https://code.earthengine.google.com/8dee3af339c0b7808414b1c300ffdd26>

AUSTRALIA :

<https://code.earthengine.google.com/2a5daff048db3d49d13d84b13cbfb896>

.csv files

<https://code.earthengine.google.com/58460b520c277f897d7a6efcd0698ddf>

calculate area

<https://code.earthengine.google.com/db310d2dda7fc7b996a3848160c138b7>