



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
Δ.Π.Μ.Σ. ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ
Εργαστήριο Τηλεπισκόπησης

**«ΕΤΗΣΙΑ ΧΑΡΤΟΓΡΑΦΗΣΗ ΚΑΛΥΨΗΣ ΓΗΣ ΚΑΙ
ΚΑΛΛΙΕΡΓΕΙΩΝ ΜΕ ΧΡΗΣΗ ΔΕΔΟΜΕΝΩΝ SENTINEL-2
ΚΑΙ ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ»**

Μεταπτυχιακή Εργασία

Αντωνίου Γεωργία

Αθήνα, Φεβρουάριος 2020



**NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF RURAL AND SURVEYING ENGINEERING**

**Department of Topography
Remote Sensing Laboratory**

**« LAND COVER AND CROP-TYPE
MAPPING AT AN ANNUAL BASIS FROM
SENTINEL-2 SATELLITE DATA WITH
MACHINE LEARNING TECHNIQUES »**

Master Thesis

Antoniou Georgia

Athens, February 2020



RSLab

**Remote Sensing Laboratory
National Technical University of Athens**

✓ Sensing ✓ Analytics ✓ Monitoring



...



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ

Τομέας Τοπογραφίας
Εργαστήριο Τηλεπισκόπησης

**« ΕΤΗΣΙΑ ΧΑΡΤΟΓΡΑΦΗΣΗ ΚΑΛΥΨΗΣ ΓΗΣ ΚΑΙ
ΚΑΛΛΙΕΡΓΕΙΩΝ ΜΕ ΧΡΗΣΗ ΔΕΔΟΜΕΝΩΝ SENTINEL-2
ΚΑΙ ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ »**

Μεταπτυχιακή Εργασία
Αντωνίου Γεωργία

Τριμελής εξεταστική επιτροπή:

Κ. Καράντζαλος

Χ. Πότσιου

Δ. Αργιαλάς

.....

.....

.....

Αν. Καθηγητής Ε.Μ.Π.
Επιβλέπων

Καθηγήτρια Ε.Μ.Π.

Καθηγητής Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2020

Αφιερώνεται στους Γονείς μου

Ευχαριστίες

Θα ήθελα να ευχαριστήσω αρχικά τον κ. Κωνσταντίνo Καράντζαλο για την καθοριστική συμβολή του στην εκπόνηση της συγκεκριμένης εργασίας. Θα ήθελα να τον ευχαριστήσω ιδιαίτέρως για την αμεσότητα του και το ευχάριστο κλίμα της συνεργασίας μας. Μέσα από τις γνώσεις και τις συμβουλές του κατάφερα να φέρω εις πέρας την παρούσα εργασία.

Στη συνέχεια, θα ήθελα να ευχαριστήσω τη Χριστίνα Καρακίζη, για τις συμβουλές της, την υπομονή της και την αμέριστη βοήθεια της, που ήταν καθοριστική, σε όλα τα επίπεδα.

Επίσης, θα ήθελα να ευχαριστήσω τον Ζαχαρία Κανδυλάκη, για την τεχνική υποστήριξη που μου παρείχε στο προγραμματιστικό κομμάτι της εργασίας μου και την συμβολή του στην επιλογή και εφαρμογή του πειραματικού πλαισίου.

Ακόμη, θα ήθελα να ευχαριστήσω τον Αλέξανδρο Φάλαγγα και τον Μεμέτ Ουζούν για την σημαντική βοήθειά τους στην προεπεξεργασία των εικόνων και τη συλλογή των δεδομένων εκπαίδευσης.

Τέλος, χρωστάω ένα μεγάλο ευχαριστώ στους γονείς μου για τη στήριξη που μου παρείχαν όλα αυτά τα χρόνια.

Αντωνίου Γεωργία

Αθήνα, Φεβρουάριος 2020

ΠΕΡΙΛΗΨΗ

Η αξιοποίηση των διαθέσιμων τηλεπισκοπικών δεδομένων για την παραγωγή χάρτη χρήσεων/ κάλυψης γης, απαιτεί το σχεδιασμό και την εφαρμογή αποδοτικών μεθόδων και τεχνικών για την αναγνώριση των ποικίλων θεματικών κατηγοριών που καλύπτουν την επιφάνεια της γης. Προς αυτή τη κατεύθυνση, στην παρούσα εργασία σχεδιάστηκε και υλοποιήθηκε ένα σύστημα επεξεργασίας με βάση τα δένδρα απόφασης, με στόχο τον εντοπισμό και το διαχωρισμό των διαφορετικών εδαφοκαλύψεων και τη δημιουργία χάρτη χρήσεων γης και καλλιεργειών. Χρησιμοποιήθηκαν δορυφορικά δεδομένα υψηλής ανάλυσης (Sentinel-2) σε περιοχή (συνολικά ~14.000 km²) της Ανατολικής Μακεδονίας και Θράκης. Τα δορυφορικά δεδομένα αφορούν το έτος 2016 και προέρχονται από 6 διαφορετικές ημερομηνίες. Ο ταξινομητής αξιοποίησε 6 πολυφασματικά κανάλια (Blue, Green, Red, NIR, SWIR1, SWIR2) και 4 φασματικούς δείκτες (NDVI, MSAVI, NDWI, NDBI) για κάθε ημερομηνία. Αρχικά διενεργήθηκε η συλλογή δεδομένων εκπαίδευσης, τα οποία προήλθαν από φωτοερμηνεία στην περιοχή μελέτης. Τα διαθέσιμα δεδομένα αναφοράς για το έτος 2016 περιελάμβαναν 31 κατηγορίες κάλυψης γης, βασισμένες στην ονοματολογία του CORINE αλλά και πολυάριθμες καλλιέργειες από γεωχωρικά δεδομένα του ΟΠΕΚΕΠΕ. Στη συνέχεια, υλοποιήθηκε η ταξινόμηση των πολυφασματικών δορυφορικών δεδομένων. Οι ταξινομήσεις ομαδοποιήθηκαν σε 3 κατηγορίες, με βάση τα tiles που χρησιμοποιούνται για εκπαίδευση και έλεγχο της ταξινόμησης. Η πρώτη κατηγορία ταξινόμησης έκανε χρήση του ίδιου tile για την εκπαίδευση και τον έλεγχο του αλγορίθμου ταξινόμησης, ενώ πεδίο διερεύνησης αποτέλεσε το βέλτιστο ποσοστό αποκοπής των δεδομένων για την εκπαίδευση και τον έλεγχο του ταξινομητή. Στην συγκεκριμένη κατηγορία υλοποιήθηκαν 15 ταξινομήσεις, 9 για κάθε tile με τα ποσοστά των δεδομένων εκπαίδευσης επί του συνόλου δεδομένων αναφοράς να κυμαίνονται από 10% έως 90%. Τα αποτελέσματα που προέκυψαν αξιολογήθηκαν με πολύ υψηλή ακρίβεια (συνολική ακρίβεια >98%). Η δεύτερη κατηγορία ταξινόμησης έκανε χρήση διαφορετικού tile για την εκπαίδευση και για τον έλεγχο του ταξινομητή, διερευνώντας έτσι, τη δυνατότητα δημιουργίας χάρτη χρήσεων γης από ταξινόμηση σε διαφορετικό μοντέλο. Υλοποιήθηκαν δύο ταξινομήσεις, αξιοποιώντας και στις δύο το ίδιο μοντέλο εκπαίδευσης του ταξινομητή στο tie 35TLF. Τα αποτελέσματα, που προέκυψαν ήταν μέτριας ακρίβειας (Συνολική ακρίβεια 64-70%). Η τρίτη κατηγορία ταξινόμησης έκανε χρήση όλων των επιλεγμένων tiles για εκπαίδευση και έλεγχο της ταξινόμησης, χτίζοντας ένα κοινό μοντέλο. Υλοποιήθηκαν 9 ταξινομήσεις, χρησιμοποιώντας σε κάθε μια διαφορετικό ποσοστό αποκοπής δεδομένων εκπαίδευσης, με σκοπό τη διερεύνηση της βέλτιστης επιλογής. Τα αποτελέσματα, που προέκυψαν ήταν πολύ υψηλής ακρίβειας (>99%). Συνολικά, παρήχθησαν, 6 χάρτες χρήσεων γης και καλλιεργειών οι οποίοι βασίστηκαν στο βέλτιστο σχήμα ταξινόμησης της κάθε κατηγορίας. Από τη συνολική ποσοτική και ποιοτική αξιολόγηση της μεθοδολογίας και των αποτελεσμάτων επιβεβαιώνεται η αποτελεσματικότητα του ταξινομητή τυχαίων δασών σε τηλεπισκοπικές εφαρμογές για την χαρτογράφηση κάλυψης γης και καλλιεργειών. Σημειώνεται ότι παρόλα τα υψηλά συνολικά ποσοστά ακρίβειας συγκεκριμένες κατηγορίες επιτυγχάνουν σχετικά χαμηλότερα ποσοστά και αναλύεται η συμπεριφορά τους.

« LAND COVER AND CROP-TYPE MAPPING AT AN ANNUAL BASIS FROM SENTINEL-2 SATELLITE DATA WITH MACHINE LEARNING TECHNIQUES »

Antoniou Georgia

Master Thesis

February 2020

ABSTRACT

The exploitation of the available abundance of remote sensing data for the production of land cover maps requires design and implementation of efficient methods and techniques for highlighting different land cover types. To this end, a framework of random forest classification was designed and implemented in order to identify and separate the different land cover types and finally create a land cover map. We used high-resolution satellite data (Sentinel-2) in a region (totally ~14.000 km²) of Eastern Macedonia & Thrace, Greece. The satellite data came from 6 different dates of 2016. The classifier used six multi-spectral bands (Blue, Green, Red, NIR, SWIR1, SWIR2) and four spectral indices (NDVI, MSAVI, NDWI, NDBI). First step was the collection of training data, which produced by photo interpretation of the study area. The available reference data for the year 2016 included 31 land cover categories, based on the CORINE nomenclature and included also numerous crops derived from Greek Payment Authority of Common Agricultural Policy (OPEKEPE) geospatial data. Subsequently, multiple experiments were conducted. The classifications were grouped into 3 categories, based on the tiles that are used for training and testing algorithm. The first category of classifications used the same tile for training and testing algorithm, changing each time the percentage of ground truth that was used for training algorithm (percentage from 10% to 90%- 27 classifications), the second used different tile for training and testing algorithm (2 classifications), and the third used all tiles for training and testing algorithm, changing each time the percentage of ground truth that was used for training algorithm (percentage from 10% to 90%- 9 classifications). The first group of experiments yielded results of very high precision (98-99%), the second results of medium precision (64-70%) and the third results of very high precision (almost 100%). Totally, 6 land cover maps were produced, based on the optimal classification scheme of each category. From the overall quantitative and qualitative evaluation of the methodology and the results, the effectiveness of the RF techniques for the classification of remote sensing data is confirmed.



RSLab

Remote Sensing Laboratory
National Technical University of Athens

✓ Sensing ✓ Analytics ✓ Monitoring



ΠΕΡΙΕΧΟΜΕΝΑ

1	ΕΙΣΑΓΩΓΗ	1
1.1	Αντικείμενο και Στόχοι της Εργασίας	1
1.2	Κίνητρο και Συνεισφορά της Εργασίας	2
2	ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΚΑΙ ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ	3
2.1	Ανασκόπηση Βιβλιογραφίας.....	3
2.2	Ηλεκτρομαγνητικό Φάσμα και Τηλεπισκόπηση	5
2.2.1	Ορισμός φωτοερμηνείας	5
2.2.2	Ορισμός τηλεπισκόπησης	5
2.2.3	Ηλεκτρομαγνητικό φάσμα.....	5
2.3	Ο Δορυφόρος Sentinel-2.....	7
2.3.1	Γενικά στοιχεία	7
2.3.2	Ιδιαίτερα γνωρίσματα και πλεονεκτήματα	7
2.3.3	Κανάλια.....	7
2.4	Μετασχηματισμοί Πολυφασματικών Εικόνων	12
2.4.1	Γενικά στοιχεία	12
2.4.2	Αριθμητικές πράξεις	12
2.4.3	Δείκτες	13
2.5	Ταξινόμηση Πολυφασματικών Εικόνων	16
2.5.1	Γενικά στοιχεία	16
2.5.2	Μέθοδοι ταξινόμησης.....	16
2.5.3	Ο αλγόριθμος Random Forest.....	19
2.6	Ακρίβεια Ταξινόμησης	23
2.6.1	Γενικά στοιχεία	23
2.6.2	Πίνακας σύγχυσης	24
2.6.3	Μεγέθη αξιολόγησης της ακρίβειας	24
3	ΠΕΡΙΟΧΗ ΜΕΛΕΤΗΣ & ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	27
3.1	Περιοχή Μελέτης.....	27
3.1.1	Γεωγραφία – δημογραφία.....	27
3.1.2	Περιβάλλον	28
3.1.3	Διοικητική διαίρεση.....	29

3.1.4	Οριοθέτηση περιοχής μελέτης.....	29
3.2	Δορυφορικά Δεδομένα και Προεπεξεργασίες	30
3.2.1	Επιλογή δορυφορικών δεδομένων	30
3.2.2	Προεπεξεργασία δορυφορικών δεδομένων	32
3.3	Συλλογή και Προ-επεξεργασία Δεδομένων Αναφοράς Error! Bookmark not defined.	
3.3.1	Οι κατηγορίες μελέτης/ ταξινόμησης.....	33
3.3.2	Συλλογή δεδομένων αναφοράς	38
4	ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΣΧΕΔΙΑΣΜΟΣ ΠΕΙΡΑΜΑΤΩΝ.....	43
4.1	Βήματα Εφαρμογής της Μεθοδολογίας Ταξινόμησης	43
4.2	Κατηγορίες Πειραμάτων	47
4.2.1	Πρώτη κατηγορία πειραμάτων – ανά tile	47
4.2.2	Δεύτερη κατηγορία πειραμάτων – ανά δύο tiles	47
4.2.3	Τρίτη κατηγορία πειραμάτων – ανά τρία tiles.....	47
5	ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ	49
5.1	Πρώτη Κατηγορία Πειραμάτων- Ανά Tile.....	49
5.1.1	Πρώτο πείραμα: ταξινόμηση στο tile 35TLF.	49
5.1.2	Δεύτερο πείραμα: ταξινόμηση στο tile 35TMF	63
5.1.3	Τρίτο πείραμα: ταξινόμηση στο tile 35TMG	76
5.2	Δεύτερη Κατηγορία Πειραμάτων- Ανά Δύο Tiles.....	85
5.2.1	Πρώτο πείραμα: εκπαίδευση στο tile 35TLF και έλεγχος στο Tile 35TMF 85	
5.2.2	Δεύτερο πείραμα: εκπαίδευση στο tile 35TLF και έλεγχος στο tile 35TMF 103	
5.3	Τρίτη Κατηγορία- Ανά Τρία Tiles	116
5.4	Συγκριτική Αξιολόγηση των Πειραμάτων	127
6	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ	135
6.1	Συμπεράσματα.....	135
6.2	Προοπτικές	138
	ΒΙΒΛΙΟΓΡΑΦΙΑ	139
	ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ	141

ΚΑΤΑΛΟΓΟΣ ΧΑΡΤΩΝ	142
ΚΑΤΑΛΟΓΟΣ ΓΡΑΦΗΜΑΤΩΝ	142
ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ.....	143
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	143

1

ΕΙΣΑΓΩΓΗ

Χάρη στην εξέλιξη της τεχνολογίας έχει γίνει δυνατή η καθημερινή και διαρκής παρατήρηση της γης από το διάστημα, εδραιώνοντας την ως το βασικό εργαλείο μελέτης του περιβάλλοντος. Το πλήθος των δορυφόρων που βρίσκονται σε τροχιά γύρω από τη γη είναι ικανό να παρέχει μεγάλο όγκο δεδομένων διαφορετικής φύσης και διακριτικής ικανότητας, τα οποία είναι δυνατό να αξιοποιηθούν πολύπλευρα σε πολλαπλές εφαρμογές που έχουν ως στόχο τη διαχρονική ανάλυση των μεταβολών του περιβάλλοντος, την πρόβλεψη των επιπτώσεων που απορρέουν από αυτές και εν συνεχεία τον κατάλληλο σχεδιασμό αναπτυξιακών προγραμμάτων για την ελαχιστοποίηση τους.

Η παραγωγή χαρτών κάλυψης γης και η παρακολούθηση των γεωργικών δραστηριοτήτων μέσα στον χρόνο αποτελούν βασικές εφαρμογές χρήσεις των δορυφορικών δεδομένων. Η γεωργία ακριβείας αποτελεί την εξέλιξη της συμβατικής γεωργίας. Η τηλεπισκόπηση αξιοποιείται ως ένα βασικό εργαλείο στη γεωργία ακριβείας, με την συμβολή της οποίας μπορούμε να παρακολουθήσουμε την εξέλιξη της ζωντανής βλάστησης κατά τη διάρκεια του βιολογικού κύκλου των φυτών, να αναγνωρίζουμε και να διαχωρίζουμε τις διάφορες καλλιέργειες και να εισάγουμε δεδομένα τηλεπισκόπισης σε αγροκλιματικά μοντέλα

Η παρούσα μεταπτυχιακή εργασία επικεντρώνεται στη διερεύνηση των δυνατοτήτων της ταξινόμησης πολυφασματικών δορυφορικών απεικονίσεων για την εξαγωγή χάρτη πολυάριθμων καλύψεων γης, σε συνδυασμό με την ανάλυση και χαρτογράφηση διαφόρων καλλιεργειών ενώ παράλληλα εξετάζεται και η απόδοση του αλγορίθμου μηχανικής μάθησης, μέσα από πολυάριθμα πειράματα με διαφορετική παραμετροποίηση.

Στις παραγράφους που ακολουθούν αρχικά καθορίζεται και αναλύεται το αντικείμενο και οι στόχοι της παρούσας μεταπτυχιακής εργασίας και μετέπειτα περιγράφεται το κίνητρο και η συνεισφορά της.

1.1 Αντικείμενο και Στόχοι της Εργασίας

Το βασικό αντικείμενο της παρούσας μεταπτυχιακής εργασίας αποτέλεσε η **εφαρμογή και αξιολόγηση** τεχνικών επιβλεπόμενης μηχανικής μάθησης και συγκεκριμένα του αλγορίθμου, Τυχαίων Δασών (RF, Random Forests), για τη **χαρτογράφηση πολυάριθμων κατηγοριών κάλυψης γης και καλλιεργειών** σε τρία διαφορετικά Sentinel-2 tiles (συνολικά ~14.000 km²).

Επιπρόσθετη επιδίωξη της παρούσας εργασίας αποτέλεσε η διερεύνηση της δυνατότητας **γενίκευσης** των μοντέλων με εκπαίδευση του αλγορίθμου από δεδομένα προερχόμενα, από διαφορετική περιοχή μελέτης (διαφορετικό tile), καθώς και η

ανάδειξη του βέλτιστου ποσοστού δεδομένων εκπαίδευσης έναντι δεδομένων ελέγχου για ταξινομήσεις RF.

Πιο συγκεκριμένα, το αντικείμενο της εργασίας αυτής ήταν η κατά το δυνατόν αυτοματοποίηση της παραγωγής ενός χάρτη καλύψεων γης και καλλιεργειών. Συλλέχθηκαν δεδομένα εκπαίδευσης για **31 διαφορετικές θεματικές κατηγορίες**, βάσει των οποίων εκπαιδεύθηκε ο αλγόριθμος.

Πραγματοποιήθηκαν εκτεταμένα **πολυάριθμα πειράματα**, υλοποιήθηκαν, πολλαπλές ταξινομήσεις, σε μερικές από τις οποίες ο αλγόριθμος εκπαιδεύτηκε και ελέγχθηκε με δεδομένα από διαφορετικά tile, ενώ στις άλλες εκπαιδεύτηκε και ελέγχθηκε με δεδομένα από το ίδιο tile.

Επιπλέον, για αυτές τις ταξινομήσεις διερευνήθηκε το **βέλτιστο ποσοστό δεδομένων εκπαίδευσης** μέσα από συγκρίσεις και αναλύσεις, των συνολικών μεγεθών ακρίβειας (συνολική ακρίβεια, δείκτης kappa) και των ατομικών μεγεθών για την κάθε κλάση (ατομική ακρίβεια και αξιοπιστία) και αναδείχθηκε ανά περίπτωση η βέλτιστη ταξινόμηση, βάσει της οποίας δομείται και ο χάρτης.

1.2 Κίνητρο και Συνεισφορά της Εργασίας

Κίνητρο για τη διερεύνηση της εφαρμογής σύγχρονων τηλεπισκοπικών μεθόδων και τεχνικών για την καταγραφή των καλύψεων γης αποτέλεσε η ραγδαία ανάπτυξη σύγχρονων τηλεπισκοπικών αισθητήρων που έχει οδηγήσει στην παραγωγή ενός διαρκώς αυξανόμενου όγκου δεδομένων υψηλής χωρικής και φασματικής ανάλυσης. Επιπλέον, κίνητρο για τη διεξαγωγή της συγκεκριμένης εφαρμογής αποτέλεσε η σύγχρονη τάση για υιοθέτηση νέων μεθόδων ταξινόμησης των δορυφορικών δεδομένων, πέρα από τις κλασικές (επιβλεπόμενη και μη επιβλεπόμενη ταξινόμηση) για τη βελτιστοποιημένη παραγωγή χαρτών χρήσεων γης.

Πιο συγκεκριμένα, η παρούσα εργασία συνεισφέρει στη διεθνή βιβλιογραφία σε τρία επίπεδα:

I. Εφαρμογή και αξιολόγηση μεθοδολογίας RF ταξινόμησης για τη χαρτογράφηση των διαφορετικών καλύψεων γης και καλλιεργειών σε πολυφασματικά δορυφορικά δεδομένα, υψηλής χωρικής ανάλυσης.

II. Μελέτη και αξιολόγηση απόδοσης των ταξινομήσεων με δεδομένα εκπαίδευσης και δεδομένα ελέγχου από διαφορετικές περιοχές μελέτης.

III. Μελέτη και ανάλυση του ποσοστού των δεδομένων εκπαίδευσης έναντι των δεδομένων ελέγχου σε σχέση με την απόδοση του RF αλγόριθμου ταξινόμησης αξιοποιώντας εικόνες από ένα tile, ή αξιοποιώντας εικόνες από περισσότερα του ενός tiles.

2

ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΚΑΙ ΣΧΕΤΙΚΗ
ΒΙΒΛΙΟΓΡΑΦΙΑ

2.1 Ανασκόπηση Βιβλιογραφίας

Στην παράγραφο αυτή γίνεται μια καταγραφή της υπάρχουσας βιβλιογραφίας που σχετίζεται με τις μεθόδους και τις τεχνικές που αναπτύχθηκαν και εφαρμόστηκαν σε αυτή τη μεταπτυχιακή εργασία. Η βιβλιογραφία που παρατίθεται αφορά εφαρμογές χαρτογράφησης καλύψεων γης, καθώς και τηλεπισκοπικές εφαρμογές, που κάνουν χρήση του SVM αλγόριθμου ταξινόμησης.

Οι Stoian et al. (2019) [1], ορμώμενοι από την ριζική μετατόπιση στο πεδίο την επιστήμης της όρασης υπολογιστών, ανέπτυξαν έναν αλγόριθμο τον «FG-UNET» βασισμένο στο U-Net μοντέλο (ένα πλήρως συνελκτικό νευρωνικό δίκτυο), ώστε να μπορέσει να αντεπεξέλθει με αραιά δεδομένα αναφοράς, ενώ διαχειρίζεται υψηλής ανάλυσης δεδομένα εισόδου, με σκοπό την ανάλυση των αποτελεσμάτων στο πλαίσιο της λειτουργικής παραγωγής των χαρτών κάλυψης γης. Περιοχές μελέτης αποτέλεσαν 11 διασκορπισμένες περιοχές σε όλη τη Γαλλία. Χρησιμοποιήθηκαν χρονοσειρές εικόνων Sentinel-2, οι οποίες διορθώθηκαν ατμοσφαιρικά. Τα χαρακτηριστικά που επιλέχθηκαν ήταν 10 -τα κανάλια χωρικής ανάλυσης 10m, τα κανάλια χωρικής ανάλυσης 20m και σε ορισμένα πειράματα προστέθηκαν οι φασματικοί δείκτες NDVI, NDWI και Brightness (μέσος όρος των δέκα καναλιών του δορυφόρου Sentinel-2). Οι κλάσεις της ταξινόμησης ήταν 17. Τα δεδομένα εκπαίδευσης διαχωρίστηκαν τυχαία ώστε να αποτελέσουν δύο διακεκριμένα σύνολα, ένα για εκπαίδευση ποσοστού 67% και ένα για επαλήθευση του αλγορίθμου ταξινόμησης ποσοστού 33%. Έγινε σύγκριση ανάμεσα στον αλγόριθμο «FG-UNET» και σε άλλους εναλλακτικών δικτύων παρόμοιους με αυτόν (UNET, MLP, LSTM, TSFG-UNET). Σε όλες τις περιπτώσεις διαπιστώθηκε ότι ο αλγόριθμος «FG-UNET» απέδωσε εξίσου ή καλύτερα ως προς την ακρίβεια της ταξινόμησης. Σχετικά με την χρησιμότητα των δεικτών δεν παρατηρήθηκε ιδιαίτερη αύξηση στην ακρίβεια της ταξινόμησης και διαπιστώθηκε ότι η πληροφορία που παρέχουν μπορεί να παρθεί και από τα δεδομένα των φασματικών καναλιών του δορυφόρου. Ιδιαίτερη μνεία έγινε για τον αλγόριθμο τυχαίου δάσους «RF», καθώς έγινε σύγκριση των αλγορίθμων ως προς τη δυνατότητα τους να ταξινομήσουν ένα tile του οποίου τα δεδομένα εκπαίδευσης δεν χρησιμοποιήθηκαν στην εκπαίδευση του ταξινομητή. Διαπιστώθηκε ότι σε κάθε περίπτωση η απόδοση πέφτει όταν ένα tile δεν είναι μέρος της εκπαίδευσης, ιδιαίτερα όταν αυτό το tile περιέχει καλύψεις γης διαφορετικές από τα άλλα tile. Ακόμη διαπιστώθηκε ότι ο αλγόριθμος «FG-UNET» πάσχει από υψηλότερη μείωση σε σχέση με τον αλγόριθμο «RF». Σε σχέση με τη γεωμετρική ακρίβεια των δύο αλγορίθμων, διαπιστώθηκε ότι ο pixel-based αλγόριθμος «RF» απέδωσε καλύτερα αποτελέσματα σε σχέση με τον αλγόριθμο «FG-UNET».

Οι Christina Karakizi et al (2018) [2], παρουσίασαν μία μεθοδολογία για την ταξινόμηση ετήσιων υψηλής ανάλυσης δεδομένων σε συγκεκριμένες εξειδικευμένες κλάσεις. Επιλέχθηκαν 27 κλάσεις για τη διεξαγωγή της ταξινόμησης, οι οποίες βασίστηκαν στην ονοματολογία τριών επιπέδων του CORINE Land Cover, με επιπλέον ανάλυση σε εξειδικευμένες κατηγορίες καλλιεργειών. Δεν χρησιμοποιήθηκαν μάσκες νεφοκάλυψης ούτε διαδικασίες παραβολής (interpolation procedure) και δομήθηκαν πειραματικά dataset εικόνων του δορυφόρου Landsat-8, με βαθμιαία αύξηση της νεφοκάλυψης, ώστε να ελεγχθεί η επίδραση της νεφοκάλυψης στα δεδομένα αναφοράς σε σχέση με την ακρίβεια της ταξινόμησης. Για την ταξινόμηση επιλέχθηκαν δύο αλγόριθμοι τεχνητής μάθησης διαφορετικής φύσης, ο αλγόριθμος μηχανών διανυσματικής υποστήριξης (SVM), ο οποίος πρόκειται για έναν αλγόριθμο εποπτικής μάθησης αβαθή πυρήνα, και μία αρχιτεκτονική συνελκτικών νευρωνικών δικτύων (CNN) βαθιάς μάθησης. Για τις ταξινομήσεις χρησιμοποιήθηκαν 11 χαρακτηριστικά, τα 7 φασματικά κανάλια του δορυφόρου Landsat-8 και οι 4 φασματικοί δείκτες οι οποίοι είναι οι NDVI, MSAVI, NDWI, NDBI. Συγκρίνοντας τα αποτελέσματα που απέδωσαν οι δύο ταξινομητές διαπιστώθηκε ότι η ακρίβεια που επιτυγχάνεται είναι της ίδιας τάξης, με την ακρίβεια όμως του ταξινομητή SVM να είναι βέλτιστη κατά 3%. Συνεπώς, ο ταξινομητής SVM κρίνεται ως καταλληλότερος σε σχέση με τη μέθοδο βαθιάς γνώσης, καθώς επιτυγχάνει καλύτερα αποτελέσματα ιδιαίτερα στα σύνορα των κλάσεων. Επιπροσθέτως έχουν μικρότερες απαιτήσεις σε υπολογιστικό κόστος. Ως προς τα επίπεδα νεφοκάλυψης, διαπιστώθηκε ότι, η ακρίβεια είναι ανώτερη (πάνω από 80%) όταν αξιοποιούνται διαχρονικά δεδομένα με νεφοκάλυψη έως 30%.

Οι Marais Sicre et al. (2016) [3], παρουσίασαν μία μέθοδο έγκαιρης αναγνώρισης των θερινών καλλιεργειών εφαρμόζοντας τεχνικές κατωφλίωσης σε χρονοσειρές του φασματικού δείκτη NDVI. Αξιοποιήθηκαν διαχρονικά πολυφασματικά δεδομένα των δορυφόρων Spot και Formosat-2. Αρχικά αναλύθηκε η φαινολογία καλλιεργειών, θερινών και χειμερινών αλλά και βοσκοτόπων, οδηγώντας έτσι στην εύρεση των τιμών κατωφλίωσης του δείκτη NDVI. Οι τιμές αυτές αποτέλεσαν τις παραμέτρους στον αλγόριθμο δέντρο απόφασης (decision tree) για την ανίχνευση των καλοκαιρινών καλλιεργειών (Απρίλιος-Ιούνιος). Η χρήση εικόνων τριών ημερομηνιών εντός της συγκεκριμένης χρονικής περιόδου, έναντι μιας αύξησε την ακρίβεια κατά 10% έως 30% ανάλογα με το έτος εξέτασης. Η συνολική ακρίβεια που επιτεύχθηκε στην ανάδειξη των καλοκαιρινών καλλιεργειών κυμαινόταν από 90% έως 99%. Παράγοντες που επηρέαζαν το αποτέλεσμα ήταν το κλίμα του πρώτου εξαμήνου και η διαθεσιμότητα εικόνων, ανεξαρτήτως δορυφόρου προέλευσης. Συμπερασματικά, διαπιστώθηκε πως η διαθεσιμότητα δεδομένων ήταν ο παράγοντας εκείνος που επηρέασε κατά το μέγιστο βαθμό το αποτέλεσμα της ταξινόμησης και συνεπώς την ανάδειξη των κλάσεων των καλλιεργειών.

Οι Inglada et al. (2015) [4], εξέτασαν 5 αλγορίθμους επιβλεπόμενης ταξινόμησης για την αυτόματη παραγωγή χαρτών καλλιεργειών σε 12 περιοχές μελέτης. Οι ταξινομητές που επιλέχθηκαν απαιτούν συλλογή δεδομένων επιτόπου κατά την εκπαίδευση του ταξινομητή ενώ η παραγωγή του χάρτη γίνεται κατά τρόπο αυτόματο. Οι πολυφασματικές εικόνες που χρησιμοποιήθηκαν προέρχονταν από τους δορυφόρους SPOT4 και Landsat 8. Οι περιοχές μελέτης που επιλέχθηκαν ήταν διασκορπισμένες σε τέσσερις ηπείρους, πιο συγκεκριμένα τέσσερις περιοχές μελέτης ήταν στην Ευρώπη, τέσσερις στην Αφρική, δύο στην Αμερική και δύο στην Ασία. Σκοπός των

πειραμάτων ήταν η σύγκριση των ταξινομητών SVM, Decision Trees, GBT και Random Forests ως προς την ακρίβεια ταξινόμησης. Τα χαρακτηριστικά της ταξινόμησης αποτέλεσαν οι φασματικές τιμές των δορυφορικών καναλιών, οι φασματικούς δείκτες, στατιστικά υψής και χαρακτηριστικά χρονοσειρών. Η μελέτη των αποτελεσμάτων ανέδειξε τον ταξινομητή Random Forest, ως τον καταλληλότερο για ταξινομήσεις με χρονοσειρές, καθώς παρουσίασε τα καλύτερα αποτελέσματα και πιο συγκεκριμένα, συνολική ακρίβεια άνω του 80% για όλες τις περιοχές.

Οι Adam et al. το 2014 [5], μελέτησαν τη χαρτογράφηση προτύπων και τη χωρική κατανομή των κάλυψης γης από δεδομένα προερχόμενα από το δορυφόρο RapidEye. Περιοχή μελέτης αποτέλεσε μιας ετερογενής παραθαλάσσια περιοχή, ενώ για την ταξινόμηση αξιοποιήθηκε ο ταξινομητής τυχαίου δάσους (RF) και ο ταξινομητής διανυσμάτων μηχανικής υποστήριξης (SVM). Η ακρίβεια της ταξινόμησης αξιολογήθηκε μέσα από τους πίνακες σύγχυσης και το δείκτη kappa. Ο ταξινομητής SVM σημείωσε ποσοστό συνολικής ακρίβειας 91,80% με συντελεστή kappa 0.92, ενώ ο ταξινομητής RF σημείωσε συνολική ακρίβεια 93.07% με συντελεστή kappa 0.92. Μέσα από την δειξαγωγή του τεστ McNemar, διαπιστώθηκε πως οι δύο αλγόριθμοι απέδωσαν παρόμοια ($Z = 1.15$). Επιπλέον, ως προς την συνεισφορά των φασματικών καναλιών, διαπιστώθηκε ότι εκείνο με την μέγιστη συνεισφορά ήταν το red edge καθώς παρέχει σημαντική πληροφορία για την ταξινόμηση των καλλιεργειών.

2.2 Ηλεκτρομαγνητικό Φάσμα και Τηλεπισκόπηση

2.2.1 Ορισμός φωτοερμηνείας

Η Φωτοερμηνεία ορίζεται ως ο οπτικός προσδιορισμός της ταυτότητας των αντικειμένων στην επιφάνεια της γης, καθώς και η ανάλυση και κρίση της σημασίας τους από αεροφωτογραφίες. Η φωτοερμηνεία αποτελεί ένα επιστημονικό και τεχνικό εργαλείο απόκτησης, συλλογής, ψηφιακής επεξεργασίας, ανάλυσης, αξιολόγησης και αξιοποίησης ολοκληρωμένων μετρητικών και ποιοτικών πληροφοριών για τη Γη και το Περιβάλλον από μακριά. Η εμπειρία, η λογική κι η επιστημονική γνώση, είναι ουσιαστικής σημασίας εργαλεία στην Φωτοερμηνευτική Μεθοδολογία [6].

2.2.2 Ορισμός τηλεπισκόπησης

Η τηλεπισκόπηση είναι η επιστήμη και η τέχνη της απόκτησης πληροφοριών για ένα αντικείμενο, μια περιοχή ή ένα φαινόμενο δια μέσου της ανάλυσης δεδομένων, τα οποία έχουν αποκτηθεί από μια συσκευή, η οποία δεν είναι σε επαφή με το υπό έρευνα αντικείμενο, περιοχή ή φαινόμενο [7]. Τα αντικείμενα του περιβάλλοντος εκπέμπουν ηλεκτρομαγνητική ακτινοβολία. Η ακτινοβολία αυτή συλλέγεται χρησιμοποιώντας δείκτες ηλεκτρομαγνητικής ενέργειας. Παράγονται έτσι δεδομένα για τα επίπεδα της ηλεκτρομαγνητικής ενέργειας στις διάφορες περιοχές του φάσματος. Η ανάλυση των δεδομένων οδηγεί στην παραγωγή πληροφοριών για τις διάφορες συνιστώσες του περιβάλλοντος, αναλόγως με τα αντικείμενα ενδιαφέροντος.

2.2.3 Ηλεκτρομαγνητικό φάσμα

Το ηλεκτρομαγνητικό φάσμα είναι η οργάνωση της ηλεκτρομαγνητικής ακτινοβολίας ανάλογα του μήκους κύματος.

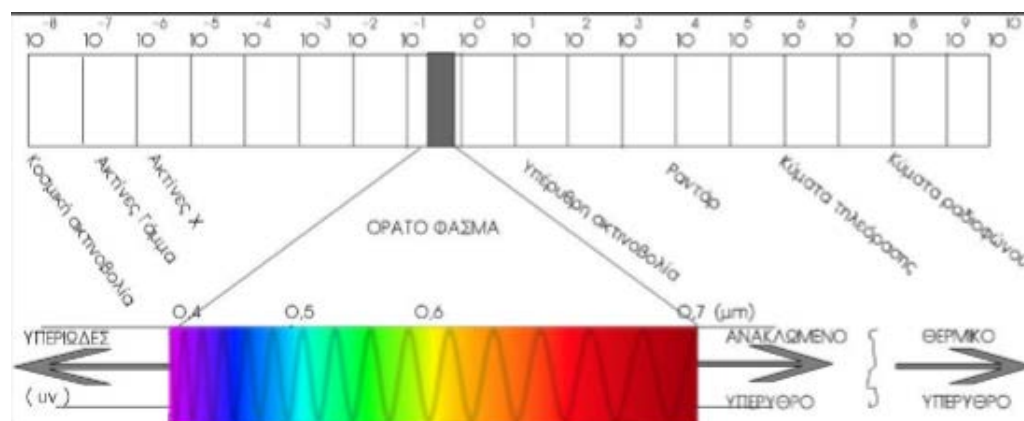
Η ηλεκτρομαγνητική ακτινοβολία (ή ηλεκτρομαγνητική ενέργεια) αποτελείται από κύματα ηλεκτρικού και μαγνητικού πεδίου. Τα κύματα αυτά διαδίδονται στον χώρο λόγω της περιοδικής διαταραχής ηλεκτρικού και μαγνητικού πεδίου με την ταχύτητα το φωτός ($c = 2.99792458 \times 10^8 \text{ m/s}$). Τα βασικά χαρακτηριστικά τους είναι η συχνότητά τους και το μήκος κύματος, το γινόμενο των οποίων ισούται με την ταχύτητα του φωτός. Τα ηλεκτρομαγνητικά κύματα ποικίλλουν ως προς το μήκος κύματος. Απόρροια αυτού είναι οι πολλαπλές κατηγορίες κατάταξής τους.

Τα ηλεκτρομαγνητικά κύματα μεταδίδονται σε κάποιο μέσο (ατμόσφαιρα, νερό, υλικά σώματα) λόγω της περιοδικής διαταραχής του ηλεκτρικού και του μαγνητικού πεδίου συγχρόνως, με την ταχύτητα του φωτός που είναι $c = 2.99792458 \times 10^8 \text{ m/s}$, και χαρακτηρίζονται από την συχνότητα και το μήκος κύματός τους, το γινόμενο των οποίων δίνει την ταχύτητα του φωτός. Το ηλεκτρομαγνητικό φάσμα περιέχει ένα μεγάλο εύρος διαφορετικών μηκών κύματος, εκ των οποίων μόνον ένα μικρό εύρος μηκών κυμάτων καταλαμβάνει το ορατό (μεταξύ 400 και 700nm).

Έχει κατηγοριοποιηθεί σε:

- Ράδιο-κύματα που έχουν μήκος κύματος από 105m έως μερικά εκατοστά.
- Μικροκύματα που έχουν μήκος κύματος από 30cm έως 1mm
- Υπέρυθρες ακτίνες που έχουν μήκος κύματος από 1mm έως $7 \times 10^{-7} \text{ m}$
- Ορατό φως το κομμάτι εκείνο του ηλεκτρομαγνητικού φάσματος που είναι ορατό από το ανθρώπινο μάτι και δυνατό να καταγραφεί από τους τηλε-απεικονιστές. Το μήκος κύματος του ορατού φωτός κυμαίνεται από 400 nm έως 700 nm
- Υπεριώδεις ακτίνες που έχουν μήκος κύματος από $3 \times 10^{-7} \text{ m}$ έως $6 \times 10^{-8} \text{ m}$
- Ακτίνες X είναι ηλεκτρομαγνητική ακτινοβολία με μήκη κύματος από 10-8 m έως 10-13 m
- Ακτίνες γ με μήκη κύματος από 10-10 m και φτάνουν ως τα 10-14 m.

Μία πιο γενική σχηματική κλίμακα του ηλεκτρομαγνητικού φάσματος γενικά και ειδικότερα του ορατού παρουσιάζεται στην εικόνα 1. Σημειώνεται ότι το μήκος κύματος μετριέται από μια κορυφή ενός κύματος έως την επόμενη και συμβολίζεται με το ελληνικό γράμμα λ . [8]



Εικόνα 2-1 Το ηλεκτρομαγνητικό φάσμα. Πηγή: Περράκης, 2016[9]

2.3 Ο Δορυφόρος Sentinel-2

2.3.1 Γενικά στοιχεία

Ο δορυφόρος Sentinel-2, είναι μια αποστολή παρακολούθησης της Γης που αναπτύχθηκε από τον Ευρωπαϊκό Οργανισμό Διαστήματος (ESA), ως μέρος του Προγράμματος Copernicus, για την εκτέλεση επίγειων παρατηρήσεων, οι οποίες είναι αναγκαίες για την υποστήριξη υπηρεσιών όπως η παρακολούθηση των δασών, η ανίχνευση αλλαγών κάλυψης γης και η διαχείριση φυσικών καταστροφών. Αποτελείται από δύο ταυτόσημους δορυφόρους, τους Sentinel-2A και Sentinel-2B. 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 180ο 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.[10]

2.3.2 Ιδιαίτερα γνωρίσματα και πλεονεκτήματα

Η αποστολή Sentinel-2 έχει τις ακόλουθες δυνατότητες:

Πολλαπλά φασματικά δεδομένα με 13 ζώνες στο ορατό, στο κοντινό υπέρυθρο και στο μικροκυματικό υπέρυθρο τμήμα του φάσματος. (υπέρυθρη ακτινοβολία μικρού κύματος)

Συστηματική συνολική κάλυψη των επιφανειών της γης από τις 56 ° N έως τις 84 ° B, τα παράκτια ύδατα και όλη τη Μεσόγειο Θάλασσα

Επανεξέταση κάθε 5 ημέρες υπό τις ίδιες γωνίες θέασης. Σε μεγάλα γεωγραφικά πλάτη, ο Sentinel-2 επικαλύπτει ορισμένες περιοχές, οι οποίες παρατηρούνται δύο φορές ή και περισσότερες κάθε 5 ημέρες, αλλά με διαφορετικές οπτικές γωνίες θέασης. Χωρική ανάλυση 10 m, 20 m και 60 m.

Πεδίο προβολής στα 290km.

Δωρεάν και ανοιχτή πολιτική δεδομένων.

2.3.3 Κανάλια

Ο σχεδιασμός του πολυφασματικού οργάνου του δορυφόρου Sentinel-2, βασίστηκε στην ανάγκη για μεγάλο εύρος πληροφορίας, υψηλής γεωμετρικής και φασματικής ακρίβειας. Έτσι, χρησιμοποιήθηκαν 13 φασματικά κανάλια. Τα κανάλια αυτά καλύπτουν μεγάλο εύρος του ηλεκτρομαγνητικού φάσματος, από τα 443.9 έως τα 2202.4nm και παρουσιάζονται αναλυτικά στην συγκεκριμένη παράγραφο.

1 Coastal Aerosol

Όπως υποδηλώνει και το όνομα του συγκεκριμένου καναλιού, το κανάλι 1 (η παράκτια ζώνη αερολύματος) είναι ιδιαίτερα χρήσιμη σε παράκτιες μελέτες, σε βαθυμετρικές μελέτες και μελέτες αερολυμάτων. Η παράκτια ζώνη αντανακλά το μπλε και το βιολετί

φως και γίνεται όλο και πιο κοινός στους δορυφορικούς αισθητήρες. Μερικές από τις εφαρμογές της φασματικής μπάντας είναι οι εξής παρακάτω:

- Μέτρηση του πληθυσμού φαλαινών: Λόγω της ικανότητάς της συγκεκριμένης ακτινοβολίας να διεισδύει στο νερό (έως 20-30m σε καθαρό νερό), το κανάλι 1, χρησιμοποιείται για τον υπολογισμό των πληθυσμών φαλαινών που βρίσκονται κάτω από την επιφάνεια.
- Βαθυμετρία -:Επειδή το κανάλι 1 απορροφάται λιγότερο από το νερό, επιτρέπει την παρατήρηση του παράκτιου νερού και των ωκεανών. Αυτό είναι επίσης χρήσιμο για τη δημιουργία φασματικών υπογραφών για τη θαλάσσια βλάστηση.
- Αερόλυμα - Επειδή το κανάλι 1 είναι πιο ευαίσθητο στα σύννεφα, τον καπνό και τη θολότητα, χρησιμοποιείται για το φιλτράρισμα των νεφών σε επεξεργασία εικόνων.

2 Μπλε

Τα μάτια μας μπορούν να δουν το ανακλώμενο μπλε φως. Ανάλογα με τη διαύγεια του νερού, το ορατό φως, του μπλε καναλιού, μπορεί να φτάσει σε βάθη μέχρι 20-30 μέτρα. Χρησιμοποιώντας τη μπλε ζώνη είναι δυνατόν να εξασφαλιστεί η φασματική υπογραφή για το βαθύ νερό, ενώ άλλες εφαρμογές στις οποίες χρησιμοποιείται είναι:

- Απεικόνιση του νερού : Λόγω του λεπτότερου επιπέδου φασματικής διείσδυσης στις παράκτιες περιοχές, υποβρύχια ύφαλοι, θολότητα νερού και ιζήματα, βυθισμένη υδρόβια βλάστηση, θολερότητα και βαθυμετρική χαρτογράφηση είναι δυνατό να ανιχνευθούν σε εφαρμογές τηλεπισκόπησης μέσα από την χρήση του καναλιού 2.
- Ανίχνευση των ακτινών καπνού, της θολότητας της ατμόσφαιρας και των σύννεφων: Η ορατή περιοχή μπλε χαρακτηρίζεται από μικρότερα μήκη κύματος (0,45-0,51μm), τα οποία είναι πιο ευαίσθητα στην ατμόσφαιρα. Αυτό καθιστά την μπλε ζώνη ιδανική για την ανίχνευση καπνών, επειδή τα μικρότερα μήκη κύματος διασπείρονται πιο εύκολα από τα μικρότερα σωματίδια.
- Διάκριση σύννεφων, χιονιού και βράχων : Με την χρήση του καναλιού 2 διαχωρίζονται τα σύννεφα από το χιόνι και το βράχο επειδή η φασματική υπογραφή της ομίχλης και των λεπτών σύννεφων είναι ιδιαίτερα ευαίσθητη σε μπλε και κόκκινα μήκη κύματος.

3 Πράσινο

Το πράσινο, όπως και στην καθημερινότητα, έτσι και στην τηλεπισκόπηση σημαίνει φύση - φυτά, δέντρα και δάση. Το πράσινο τμήμα καλύπτει την κορυφή ανάκλασης από τις επιφάνειες των φύλλων. Αυτό σημαίνει ότι η φασματική ανακλαστικότητα είναι χαμηλή στις μπλε και κόκκινες περιοχές του φάσματος λόγω της απορρόφησης από τη χλωροφύλλη κατά τη διάρκεια της φωτοσύνθεσης. Το πράσινο κανάλι χρησιμοποιείται ευρέως και σε διάφορες εφαρμογές όπως:

- Εκτίμηση της υγείας και ακμαιοτήτας των φυτών: Το πράσινο κανάλι χρησιμοποιείται για να διακρίνει τις μεγάλες κατηγορίες βλάστησης, το φυτικό υλικό, αλλά και την υγεία του φυτού. Η φασματική υπογραφή της κορυφής ανάκλασης από τις επιφάνειες των φύλλων επισημαίνεται στην πράσινη ζώνη.

•Διάκριση φυτών αλγών και κυανοβακτηρίων: Το καθαρό νερό γενικά δεν αντανακλάται σε κανένα μέρος του φάσματος, εκτός από το μπλε άκρο του φάσματος, εξ ου και το μπλε χρώμα που βλέπουμε τα μάτια μας. Αλλά όταν είναι παρόντα ανθοφόρα φυτά, στο πράσινο εντοπίζεται η μέγιστη ανακλαστικότητα.

•Χωροθέτηση σημείων αστικής αναψυχής : Οι χλοοτάπητες και τα χόρτα στις αστικές περιοχές ενδείκνυνται για χώρους αναψυχής όπως πάρκα, γήπεδα γκολφ και νεκροταφεία.

4 Κόκκινο

Στα τροπικά εδάφη, στο δομημένο περιβάλλον και στα γεωλογικά χαρακτηριστικά γενικά υπάρχει μεγάλη ανάκλαση της ηλεκτρομαγνητικής ακτινοβολίας του κόκκινου χρώματος, με αποτέλεσμα να αποτελεί βασική συνιστώσα στο διαχωρισμό και την ανάδειξή τους. Επιπλέον, το κόκκινο χρησιμοποιείται σε τύπους όπως ο Κανονικοποιημένος Δείκτης Βλάστησης Διαφοράς (NDVI), διότι τα φυτά απορροφούν σε μεγάλο βαθμό το κόκκινο (αλλά και το μπλε) φως. Πιο συγκεκριμένα, το κόκκινο κανάλι χρησιμοποιείται σε διάφορες εφαρμογές, όπως:

•Διάκριση εδαφών και Πρόβλεψη των ιδιοτήτων τους: Οι διαφορετικοί τύποι εδαφών έχουν και διαφορετικά χαρακτηριστικά. Η ανακλαστικότητα του γυμνού εδάφους εξαρτάται γενικά από τη σύνθεσή του π.χ. τα εδάφη που είναι πλούσια σε οξείδιο του σιδήρου έχουν υψηλή ανάκλαση στην κόκκινη φασματική ζώνη.

•Διάκριση του ανθρωπογενούς από το φυσικό περιβάλλον: Η κόκκινη ζώνη έχει χρησιμοποιηθεί για τη διάκριση μεταξύ τεχνητών αντικειμένων και βλάστησης, καθώς οι δρόμοι και οι ανθρωπίνι οικισμοί εντοπίζονται καλύτερα στην κόκκινη ζώνη.

•Απορρόφηση χλωροφύλλης: Η υγιής βλάστηση ανακλά περισσότερο την υπέρυθη ακτινοβολία (NIR) και το πράσινο φως σε σύγκριση με τα άλλα μήκη κύματος, ενώ απορροφά περισσότερο κόκκινο και μπλε φως. Έτσι, το κόκκινο κανάλι χρησιμοποιείται για την ποσοτικοποίηση της βλάστησης στο NDVI.

5, 6 & 7 Red-Edge

Τα κανάλια «κόκκινου άκρου» βρίσκονται ανάμεσα στην κοντινή υπέρυθη και την κόκκινη ζώνη. Για την χλωροφύλλη, το κόκκινο τμήμα απορροφά έντονα το κόκκινο φως, ενώ ανακλά την ηλεκτρομαγνητική ακτινοβολία του εγγύς υπέρυθρου. Στη μετάβαση μεταξύ αυτών των δύο φασματικών ζωνών βρίσκεται η ζώνη κόκκινου άκρου, η οποία χρησιμοποιείται σε διάφορες εφαρμογές:

•Εύρεση της φυτικής υγείας και κατάταξη σε ηλικιακές τάξεις : Καθώς τα υγιή και ανεπτυγμένα φυτά, διαθέτουν μεγαλύτερη ποσότητα χλωροφύλλης, ανακλούν πιο έντονα από τα ασθενικά φυτά στο τμήμα Red-Edge, του ηλεκτρομαγνητικού φάσματος.

•Παρακολούθηση ανάπτυξης καλλιεργειών: Μέσα από την χρήση του καναλιού Red-Edge, είναι δυνατή η παρακολούθηση της ανάπτυξης των φυτών, καθώς διακρίνεται η ηλικία και η κατάσταση υγείας των φυτών, με αποτέλεσμα να μπορούν να διακριθούν οι διαφορετικοί τύποι φυτών, οι διαφορετικές καταστάσεις υγείας και να επιλεγθούν τα κατάλληλα για τις διατροφικές ανάγκες.

8 & 8A Εγγύς υπέρυθρο – NIR

Η ανάκλαση της εγγύς υπέρυθρης ακτινοβολίας είναι ένας από τους ισχυρότερους τρόπους ταξινόμησης της υγιούς βλάστησης. Ο διαχωρισμός κλάσεων όπως το νερό και η βλάστηση είναι πάντα ευκολότερος στην περιοχή της εγγύς υπέρυθρης ακτινοβολίας. Αυτό οφείλεται στο γεγονός ότι τα υγιή φυτά αντανακλούν την υπέρυθρη ακτινοβολία, ενώ το νερό την απορροφά. Η συγκεκριμένη ιδιότητα οδηγεί σε πολλές εφαρμογές του καναλιού NIR:

- Ανάδειξη της υγιούς βλάστησης και του περιεχομένου της βιομάζας : Η εσωτερική δομή της υγιούς χλωροφύλλης ανακλά την εγγύς υπέρυθρη ακτινοβολία. Αλλά όταν τα φυτά μαραίνονται και τελικά πεθαίνουν, η απόκριση στην εγγύς υπέρυθρη ακτινοβολία μειώνεται έως ότου να εκμηδενιστεί.

- Ανάδειξη αρχαιολογικών χώρων: Μέσα από την χρήση του καναλιού NIR, είναι ευκολότερη η ανεύρεση αρχαίων αρχαιολογικών χώρων ερμηνεύοντας τα πυκνότερα τούβλα λάσπης, τα σημάδια των καλλιεργειών και τις λεπτές διαφορές στη βλάστηση, το έδαφος και τη γεωλογία.

- Κανονικοποιημένος Δείκτης Βλάστησης Διαφοράς (NDVI) : Η εγγύς υπέρυθρη ακτινοβολία χρησιμοποιείται για τη μέτρηση της υγείας των φυτών. Αυτό μπορεί να γίνει με τη χρήση του λόγου NDVI, που αποτελείται από εγγύς υπέρυθρες και κόκκινες ζώνες. Το NDVI εξάγει τις διαφορές της βλάστησης με μεγαλύτερη ακρίβεια απ' ό, τι εάν εξεταστεί μόνο το ορατό πράσινο χρώμα.

9 Water Vapor (Υδρατμοί)

Υπάρχει σημαντική αλληλεπικάλυψη μεταξύ των καναλιών 8 και 9 (εγγύς υπέρυθρης ακτινοβολίας και υδρατμών). Η κύρια διαφορά έγκειται στην μικρότερη επίδραση του καναλιού 9 από την ατμοσφαιρική επιρροή. Αυτή η ιδιότητα του συγκεκριμένου καναλιού, το καθιστά απαραίτητο σε ορισμένες εφαρμογές:

- Μελέτες βλάστησης : Η χρήση του καναλιού 9, παρέχει πιο εξελιγμένη ανάλυση βλάστησης και μελέτη της βιομάζας, καθώς η επίδραση από την ατμόσφαιρα είναι μικρότερη σε σχέση με το κανάλι του εγγύς υπέρυθρου.

- Ανάδειξη των ορίων μεταξύ εδάφους και νερού : Στο συγκεκριμένο κανάλι το νερό είναι ένας ισχυρός απορροφητής, ενώ η βλάστηση είναι ένας ισχυρός ανακλαστήρας, με αποτέλεσμα να είναι ευκολότερη η οριοθέτηση των δασών και η ποσοτικοποίηση της απώλειας ή της αύξησης δασικών εκτάσεων.

10 SWIR –Cirrus

Το κανάλι cirrus κέρδισε το όνομά του, επειδή είναι ειδικό στην ανίχνευση σύννεφων. Η ατμόσφαιρα απορροφά σχεδόν όλη την ηλεκτρομαγνητική ακτινοβολία της συγκεκριμένης ζώνης του φάσματος, καθώς το έδαφος είναι μόλις ορατό. Τα σύννεφα υψηλού υψομέτρου ανιχνεύονται και διαχωρίζονται. Η ιδιότητα αυτή εντοπίζεται μόνο στο συγκεκριμένο κανάλι και έτσι χρησιμοποιείται σε αντίστοιχες εφαρμογές:

- Ανίχνευση σύννεφων: Τα σύννεφα αντανακλούν έντονα στο κανάλι 10, ενώ οι περισσότερες επιφάνειες γης θα εμφανιστούν σκοτεινές. Η βελτιωμένη ανίχνευση του

νέφους κύκλων βοηθά στην αφαίρεση του νέφους και στις βελτιώσεις των χάρτες απεικόνισης.

11 & 12 Μικροκυματικό Υπέρυθρο- SWIR

Το υπέρυθρο βραχέων κυμάτων έχει τις ομοιότητές του με το SWIR-Cirrus. Το SWIR χρησιμοποιείται κυρίως για την απεικόνιση τύπων εδάφους, γεωλογικών χαρακτηριστικών και ορυκτών, όπως ο χαλκός και τα θειικά άλατα. Είναι επίσης ευαίσθητο στις διακυμάνσεις της βλάστησης και της υγρασίας του εδάφους. Επιπλέον, το χιόνι, ο πάγος και τα σύννεφα φαίνονται πιο σκούρα και αναδεικνύονται. Οι συγκεκριμένες ιδιότητες του καναλιού SWIR, το καθιστούν χρήσιμο σε διάφορες εφαρμογές όπως:

- Εφαρμογές ανάδειξη των χαρακτηριστικών και των ιδιοτήτων του νερού: Το νερό απορροφάει έντονα την SWIR ακτινοβολία. Αυτό βοηθάει στην ανάδειξη του αλλά και στη δημιουργία των φασματικών υπογραφών της θαλάσσιας βλάστησης και των θολών υδάτων.

- Εύρεση των κατάλληλων πρακτικών άρδευσης: Η ανάκλαση των υπέρυθρων ακτινών βραχέων κυμάτων επηρεάζεται από την περιεκτικότητα σε νερό των φύλλων, καθιστώντας την ιδανική για να διαπιστώσουμε το βαθμό ξηρασίας των καλλιεργειών και να στοχεύσουμε στις κατάλληλες τεχνικές άρδευσης.

- Χαρτογράφηση ορυκτών: Το κανάλι SWIR καλύπτει ένα εύρος ηλεκτρομαγνητικής ακτινοβολίας, όπου οι διαφορετικοί τύποι ορυκτών αργίλου παράγουν μια εντονότερη και λεπτομερέστερη φασματική υπογραφή, με αποτέλεσμα να μπορούν να αναδειχθούν και να χαρτογραφηθούν τα διάφορα ορυκτά. Ακόμη, τα υδρόφιλα ορυκτά εμφανίζονται πιο σκούρα, με αποτέλεσμα να είναι δυνατός ο μεταξύ των διαφορετικών τύπων καολινίτη και κοκκινίτη διαχωρισμός.

Πίνακας 2-1 Κανάλια Sentinel-2

Κανάλια	S2A		S2B		Χωρική Ανάλυση(m)
	Κεντρικό Μήκος Κύματος(nm)	Εύρος Μήκους Κύματος	Κεντρικό Μήκος Κύματος(nm)	Εύρος Μήκους Κύματος	
Band 1 – Coastal aerosol	443.9	27	442.3	45	60
Band 2 – Blue	496.6	98	492.1	98	10
Band 3 – Green	560.0	45	559	46	10
Band 4 – Red	664.5	38	665	39	10
Band 5 – Vegetation Red Edge	703.9	19	703.8	20	20
Band 6 – Vegetation Red Edge	740.2	18	739.1	18	20

Band 7 – Vegetation Red Edge	782.5	28	779.7	28	20
Band 8 – NIR	835.1	145	833	133	10
Band 8A – Narrow NIR	864.8	33	864	32	20
Band 9 – Water vapour	945.0	26	943.2	27	60
Band 10 – SWIR – Cirrus	1373.5	75	1376.9	76	60
Band 11 – SWIR	1613.7	143	1610.4	141	20
Band 12 – SWIR	2202.4	242	2185.7	238	20

2.4 Μετασχηματισμοί Πολυφασματικών Εικόνων

2.4.1 Γενικά στοιχεία

Οι μετασχηματισμοί πολυφασματικών εικόνων είναι τεχνικές επεξεργασίας των δορυφορικών δεδομένων με σκοπό τη δημιουργία νέων εικόνων οι οποίες να παρέχουν περισσότερες πληροφορίες από τις αρχικές εικόνες. Όπως αναφέρει ο Αργιαλάς (1999) [11], πρόκειται για τα νέα στοιχεία ενός διανύσματος *pixel*, που σχετίζονται με τις παλιές τιμές φωτεινότητας της αρχικής ομάδας φασματικών καναλιών, μέσω ενός γραμμικού μετασχηματισμού. Ουσιαστικά οι μετασχηματισμένες εικόνες επιδιώκουν μία εναλλακτική περιγραφή των δεδομένων με τις νέες συνιστώσες και μπορούν να αναδείξουν χαρακτηριστικά του χώρου τα οποία δεν ήταν διακριτά στις αρχικές εικόνες ή να συμπιέσουν μεγάλο όγκο πληροφοριών σε μικρότερο αριθμό καναλιών. Οι μετασχηματισμοί αυτοί, σύμφωνα με τον Περάκη[9] (όπως αναφέρει ο Richards 1986), εφαρμόζονται είτε σε μία πολυφασματική εικόνα είτε σε δύο ή περισσότερα φασματικά κανάλια, της ίδιας περιοχής, που έχουν αποκτηθεί σε διαφορετικές χρονικές στιγμές (διαχρονικά δεδομένα). Η ψηφιακή επεξεργασία εικόνας παρέχει μεγάλο εύρος δυνατών μετασχηματισμών των πολυφασματικών δορυφορικών δεδομένων. Μερικές από τις πιο γνωστές τεχνικές μετασχηματισμού των εικόνων είναι:

Οι αριθμητικές πράξεις. Αναφέρονται στην πρόσθεση, στην αφαίρεση, στον πολλαπλασιασμό και στην διαίρεση (λόγος) των εικόνων.

Η ανάλυση σε κύριες συνιστώσες (*principal component analysis-PCA*).

Ο μετασχηματισμός *Kauth-Thomas*.

2.4.2 Αριθμητικές πράξεις

Οι αριθμητικές πράξεις της πρόσθεσης, της αφαίρεσης, του πολλαπλασιασμού και της διαίρεσης μπορούν να εφαρμοστούν σε δύο ή περισσότερες εικόνες που έχουν το ίδιο σύστημα αναφοράς και απεικονίζουν την ίδια γεωγραφική περιοχή. Οι πράξεις υλοποιούνται μεταξύ των ψηφιακών τιμών των εικονοστοιχείων των φασματικών καναλιών χωρίς να εμπλέκονται τα γειτονικά. Οι αριθμητικές πράξεις είναι δυνατόν να

τελεστούν σε φασματικά κανάλια της ίδιας δορυφορικής εικόνας ή σε φασματικά κανάλια εικόνων διαφορετικής ημερομηνίας[9].

2.4.3 Δείκτες

Οι δείκτες αποτελούν τεχνικές φασματικής ενίσχυσης των εικόνων. Είναι εικόνες οι οποίες προκύπτουν μετά από μαθηματικές πράξεις μεταξύ των φασματικών καναλιών της ίδιας εικόνας ή διαφορετικών εικόνων[9]. Οι μαθηματικές πράξεις μπορεί να είναι είτε απλές όπως «Κανάλι 1 – Κανάλι 2» είτε πιο σύνθετες όπως λόγος καναλιών «Κανάλι 1 / Κανάλι 2». Οι λόγοι καναλιών βασίζονται ιδιαίτερα στις φασματικές ιδιότητες των υλικών του εδάφους, καθώς τα υλικά απορροφούν ή ανακλούν ηλεκτρομαγνητική ακτινοβολία διαφορετικά στα διάφορα μήκη κύματος. Η απορρόφηση εξαρτάται από τη μοριακή δομή της επιφάνειας που προσπίπτει η ηλεκτρομαγνητική ακτινοβολία. Ως αποτέλεσμα προκύπτει ότι οι λόγοι των φασματικών καναλιών δίνουν πληροφορίες για τη χημική σύσταση του κάθε στόχου και είναι ικανοί μέσα από την εφαρμογή τους να αναδείξουν διαφορετικά είδη βλάστησης, να εντοπίσουν ορυκτά ή αστικές περιοχές, περιοχές με ξηρασία, καμένη βλάστηση, ή την ύπαρξη νερού, χιονιού και πάγου.

Οι δείκτες, ανάλογα με την μέθοδο που εφαρμόζεται για τη δημιουργία τους, ταξινομούνται σε 4 κατηγορίες:

Απλοί δείκτες. Οι δείκτες αυτοί δημιουργούνται από τους λόγους των φασματικών καναλιών και τονίζουν τις φασματικές ιδιότητες των καλύψεων γης.

Δείκτες που ελαχιστοποιούν τα προβλήματα που δημιουργούνται εξαιτίας του εδάφους σε σχέση με τις καλύψεις (βλάστηση, νερό κ.ά.).

Δείκτες που ελαχιστοποιούν τις ατμοσφαιρικές επιδράσεις.

Εξειδικευμένοι δείκτες εκτίμησης φαινομένων που συμβαίνουν στο χώρο. Χαρακτηριστικό παράδειγμα αποτελούν οι δείκτες: πυκνότητας βλάστησης, εκτίμησης της δριμύτητας της πυρκαγιάς, κ.ά.

Ένας δείκτης για να είναι λειτουργικός πρέπει να ακολουθεί μια σειρά από κριτήρια όπως να μεγιστοποιεί την ευαισθησία σε επιφάνειες με συγκεκριμένα χαρακτηριστικά (π.χ. στη χλωροφύλλη της βλάστησης), να μεταβάλλεται γραμμικά στις αλλαγές που συντελούνται και να έχει εφαρμογή σε όσο το δυνατόν μεγαλύτερο εύρος περιοχών. Ακόμη θα πρέπει να εξομαλύνει και να ελαττώνει τις επιπτώσεις εξαιτίας της γωνίας του ήλιου, της τοπογραφίας, των ατμοσφαιρικών συνθηκών, των σφαλμάτων των αισθητήρων, επιτρέποντας διαχρονικές συγκρίσεις στην περιοχή μελέτης. Τέλος, χρήσιμη είναι η σύνδεσή του με διαδικασίες υπολογισμού ποσοτικών παραμέτρων όπως : ο υπολογισμός της βιομάζας, του δείκτη φυλλικής επιφάνειας κ.ά.

Στην συνέχεια παραθέτονται οι δείκτες οι οποίοι χρησιμοποιήθηκαν στην παρούσα εργασία. Οι δείκτες που επιλέχθηκαν ήταν 4, ο NDVI, ο MSAVI, ο NDWI και ο NDBI. Οι δύο δείκτες από τους τέσσερεις είναι δείκτες βλάστησης και βασίζονται στην αλληλεπίδραση της ηλεκτρομαγνητικής ακτινοβολίας με τα φύλλα των φυτών. Ο τρίτος δείκτης είναι ένας δείκτης νερού, που διαχωρίζει τα υδάτινα σώματα από άλλα χαρακτηριστικά κάλυψης γης και ο τέταρτος δείκτης είναι ένας δείκτης οικιστικής ανάπτυξης. Οι δείκτες αυτοί υπαγορεύτηκαν από παρόμοιες εργασίες οι οποίες απέδωσαν ακριβή αποτελέσματα [12], [13].

Δείκτης Βλάστησης NDVI

Ο πιο δημοφιλής και αξιόπιστος δείκτης βλάστησης είναι ο δείκτης NDVI. Όπως αναφέρει ο Περάκης[9], ο κανονικοποιημένος δείκτης βλάστησης διαφοράς, (NDVI) δημιουργήθηκε το 1974 από τους Rouse et al, με στόχο τον διαχωρισμό της βλάστησης από τη φωτεινότητα του εδάφους χρησιμοποιώντας δορυφορικά δεδομένα Landsat MSS. Αποτελεί τον πιο συχνά χρησιμοποιούμενο δείκτη παγκοσμίως. Ανήκει στην κατηγορία δεικτών που προκύπτει από αριθμητικές πράξεις. Στα πλεονεκτήματα του δείκτη είναι η ελαχιστοποίηση των τοπογραφικών επιδράσεων. Επίσης είναι σχεδόν αναλλοίωτος από τις διάφορες συνθήκες εξαιτίας των κανονικοποιημένων τιμών. Το εύρος τιμών είναι από -1 έως και +1 με το 0 να εκφράζει απουσία βλάστησης ενώ οι αρνητικές τιμές να περιγράφουν καλύψεις γης όπως νερό, ανθρωπογενείς κατασκευές κ.ά. Ο δείκτης βλάστησης περιγράφει πολλές ιδιότητες των φυτών με αποτέλεσμα να χρησιμεύει στην παρακολούθηση της υγείας τους, στην εκτίμηση της βιομάζας, στην απόδοση των καλλιεργειών κ.α.. Στα μειονεκτήματα του δείκτη είναι ότι παρουσιάζει δείγματα κορεσμού (saturation) σε πολύ υψηλές συγκεντρώσεις βλάστησης [9]. Ακόμη, όπως επισημαίνουν οι Bannari et al (σύμφωνα με τους Jackson et al)[14], ο δείκτης παρουσιάζει ευαισθησία στις ατμοσφαιρικές επιδράσεις, ενώ η δυναμική του εξασθενεί όταν η φυτοκάλυψη είναι αραιή (συγκέντρωση μικρότερη του 50%), εξαιτίας της ανακλαστικότητας του εδάφους.

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

Ο NDVI για τις εικόνες Sentinel-2 ορίζεται από τον τύπο:

$$NDVI = \frac{(B8 - B4)}{(B8 + B4)}$$

Όπου B8 είναι το κοντινό υπέρυθρο (0.842 μm) και B4 είναι το κόκκινο (0.665μm) φασματικό κανάλι. Στον NDVI οι αρδευόμενες εκτάσεις (άρα υψηλά ποσοστά πράσινης βιομάζας) εμφανίζονται με ανοικτότερους τόνους του γκρι σε σχέση με τις μη αρδευόμενες εκτάσεις καθώς και τις υδάτινες επιφάνειες. Ενδεικτικά παρουσιάζεται στον παρακάτω πίνακα η συσχέτιση ανάμεσα στις τιμές του δείκτη, τους τόνους του γκρι και την κάλυψη γης.

Πίνακας 2-2 Συσχέτιση των Τιμών του Δείκτη NDVI και των Καλύψεων Γης

Τιμές	Κάλυψη γης
Αρνητικές τιμές (Πολύ Σκούροι Τόνοι)	Νερό
Κοντά στο μηδέν	Βράχος, Άμμος, Χιόνι
Χαμηλές Θετικές Τιμές	Θάμνους, Λειμώνες
Υψηλές Θετικές Τιμές (Ανοικτοί Τόνοι)	Εύκρατα και Τροπικά Δάση

Δείκτης Βλάστησης MSAVI (Modified Soil Adjusted Vegetation Index)

Ο «τροποποιημένος εδαφικά προσαρμοσμένος δείκτης βλάστησης » (Modified Soil Adjusted Vegetation Index, MSAVI), δημιουργήθηκε από τους Qi et al το 1994 [15] με

στόχο τη διόρθωση της φωτεινότητας του εδάφους σε διαφορετικές καλύψεις βλάστησης. Αναλυτικότερα πετυχαίνει: (α) την απομάκρυνση του «θορύβου» από την ανάκλαση του εδάφους όταν δεν το κατορθώνουν άλλοι δείκτες βλάστησης όπως ο NDVI και (β) την κανονικοποιημένη διόρθωση ώστε οι τιμές να μην είναι εκτός του διαστήματος [-1, 1], γεγονός που παρατηρείται σε άλλους δείκτες. Το μεγαλύτερο πλεονέκτημα του δείκτη MSAVI είναι ότι ενώ ανήκει στους δείκτες που μετρούν την ραδιομετρική απόσταση από την γραμμή του εδάφους, πρακτικά δεν απαιτείται ο εμπειρικός προσδιορισμός της γραμμής του εδάφους. Οι τιμές κάτω από 0 υποδηλώνουν απουσία βλάστησης ενώ πάνω από 0 παρουσία βλάστησης[15].

Ο MSAVI ορίζεται από τον τύπο:

$$MSAVI = \frac{1}{2} \left((2 * NIR) + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)} \right)$$

Για το δορυφόρο Sentinel-2, ο τύπος είναι ως εξής:

$$MSAVI = \frac{1}{2} \left((2 * B8) + 1 - \sqrt{(2 * B8 + 1)^2 - 8 * (B8 - B4)} \right)$$

Πίνακας 2-3 Συσχέτιση των Τιμών του Δείκτη MSAVI και των Καλύψεων Γης

Τιμές	Κάλυψη Γης
Αρνητικές Τιμές (Σκούροι Τόνοι του Γκρι)	Νερά, Δρόμοι (Απουσία Βλάστησης)
Θετικές Τιμές (Ανοιχτοί Τόνοι του Γκρι)	Καλλιεργούμενες Εκτάσεις, Φυσική Βλάστηση (Υψηλή Βιομάζα)

Δείκτης Νερού NDWI

Ο κανονικοποιημένος δείκτης διαφοράς νερού (NDWI), είναι ένας αριθμητικός δείκτης. Ο τύπος του NDWI, όπως αναφέρει ο Xu (2007) [16] δημοσιεύθηκε αρχικά από τον McFeeters το 1996. Ο NDWI του McFeeters υπολογίζεται ως εξής:

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR}$$

Αυτή η διατύπωση του NDWI παράγει μια εικόνα στην οποία οι θετικές τιμές δεδομένων είναι συνήθως ανοικτές περιοχές νερού, ενώ οι αρνητικές τιμές είναι χαρακτηριστικά μη υδάτινα (δηλαδή, χερσαία βλάστηση, γυμνό έδαφος κ.τ.λ.). Όπως και ο NDVI, ο δείκτης NDWI έχει μετρητική κλίμακα από -1 έως +1. Προσφέρει την ανάδειξη των υδάτινων σωμάτων και μπορεί να χρησιμοποιηθεί για την εκτίμηση των επιπτώσεων πλημμύρας. Ο NDWI, χρησιμοποιήθηκε και με άλλες μορφές (π.χ. χρήση των καναλιών NIR και SWIR), σε μία προσπάθεια μείωσης του θορύβου. Ο δείκτης NDWI, μπορεί να χρησιμοποιηθεί συνδυαστικά με τον δείκτη NDVI, με σκοπό την αξιολόγηση των αλλαγών της περιοχής.

Ο τύπος του NDWI, για τον Sentinel-2, είναι ο εξής:

$$NDWI = \frac{B3 - B8}{B3 + B8}$$

Δείκτης Περιοχής Οικιστικής Ανάπτυξης NDBI

Ο κανονικοποιημένος δείκτης περιοχών οικιστικής ανάπτυξης όπως αναφέρει ο Xu (2007) [16], προτάθηκε από τους Zha, Gao, Ni το 2003, με σκοπό την υπογράμμιση των αστικών περιοχών, όπου συνηθέστερα υπάρχει υψηλότερη ανάκλαση στην μικροκυματική περιοχή SWIR του ηλεκτρομαγνητικού φάσματος, σε σύγκριση με την περιοχή της υπέρυθρης ακτινοβολίας NIR. Χρησιμοποιείται ευρέως σε εφαρμογές απορροής ποταμών και χωροταξικού σχεδιασμού. Αρχικά αναπτύχθηκε με βάση τα κανάλια 4 και 5, του Landsat TM 5.

Ο γενικός τύπος είναι:

$$NDBI = \frac{NIR - SWIR}{NIR + SWIR} \quad \text{ή} \quad NDBI = \frac{SWIR - NIR}{SWIR + NIR}$$

Για τον Sentinel-2, ο τύπος είναι:

$$NDBI = \frac{B11 - B8}{B11 + B8}$$

2.5 Ταξινόμηση Πολυφασματικών Εικόνων

2.5.1 Γενικά στοιχεία

Τα τηλεπισκοπικά δεδομένα χρησιμοποιούνται σε πολλές εφαρμογές. Συνηθέστερα, εφαρμόζεται μια διαδικασία ταξινόμησης εικόνων για τη μετατροπή των πολυφασματικών δεδομένων σε σημαντικές πληροφορίες.

Ο σκοπός της ταξινόμησης πολυφασματικών εικόνων είναι η αυτόματη κατηγοριοποίηση όλων των εικονοστοιχείων μίας εικόνας σε διάφορες κατηγορίες κάλυψης γης. Η κατηγοριοποίηση αυτή βασίζεται στα φασματικά πρότυπα, που ενυπάρχουν στα δεδομένα του κάθε εικονοστοιχείου, καθώς χρησιμοποιούνται ως η ψηφιακή βάση της κατηγοριοποίησης [11]. Οι κατηγορίες κάλυψης γης διαφοροποιούνται φασματικά, χάρη στους διαφορετικούς συνδυασμούς ψηφιακών τιμών στα κανάλια μίας πολυφασματικής εικόνας, ανάλογα με τις εγγενείς ιδιότητες της φασματικής ανακλαστικότητας και εκπομπής. Δημιουργείται, λοιπόν, ένα φασματικό πρότυπο, για την κάθε κλάση, το οποίο υπαγορεύεται από τον συνδυασμό μετρήσεων έντασης ακτινοβολίας, από τα διάφορα φασματικά κανάλια, για το κάθε εικονοστοιχείο των δεδομένων εκπαίδευσης της κάθε κλάσης. Η φασματική αναγνώριση προτύπων, αναφέρεται στην οικογένεια των τεχνικών ταξινόμησης, οι οποίες χρησιμοποιούν αυτήν την φασματική πληροφορία, ανά εικονοστοιχείο, ως τη βάση για την αυτόματη ταξινόμηση κάλυψης γης.

2.5.2 Μέθοδοι ταξινόμησης

Όπως αναφέρουν οι Al-doski et al (2013) [17], υπάρχουν διάφορες προσεγγίσεις ταξινόμησης που έχουν αναπτυχθεί και ευρέως χρησιμοποιηθεί για την παραγωγή χαρτών κάλυψης γης (με βάση τους Arlin, Atkinson 2004). Η κάθε μία μέθοδος στηρίζεται σε διαφορετική λογική ταξινόμησης, που κυμαίνεται από την επιβλεπόμενη

έως τη μη επιβλεπόμενη ταξινόμηση, την παραμετρική στην μη παραμετρική έως την μη-μετρική ή από την «σκληρή» στην «μαλακή» ταξινόμηση (ταξινόμηση ασαφής γνώσης) ή σε ταξινομήσεις στο επίπεδο του εικονοστοιχείου (ανά pixel) , στο επίπεδο κάτω από το εικονοστοιχείο και στο επίπεδο του πεδίου (per field) (Al-doski et al, 2013, σύμφωνα με τους Keuchel et al. 2003a, Jensen 2005) [18]. Παρακάτω παρουσιάζονται οι βασικές μέθοδοι ταξινόμησης σε μορφή πίνακα:

Πίνακας 2-4 Μέθοδοι Ταξινόμησης, Παραδείγματα και Χαρακτηριστικά

Μέθοδοι	Παραδείγματα	Χαρακτηριστικά
Παραμετρική	Αλγόριθμος Μέγιστης Πιθανοφάνειας, Μη επιβλεπόμενη Ταξινόμηση	Υποθέσεις: Περιοχή δεδομένων που κατανέμεται κανονικά Προηγούμενη Γνώση των συναρτήσεων πυκνότητας κλάσης
Μη Παραμετρική	Κοντινότερου Γείτονα, Ασαφείς Ταξινομήσεις, Νευρωνικά Δίκτυα, SVM	Δεν γίνονται προηγούμενες υποθέσεις
Μη Μετρική	Δέντρο αποφάσεων βασισμένο σε κανόνες Ταξινόμηση	Μπορεί να λειτουργεί τόσο σε δεδομένα πραγματικής αποτίμησης όσο και σε ονομαστικά κλιμακούμενα δεδομένα στατιστικής ανάλυσης
Επιβλεπόμενη	Μέγιστης Πιθανοφάνειας, Μικρότερης Απόστασης, Ταξινόμησης Πραλληλεπιπέδου	Ταξινομητής αναγνωρίζει τα πολύγωνα εκπαίδευσης που εκπροσωπούν την κάθε κλάση και ταξινομεί κάθε εικονοστοιχείο με βάση την στατιστική ανάλυση
Μη Επιβλεπόμενη	ISODATA, K-means	Οι προηγούμενες πληροφορίες εδάφους δεν είναι γνωστές. Τα εικονοστοιχεία με παρόμοια φασματικά χαρακτηριστικά ομαδοποιούνται σύμφωνα με συγκεκριμένα στατιστικά κριτήρια
Σκληρή (Παραμετρική)	Επιβλεπόμενες και Μη Επιβλεπόμενες Ταξινομήσεις	Ταξινόμηση χρησιμοποιώντας διακριτές κατηγορίες
Μαλακή (Μη Παραμετρική)	Ταξινόμησης ασαφής γνώσης	Βασίζεται στην θεωρία της ετερογενούς φύσης του πραγματικού κόσμου. Σε κάθε εικονοστοιχείο δίνεται ένα ποσοστό της γης τύπος καλύμματος που βρίσκεται μέσα στο εικονοστοιχείο
Ανά Εικονοστοιχείο		Ταξινόμηση της εικόνας εικονοστοιχείο προς εικονοστοιχείο
Αντικειμενοστραφής		Η εικόνα ταξινομείται σε ομογενή αντικείμενα. Η ταξινόμηση εκτελείται σε κάθε αντικείμενο και σε κάθε εικονοστοιχείο
Υβριδικές Προσεγγίσεις		Περιλαμβάνει έμπειρα συστήματα και τεχνητή νοημοσύνη

Εντούτοις, δύο είναι οι ευρείς κατηγορίες ταξινόμησης, που χρησιμοποιούνται σε διαδικασίες ταξινόμησης πολυφασματικών εικόνων, οι επιβλεπόμενες και οι μη επιβλεπόμενες ταξινομήσεις. Οι δύο αυτές προσεγγίσεις, μπορούν να συσχετιστούν και να αποδώσουν νέες υβριδικές μεθοδολογίες (Al-doski et al, 2013, σύμφωνα με τους Richards, Jia, 2006)[17].

Επιβλεπόμενη Ταξινόμηση

Η πρώτη κατηγορία ταξινομήσεων γνωστή ως Επιβλεπόμενες Ταξινομήσεις αναπτύχθηκε βάσει της εκ των προτέρων γνώσης αλλά και της δυνατότητας πρόσβασης στις υπάρχουσες κατηγορίες καλύψεις γης ώστε να εντοπισθούν χωρικά δειγματοληπτικές περιοχές. Ο όρος «επιβλεπόμενες» προέκυψε από το γεγονός ότι ο αλγόριθμος εύρεσης των ραδιομετρικών τιμών κάθε κάλυψης γης εκπαιδεύεται σύμφωνα με τις ομογενείς, ή όχι, δειγματοληπτικές περιοχές που έχουν επιλεγεί από πριν για κάθε κατηγορία κάλυψης γης [9]. Οι περιοχές αυτές χρησιμοποιούνται για να διαμορφωθεί ένα φωτοερμηνευτικό κλειδί, το οποίο θα περιγράφει τις φασματικές ιδιότητες της κάθε κλάσης, φασματικές ιδιότητες, που πηγάζουν από τις ραδιομετρικές τιμές της εικόνας [11]. Στην συνέχεια, κάθε εικονοστοιχείο της εικόνας, συγκρίνεται αριθμητικά με όλες της κατηγορίες του κάθε φωτοερμηνευτικού κλειδιού και κατατάσσεται στην κλάση, στην οποία το φωτοερμηνευτικό κλειδί μοιάζει περισσότερο. Οι αριθμητικές στρατηγικές που μπορούν να χρησιμοποιηθούν για να διεξαχθεί η σύγκριση ανάμεσα στα εικονοστοιχεία εκπαίδευσης και τα άγνωστα εικονοστοιχεία της εικόνας αυτή είναι πολυάριθμες. Δυστυχώς, η ταξινόμηση εικόνων δεν είναι μια τετριμμένη εργασία, καθώς τα περισσότερα από τα συστήματα επιβλεπόμενης ταξινόμησης απαιτούν αρκετά μεγάλο αριθμό δειγμάτων εκπαίδευσης, των οποίων ο ορισμός και η απόκτηση είναι συχνά ένα κρίσιμο πρόβλημα.



Εικόνα 2-2 Πορεία Επιβλεπόμενης Ταξινόμησης (Πηγή: Al-doski, 2013, p.p. 143)[17]

Μη Επιβλεπόμενη Ταξινόμηση

Η δεύτερη κατηγορία ταξινομήσεων είναι γνωστή ως «Μη Επιβλεπόμενες Ταξινομήσεις». Και σε αυτήν την κατηγορία ταξινομήσεων, η διαδικασία πραγματοποιείται σε δύο στάδια. Η κύρια διαφοροποίηση ανάμεσα στις δύο κατηγορίες ταξινομήσεων, είναι η διαδικασία της επίβλεψης. Στην κατηγορία των μη επιβλεπόμενων ταξινομήσεων, τα δεδομένα πρώτα ταξινομούνται σε κλάσεις άγνωστες, συναθροίζοντας τα δεδομένα της εικόνας σε φυσικές φασματικές ομάδες (συσσωρεύσεις) και στην συνέχεια ο φωτοερμηνευτής συγκρίνοντας τον θεματικό χάρτη, που προέκυψε ως αποτέλεσμα της ταξινόμησης, με επίγεια δεδομένα αναφοράς, χάρτες και αεροφωτογραφίες, αντιστοιχίζει τις κλάσεις που προέκυψαν με καλύψεις γης [11]. Η ταξινόμηση στις μη επιβλεπόμενες ταξινομήσεις υλοποιείται με διάφορους αλγόριθμους, όπως είναι η μέθοδος της ιεραρχικής ή της μη ιεραρχικής ταξινόμησης κατά συστάδες[9].



Εικόνα 2-3 Πορεία Μη Επιβλεπόμενης Ταξινόμησης (Πηγή: Al-doski, 2013, p.p. 143)[17]

Υβριδική Ταξινόμηση

Διάφορες μέθοδοι ταξινόμησης έχουν αναπτυχθεί εκτός των παραπάνω, που πολλές από αυτές αποτελούν συνδυασμό των παραπάνω ή παραλλαγές τους. Ο Περάκης [9] επισημαίνει ότι (σύμφωνα με τους Lillesand και Kiefer, 1994) οι συγκεκριμένες μέθοδοι αναφέρονται στη βιβλιογραφία ως υβριδικές. Περιπτώσεις υβριδικών μεθόδων χρησιμοποιούν μάσκες μικρών διαστάσεων για την σάρωση της εικόνας, μέσα από την οποία ελέγχεται η διακύμανση των ραδιομετρικών τιμών για κάθε θέση. Ορίζεται από τον χρήστη ένα κατώφλι διακύμανσης το οποίο χαρακτηρίζει την διακύμανση μικρή εάν δεν το υπερβαίνει και μεγάλη εάν το υπερβαίνει. Η μέση τιμή της μάσκας για την πρώτη θέση, αποτελεί και το πρώτο κέντρο συστάδας, η μέση τιμή της μάσκας για τη δεύτερη το δεύτερο κ.ο.κ. Ορίζεται από τον χρήστη ο μέγιστος αριθμός συστάδων και έτσι ο αλγόριθμος προχωρά σε συνενώσεις συστάδων, έως ότου να ισούνται με τον επιθυμητό αριθμό συστάδων.

Αντικειμενοστραφής Ταξινόμηση

Σε αυτές τις μεθόδους ταξινόμησης, τα εικονοστοιχεία ομαδοποιούνται ώστε να σχηματίζουν αντιπροσωπευτικά σχήματα και μεγέθη αντικειμένων. Αυτή η διαδικασία πρόκειται για την κατάτμηση πολλαπλών αναλύσεων ή αλλιώς την μέση μετατόπιση τμήματος. Η πολυδιάστατη κατάτμηση παράγει ομοιογενή αντικείμενα εικόνας με ομαδοποίηση εικονοστοιχείων. Δημιουργεί ταυτόχρονα αντικείμενα με διαφορετικές κλίμακες σε μια εικόνα. Αυτά τα αντικείμενα έχουν ιδιαίτερη σημασία επειδή αντιπροσωπεύουν χαρακτηριστικά στην εικόνα. Αλλά το πιο σημαντικό, είναι ότι σε τέτοιου τύπου ταξινομήσεις ο χρήστης μπορεί να ταξινομήσει τα αντικείμενα με βάση την υφή τους, το πλαίσιο και τη γεωμετρία τους.

Στην συνέχεια αναλύεται διεξοδικά ο αλγόριθμος τυχαίων δασών (Random Forest), που αποτελεί και την μέθοδο που εφαρμόστηκε για τη διεξαγωγή των ταξινομήσεων των πειραμάτων της συγκεκριμένης μεταπτυχιακής εργασίας.

2.5.3 Ο αλγόριθμος Random Forest

Ο αλγόριθμος τυχαίων δασών (Random Forest), που προτάθηκε το 2001 από τον Leo Breiman[19] πρόκειται για έναν ομοιογενή συνδυαστικό ταξινομητή. Αυτός ο αλγόριθμος χαρακτηρίζεται από μία επιβλεπόμενη ταξινόμηση από ένα μεγάλο σετ από ασυσχέιστα δένδρα αποφάσεων και έχει ως στόχο την αποφυγή του φαινομένου «overfitting» μέσα από την ελαχιστοποίηση της συσχέτισης ανάμεσα στα επιμέρους δένδρα. Οδηγεί λοιπόν σε καλύτερα αποτελέσματα συγκρινόμενος με τον δομικό του αλγόριθμο των δένδρων απόφασης[20], καθώς κάθε κόμβος διαιρείται/ διαχωρίζεται χρησιμοποιώντας το βέλτιστο συνδυασμό από όλες τις μεταβλητές.

Σημείωση: *Overfitting* είναι όταν το εξαγόμενο από μία ανάλυση ανταποκρίνεται τόσο κοντά ή ακριβώς στο συγκεκριμένο σετ δεδομένων, που αδυνατεί να ανταποκριθεί να ταιριάξει σε διαφορετικά δεδομένα ή να προβλέψει μελλοντικές παρατηρήσεις αξιόπιστα. Είναι ένα στατιστικό μοντέλο που περιέχει περισσότερες παραμέτρους από ότι τα δεδομένα μπορούν να δικαιολογήσουν. Είναι φαινόμενο που μπορεί να συμβεί ιδίως σε εφαρμογές *machine learning* και υπάρχει επειδή το κριτήριο που χρησιμοποιείται για την επιλογή του μοντέλου δεν είναι το ίδιο με το κριτήριο που χρησιμοποιείται για να κρίνει την καταλληλότητα του μοντέλου. Για παράδειγμα ένα μοντέλο μπορεί να επιλέγεται επειδή μεγιστοποιεί την απόδοση σε κάποια σετ δεδομένων εκπαίδευσης και στην συνέχεια η καταλληλότητά του προσδιορίζεται από την ικανότητα του να ανταποκρίνεται καλά στα *unseen data*. Τότε το *overfitting* αναμένεται όταν ένα μοντέλο ξεκινά να αποστηθίζει τα δεδομένα εκπαίδευσης από τα να εκπαιδεύεται ώστε να γενικεύει μια τάση. Απόρροια του φαινομένου είναι η προσαρμογή του μοντέλου στο θόρυβο που υπάρχει στα δεδομένα.

Αλγόριθμος

Ο αλγόριθμος κατασκευάζει ένα δάσος από δένδρα απόφασης, μία διαδικασία που ο ταξινομητής επιτυγχάνει σε δύο στάδια.

Στο πρώτο στάδιο, κάθε δένδρο δομείται με τη μέθοδο «bagging», χρησιμοποιώντας δηλαδή τυχαία επιλεγμένα δείγματα με αντικατάσταση[19]. Συνεπώς τα δεδομένα εκπαίδευσης που αξιοποιούνται στην εκπαίδευση ενός δένδρου, μπορούν να αξιοποιηθούν ξανά όλα ή μέρος αυτών για την εκπαίδευση άλλων δένδρων του δάσους, ενώ τμήμα των δεδομένων εκπαίδευσης μπορεί να μην αξιοποιηθεί σε κανένα δένδρο. Κάθε δένδρο στο δάσος εκπαιδεύεται με διαφορετικά δείγματα, αλλά με την ίδια διανομή, δηλαδή με δείγματα ίδιου μεγέθους, οδηγούμαστε έτσι σε στατιστικά ανεξάρτητα δεδομένα εκπαίδευσης των δένδρων[19], [21]. Τα δύο τρίτα από τα δείγματα εκπαίδευσης (δηλαδή τα «in bag samples») χρησιμοποιούνται για να εκπαιδεύσουν τα δένδρα, ενώ το υπολειπόμενο ένα τρίτο (δηλαδή τα «out of bag samples») χρησιμοποιούνται στη διασταυρωμένη επικύρωση για να προβλέψει πόσο καλά εκτελείται το μοντέλο *random forest* που προκύπτει. Η βασική φιλοσοφία σε αυτήν την προσέγγιση είναι ότι η δύναμη των δένδρων διατηρείται καθώς μειώνεται η συσχέτιση μεταξύ των δένδρων [20].

Καθώς τα δέντρα αποφάσεων μεγαλώνουν/ αναπτύσσονται, μία επιπλέον πηγή τυχαιότητας εμπλουτίζει το μοντέλο, στοχεύοντας στη περαιτέρω ελαχιστοποίηση της συσχέτισης δύο δένδρων και συνεπώς στην ελαχιστοποίηση της αστάθειας, ιδιότητα που χαρακτηρίζει τους δένδροειδείς ταξινομητές [22]. Αυτή είναι η τυχαία επιλογή των μεταβλητών (στην προκειμένη περίπτωση των καναλιών) για την φασματική οριοθέτηση διαχωρισμού σε κάθε κόμβο. Το πλήθος των μεταβλητών απορρέει από το σύνολό τους, που χρησιμοποιείται στο μοντέλο και συνήθως ισούται με $m_{try} = \sqrt{V}$ [23].

Για να δομηθεί, επομένως το συγκεκριμένο μοντέλο, απαιτείται ο καθορισμός από τον χρήστη του αριθμού των στοιχείων που θα χρησιμοποιηθούν σε κάθε κόμβο, η τυχαία επιλογή του αριθμού των μεταβλητών για την εκπαίδευση των δένδρων, καθώς και μία τρίτη παράμετρος, που δεν είναι άλλη από τον αριθμό των δένδρων που θα αναπτυχθούν στο δάσος. Η παράμετρος αυτή καθορίζεται από τον χρήστη. Η αύξηση των δένδρων οδηγεί σε ομαλότερη μετάβαση ανάμεσα στις κλάσεις και συνεπώς σε βέλτιστο αποτέλεσμα ταξινόμησης. Όμως, μέσα από πληθώρα μελετών, έχει διαπιστωθεί ότι τα σφάλματα σταθεροποιούνται πριν το 200. Συνεπώς, ορίζεται ως τιμή το 200.

Στο δεύτερο βήμα, αποφασίζονται οι συνθήκες διαχωρισμού για τον προσδιορισμό των κλάσεων, σε κάθε κόμβο σε ένα δένδρο. Ο ταξινομητής, χρησιμοποιεί το δείκτη «GINI» σαν ένα μέτρο επιλογής χαρακτηριστικών για να μετρήσει την ασυμβατότητα ενός χαρακτηριστικού σε σχέση με τις κλάσεις. Η διαδικασία που ακολουθείται είναι μία επαναληπτική εξέταση και αξιολόγηση των διαφόρων ορίων αξιοποιώντας τον προαναφερθέν δείκτη. Το όριο που επιλέγεται είναι εκείνο με την χαμηλότερη τιμή του δείκτη, καθώς το όριο αυτό χαρακτηρίζεται από υπολογιστική ταχύτητα και συνεπώς από ταχύτερη λήψη της τελικής απόφασης, η οποία λαμβάνεται βάσει της ψήφου πλειοψηφίας όλων αποφάσεων των δένδρων.

Ο δείκτης «GINI» μπορεί να υπολογιστεί από την εξής εξίσωση.

$$\sum_{j \neq i} \sum (f(C_i, T)/|T|) (f(C_j, T)/|T|)$$

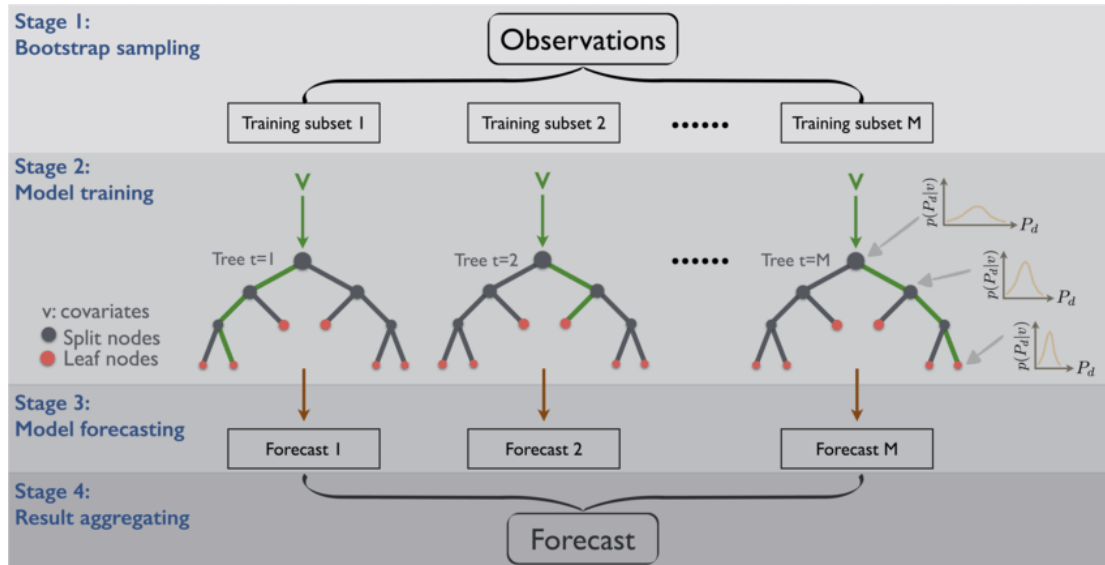
Όπου T είναι το δοθέν σύνολο δεδομένων,

C είναι η κλάση στην οποία ανήκει μία τυχαία επιλεγμένη περίπτωση (Εικονοστοιχεία ή αντικείμενα της εικόνας) και

$(f(C_i, T)/|T|)$ είναι η πιθανότητα ότι η επιλεγμένη περίπτωση ανήκει στην κλάση C_i.

Όταν ο εκτιμώμενος δείκτης αυξάνεται, αυξάνεται και η ετερογένεια των κλάσεων. Εάν ο δείκτης «GINI» ενός κόμβου-παιδιού είναι μικρότερος από τον δείκτη «GINI» του κόμβου γονέα, τότε ο διαχωρισμός είναι επιτυχημένος. Όταν όλα τα δένδρα μεγαλώσουν στο δάσος με την παραπάνω θεώρηση, τότε ο ταξινομητής πραγματοποιείται στο νέο σετ δεδομένων.

Συμπερασματικά, ο αλγόριθμος των τυχαίων δένδρων αποτελείται από τέσσερα στάδια. Στο πρώτο βήμα επιλέγονται τα υποσύνολα των δεδομένων εκπαίδευσης για τα δένδρα. Στο δεύτερο βήμα εκπαιδεύονται τα μεμονωμένα μοντέλα. Στο τρίτο βήμα αποδίδονται τα αποτελέσματα από το κάθε δένδρο, οδηγώντας σε πολλαπλά μοντέλα πρόβλεψης, ενώ στο τέταρτο βήμα συγκεντρώνονται τα αποτελέσματα και προκύπτει το τελικό μοντέλο για την πρόβλεψη. Η συγκεκριμένη διαδικασία, παρουσιάζεται σχηματικά στην παρακάτω εικόνα.



Εικόνα 2-4 Σχηματική απεικόνιση του «random forest» αλγορίθμου, βασισμένο στην μέθοδο «bagging».

Μία άλλη σημαντική πτυχή που πρέπει να ληφθεί υπόψη είναι ότι το overfitting* και η ευαισθησία ανάλυσης των μεταβλητών. Επειδή υπάρχουν αρκετά δένδρα στο δάσος, ο ταξινομητής δεν αναμένεται να κάνει overfit το μοντέλο. Παρέχοντας πληροφορίες σχετικά με το πόσο σχετίζεται κάθε στοιχείο με το αποτέλεσμα του μοντέλου (εξαρτημένη μεταβλητή), είναι δυνατόν να επιλεγθούν ποια θα είναι τα χαρακτηριστικά που θα χρησιμοποιηθούν στο μοντέλο, καθώς χαρακτηριστικά με χαμηλή σημασία μπορούν να οδηγήσουν συχνά σε αρνητικά αποτελέσματα [24]. Το τελικό αποτέλεσμα της ταξινόμησης παράγεται με βάση της πλειοψηφίας των δένδρων στο δάσος. Θα πρέπει να υπογραμμιστεί ότι τα νέα δείγματα εκπαίδευσης επιλέγονται από το αρχικό σετ δεδομένων με αντικατάσταση. Ουσιαστικά πρόκειται για μία βελτιωμένη έκδοση της μεθόδου «bagging ensemble» [20].

Παρόλο που ο ταξινομητής έχει αναπτυχθεί για την κοινότητα του «machine learning», είναι ιδιαίτερως δημοφιλής στην κοινότητα της ψηφιακής τηλεπικοινωνίας, εξαιτίας της ακρίβειας που πετυχαίνει, της ταχύτητας και της εύκολης παραμετροποίησης του. Σύμφωνα με τους Joellsson et al [25], η μοναδικότητα των δεδομένων καθορίζεται από τρεις παράγοντες.

Κάθε δένδρο εκπαιδεύεται χρησιμοποιώντας ένα τυχαίο υποσύνολο από τα δεδομένα εκπαίδευσης.

Κατά τη διάρκεια ανάπτυξης του δένδρου ο βέλτιστος διαχωρισμός σε κάθε κόμβο στο δένδρο βρίσκεται αναζητώντας μέσα από m τυχαία επιλεγμένα χαρακτηριστικά. Για ένα σετ M δεδομένων, m επιλέγεται από τον χρήστη και κρατείται πολύ μικρότερο από το M .

Κάθε δένδρο αναπτύσσεται στο έπακρο για να διαφοροποιήσει τα δένδρα μεταξύ τους, ώστε να μην υπάρξει κλάδεμα.

Οι Breiman και Cutler(2001) [19], υπογραμμίζουν ότι ο συγκεκριμένος ταξινομητής, είναι ανυπέρβλητος σε ακρίβεια απέναντι στους συμβατικούς ταξινομητές και αλλά και σε αλγορίθμους «machine learning».

Όπως αναφέρει ο Kanvzoglu (2017) [20] μερικά από τα πλεονεκτήματα που χαρακτηρίζουν τον συγκεκριμένο ταξινομητή είναι:

Η καταλληλότητά του για χρήση σε συνεχή και διαφόρων κατηγορών δεδομένα, καθώς δεν είναι παραμετρικός.

Η εύκολη παραμετροποίηση του αλγορίθμου.

Η σταθερότητά του απέναντι στο overfitting και στον θόρυβο των δεδομένων.

Η δυνατότητα πρόβλεψης των σημαντικών μεταβλητών και της εκτίμησης της πιθανότητας σφάλματος, αξιοποιώντας τα «out of bag samples»

Κάθε δένδρο στο δάσος εκπαιδεύεται με διαφορετικά δείγματα, αλλά με την ίδια διανομή, δηλαδή με δείγματα ίδιου μεγέθους, οδηγούμαστε έτσι σε στατιστικά ανεξάρτητα δεδομένα εκπαίδευσης των δένδρων

Η τυχαία επιλογή των δεδομένων εκπαίδευσης κάθε δένδρου οδηγεί σε ασυσχέτιστα αποτελέσματα.

Κάθε δένδρο αναπτύσσεται στο έπακρο χωρίς κλάδεμα, με αποτέλεσμα τα δένδρα που δημιουργούνται να είναι πλήρως διαφοροποιημένα.

Ο αλγόριθμος ολοκληρώνεται σε σταθερό αριθμό βημάτων.

Πέρα από τα θετικά γνωρίσματα του αλγορίθμου, οφείλουμε να παραθέσουμε και μερικά από τα λιγοστά μειονεκτήματα του αλγορίθμου. Αυτά είναι:

Το υψηλό υπολογιστικό κόστος, καθώς η εκπαίδευσης δένδρων με μεγάλο διαχωρισμό οδηγούν σε χρήση μεγάλων επιπέδων μνήμης.

Πρόκειται για έναν αλγόριθμο χρονοβόρο κατά τη διαδικασία πρόβλεψης

Τα νέα παραδείγματα, πρέπει να περάσουν από όλα τα δένδρα για να επιτευχθεί η εκτίμηση.

Σε περίπτωση μεταβολής των κλάσεων απαιτείται η ανακατασκευή του μοντέλου.

Κάθε νέο παράδειγμα πρέπει να διασχίσει όλα τα δέντρα του δάσους για την εκτίμηση της κατηγορίας του.

2.6 Ακρίβεια Ταξινόμησης

2.6.1 Γενικά στοιχεία

Εφαρμόζοντας οποιονδήποτε αλγόριθμο ταξινόμησης για την ερμηνεία μιας εικόνας, μας ενδιαφέρει πάντα η ακρίβεια των αποτελεσμάτων. Σύμφωνα με τους Xie et al (2008) [26], τέσσερα σημαντικά στάδια έχουν παρατηρηθεί σε μεθόδους αξιολόγησης της ακρίβειας. Ο απλούστερος τρόπος αξιολόγησης, που αποτελεί το πρώτο στάδιο αξιολόγησης, είναι η οπτική αξιολόγηση των παραγόμενων χαρτών. Συγκρίνοντας την εικόνα με τα αποτελέσματα της ερμηνείας της, μπορούμε να δούμε λάθη και να υπολογίσουμε το μέγεθος τους. Η συγκεκριμένη μέθοδος θεωρείται ιδιαίτερα υποκειμενική και συχνά δεν είναι ακριβής. Συνεπώς, εάν χρειαζόμαστε αξιόπιστη αξιολόγηση της ακρίβειας, δεν μπορούμε να το κάνουμε χωρίς ποσοτικές μεθόδους αξιολόγησης.

Το δεύτερο στάδιο αξιολόγησης της ακρίβειας της ταξινόμησης, είναι η σύγκριση των εκτάσεων της περιοχής των παραγόμενων χαρτών καλύψεων γης με τις αντίστοιχες εκτάσεις στο έδαφος ή σε κάποιο άλλο σύνολο δεδομένων αναφοράς (ground truth). Η μέθοδος αυτή για την αξιολόγηση των αποτελεσμάτων αποτελεί μία πιο αντικειμενική μέθοδο σε σχέση με την πρώτη, αλλά χαρακτηρίζεται από το μείζον πρόβλημα της τοποθεσίας, καθώς η ανάδειξη σωστών αναλογιών των διαφόρων καλύψεων γης, δεν συνάδει απαραίτητα με τις σωστές θέσεις εντοπισμού τους.

Στην συνέχεια, το τρίτο στάδιο αξιολόγησης της ακρίβειας ταξινόμησης, αποτελεί η σύγκριση των ετικετών κλάσεων του θεματικού χάρτη, με τα δεδομένα εδάφους για τις ίδιες θέσεις. Για τη διεξαγωγή της συγκεκριμένης μεθόδου, χρησιμοποιούνται διάφορα μεγέθη, όπως τα ποσοστά των σωστών αλλά και των εσφαλμένων ταξινομημένων εικονοστοιχείων.

Τέλος, η αξιολόγηση της ακρίβειας στο τέταρτο στάδιο, βελτιώνοντας την κεντρική ιδέα της προηγούμενης μεθόδου, υπολογίζεται κατασκευάζοντας έναν πίνακα σύγχυσης, ο οποίος περιγράφει την καταλληλότητα μεταξύ των κλάσεων και των δεδομένων αναφοράς χρησιμοποιώντας μεγάλη ποικιλία διάφορων μέτρων όπως η συνολική ακρίβεια, ο συντελεστής kappa κ.ά..

2.6.2 Πίνακας σύγχυσης

Οι πίνακες σύγχυσης χρησιμοποιούνται ευρέως σε εφαρμογές μηχανικής μάθησης και ειδικότερα του προβλήματος της στατιστικής ταξινόμησης. Πρόκειται για έναν πίνακα συγκεκριμένης διάταξης, τετραγωνικής διάστασης, που επιτρέπει την απεικόνιση της απόδοσης ενός αλγορίθμου, συνήθως επιβλεπόμενης εκπαίδευσης (στην μη επιβλεπόμενη εκπαίδευση ονομάζεται πίνακας αντιστοιχίας). Κάθε σειρά του πίνακα αντιπροσωπεύει τις περιπτώσεις της προβλεπόμενης κλάσης, ενώ κάθε στήλη αντιπροσωπεύει τις περιπτώσεις της πραγματικής κλάσης (ή αντίστροφα) και κάθε συνδυασμός διαστάσεων και κλάσης είναι μία μεταβλητή στον πίνακα ενδεχομένων (Visa S. et al, 2011) [27]. Το όνομα πηγάζει από το γεγονός ότι καθιστά εύκολο να δούμε αν το σύστημα προκαλεί σύγχυση σε δύο κατηγορίες (δηλαδή να ταξινομήσει λανθασμένα την μία κλάση ως τη δεύτερη).

2.6.3 Μεγέθη αξιολόγησης της ακρίβειας

Ο πίνακας σύγχυσης, χρησιμοποιείται συχνά για την οργάνωση και την εμφάνιση των πληροφοριών που χρησιμοποιούνται στην αξιολόγηση της ακρίβειας ενός χάρτη κάλυψης γης, ενώ έχουν προταθεί πολυάριθμα μεγέθη αξιολόγησης της ακρίβειας για τη συνοπτική παρουσίαση των πληροφοριών που περιέχονται σε αυτόν τον πίνακα σφαλμάτων. Κανένα μέγεθος δεν μπορεί να θεωρηθεί ως το καλύτερο καθολικά, καθώς τα μεγέθη που επιλέγονται για την αξιολόγηση της ακρίβειας της ταξινόμησης εξαρτώνται από τους στόχους της κάθε εφαρμογής. Συνεπώς, η επιλογή των κατάλληλων μέτρων ακρίβειας που αφορούν τους στόχους του εκάστοτε έργου χαρτογράφησης είναι κρίσιμη (Stehman, Stephen V., 1997) [28].

Τα πιο διαδεδομένα μεγέθη αξιολόγησης της ακρίβειας, τα οποία προκύπτουν από τα στοιχεία του πίνακα σύγχυσης, είναι τα εξής:

Ακρίβεια (επίσης γνωστή ως ακρίβεια του παραγωγού): Το μέγεθος που αντιπροσωπεύει την ακρίβεια της ταξινόμησής για την κάθε κλάση. Προκύπτει από τον

λόγο των σωστά ταξινομημένων εικονοστοιχείων σε σχέση με όλα τα εικονοστοιχεία της εν λόγω κλάσης του ground truth. Για κάθε κλάση εικονοστοιχείων του ground truth (σειρά), ο αριθμός των σωστά ταξινομημένων εικονοστοιχείων διαιρείται με τον συνολικό αριθμό των εικονοστοιχείων του ground truth αυτής της κλάσης.

Αξιοπιστία (επίσης γνωστή ως ακρίβεια χρήση): Το μέγεθος που αντιπροσωπεύει την αξιοπιστία της κάθε κλάσης στην ταξινομημένη εικόνα. Πρόκειται για το κλάσμα των σωστά ταξινομημένων ρixel σε σχέση με όλα τα εικονοστοιχεία που ταξινομούνται σε αυτήν την κατηγορία στην ταξινομημένη εικόνα. Για κάθε κλάση στην ταξινομημένη εικόνα (στήλη), ο αριθμός των σωστά ταξινομημένων εικονοστοιχείων διαιρείται με τον συνολικό αριθμό των εικονοστοιχείων που ταξινομήθηκαν σε αυτήν την κατηγορία.

Η μέση ακρίβεια υπολογίζεται ως το άθροισμα της ακρίβειας όλων των κλάσεων διαιρούμενο με τον αριθμό των κλάσεων του ground truth.

Η μέση αξιοπιστία υπολογίζεται ως το άθροισμα της αξιοπιστίας όλων των κλάσεων διαιρούμενο με τον αριθμό των κλάσεων του ground truth.

Η συνολική ακρίβεια υπολογίζεται ως ο συνολικός αριθμός των σωστά ταξινομημένων εικονοστοιχείων (διαγώνιων στοιχείων) διαιρούμενος με τον συνολικό αριθμό των εικονοστοιχείων ελέγχου.

Ο δείκτης Fisher (F1), πρόκειται για ένα μέτρο συσχέτισης της ακρίβειας και τη αξιοπιστίας κάθε μιας κατηγορίας και υπολογίζεται σύμφωνα με τον τύπο:

$$F1 = \frac{(2 * UA * PA)}{(UA + PA)}$$

Όπου:

UA= η αξιοπιστία- ακρίβεια χρήση

PA = η ακρίβεια- ακρίβεια παραγωγού

Η στατιστική τιμή KHAT ή Kappa, είναι ένα μέτρο της διαφοράς μεταξύ της πραγματικής συμφωνίας μεταξύ των επίγειων δεδομένων και ενός αυτόματου αλγορίθμου ταξινόμησης και της τυχαίας συμφωνίας μεταξύ των επίγειων δεδομένων και ενός τυχαίου αλγορίθμου ταξινόμησης. Αποτελεί μία ένδειξη του βαθμού κατά τον οποίον οι ποσοστιαίες ορθές τιμές ενός πίνακα σφαλμάτων οφείλονται σε αληθή συμφωνία έναντι τυχαίας συμφωνίας. Κυμαίνεται μεταξύ 0 και 1, με τη βέλτιστη τιμή την μονάδα και την χειρίστη το μηδέν, το οποίο αντικατοπτρίζει μία τυχαία καταχώρηση των εικονοστοιχείων [11]. Υπολογίζεται σύμφωνα με τον τύπο:

$$k = \frac{N * \sum_{i=1}^r x_{ij} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})}$$

Όπου:

r= ο αριθμός των γραμμών του πίνακα σύγκρισης

x_{ij}= ο αριθμός των παρατηρήσεων στη γραμμή *i* και στη στήλη *j* (στην κύρια διαγώνιο)

x_{i+}=το σύνολο των παρατηρήσεων στη γραμμή *i* (παρουσιάζεται ως περιθωριακό σύνολο στο δεξί μέρος του πίνακα)

x₊= το σύνολο των παρατηρήσεων στην στήλη *i* (παρουσιάζεται ως περιθωριακό σύνολο στο κάτω μέρος του πίνακα)

N= ο συνολικός αριθμός των παρατηρήσεων που περιλαμβάνονται στον πίνακα

Τα μέτρα ακρίβειας που μπορούν να ερμηνευθούν απευθείας ως πιθανότητες εμφάνισης ορισμένων τύπων σφαλμάτων ταξινόμησης ή ορθών ταξινομήσεων θα πρέπει να επιλέγονται σε σχέση με μέτρα που δεν μπορούν να ερμηνευθούν ως τέτοια. Η ακρίβεια του χρήστη και του παραγωγού για την κάθε κατηγορία της ταξινόμησης καθώς και η συνολική ακρίβεια της ταξινόμησης είναι παραδείγματα μέτρων ακρίβειας, που διαθέτουν την επιθυμητή πιθανολογική ερμηνεία (Stehman, Stephen V., 1997) [28].

Ο συντελεστής συμφωνίας kappa δεν έχει τέτοια πιθανολογική ερμηνεία λόγω της προσαρμογής για την υποθετική τυχαία συμφωνία που ενσωματώνεται σε αυτό το μέτρο και η ισχυρή εξάρτηση του kappa από τις οριακές αναλογίες του πίνακα σφαλμάτων καθιστά την χρησιμότητα του στις συγκρίσεις ύποπτη (Stehman, Stephen V., 1997) [28].

Ακόμη, η κανονικοποίηση ενός πίνακα σφάλματος οδηγεί σε εκτιμήσεις που δεν είναι συνεπείς ως προς τις παραμέτρους ακριβείας του χάρτη που αξιολογείται, με αποτέλεσμα η διαδικασία αυτή να μην είναι εγγυημένη αναφορικά με τις περισσότερες εφαρμογές (Stehman, Stephen V., 1997) [28].

3

ΠΕΡΙΟΧΗ ΜΕΛΕΤΗΣ & ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Η περιοχή μελέτης της παρούσας μεταπτυχιακής εργασίας, απαντάται στην περιφέρεια Ανατολικής Μακεδονίας και Θράκης. Η Περιφέρεια Ανατολικής Μακεδονίας και Θράκης (ΠΑΜΘ) καλύπτει το βορειοανατολικό άκρο της Ελλάδας και αποτελείται από την Ανατολική Μακεδονία και τη Θράκη. Οι νομοί που την απαρτίζουν είναι οι νομοί Έβρου, Ροδόπης, Ξάνθης, Καβάλας και Δράμας. Στην συγκεκριμένη περιφέρεια ανήκουν τα νησιά Θάσος και Σαμοθράκη του θρακικού πελάγους. Πρωτεύουσά της είναι η Κομοτηνή, ενώ η μεγαλύτερη πόλη που απαντάται στην συγκεκριμένη περιφέρεια είναι η Αλεξανδρούπολη. Είναι η 4^η σε κατάταξη περιφέρεια με συνολική έκταση 14157 χλμ. Ως προς τον πληθυσμό της βάσει της απογραφής του 2011 κατατάσσεται στην 6^η θέση με πληθυσμό 608182 ,κάτοικοι, ενώ σε σχέση με την πληθυσμιακή πυκνότητα κατατάσσεται στην 9^η θέση, καθώς η πληθυσμιακή της πυκνότητα χαρακτηρίζεται ως χαμηλή 43,16 κάτοικοι/χλμ [29].

3.1 Περιοχή Μελέτης

3.1.1 Γεωγραφία - δημογραφία

Η περιφέρεια Ανατολικής Μακεδονίας και Θράκης είναι η 4^η μεγαλύτερη περιφέρεια σε έκταση, καθώς καταλαμβάνει 14157 km² και ο πληθυσμός της ανέρχεται σε 608182 κατοίκους, σύμφωνα με την απογραφή της ΕΣΥΕ το 2011, κατατάσσοντάς την στην 6^η μεγαλύτερη πληθυσμιακά περιφέρεια. Σε σχέση με την πληθυσμιακή πυκνότητα κατατάσσεται στην 9^η θέση (από τις 13), καθώς η πληθυσμιακή πυκνότητά της είναι χαμηλή και συγκεκριμένα 43.16 κάτοικοι/ km².

Προς βορρά συνορεύει με την Βουλγαρία και ανατολικά με την Τουρκία. Συνορεύει επίσης με τις περιφέρειες Κεντρικής Μακεδονίας στα δυτικά ενώ βρέχεται από το Θρακικό πέλαγος στο νότο. Πρωτεύουσα της περιφέρειας είναι η Κομοτηνή, ενώ μεγαλύτερη πόλη είναι η Αλεξανδρούπολη.

Η περιφέρεια είναι ημι-ορεινή. Απλώνεται στον νότιο ορεινό όγκο της οροσειράς Ροδόπης. Στη έκτασή της απαντώνται τόσο λοφώδεις εξάρσεις όσο και μεμονωμένα βουνά, όπως ο Παγγαίος. Επιπλέον διαθέτει πεδιάδες, οι περισσότερες από τις οποίες είναι είτε παραθαλάσσιες είτε βρίσκονται πλησίον εκβολών ποταμών. Διατρέχεται από σημαντικά ποτάμια (Έβρος, Νέστος, Άρδας) που οι πηγές τους είναι στη Βουλγαρία, όμως δεν διαθέτει πολλές λίμνες. Η πιο σημαντική είναι η λιμνοθάλασσα της Βιστωνίδας και οι λιμνοθάλασσες του Έβρου.

Είναι ιδιαίτερα αναπτυγμένη, τόσο στον πρωτογενή όσο και στον δευτερογενή τομέα. Στην κεντρική Μακεδονία βρίσκεται η μεγαλύτερη πεδιάδα της Ελλάδας. Η Θεσσαλονίκη αποτελεί τη φυσική διέξοδο των γειτονικών κρατών προς το Αιγαίο και φιλοδοξεί να αναδειχθεί και πάλι σε μητρόπολη των Βαλκανίων.

Τα ψηλότερα βουνά της κεντρικής Μακεδονίας είναι το Φαλακρό (2.232 m), ο Όρβηλος (2.212 m), το Μενοίκιο (1963 m), το Παγγαίο (1956 m), το όρος Φρακτό (1953 m), τα όρη Βροντούς (1849 m), η Χαϊντού ή Ερύμανθος (1828 m) Κι το Παπίκιο (1510 m). Μεγαλύτερος ποταμός είναι ο Έβρος -που είναι και ο μεγαλύτερος ποταμός της Ελλάδας - με μήκος 230 km εντός Ελλάδας. Ο Έβρος αποτελεί γεωγραφικό σύνορο ανάμεσα στην Ελλάδα και την Τουρκία και σχηματίζει ένα τεράστιο δέλτα, το οποίο αποτελεί τον σημαντικότερο υδροβιότοπο της Ελλάδας και παράλληλα πρόσφορο έδαφος για την ανάπτυξη καλλιεργειών. Εκτός από τον Έβρο δύο ακόμα σημαντικά ποτάμια διατρέχουν την περιφέρεια Ανατολικής Μακεδονίας και Θράκης ο Νέστος και ο Κομψάτος. Ο Νέστος οριοθετεί τα σύνορα ανάμεσα στην Μακεδονία και την Θράκη και το μήκος του εντός Ελλαδικού χώρου είναι 130 km, ενώ καταλήγει στο νότιο τμήμα του νομού Καβάλας και του νομού Θράκης. Ο Κομψάτος είναι μικρότερος ποταμός με μήκος 68 km, πηγάζει από την οροσειρά της Ροδόπης (περιοχή Διάσπαρτου) και καταλήγει στην λίμνη Βιστωνίδα. Αποτελεί σημαντικό βιότοπο για πολλά είδη πουλιών υπό εξαφάνιση.

Οι ακτές είναι συνεχείς, ομαλές και αμμώδεις. Όλες προσφέρονται για κολύμβηση, εκτός από τις εκβολές των ποταμών και τις ακτές του πολεοδομικού συγκροτήματος της Αλεξανδρούπολης και της Καβάλας.

Οι κλιματικές συνθήκες στην Ανατολική Μακεδονία και Θράκη διαφέρουν κατά τόπους. Λόγω της γεωγραφικής θέσης, χαρακτηρίζεται γενικά από ήπιες θερμοκρασίες. Το κλίμα διαφοροποιείται από τα παράλια προς την ενδοχώρα. Οι παραλίες της Μακεδονίας και της Θράκης διαθέτουν κλίμα μεσογειακό καθώς χαρακτηρίζονται από ήπιες θερμοκρασίες και καιρό αίθριο, ενώ στην ενδοχώρα συναντάται κλίμα ηπειρωτικό με εξαίρεση τις ορεινές περιοχές, όπου ο χειμώνας είναι βαρύς με πολλές χιονοπτώσεις. Ως προς τις βροχοπτώσεις, η συχνότητά τους είναι αυξημένη κατά τους φθινοπωρινούς μήνες [30].

3.1.2 Περιβάλλον

Η ανατολική Μακεδονία και Θράκη, όπως άλλωστε και ολόκληρη η Βαλκανική χερσόνησος, βρίσκονταν αρχικά στο βυθό μιας τεράστιας προϊστορικής θάλασσας, που οι επιστήμονες της έχουν δώσει το συμβολικό όνομα Τηθύς. Η "ανάδειξη" της σε ξηρά άρχισε περίπου πριν από 30.000.000 χρόνια, στην αρχή της γεωλογικής περιόδου που ονομάζεται Μειόκαινο.

Οι βασικές γεωλογικές ζώνες στην περιοχή είναι οι εξής:

Η μάζα της Ροδόπης διαιρείται σε δύο ενότητες στην κατώτερη ενότητα Παγγαίου, που αποτελείται από γνεύσιους, αμφιβολίτες, μαρμαρυγιακούς σχιστόλιθους και μάρμαρα και στην ανώτερη ενότητα Σιδηρόνερου που είναι παρόμοιας σύστασης ενώ διαθέτει επιπλέον και μιγματίτες, ενώ σε όλη την έκταση της μάζας της Ροδόπης συναντώνται διάσπαρτα υπερβασικά σερπεντινιτικά σώματα.

Η Σερβομακεδονική μάζα η οποία αποτελείται κυρίως από γνεύσιους, μαρμαρυγιακούς και αμφιβολιτικούς σχιστόλιθους, μάρμαρα, μεταγάββρους – μεταδιαβάσες, αμφιβολίτες, καθώς και σερπεντινιτικά σώματα.

Η Περιροδοπική ζώνη της Θράκης, η οποία περιλαμβάνει μεταμορφωμένα πετρώματα Μεσοζωικής ηλικίας που ανήκουν στην ενότητα Μάκρης (μάρμαρα, δολομίτες, ασβεστικοί σχιστόλιθοι, φυλλίτες, πρασινοσχιστόλιθοι) και στην ενότητα Δρυμού-

Μελίας (μεταγραουβάκες, μεταχαλαζίτες, μεταψαμμίτες, μετατόφφους, μεταλάβες, αργιλικόι σχιστόλιθοι), καθώς και σερπεντινιτικά και μεταγαββρικά πετρώματα.

Στην περιοχή υπάρχουν επίσης αρκετά Τριτογενή πλουτωνικά, υποηφαιστειακά και ηφαιστειακά σώματα που διεισδύουν στις παραπάνω γεωτεκτονικές ζώνες, ενώ αρκετά μεγάλες εκτάσεις καλύπτονται από ιζηματογενείς αποθέσεις του Ηωκαίνου (ασβεστόλιθοι, κροκαλοπαγή, ψαμμίτες), του Ολιγοκαίνου – Μειοκαίνου (παράκτιες αποθέσεις, κροκάλες, ψαμμίτες, άργιλοι, αμμοάργιλοι, άμμοι) και του Τεταρτογενούς (αλλουβιακές αποθέσεις: άμμοι, χαλαρά κροκαλοπαγή).

Το υπέδαφός της φιλοξενεί ένα πολύ μεγάλο αριθμό μεταλλευμάτων. Υπάρχουν αποδεδειγμένα κοιτάσματα ή ενδείξεις για κοιτάσματα χρυσού, χαλκού, μολύβδου και ψευδαργύρου, σιδήρου και μαγγανίου, μολύβδου και βολφραμίου. Αξιόλογο δυναμικό έχει να επιδείξει η περιοχή της Θράκης και στον τομέα των βιομηχανικών ορυκτών, καθώς στο υπέδαφος απαντώνται ορυκτά όπως ο ζεόλιθος, ο γρανίτης ο γραφίτης κ.λπ. [29]

3.1.3 Διοικητική διαίρεση

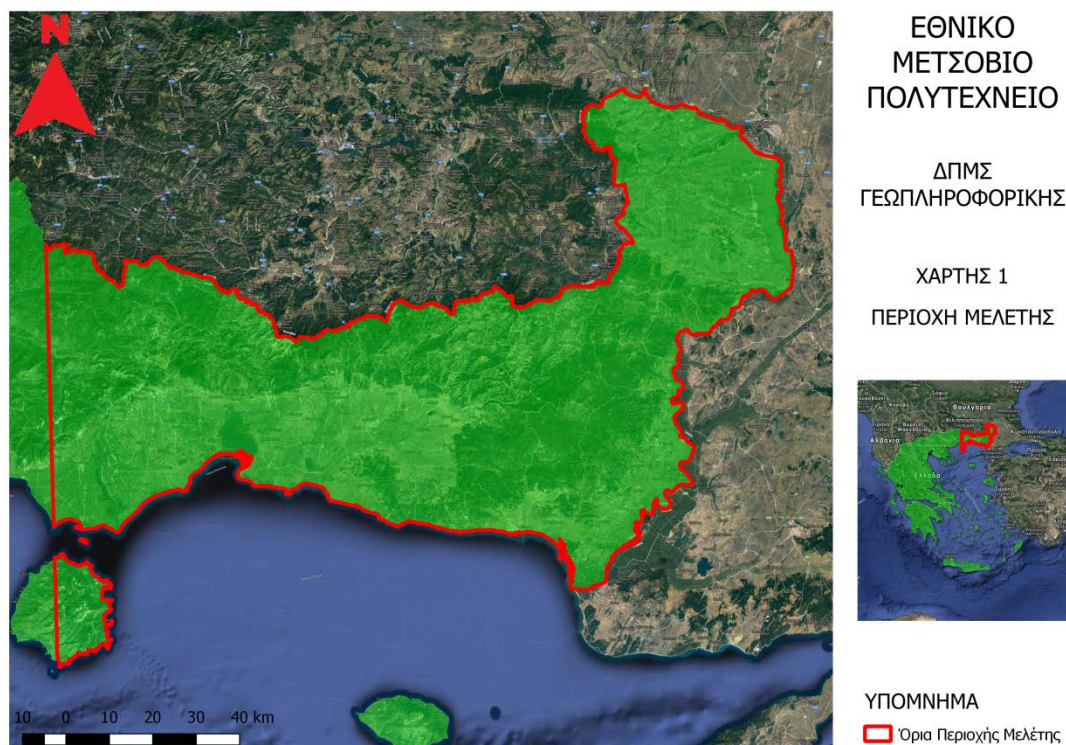
Η περιφέρεια διαιρείται σε πέντε περιφερειακές ενότητες, οι οποίες ταυτίζονται γεωγραφικά με τους αντίστοιχους νομού, δηλαδή τους νομούς Έβρου, Ροδόπης, Ξάνθης, Δράμας και Καβάλας. Στις ΠΕ υπάγονται οι 22 δήμοι που συνιστούν την περιφέρεια [29].

3.1.4 Οριοθέτηση περιοχής μελέτης

Η περιοχή μελέτης ανήκει στην Περιφέρειας Ανατολικής Μακεδονίας και Θράκης, και περικλείει τους πρώην νομούς Έβρου, Ροδόπης, Ξάνθης, ενώ τμήμα της εκτείνεται και στο δυτικό τμήμα του νομού Καβάλας. Ακόμη, εντός της περιοχής μελέτης βρίσκεται τμήμα της νήσου Θάσου, όμως δεν συμπεριλαμβάνεται η νήσος Σαμοθράκη. Η περιοχή μελέτης απεικονίζεται στα 3 υπό-μελέτη Sentinel-2 tiles: 35TLF, 35TMG και 35TMF.

Πιο συγκεκριμένα, το βορειοδυτικό άκρο της βρίσκεται 15 km βορειοδυτικά από την κορυφή Χαϊντου, το νοτιοδυτικό άκρο της περιοχής μελέτης βρίσκεται στην περιοχή Αστρίδα της νήσου Θάσου, το βορειοανατολικό άκρο της περιοχής μελέτης εντοπίζεται πλησίον του χωριού Ορμένιον στα σύνορα Ελλάδας- Βουλγαρίας και το νοτιοανατολικό άκρο της, εντοπίζεται στο δέλτα του ποταμού Έβρου, στην λίμνη Νυμφών.

Η συγκεκριμένη περιοχή επιλέχθηκε με κριτήριο την ποικιλομορφία της, ως προς τις καλύψεις γης, καθώς διαθέτει ταυτοχρόνως, σημαντικές περιοχές αστικού ιστού, φυσικής βλάστης και καλλιεργήσιμης γης, στην οποία ευδοκούν διάφορων ειδών καλλιέργειες.



Χάρτης 3.1.: Περιοχή Μελέτης.

3.2 Δορυφορικά Δεδομένα και Προεπεξεργασίες

Στην παρούσα μελέτη, έγινε μία προσπάθεια παραγωγής χάρτη καλύψεων γης για την περιοχή της Βορειοανατολικής Μακεδονίας και Θράκης. Χρησιμοποιήθηκαν πολυφασματικά δεδομένα του δορυφόρου Sentinel-2, για το έτος 2016. Επιλέχθηκε να χρησιμοποιηθούν δεδομένα του συγκεκριμένου δορυφόρου, καθώς παρέχει αξιόλογη χωρική και φασματική ακρίβεια, ενώ ταυτόχρονα διατίθεται σημαντικό πλήθος δεδομένων πολλαπλών ημερομηνιών, τα οποία παρέχονται ελεύθερα μέσω της ανοιχτής πολιτικής δεδομένων (βλ. Παράγραφο 2.3.)

3.2.1 Επιλογή δορυφορικών δεδομένων

Οι εικόνες που χρησιμοποιήθηκαν προήλθαν από 6 διαφορετικές ημερομηνίες του έτους 2016. Οι πολλαπλές ημερομηνίες βελτιώνουν την αποδοτικότητα του ταξινομητή. Τα κριτήρια επιλογής τους ήταν η χαμηλή νεφοκάλυψη και η ημερομηνία λήψης. Κρίθηκε απαραίτητο οι χρονοσειρές να απαρτίζονται από πολυφασματικές εικόνες προερχόμενες από διαφορετικές εποχές. Αυτές είναι:

Πίνακας 3-1 Εικόνες Ταξινόμησης

Tile Εποχή	35TLF	35TMF	35TMG
	Ημερομηνία Λήψης	Ημερομηνία Λήψης	Ημερομηνία Λήψης
Άνοιξη	21-04-2016	21-04-2016	21-04-2016
Άνοιξη προς Καλοκαίρι	20-06-2016	20-06-2016	10-06-2016
Καλοκαίρι	19-08-2016	Σύνθεση 16-08-2016 19-08-2016	19-08-2016
Φθινόπωρο	28-09-2016	28-09-2016	18-09-2016
Φθινόπωρο	17-11-2016	17-11-2016	17-11-2016
Φθινόπωρο προς Χειμώνας	17-12-2016	17-12-2016	17-12-2016

Λόγω υψηλής νεφοκάλυψης για την εποχή του χειμώνα δεν επιτεύχθηκε η εύρεση εικόνες κατά τους μήνες Ιανουάριο έως Μάρτιο. Όμως, καθώς στους συγκεκριμένους μήνες δεν εμφανίζονται αξιοσημείωτες μεταβολές στην βλάστηση και ιδίως τις καλλιέργειες θεωρήθηκε μικρής σημασίας.

Έγινε προσπάθεια επιλογής εικόνων με ίδια ημερομηνία λήψης για κάθε περιοχή. Σε ορισμένες περιπτώσεις η παράμετρος αυτή δεν ήταν δυνατόν να ικανοποιηθεί λόγω μεγάλης νεφοκάλυψης στις εικόνες. Συνεπώς επιλέχθηκαν εικόνες με ανεκτά επίπεδα νεφοκάλυψης και παραπλήσιων κατά το δυνατόν ημερομηνιών ώστε να μην μεταβάλλεται η παρεχόμενη πληροφορία των χρονοσειρών από tile σε tile.

Ιδιαίτερα για το tile 35TMF η εύρεση εικόνας κατά τους θερινούς μήνες αποδείχθηκε ιδιαίτερα επίπονη διαδικασία, λόγω υψηλής νεφοκάλυψης σε εικόνες οι οποίες περιείχαν όλη την περιοχή μελέτης και λόγω κομμένων εικόνων βάσει των γραμμών τροχιάς του δορυφόρου σε ημερομηνίες με χαμηλή νεφοκάλυψη. Οι γραμμές τροχιάς του δορυφόρου είναι τέτοιες που αποδίδουν ελλειπείς εικόνες. Οι περιοχές που δεν λαμβάνονταν από το δορυφόρο, περιέχονταν στην περιοχή μελέτης. Αποτέλεσμα ήταν η απουσία δεδομένων για την καλοκαιρινή περίοδο στο συγκεκριμένο tile. Το εμπόδιο αυτό προσπελάστηκε με σύνθεση δύο εικόνων. Πιο συγκεκριμένα, επιλέχθηκε η εικόνα με ημερομηνία λήψης 16-08-2016, στην οποία έλειπαν δεδομένα για ένα τμήμα της. Η εικόνα αυτήν ήταν καθαρή από νεφοκάλυψη. Επιπλέον επιλέχθηκε η εικόνα με ημερομηνία λήψης 19-08-2016, στην οποία το μεγαλύτερο της τμήμα ήταν καλυμμένο με σύννεφα, όχι όμως και η περιοχή στην οποία είχαμε απώλεια δεδομένων στην

εικόνα που λήφθηκε στις 16-08. Συνθέσαμε έτσι μία νέα εικόνα, αντικαθιστώντας το μηδέν στον πίνακα των χαρακτηριστικών με τις τιμές του πίνακα της εικόνας με πληροφορία (19-08). Η σύνθεση αυτή ήταν εφικτή καθώς οι ημερομηνίες λήψης ήταν ιδιαίτερας κοντινές, ενώ ο χρόνος της λήψης ήταν σχεδόν ο ίδιος (Η εικόνα 16-08 λήφθηκε στις 9:00 ενώ η εικόνα 19-08 στις 09:10).

3.2.2 Προεπεξεργασία δορυφορικών δεδομένων

Για την εκτέλεση του αλγορίθμου, απαιτείται μία σειρά διορθώσεων και βελτιστοποιήσεων των πρωτογενών δεδομένων. Αυτές είναι η ατμοσφαιρική διόρθωση, η βελτιστοποίηση της χωρικής ανάλυσης και η κανονικοποίηση των φασματικών τιμών.

Ατμοσφαιρική διόρθωση:

Η ατμοσφαιρική διόρθωση είναι η διαδικασία εκείνη κατά την οποία είναι δυνατή η εξάλειψη της αλλοίωσης που προκαλεί στην ακτινοβολία η επίδραση της ατμόσφαιρας, η οποία παρεμβάλλεται ανάμεσα στην κάμερα φωτοληψίας και στα αντικείμενα παρατήρησης.

Οι δορυφορικές εικόνες που παρέχει της πλατφόρμας της ESA βρίσκονται σε δύο επίπεδα επεξεργασίας στο Level 1C ή στο Level 2A (για περιορισμένα tiles. Στο επίπεδο L1C έχει προηγηθεί γεωμετρική και ραδιομετρική διόρθωση ώστε το προϊόν αποτελεί μια ορθοεικόνα σε σύστημα προβολής UTM/WGS84, καλύπτει έκταση 100x100 χλμ² και έχει όγκο περίπου 600 MB. Στο επίπεδο L2C έχουν πραγματοποιηθεί και οι απαραίτητες ατμοσφαιρικές διορθώσεις.

Κατά την περίοδο προεπεξεργασίας των εικόνων της παρούσας εργασίας, δεν είχε ξεκινήσει η παροχή ατμοσφαιρικά διορθωμένων εικόνων. Όμως, παρέχονταν ο αλγόριθμος με τον οποίο είναι δυνατός ο μετασχηματισμός τους. Ο αλγόριθμος αυτός ονομάζεται Sen2Cor. Στην παρούσα εργασία, εκτελέστηκε μέσα από το περιβάλλον Anaconda σε γλώσσα Python 2.7.

Ο αλγόριθμος Sen2Cor βασίζεται στο αριθμητικό μοντέλο ATCOR (Atmospheric and Topographic Correction) το οποίο ανήκει στις απόλυτες ατμοσφαιρικές διορθώσεις, είναι πολύπλοκο αλλά παρέχει μεγάλη ακρίβεια στην μετατροπή των τιμών ανακλαστικότητας από την επιφάνεια της ατμόσφαιρας (Top of Atmosphere, TOA) στην επιφάνεια της Γης (Bottom of Atmosphere, BOA) ή αλλιώς σε τιμές ανάκλασης εδάφους (Surface Reflectance). Περιγράφει με μαθηματικό μοντέλο τη συμπεριφορά της ακτινοβολίας κατά την είσοδό της στην ατμόσφαιρα και την αλληλεπίδρασή της με τα μόρια και τα σωματίδια του αέρα, προβαίνει στην απαλοιφή των ατμοσφαιρικών επιδράσεων εξαιτίας της απορρόφησης και σκέδασης της ηλεκτρομαγνητικής ακτινοβολίας και διαχειρίζεται δύσκολες ατμοσφαιρικές συνθήκες όπως για παράδειγμα την ύπαρξη νεφοκάλυψης [9].

Επιλογή Φασματικών Καναλιών:

Ως προς τα χαρακτηριστικά που θα συμμετείχαν στη ταξινόμηση επιλέχθηκαν τα κανάλια στο ορατό φάσμα (Blue, Green, Red), το κανάλι στο εγγύς υπέρυθρο (NIR) και τα δύο στο μικροκυματικό υπέρυθρο (SWIR). Ακόμη, χρησιμοποιήθηκαν και 4 δείκτες (NDVI, MSAVI, NDWI, NDBI), επιλογές που υπαγορεύτηκαν από παρόμοιες εργασίες, οι οποίες απέδωσαν ακριβή αποτελέσματα [2], [12].

Μετατροπή τύπου εικόνας:

Οι πολυφασματικές εικόνες που παρέχονται από την ESA είναι σε format .jp2. Ο συγκεκριμένος τύπος αρχείου δεν είναι εύκολα διαχωρίσιμος κατά την εκτέλεση αλγορίθμων, ιδίως κάνοντας χρήση της βιβλιοθήκης «GDAL». Κρίνεται συνεπώς απαραίτητη η μετατροπή τους σε μορφή .tiff. Η μετατροπή αυτή εφαρμόστηκε κάνοντας χρήση της βιβλιοθήκης «GDAL» με αλγόριθμο γραμμένο σε γλώσσα python 2.7.

Μετατροπή χωρικής ανάλυσης:

Τα φασματικά κανάλια που επιλέχθηκαν δεν χαρακτηρίζονταν από κοινή χωρική ανάλυση. Συνεπώς απαραίτητη ήταν η μετατροπή της χωρικής ανάλυσης εκείνων των καναλιών με μειωμένη, ώστε να καταστούν δυνατές οι πράξεις μεταξύ τους (διανύσματα με κοινή διάσταση). Για να επιτευχθεί αυτή η μετατροπή αξιοποιήθηκε η δικυβική παρεμβολή από τα 20m στα 10m.

Υπολογισμός δεικτών και δημιουργία υπερκύβων:

Μετά τις απαραίτητες μετατροπές και βελτιστοποιήσεις των πρωτογενών δεδομένων, επόμενο βήμα της διαδικασίας αποτελεί ο υπολογισμός των δεικτών και η δημιουργία των υπερκύβων. Οι δείκτες υπολογίστηκαν βάσει των γνωστών τύπων και κανονικοποιήθηκαν από 0-1. Στην συνέχεια τα επιλεγμένα κανάλια και οι επιλεγμένοι φασματικοί δείκτες καθεμίας ημερομηνίας λήψης αποτέλεσαν και έναν τρισδιάστατο πίνακα, έναν κύβο, με διαστάσεις $m*n*bnd$, όπου $m*n$ οι διαστάσεις κάθε φασματικού χαρακτηριστικού, όπως προκύπτει από την κάμερα λήψης και bnd το πλήθος των φασματικών χαρακτηριστικών, προερχόμενα από συγκεκριμένη ημερομηνία. Οι τρισδιάστατοι αυτοί πίνακες δεδομένων διαστάσεων, στοιβάζονται ο ένας πίσω από τον άλλον δημιουργώντας έτσι υπερκύβους, με βάθος ίσο με το πλήθος των διαφορετικών ημερομηνιών, δηλαδή ίσο με το πλήθος των κύβων.

3.3 Συλλογή και Προ-επεξεργασία Δεδομένων Αναφοράς

Πριν αναλυθεί η πορεία συλλογής των δεδομένων εκπαίδευσης, κρίνεται απαραίτητη η αναλυτική παρουσίαση και καταγραφή των κλάσεων που χρησιμοποιήθηκαν στην ταξινόμηση. Στην συγκεκριμένη παράγραφο, παρουσιάζονται αρχικά οι κλάσεις του επιπέδου 3 του CORINE Land Cover, οι οποίες αποτελούν και το υπόβαθρο των κλάσεων της ταξινόμησης. Στη συνέχεια, αναλύεται η μεθοδολογία επιλογής των κλάσεων και παρουσιάζονται οι κλάσεις που χρησιμοποιήθηκαν τελικά στην ταξινόμηση

3.3.1 Οι κατηγορίες μελέτης/ ταξινόμησης

Για τη παραγωγή του χάρτη χρήσεων γης και καλλιεργειών της περιοχής της Ανατολικής Μακεδονίας και Θράκης, χρησιμοποιήθηκαν 31 κλάσεις, οι οποίες απαρτίζουν τις βασικές καλύψεις γης της περιοχής μελέτης.

Η επιλογή των κλάσεων βασίστηκε στη λογική που ακολουθήθηκε κατά την εκπόνηση προηγούμενων εργασιών ([31], [2], [13]). Πιο συγκεκριμένα, η ονοματολογία που αναπτύχθηκε προήλθε με αρχική βάση τη δομή της ονοματολογίας του CORINE Land Cover (Πίνακας 3.2.), ενώ οι κατηγορίες των αρώσιμων καλλιεργειών βασίστηκαν στα γεωχωρικά δεδομένα, που μας παραχωρήθηκαν από τον Ελληνικό Οργανισμό Πληρωμών των κοινοτικών ενισχύσεων που εποπτεύεται από τον Υπουργό Αγροτικής

Ανάπτυξης και Τροφίμων. Η προτεινόμενη ονοματολογία στην συγκεκριμένη εργασία δομήθηκε ξεκινώντας από τις υποκατηγορίες του τρίτου επιπέδου του CORINE. Καθώς οι κατηγορίες αυτές αποτελούν κατηγορίες χρήσεων/ καλύψεων γης και εφαρμόζονται σε σχετικά μεγάλες μονάδες, δηλ. 25 εκτάρια ή περισσότερο, απαιτείται σε πολλές περιπτώσεις περαιτέρω διαχωρισμός τους (π.χ. Πυκνή και Αραιή Σκληροφυλλική Βλάστηση) ή και αποκλεισμός ορισμένων κατηγοριών που αφορούν χρήση γης (π.χ. Εγκαταστάσεις αθλητισμού και αναψυχής). Ακόμη, λόγω του μικρού ποσοστού εδαφοκάλυψης ορισμένων κλάσεων, κρίθηκε χρήσιμη η ενοποίηση τους, με άλλες παρόμοιες κλάσεις, ώστε να μην προκληθεί σύγχυση κατά την φάση της ταξινόμησης.

Πίνακας 3-2 Ελληνική Ονοματολογία CORINE LAND COVER. Πηγή: <http://portal.survey.ntua.gr>[32]

Ελληνική ονοματολογία CORINE LAND COVER

Πρώτο επίπεδο	Δεύτερο επίπεδο	Τρίτο επίπεδο
1. Τεχνητές επιφάνειες	1.1 Αστικός ιστός	1.1.1 Συνεχής αστικός ιστός 1.1.2 Ασυνεχής αστικός ιστός
	1.2 Βιομηχανικές-εμπορικές ζώνες και δίκτυα μεταφορών	1.2.1 Βιομηχανικές και εμπορικές ζώνες 1.2.2 Οδικά και σιδηροδρομικά δίκτυα 1.2.3 Ζώνες λιμένων 1.2.4 Αεροδρόμια
	1.3 Ορυχεία, χώροι απορρίψεως απορριμμάτων και χώροι οικοδόμησης	1.3.1 Χώροι εξορύξεως ορυκτών 1.3.2 Χώροι απορρίψεως απορριμμάτων 1.3.3 Χώροι οικοδόμησης
	1.4 Τεχνητές μη γεωργικές ζώνες πράσινου	1.4.1 Περιοχές αστικού πρασίνου 1.4.2 Εγκαταστάσεις αθλητισμού και αναψυχής
2. Γεωργικές περιοχές	2.1 Αρόσιμη γη	2.1.1 Μη αρδευόμενη αρόσιμη γη 2.1.2 Μόνιμα αρδευόμενη γη 2.1.3 Ορυζώνες
	2.2 Μόνιμες καλλιέργειες	2.2.1 Αμπελώνες 2.2.2 Οπωροφόρα δένδρα και φυτείες με σαρκώδεις καρπούς 2.2.3 Ελαιώνες
	2.3 Λιβάδια	2.3.1 Λιβάδια
	2.4 Ετερογενείς γεωργικές περιοχές	2.4.1 Ετήσιες καλλιέργειες που σχετίζονται με μόνιμες καλλιέργειες 2.4.2 Σύνθετες καλλιέργειες 2.4.3 Γη που χρησιμοποιείται κυρίως για γεωργία μαζί με σημαντικά τμήματα φυσικής βλάστησης 2.4.4 Γεωργο-δασικές περιοχές
3. Δάση και ημι-φυσικές περιοχές	3.1 Δάση	3.1.1 Δάσος πλατύφυλλων 3.1.2 Δάσος κωνοφόρων 3.1.3 Μικτό δάσος
	3.2 Συνδυασμοί θαμνώδους ή/και ποώδους βλάστησης	3.2.1 Φυσικοί βοσκότοποι 3.2.2 Θάμνοι και χερσότοποι 3.2.3 Σκληροφυλλική βλάστηση 3.2.4 Μεταβατικές δασώδεις και θαμνώδεις εκτάσεις
	3.3 Ανοιχτοί χώροι με λίγη ή καθόλου βλάστηση	3.3.1 Παραλίες, αμμόλοφοι, Αμμουδιές 3.3.2 Απογυμνωμένοι βράχοι 3.3.3 Εκτάσεις με αραιή βλάστηση 3.3.4 Αποτεφρωμένες εκτάσεις 3.3.5 Παγετώνες και αένας χιόνι
4. Υγρότοποι	4.1 Υγρότοποι ενδοχώρας	4.1.1 Βάλτοι στην ενδοχώρα 4.1.2 Τυρφώνες
	4.2 Παραθαλάσσιοι υγρότοποι	4.2.1 Παραθαλάσσιοι βάλτοι 4.2.2 Αλυκές 4.2.3 Ζώνες που καλύπτονται από παλιρροιακά ύδατα
5. Υδάτινες επιφάνειες	5.1 Χερσαία ύδατα	5.1.1 Υδατορρέματα 5.1.2 Επιφάνειες στάσιμου ύδατος
	5.2 Θαλάσσια ύδατα	5.2.1 Παράκτιες λιμνοθάλασσες 5.2.2 Εκβολές ποταμών 5.2.3 Θάλασσες και ωκεανοί

Επιπλέον, ορισμένες κλάσεις (π.χ. αέναο χιόνι), δεν απαντώνται στην συγκεκριμένη περιοχή, συμπέρασμα που οδήγησε στην παράλειψη των κλάσεων αυτών, κατά τη διαδικασία ταξινόμησης των πολυφασματικών εικόνων.







Τέλος, η ταξινόμηση σε αυτό το επίπεδο κλίμακας, οδήγησε στην ανάγκη συλλογής δεδομένων εκπαίδευσης, για κλάσεις, που απαρτίζονται από αντικείμενα μικρού εμβαδού (π.χ. φωτοβολταϊκά), κλάσεις οι οποίες δεν είναι δυνατόν να συμπεριληφθούν στην ιεραρχία του CORINE, όταν αυτό είναι κατασκευασμένο για συλλογή δεδομένων εκπαίδευσης, έκτασης τουλάχιστον 25 εκταρίων.







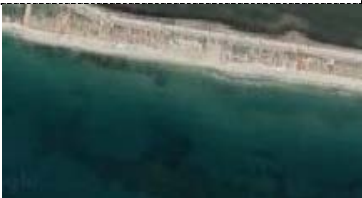


Έτσι, μέσα από τις κατάλληλες μεταβολές της ονοματολογίας του CORINE και με την συμπερίληψη των κατηγοριών καλλιεργειών του ΟΠΕΚΕΠΕ, προέκυψαν 31 κλάσεις, διαφορετικών καλύψεων γης, φασματικά διαχωρίσιμων, οι οποίες παρουσιάζονται στον παρακάτω πίνακα:

Πίνακας 3-3 Κλάσεις της Ταξινόμησης

Αριθμός και Όνομα Κλάσης		Εικονιστικό Παράδειγμα	Αριθμός και Όνομα Κλάσης		Εικονιστικό Παράδειγμα
DUF	Συνεχής Αστικός Ιστός		GRN	Θερμοκήπια	
SUF	Ασυνεχής αστικός ιστός		BLF	Δάσος πλατύφυλλων	
RAN	Οδικά και σιδηροδρομικά δίκτυα		CNF	Δάσος Κωνοφόρων	
ICU	Βιομηχανικές και εμπορικές ζώνες		NGR	Φυσικοί βοσκότοποι	
MES	Χώροι εξόρυξης ορυκτών		DSV	Πυκνή Σκληροφυλλική βλάστηση	

Περιοχή Μελέτης & Προεπεξεργασία Δεδομένων

PHT	Φωτοβολταϊκά Πάνελ		SPSV	Αραιή Σκληροφυλλική Βλάστηση	
RCK	Απογυμνωμένα Βράχια		CRL	Δημητριακά	
SND	Παραλίες, αμμόλοφοι, αμμουδιές		CTN	Βαμβάκι	
WCR	Υδατορρεύματα		GRF	Χορτονομές	
WBD	Επιφάνειες στάσιμου ύδατος		OLG	Ελαιώνες	
MRS	Βάλτοι		MAZ	Καλαμπόκι	

VNY	Αμπελώνες		VEG	Λαχανικά (διάφορα είδη)	
FRT	Οπωροφόρα Δένδρα		TBC	Καπνός	
NGR	Φυσιικοί Βοσκότοποι		RCF	Ρύζι	
CWT	Παράκτια Θαλάσσια Ύδατα		SFL	Ηλίανθος	
KWP	Ακτινίδια		SVG	Αραιή Βλάστηση	

3.3.2 Συλλογή δεδομένων αναφοράς

Η διαδικασία συλλογής των δεδομένων εκπαίδευσης ή αλλιώς του ground truth, αποτέλεσε μία επίπονη και χρονοβόρα διαδικασία, καθώς χρειάστηκε περισσότερο από 2 μήνες σκληρής καθημερινής εργασίας για την καταγραφή του. Έχοντας υπόψη ότι όσο περισσότερο ground truth συλλέξουμε, τόσο πιο επιτυχημένη θα είναι και η ταξινόμηση, καθώς θα επιτευχθεί πιο καλή αναπαράσταση των στατιστικών στοιχείων της κάθε φασματικής κατηγορίας [11] έγινε προσπάθεια πολυπληθούς συλλογής ground truth. Το ground truth πάρθηκε με τέτοιο τρόπο, έτσι ώστε οι καλύψεις γης, που απαντώνται εντονότερα στην περιοχή μελέτης, να διαθέτουν αντίστοιχα και περισσότερο ground truth. Έτσι, ψηφιοποιήθηκε πληθώρα πολυγώνων εκπαίδευσης για τις κλάσεις BLF, CNF, WBD, CRL, MAZ, CWT, καθώς και για τις κλάσεις DSV, SPSV, CTN, GRF, OLG

Στον αντίποδα, για τις κλάσεις SVA, RCK, NGR, GRN, PHT, TBC δεν ήταν δυνατόν να ψηφιοποιηθούν πολλά και μεγάλης έκτασης πολύγωνα εκπαίδευσης, είτε διότι οι κλάσεις αυτές απαντώνται σε μικρό βαθμό στην περιοχή μελέτης, είτε επειδή τα αντικείμενα των συγκεκριμένων εδαφοκαλύψεων είναι πολύ μικρού εμβαδού, που στην χωρική ανάλυση των 10m, δεν είναι δυνατόν να εντοπιστούν (π.χ. φωτοβολταϊκά πάνελ).

Έτσι, συλλέξαμε ανισόποσο ground truth, για 31 κλάσεις. Όπως αναφέρει ο Αργιαλός (1999) [11], η κατανομή των περιοχών εκπαίδευσης στην εικόνα αυξάνει την πιθανότητα των δεδομένων εκπαίδευσης να είναι αντιπροσωπευτικά όλης της μεταβλητότητας, η οποία εμπεριέχεται στις θεματικές κατηγορίες, που βρίσκονται στην περιοχή μελέτης. Συνεπώς, τα πολύγωνα εκπαίδευσης πάρθηκαν σε πολλά και διαφορετικά σημεία της περιοχής μελέτης, έτσι ώστε να καλυφθούν όλες οι περιπτώσεις και να συλλεχθεί λεπτομερής πληροφορία για την κάθε κλάση. Το πλήθος των εικονοστοιχείων, τα οποία συλλέχθηκαν ανά κλάση παρουσιάζονται στον πίνακα 3.3.

Πίνακας 3-4 Πλήθος Εικονοστοιχίων Δεδομένων Εκπαίδευσης ανά Κατηγορία Ταξινόμησης.

		35TLF	35TMF	35TMG
		Pixels	Pixels	Pixels
1	DUF	3829	3601	6
2	SUF	3729	7177	3709
3	RAN	5476	4729	2790
4	ICU	4374	1570	210
5	MES	11447	1349	147
6	PHT	276	226	700
7	GRN	72	87681	56
8	BLF	47379	42584	54291
9	CNF	123740	27943	1481
10	DSV	43190	6453	968
11	SPSV	23435	76	754
12	SVA	1603	44	8
13	RCK	9329	1569	23
14	SND	3457	21896	326
15	WCR	2667	7433	8221
16	WBD	248491	47669	7
17	MRS	6667	959	574
18	VNY	6192	6736	135
19	FRT	2607	8763	10722
20	CRL	89530	25538	
21	CTN	43580	939	
22	GRF	21372	4678	
23	OLG	35601	299	
24	MAZ	61316	1188	
25	VEG	1748	251	
26	NGR	3284	66883	
27	CWT	199298	180584	
28	KWP	7136		
29	TBC	1613		
30	RCF	4339		
31	SFL	5297		
TOTAL		1022074	558818	74406

Για να εντοπιστούν οι διάφορες κατηγορίες κάλυψης γης, εφαρμόστηκε φωτοερμηνεία στις πολυφασματικές εικόνες και επικουρικά χρησιμοποιήθηκε το υπόβαθρο του CORINE. Χρησιμοποιήθηκαν 4 πολυφασματικές εικόνες, διαφορετικών εποχών (μία για την άνοιξη, μία για το καλοκαίρι, μία για το φθινόπωρο και μία για τον χειμώνα). Καθώς οι εικόνες του Sentinel-2, έχουν διακριτική ικανότητα ίση με 10m, δεν είναι εύκολη η φωτοερμηνεία σε αυτό το επίπεδο. Συνεπώς, χρησιμοποιήθηκε και το επίπεδο του Google Earth, του bing και άλλων χαρτογραφικών υποβάθρων, αξιοποιώντας το διαθέσιμο plugin «OpenLayers» στο QGIS . Επειδή υπάρχουν μικρομετακινήσεις ανάμεσα στις εικόνες του Sentinel-2 και στο επίπεδο του Google Earth, οι ψηφιοποιήσεις έγιναν αυστηρά στις πολυφασματικές εικόνες. Ακόμη, καθώς οι εικόνες και το υπόβαθρο του Google Earth, προέρχονται από διαφορετικές ημερομηνίες, κατά τη συλλογή του ground truth, γινόταν εποπτικός έλεγχος για να διαπιστωθεί εάν το αντικείμενο που εμφανίζεται στο Google Earth, ανήκει στην ίδια εδαφοκάλυψη με αυτό των πολυφασματικών εικόνων του Sentinel-2. Τέτοιες μεταβολές εμφανίστηκαν σε εργοτάξια τα οποία έχουν πλέον ολοκληρωθεί, ή σε δάση τα οποία στο υπόβαθρο του Google Earth, φαίνονταν ως νεόφυτα ενώ στις εικόνες του Sentinel-2, φαίνονται ως ανεπτυγμένα. Ακόμη, χρήσιμο αναδείχθηκε το εργαλείο ρυθμιστικό χρόνου του Google Earth, καθώς μέσα από αυτό ήταν δυνατή η περιήγηση σε διαφορετικές ημερομηνίες, δίνοντάς μας τη δυνατότητα να διαπιστώσουμε σε ποια κλάση ανήκουν διφορούμενα αντικείμενα (όπως κάποιες καλλιέργειες αμπέλου). Το υπόβαθρο του CORINE, χρησιμοποιήθηκε με φειδώ, καθώς δεν μπορεί να θεωρηθεί αξιόπιστο για το επίπεδο λεπτομέρειας της παρούσης εργασίας, λόγω των γενικεύσεων που εφαρμόζονται στις χαρτογραφήσεις του CORINE. Ακόμη, δεν θα είχε κάποιο νόημα, εάν το ground truth, προερχόταν ως προϊόν αντιγραφής και όχι ως προϊόν προσωπικής εκτίμησης και φωτοερμηνείας. Η συμβολή του CORINE περιορίστηκε στην εύρεση των γενικών κατηγοριών (π.χ. δάση, αστικές περιοχές), στην εύρεση συγκεκριμένων καλύψεων γης οι οποίες δεν βρέθηκαν με φωτοερμηνεία (π.χ. αεροδρόμιο), ενώ χρησιμοποιήθηκε και ως γνωμοδοτικό εργαλείο σε μία προσπάθεια ψηφιοποίησης διφορούμενων περιοχών.

Τα δεδομένα αναφοράς για τις αρόσιμες καλλιέργειες (CRL, CTN, GRF, MAZ, VEG, TBC, RCF, SFL) προήλθαν από επιστάμενη εργασία φωτοερμηνείας και ψηφιοποίησης των πολυγώνων εφαρμόζοντας τις προαναφερθείσες αρχές και με βάση κυρίως τα γεωχωρικά δεδομένα του ΟΠΕΚΕΠΕ, δηλαδή τα RFVs και τα Ilots. RFVs ονομάζονται οι γρήγορες επισκέψεις πεδίου όλων των αγροτεμαχίων, για τα οποία η φωτοερμηνεία δεν επιτρέπει να επαληθευθεί η ακρίβεια της δήλωσης κατά τρόπο που κρίνεται ικανοποιητικός από την αρμόδια αρχή. Οι επισκέψεις αυτές πραγματοποιούνται μόνο στις περιπτώσεις υποβολής ενστάσεων των παραγωγών κατά των αποτελεσμάτων της τηλεπισκόπησης. Η παραπάνω διαδικασία αποτελεί την πιο αξιόπιστη μέθοδο καταγραφής των καλλιεργειών σε σχέση με την απλή δήλωση του κάθε παραγωγού μέσω του Συστήματος Αναγνώρισης Αγροτεμαχίων (LPIS). Ilots ονομάζονται από τον ΟΠΕΚΕΠΕ οι Ενότητες Ελέγχου και αντιστοιχίζονται στα πολύγωνα που οριοθετούνται από φυσικά όρια (δρόμοι, ποτάμια κλπ). Η κωδικοποίηση στο κάθε ILOT δίνεται όταν το ποσοστό κάλυψης είναι πάνω από 50%. Μέσα σε ένα ILOT μπορεί να υπάρχουν και μη επιλέξιμες περιοχές τα sub-ilots ή δασικές περιοχές οι οποίες χαρακτηρίζονται με αντίστοιχους κωδικούς [33] .

Τα συγκεκριμένα δεδομένα αναφοράς για τις αρώσιμες καλλιέργειες πρόεκυψαν και σε συνδυασμό με αυτά που προήλθαν στο πλαίσιο εκπόνησης διπλωματικής εργασίας της ΣΑΤΜ, του Μεμέτ Ουζούνογλου.

Φωτοερμηνευτικά Χαρακτηριστικά Κλάσεων

Για την αναγνώριση και το διαχωρισμό των καλύψεων γης χρησιμοποιήθηκαν τα φωτοερμηνευτικά χαρακτηριστικά κάθε κλάσης. Ακολούθως αναλύονται για την κάθε κατηγορία τα ιδιαίτερα φωτοερμηνευτικά χαρακτηριστικά της, που οδηγούν στο διαχωρισμό της από τις υπόλοιπες.

Η κατηγορία BLF ξεχωρίζει από το έντονα πράσινο χρώμα της στο έγχρωμο σύνθετο RGB-432, κατά τους θερινούς μήνες και για την στρογγυλή κόμη των δένδρων, τα οποία χαρακτηρίζονται από μαλακή υφή.

Η κατηγορία CNF ξεχωρίζει από το σκούρο πράσινο χρώμα της στο έγχρωμο σύνθετο RGB-432, καθ' όλη τη διάρκεια του χρόνου και για την μυτερή κόμη των δένδρων, η οποία οδηγεί σε οπτικό διαχωρισμό των δένδρων (φαίνονται τα μεμονωμένα δένδρα) και σε εντονότερη υφή.

Η κατηγορία NGR, διακρίνεται για το έντονο πράσινο χρώμα της, ιδιαίτερα τους μήνες της άνοιξης και του φθινοπώρου και τους καφέ τόνους, κατά τη διάρκεια του καλοκαιριού και την απουσία υφής. Για να καταγραφεί ως βοσκότοπος μία περιοχή, θα πρέπει να μην εμφανίζονται δένδρα και θάμνοι, ούτε βράχια, και να βρίσκεται σε περιοχές με φυσική βλάστηση.

Οι κατηγορίες DSV και SPSV διακρίνονται για το σκούρο χρώμα της (καφετί πράσινο) καθ' όλη τη διάρκεια του χρόνου και την έντονη υφή της (ακανθώδεις θάμνοι όπως τα πουρνάρια, τα σκίνα).

Η κατηγορία SND, απαντάται σε παραλιακές περιοχές, ή περιοχές δίπλα σε ποτάμια και λίμνες. Διακρίνεται εύκολα, λόγω χρώματος (μπλεζ- καφέ τόνοι) και λόγω των τοποθεσιών στις οποίες απαντάται.

Η κατηγορία RCK, χαρακτηρίζεται από απουσία βλάστησης και από γκρι χρώμα, απαντάται συνήθως σε υψηλά υψόμετρα και διαθέτει κάποια γραμμικά στοιχεία που υποδηλώνουν την υφή του βράχου.

Η κατηγορία SVA, χαρακτηρίζεται από μικρή και ασήμαντη βλάστηση σε βραχώδη εδάφη και εντοπίζεται εύκολα σε μεγάλα υψόμετρα.

Η κατηγορία DUF, χαρακτηρίζεται από το γκρι- λευκό χρώμα της, καθώς απαρτίζεται από κτήρια και δρόμους σε περιοχές αστικού κέντρου.

Η κατηγορία SUF, χαρακτηρίζεται από εικονοστοιχεία πολλαπλών χρωμάτων (λευκό- γκρι-κόκκινο σκούρο-πράσινο), καθώς απαρτίζεται από κτήρια με πλάκα ή κεραμίδια, δρόμους αλλά και αδόμητο αστικό χώρο, ο οποίος βρίσκεται ενδιάμεσα στα κτίσματα.

Η κατηγορία ICU, ξεχωρίζει λόγω έντονου λευκού χρώματος και του μοτίβου των βιομηχανικών και εμπορικών περιοχών. Τα κτήρια, που την απαρτίζουν, διαφέρουν από τα αστικά σε μέγεθος και χρώμα. Η κατηγορία αυτή είναι εύκολο να ανιχνευθεί

χάρη στην χωροθέτησή της, αφού συνήθως οι βιομηχανικές περιοχές απαντώνται στις παρυφές των αστικών κέντρων.

Η κατηγορία RAN, ξεχωρίζει από το γραμμικό της σχήμα και από το σκούρο γκρι χρώμα.

Η κατηγορία MES, ξεχωρίζει χάρη στο έντονο λευκό- γκρι χρώμα της, τις γραμμές εξόρυξης, καθώς και την ύπαρξη μικρών τεχνητών λιμνών, ενώ εντοπίζεται ως επί το πλείστον σε περιοχές φυσικής βλάστησης.

Η κατηγορία PHT αναδεικνύεται από το ιδιαίτερο μοτίβο της (πάνελ σε γραμμική κατανομή) και από το γκρι-μπλε χρώμα της, ενώ εντοπίζεται συνήθως σε αγροτικές περιοχές.

Η κατηγορία WCR, εντοπίζεται εύκολα, καθώς σε ηπειρωτικές περιοχές εμφανίζονται γραμμικά στοιχεία μπλε χρώματος.

Η κατηγορία WBD, αναδεικνύεται εύκολα, καθώς σε ηπειρωτικές περιοχές εμφανίζεται υδάτινος όγκος μπλε χρώματος.

Η κατηγορία MRS, αναδεικνύεται εύκολα, καθώς εμφανίζεται υδάτινος όγκος, πράσινου-καφέ χρώματος, με την ύπαρξη λάσπης και βλάστησης.

Η κατηγορία CWT, είναι η ευκολότερη κατηγορία να ανιχνευθεί, καθώς απαντάται στις παρόχθιες αβαθείς θαλάσσιες εκτάσεις.

Η κατηγορία VNY, ξεχωρίζει από το γραμμικό μοτίβο –καλλιέργειας της αμπέλου σε λωρίδες με μεγάλες αποστάσεις μεταξύ τους- και από την μεταβολή του πάχους του φυλλώματος των φυτών από εποχή σε εποχή -ανύπαρκτο το χειμώνα, πυκνό το καλοκαίρι.

Η κατηγορία OLG, αναδεικνύεται χάρη την ύπαρξη δένδρων, πρασινωπού φυλλώματος με μεταβολές στην κόμη τους από εποχή σε εποχή.

Για τις υπόλοιπες κατηγορίες των καλλιεργειών δεν παρατίθενται φασματικά χαρακτηριστικά καθώς, οι αρόσιμες καλλιέργειες χαρακτηρίζονται από παρόμοιο χρώμα, κύκλο σποράς- συγκομιδής, αλλά και από γραμμές σποράς. Ο διαχωρισμός τους κατά τη συλλογή δεδομένων εκπαίδευσης επιτεύχθηκε χάρη στην αξιοποίηση των δεδομένων ΟΠΕΚΕΠΕ, όπως αναφέρθηκε στην παραπάνω παράγραφο.

4

ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΣΧΕΔΙΑΣΜΟΣ ΠΕΙΡΑΜΑΤΩΝ

Στο συγκεκριμένο κεφάλαιο, αναλύεται η μεθοδολογία της διεξαγωγής των πειραμάτων ταξινόμησης και παρουσιάζεται η πορεία εφαρμογής του αλγορίθμου.

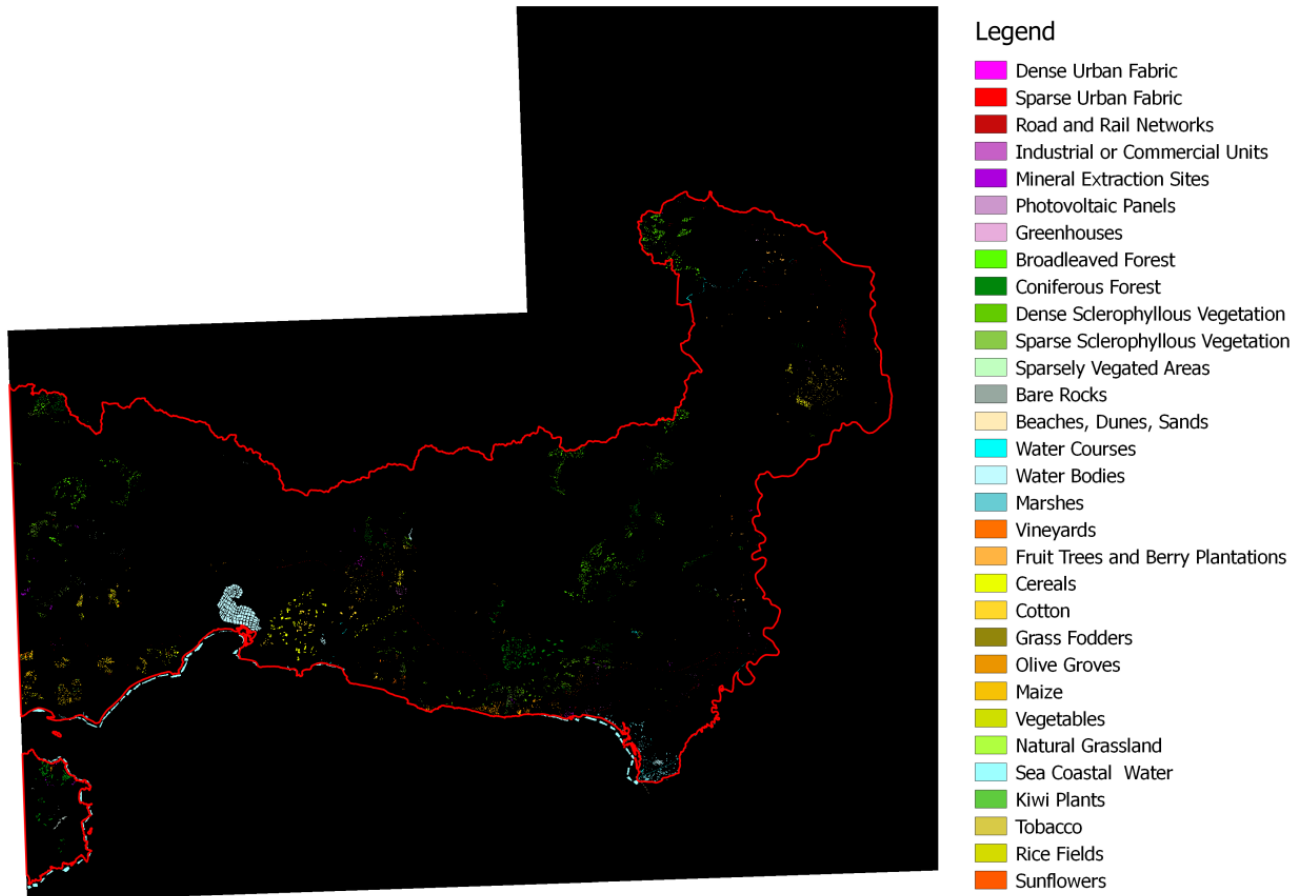
Μέσα από φωτοερμηνεία στις εικόνες του Sentinel-2, αλλά και στο υπόβαθρο του Google Earth, ψηφιοποιήθηκαν τα πολύγωνα εκπαίδευσης, για τις 31 κλάσεις της ταξινόμησης (βλ. Κεφάλαιο 3). Επόμενο βήμα περάτωσης της εργασίας αποτέλεσε η ταξινόμηση των πολυφασματικών δεδομένων. Για τη διεξαγωγή των ταξινομήσεων, επιλέχθηκε ο αλγόριθμος Random Forest. Ο συγκεκριμένος αλγόριθμος, έχει επιφέρει αποτελέσματα υψηλής ακρίβειας σε πρόσφατες εργασίες χαρτογράφησης καλύψεων γης και καλλιεργειών, με αποτέλεσμα να θεωρηθεί κατάλληλος για τη συγκεκριμένη εφαρμογή ([1], [4], [5]).

4.1 Βήματα Εφαρμογής της Μεθοδολογίας Ταξινόμησης

Όπως έχει ήδη αναφερθεί, για την εφαρμογή των ταξινομήσεων στην παρούσα μεταπτυχιακή εργασία, χρησιμοποιήθηκε ο Random Forest ταξινομητής, υλοποιημένος σε γλώσσα python. Ο RF ταξινομητής παρέχεται μέσα από την εργαλειοθήκη «sci-kit learn».

Για την εφαρμογή της ταξινόμησης, απαιτούνταν πολυάριθμα βήματα, προεπεξεργασίας του αλγορίθμου.

Αρχικά, απαραίτητη ήταν η μετατροπή του vector αρχείου των πολυγώνων εκπαίδευσης σε raster εικόνα (ή αλλιώς το «κάψιμο» του ground truth πάνω στην εικόνα). Η διαδικασία αυτή επιτυγχάνεται εύκολα στο περιβάλλον του QGIS χάρη στο εργαλείο Raster Calculator, ορίζοντας τις τιμές ενός καναλιού (οποιουδήποτε) ίσες με το μηδέν και κατόπιν μετατρέποντας το διανυσματικό αρχείο σε raster αρχείο. Δημιουργείται έτσι μία εικόνα με γκρι τόνους, στην οποία οι διαφορετικοί τόνοι του γκρι αντιπροσωπεύουν τις διαφορετικές κλάσεις των δεδομένων εκπαίδευσης.



Εικόνα 4-1 Περιοχές Ground Truth σε απεικόνιση raster εικόνας

Στην συνέχεια, μετατρέπουμε τις εικόνες των υπερκύβων των 3 tile από μορφή .jp2 σε μορφή .tif χρησιμοποιώντας την βιβλιοθήκη «GDAL».

Επόμενο βήμα αποτελεί η σύνθεση των εικόνων κάθε καναλιού σε μία εικόνα πολλών καναλιών. Και πάλι αξιοποιείται η βιβλιοθήκη «GDAL». Για τα κανάλια που έχουν ρixel μεγαλύτερο των 10m, εφαρμόζεται μετατροπή του ρixel στα 10m κάνοντας χρήση της βιβλιοθήκης «OpenCV» και συγκεκριμένα της εντολής «resizing».

Σειρά έχει η δημιουργία των stack για κάθε πολυφασματική εικόνα των διαφορετικών ημερομηνιών. Το stack είναι μία στοιβή αποτελούμενη από τα πολυφασματικά κανάλια και τους επιλεγμένους δείκτες, υπό την μορφή ενός πολυδιάστατου πίνακα, στην περίπτωση μας ενός πίνακα (n,m,bnd), όπου n και m, οι διαστάσεις οποιουδήποτε καναλιού της πολυφασματικής εικόνας και bnd το άθροισμα των καναλιών και / ή των δεικτών.

Για τη δημιουργία των stack, απαραίτητη είναι η αρχικοποίηση του πίνακα, η εκχώρηση των τιμών των καναλιών στους 5 πρώτους πίνακες και ο υπολογισμός και η εκχώρηση των τιμών των δεικτών στους 5 επόμενους πίνακες.

Οι δείκτες που επιλέχθηκαν είναι 4 (NDVI, MSAVI, NDBIA, NDWI) και υπολογίστηκαν από τους τύπους τους, όπως παρατέθηκαν στην παράγραφο 2.4.3, αξιοποιώντας τη βιβλιοθήκη «GDAL». Επιπλέον εφαρμόζεται κανονικοποίηση από 0-1 με σκοπό τη βέλτιστη εκτέλεση της ταξινόμησης. Οι δείκτες αποθηκεύονται με την μορφή νέων «καναλιών» και οι πολυφασματικές εικόνες αποτελούνται πλέον από 10 μπάντες.

Επόμενο βήμα της διαδικασίας αποτελεί η δημιουργία των υπερκύβων των τριών tiles. Χρησιμοποιώντας τα stack των εικόνων, δημιουργούμε τους υπερκύβους, οι οποίοι απαρτίζονται από το σύνολο των stack. Για να τον δημιουργήσουμε πρέπει να τον αρχικοποιήσουμε, ορίζοντας το μέγεθος, την γεωαναφορά του κύβου και την χωρική μετατροπή του raster ως προς τις χωρικές συντεταγμένες. Τις πληροφορίες αυτές τις ορίζουμε βάσει μιας εικόνας ενός καναλιού. Το βάθος του υπερκύβου προκύπτει από τον αριθμό των stack.

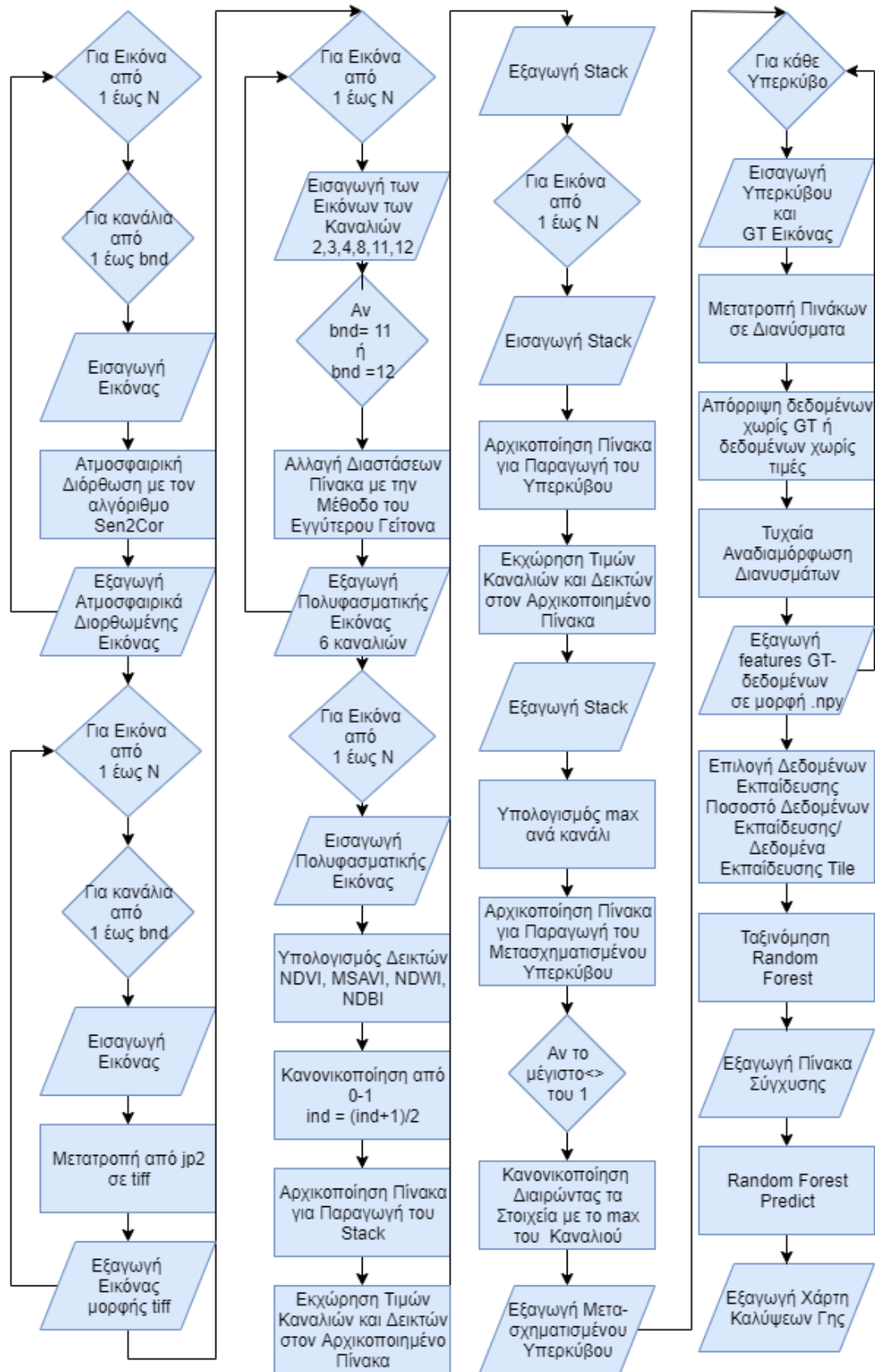
Καθώς τα δεδομένα από τον δορυφόρο Sentinel περιέχουν ορισμένα μη έγκυρα εικονοστοιχεία με πολύ μικρές ή πολύ μεγάλες τιμές, προερχόμενα από κενά στη λήψη, από σφάλμα στο δορυφορικό δέκτη κλπ, υπήρχε η ανάγκη για μία κανονικοποίηση στις τιμές των καναλιών. Η κανονικοποίηση αυτή εφαρμόστηκε ανά κανάλι, βρίσκοντας για κάθε εικόνα για κάθε κανάλι τις μέγιστες τιμές του υπολογίζοντας τον μέσο όρο τους και διαιρώντας κάθε τιμή των εικονοστοιχείων της εικόνας με τη γενικευμένη μέγιστη τιμή. Καθώς οι δείκτες είχαν ήδη κανονικοποιηθεί εξαιρέθηκαν από την συγκεκριμένη διαδικασία αναπροσαρμογής τιμών.

Κατόπιν, εξάγονται τα χαρακτηριστικά από το ground truth. Φορτώνεται ο υπερκύβος και το ground truth. Επιλέγονται με τυχαία σειρά δορυφορικά δεδομένα και ground truth data. Αγνοούμε τα εικονοστοιχεία χωρίς τιμή στη δορυφορική εικόνα και τα εικονοστοιχεία χωρίς ground truth και εφαρμόζουμε τυχαία μετάθεση. Εξάγουμε έτσι την πληροφορία του Ground Truth.

Τέλος, τα εικονοστοιχεία δειγματοληψίας χωρίζονται σε δύο σύνολα. Το πρώτο σύνολο αντιπροσωπεύει τα δεδομένα εκπαίδευσης του αλγορίθμου, ενώ το δεύτερο σύνολο αποτελεί τα δεδομένα ελέγχου του αλγορίθμου. Ανάλογα το πείραμα, αυτά τα σύνολα διαφέρουν. Εισάγονται ως ορίσματα στο μοντέλο RF της βιβλιοθήκης «scikit», το μοντέλο τρέχει και εξάγει ως αποτελέσματα ο πίνακας σύγχυσης της ταξινόμησης.

Τέλος, εφαρμόζοντας την εντολή predict, ο ταξινομητής αποδίδει τιμή σε κάθε pixel της εικόνας, βάση των προτύπων που χτίστηκαν κατά τη διαδικασία της ταξινόμησης.

Στην συνέχεια παρουσιάζεται το διάγραμμα ροής του αλγορίθμου ταξινόμησης, όπως εφαρμόστηκε στο προγραμματιστικό περιβάλλον της python3 .



Διάγραμμα 4-1 Διάγραμμα Ροής του Προγράμματος Ταξινόμησης, στην Python3

4.2 Κατηγορίες Πειραμάτων

Για την εκπόνηση της παρούσας εργασίας αποπερατώθηκαν 22 συνολικά πειράματα. Ο συγκεκριμένος αριθμός πειραμάτων σχετίζεται με τη συλλογιστική διαδικασία, τις παραμέτρους που τέθηκαν, αλλά και την απαιτούμενη υπολογιστική ισχύ για κάθε πείραμα. Τα πειράματα που υλοποιήθηκαν προέκυψαν με γνώμονα τις εικόνες/ τα tiles που θα αξιοποιηθούν κάθε φορά για την εκπαίδευση του αλγορίθμου αλλά και για τον έλεγχο αυτού. Πιο συγκεκριμένα αποπερατώθηκαν πειράματα ακολουθώντας τον εξής βασικό άξονα:

Χρήση μίας εικόνας για εκπαίδευση και έλεγχο στην ίδια
Χρήση μίας εικόνας για εκπαίδευση και έλεγχο σε μία δεύτερη
Χρήση όλων των εικόνων για εκπαίδευση και έλεγχο σε όλες.

Συνεπώς, τα πειράματα ομαδοποιήθηκαν σε τρεις κατηγορίες. Ανά κατηγορία έγιναν πολυάριθμα πειράματα, ώστε να καλύψουμε όλους τους δυνατούς συνδυασμούς επιλογής παραμέτρων.

4.2.1 Πρώτη κατηγορία πειραμάτων - ανά tile

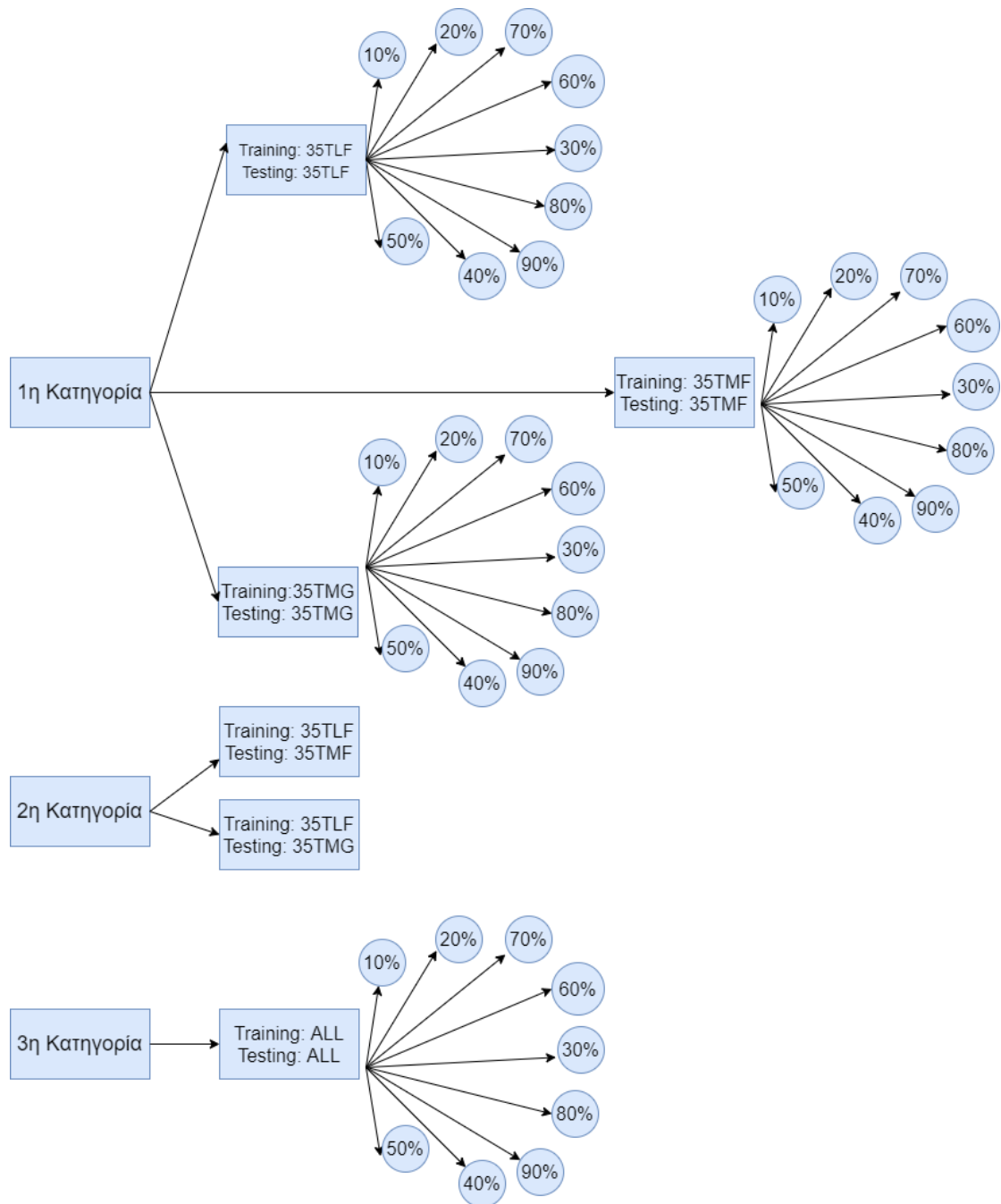
Στη πρώτη ομάδα πειραμάτων διεξήχθησαν 27 πειράματα. Οι ταξινομήσεις υλοποιήθηκαν χρησιμοποιώντας ένα tile για εκπαίδευση του αλγορίθμου και έλεγχο των αποτελεσμάτων αυτού. Για κάθε tile έγιναν 9 ταξινομήσεις, επιλέγοντας κάθε φορά διαφορετικό ποσοστό δεδομένων εκπαίδευσης, από 10 έως 90%.

4.2.2 Δεύτερη κατηγορία πειραμάτων - ανά δύο tiles

Έχοντας επιλέξει τρία tiles για την περαίωση ταξινόμησης δορυφορικών, κρίθηκε σκόπιμο να αξιοποιηθεί μία εικόνα για την εκπαίδευση του αλγορίθμου και μία από τις υπόλοιπες εικόνες να αξιοποιηθεί για τον έλεγχο του αλγορίθμου. Καθώς δεδομένα εκπαίδευσης για κάθε κλάση της ταξινόμησης απαντώνται μόνο σε ένα tile και συγκεκριμένα στο 35TLF, το συγκεκριμένο tile ήταν το μόνο που μπορούσε να αξιοποιηθεί στην εκπαίδευση του αλγορίθμου. Έτσι σχεδιάστηκαν δυο πειράματα. Στο πρώτο αξιοποιήθηκε για έλεγχο το tile 35TMG και στο δεύτερο το 35TMF.

4.2.3 Τρίτη κατηγορία πειραμάτων - ανά τρία tiles

Στην τρίτη κατηγορία πειραμάτων υλοποιήθηκε ένα πείραμα. Αξιοποιήθηκαν και τα τρία tiles για εκπαίδευση του αλγορίθμου και έλεγχο, ενώ έγιναν 9 ταξινομήσεις, επιλέγοντας κάθε φορά διαφορετικό ποσοστό δεδομένων εκπαίδευσης, από 10 έως 90%.



Διάγραμμα 4.1. Σχηματική απεικόνιση των πειραμάτων

5

ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ

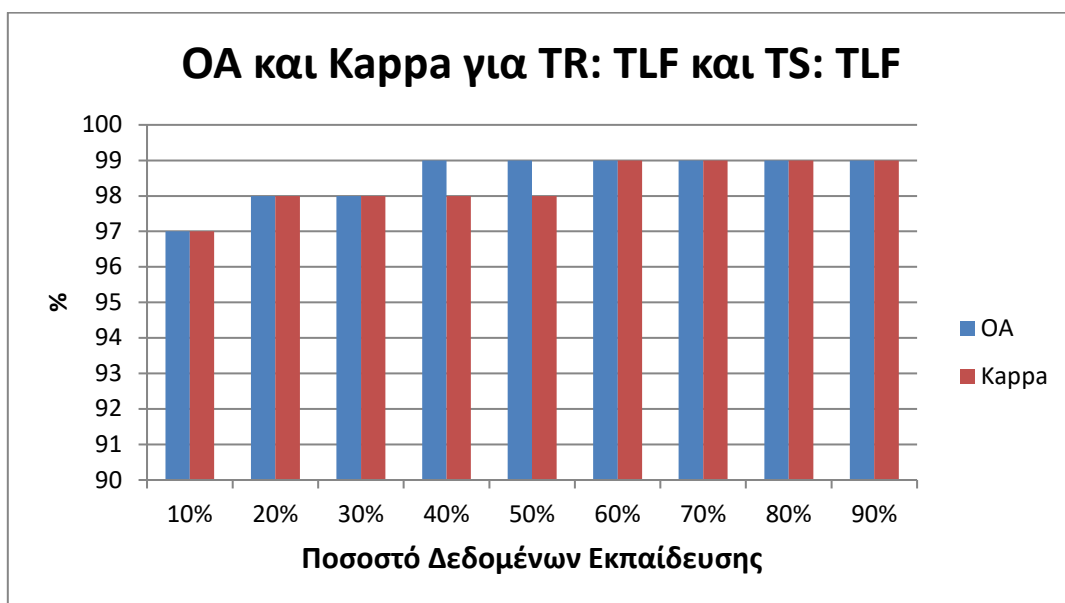
Στο κεφάλαιο που ακολουθεί, παρουσιάζονται αναλυτικά τα αποτελέσματα των πειραμάτων των τριών κατηγοριών, η συγκριτική τους αξιολόγηση και τέλος η παρουσίαση ορισμένων χαρακτηριστικών περιπτώσεων και η ανάλυση τους με εικόνες-παραδείγματα.

5.1 Πρώτη Κατηγορία Πειραμάτων- Ανά Tile

Στην πρώτη κατηγορία πειραμάτων, χρησιμοποιήθηκε το ίδιο tile τόσο για εκπαίδευση όσο και για έλεγχο του αλγορίθμου της ταξινόμησης. Υλοποιήθηκαν, πολλαπλές ταξινομήσεις για το κάθε tile, με σκοπό τη διερεύνηση του βέλτιστου ποσοστού δεδομένων εκπαίδευσης. Ελέγχοντας τα μεγέθη αξιολόγησης της ταξινόμησης, καταλήξαμε στις βέλτιστες επιλογές.

5.1.1 Πρώτο πείραμα: ταξινόμηση στο tile 35TLF.

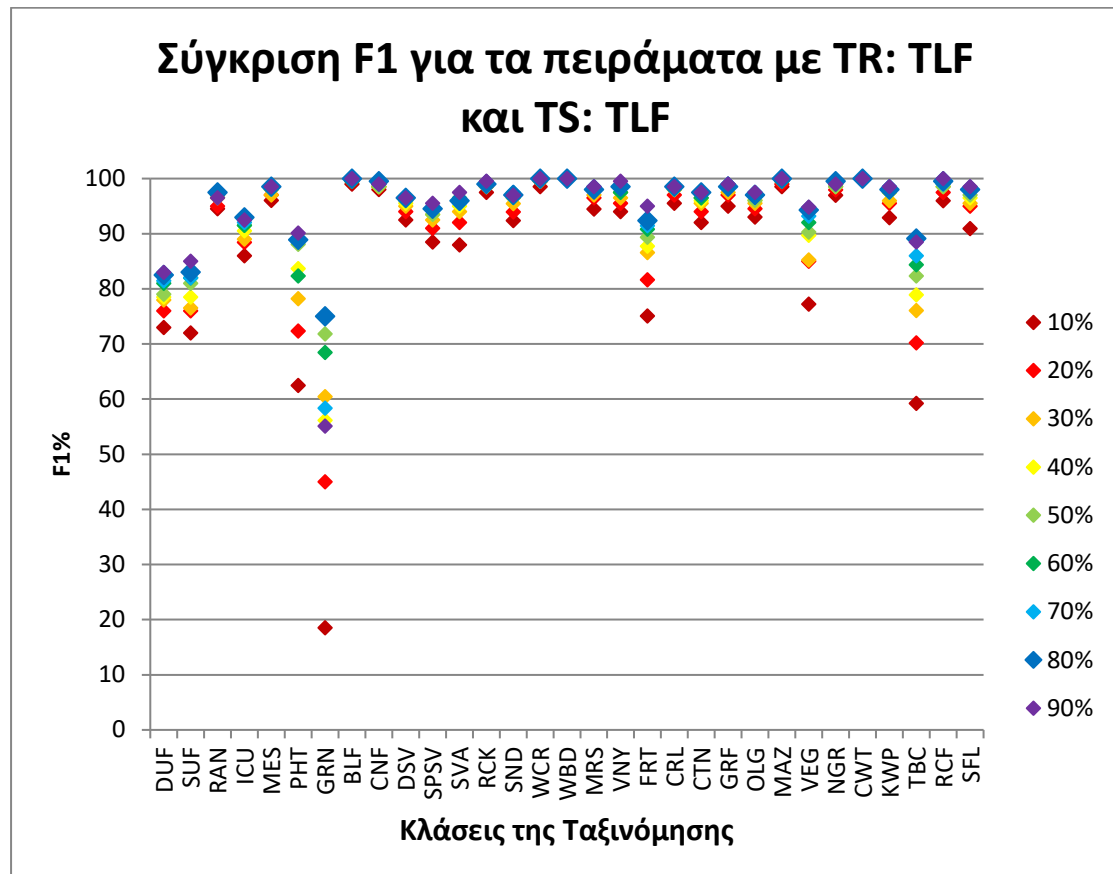
Στο πρώτο πείραμα που εφαρμόστηκε επιλέχθηκε να χρησιμοποιηθεί το Tile 35TLF για να εκπαιδευτεί ο αλγόριθμος, ενώ ο έλεγχος της ταξινόμησης επιλέχθηκε να γίνει στο ίδιο Tile. Χρησιμοποιήθηκαν και οι έξι επιλεγμένες εικόνες και από αυτές τα κανάλια 2,3,4,8,11,12, καθώς και οι δείκτες NDVI, MSAVI, NDWI, NDBI. Έτρεξαν 9 ταξινομήσεις, για εκπαίδευση του ταξινομητή με 10, 20, 30, 40, 50, 60, 70, 80 και 90% των δεδομένων αναφοράς. Τα αποτελέσματα που προέκυψαν από τις ταξινομήσεις δεν διέφεραν σημαντικά. Η συνολική ακρίβεια της ταξινόμησης και ο δείκτης Καρρα, για τα εννέα πειράματα, παρουσιάζονται στο παρακάτω γράφημα.



Γράφημα 5-1 Συνολική ακρίβεια και δείκτης Καρρα για την ταξινόμηση στην εικόνα 35TLF

Μελετώντας το γράφημα, διαπιστώνεται ότι και οι 9 ταξινομήσεις έχουν αποδώσει εξίσου καλά αποτελέσματα. Σημειώνεται ότι όσο περισσότερο ποσοστό των δεδομένων εκπαίδευσης χρησιμοποιούνται στην ταξινόμηση, τόσο υψηλότερης ακρίβειας αποτελέσματα αποδίδει ο αλγόριθμος, ενώ οι ταξινομήσεις στις οποίες χρησιμοποιήθηκε ποσοστό δεδομένων εκπαίδευσης από 50% και πάνω, αποδίδουν παρόμοια αποτελέσματα. Ελέγχοντας τα αποτελέσματα ακρίβειας και αξιοπιστίας της κάθε κλάσης της ταξινόμησης καθώς και το δείκτη F1, διαπιστώνεται ότι σχεδόν όλες οι κλάσεις αποδίδουν καλύτερα αποτελέσματα όσο αυξάνεται το ποσοστό των δεδομένων εκπαίδευσης που χρησιμοποιείται για εκπαίδευση. Όμως, τα αποτελέσματα της ακρίβειας και της αξιοπιστίας των κλάσεων δεν διαφέρουν σημαντικά στις ταξινομήσεις, που χρησιμοποιείται ποσοστό δεδομένων εκπαίδευσης από 50% και πάνω.

Στην συνέχεια παρατίθεται το γράφημα για το δείκτη F1 για την κάθε κλάση της ταξινόμησης.



Γράφημα 5-2 Δείκτης F1 κάθε κλάσης για την ταξινόμηση στην εικόνα 35TLF.

Επιλέγεται να παρατεθεί και αξιολογηθεί η ταξινόμηση με ποσοστό δεδομένων εκπαίδευσης ίσο με 50%, καθώς σημειώνει υψηλά ποσοστά στα μεγέθη αξιολόγησης της ταξινόμησης και των κλάσεων αυτής, τα οποία δεν μεταβάλλονται σημαντικά στις ταξινομήσεις με ποσοστό δεδομένων εκπαίδευσης πάνω από 50%.

Παρακάτω παρατίθεται ο πίνακας σύγκρισης της ταξινόμησης. Εν γένει η ταξινόμηση θεωρείται πετυχημένη, καθώς, η γενική ακρίβεια και ο δείκτης kappa ξεπερνούν το 98% και 99%,αντιστοίχως.

Πίνακας 5-1 Πίνακας Σύγκρισης για ταξινόμηση στην εικόνα 35TLF.

		ΚΛΑΣΕΙΣ ΤΟΥ ΑΠΟΤΕΛΕΣΜΑΤΟΣ ΤΗΣ ΤΑΞΙΝΟΜΗΣΗΣ																																	
		DU	SUF	RA	ICU	ME	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SN	WC	WB	MR	VN	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL	su	PA	
ΚΛΑΣΗΣ ΤΟΥ ΧΑΡΗΤΗ ΤΟΥ GROUND TRUTH	DU	152	209	20	103	12	0	0	0	0	0	3	0	3	4	0	0	0	7	0	9	16	0	6	0	0	0	0	0	0	0	0	0	191	0,8
	SUF	243	150	7	6	0	0	0	0	0	0	13	0	0	1	0	0	1	9	3	38	17	5	20	0	1	0	0	1	0	0	0	186	0,8	
	RA	19	6	265	19	1	1	0	0	0	0	6	1	1	5	0	0	1	0	1	7	4	0	9	0	0	0	0	0	0	0	273	0,9		
	ICU	116	9	22	200	29	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	218	0,9		
	ME	6	1	4	57	561	0	0	0	0	0	1	2	20	16	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	572	0,9		
	PHT	0	1	6	0	0	111	0	0	0	0	0	0	1	0	0	0	1	0	0	10	4	0	0	1	0	0	0	3	0	0	138	0,8		
	GR	6	2	1	1	1	0	20	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	36	0,5		
	BLF	0	0	0	0	0	0	0	236	10	38	10	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	8	0	0	0	236	1	
	CNF	0	0	0	0	0	0	0	39	614	260	48	0	0	0	0	0	1	0	6	0	0	1	33	0	0	1	0	0	0	0	0	618	0,9	
	DSV	0	0	0	0	0	0	0	65	715	205	243	0	0	0	0	0	1	0	0	0	0	0	35	0	0	0	0	0	0	0	0	215	0,9	
	SPS	0	2	3	0	0	0	0	1	49	335	109	3	2	0	0	0	1	1	1	0	0	0	379	0	0	0	0	0	0	0	0	117	0,9	
	SVA	1	7	1	0	0	0	0	1	0	0	19	765	7	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	802	0,9	
	RCK	9	7	2	10	13	0	0	0	0	0	2	6	461	2	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	466	0,9	
	SN	3	3	2	12	41	1	0	0	0	0	5	0	7	164	0	0	0	0	0	4	1	0	5	0	0	0	1	0	0	0	172	0,9		
	WC	0	0	0	0	0	0	0	0	1	0	1	0	0	1	132	1	3	0	0	0	0	0	0	0	0	0	1	0	0	0	133	0,9		
	WB	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	1E+	38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1E+	1	
	MR	2	7	5	0	0	0	0	0	1	0	6	0	1	1	0	8	327	0	3	6	3	1	13	0	0	0	0	1	0	0	0	333	0,9	
	VN	0	10	2	0	0	0	0	0	0	0	7	0	0	0	0	1	298	3	24	24	1	29	9	0	0	0	1	0	0	0	309	0,9		
	FRT	0	5	0	0	0	0	0	8	25	4	28	0	0	0	0	5	0	111	15	7	18	58	3	0	0	0	8	0	0	1	130	0,8		
	CRL	4	10	0	0	0	0	0	0	0	2	14	2	0	0	0	1	5	8	442	299	10	18	88	0	0	0	0	1	0	6	447	0,9		
CTN	8	19	2	0	0	0	0	3	0	2	3	0	0	0	0	1	9	2	856	208	13	23	2	2	0	0	0	6	0	11	217	0,9			
GRF	0	1	0	0	0	0	0	2	0	1	8	3	0	0	0	0	2	8	33	7	105	6	13	2	0	0	25	0	8	15	106	0,9			
OLG	3	8	5	0	0	0	0	1	154	79	294	0	0	0	0	8	3	24	55	18	3	171	0	0	0	3	0	0	0	178	0,9				
MA	0	2	0	0	0	0	0	1	0	0	3	0	0	0	0	1	0	9	8	8	34	2	305	1	0	0	4	1	7	6	306	1			
VEG	0	13	0	0	0	0	0	0	0	0	1	0	0	0	0	3	2	45	48	10	2	15	728	0	0	0	3	0	4	874	0,8				
NG	0	0	0	0	0	0	0	32	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	164	0,9			
CW	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	996	1			
KW	1	3	0	0	0	0	0	0	0	0	8	0	0	0	0	0	2	1	8	1	7	60	6	33	0	0	0	0	1	356	0,9				
TBC	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	131	0	2	0	0	0	0	0	0	0	807	0,7				
RCF	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	1	1	2	17	1	9	1	0	3	0	212	0	217	0,9			
SFL	0	2	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	1	10	15	46	0	33	0	0	2	0	1	253	264	0,9				
su	194	183	273	221	571	113	20	237	624	212	116	787	465	167	132	1E+	334	302	120	455	214	107	177	307	736	161	996	348	581	214	257	kap	0,9		
UA	0,7	0,8	0,9	0,9	0,9	0,9	1	0,9	0,9	0,9	0,9	0,9	0,9	0,9	1	1	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	1	0,9	0,9	0,9	0,9	0,9	0A:	0,9	

Μελετώντας επιμέρους την ακρίβεια και την αξιοπιστία ανά κατηγορία, διαπιστώνεται ότι όλες οι κατηγορίες έχουν ταξινομηθεί επιτυχώς μιας και τα ποσοστά των μετρικών ακριβείας αγγίζουν σε κάθε περίπτωση το 80%.

Πιο συγκεκριμένα, εστιάζοντας στις κλάσεις των τεχνητών επιφανειών, διαπιστώνεται ότι οι κλάσεις DUF και SUF, εμπλέκονται, καθώς τα σφάλματα παράλειψης και συμπερίληψης είναι της τάξης του 12%. Επιπλέον η κατηγορία DUF εμφανίζει μικρότερη σύγχυση και με την κατηγορία ICU, με σφάλματα παράλειψης και συμπερίληψης 6% και 5% αντίστοιχα. Οι συγχύσεις ανάμεσα στις κατηγορίες DUF και SUF είναι αναμενόμενες, καθώς πρόκειται για υποκατηγορίες της γενικής κατηγορίας, παρουσιάζουν κοινά φασματικά .

Επιπλέον η σύγχυση της DUF με την ICU είναι επίσης ως ένα βαθμό αναμενόμενη, καθώς, τα κτήρια που απαρτίζουν την κατηγορία ICU είναι δύσκολο να διαχωριστούν πλήρως από αυτά που απαρτίζουν την DUF. Ακόμη, η ICU εμφανίζει μικρή σύγχυση με την MES. Η κατηγορία PHT, χαρακτηρίζεται από υψηλή ακρίβεια παραγωγής (98%), όμως η ακρίβεια χρήστη είναι σημαντικά χαμηλότερη (80%), αφού 7% των εικονοστοιχείων της κλάσης CRL και 5% των εικονοστοιχείων της κλάσης RAN, συμπεριλαμβάνονται εσφαλμένα στην PHT. Τέλος η κλάση GRN, εμφανίζει άριστη ακρίβεια και μέτρια αξιοπιστία, καθώς 17% των pixel της DUF και 5% της SUF, ταξινομούνται στην GRN. Οι συγχύσεις που παρουσιάζονται στις κλάσεις PHT και GRN είναι αποτέλεσμα του μικρό αριθμού δεδομένων εκπαίδευσης που παρέχονται στον ταξινομητή, καθώς δεν υπάρχει επαρκές δείγμα ώστε ο αλγόριθμος να εκπαιδευτεί κατάλληλα και να κατόπιν να διαχωρίσει αυτές τις κατηγορίες.

Οι κατηγορίες της φυσικής βλάστησης ταξινομήθηκαν στο σύνολό τους επιτυχημένα από τον αλγόριθμο ταξινόμησης. Μικρές συγχύσεις παρουσιάζονται ανάμεσα στις κλάσεις DSV και SPSV. Πιο συγκεκριμένα, η κλάση DSV, παρουσιάζει μικρά σφάλματα παράλειψης και συμπερίληψης με τις κατηγορίες CNF και SPSV. Αυτές οι συγχύσεις έχουν παρατηρηθεί σε όλα τα πειράματα που έχουμε ήδη εξετάσει και οφείλονται στα παρόμοια φασματικά χαρακτηριστικά των κατηγοριών αυτών. Τα pixel της κατηγορία CNF απαρτίζονται από κωνοφόρα δένδρα, που το φύλλωμά τους εμφανίζει παρόμοια χαρακτηριστικά με τα δενδρύλλια σκληροφυλλικής βλάστησης. Σχετικά με τις κλάσεις DSV και SPSV αποτελούν υποκατηγορίες της γενικότερης κλάσης σκληροφυλλική βλάστηση και η διαφοροποίησή τους έγκειται στην πυκνότητα της βλάστησης, γεγονός που καθιστά δύσκολο το διαχωρισμό τους από τον ταξινομητή Τέλος, μικρή σύγχυση, που διαφαίνεται από το σφάλμα συμπερίληψης 2%, εμφανίζεται στην κλάση NGR, με την κλάση BLF. Η μικρή αυτή σύγχυση, θεωρείται απόρροια της παρόμοιας ανάκλασης των δύο κατηγοριών στην περιοχή του πράσινου, του κόκκινου και του εγγύς υπέρυθρου.

Οι κατηγορίες του γυμνού εδάφους διαχωρίζονται επιτυχώς, από τον ταξινομητή καθώς όλες οι κατηγορίες πετυχαίνουν ακρίβειες χρήστη και παραγωγού μεγαλύτερες του 95%. Σημαντικές συγχύσεις συνεπώς δεν παρατηρούνται. Αξιοσημείωτο είναι όμως το γεγονός, ότι στην κλάση SVA, παρατηρείται μικρή σύγχυση (σφάλμα συμπερίληψης περί τα 2%) με την κλάση SPSV. Η συμπερίληψη των εικονοστοιχείων της SPSV στην SVA είναι δικαιολογημένη, επειδή και οι δύο κατηγορίες απαρτίζονται

από εικονοστοιχεία που περιέχουν γυμνό έδαφος και βλάστηση. Έτσι, τα φασματικά τους χαρακτηριστικά ως ένα βαθμό συγκλίνουν και δυσκολεύουν τον διαχωρισμό τους. Μία ακόμη μικρή αλλά αξιοσημείωτη σύγχυση είναι εκείνη της κατηγορία SND με την κατηγορία MES. Και η συγκεκριμένη σύγχυση έγκειται στο γεγονός ότι στα εικονοστοιχεία της MES απαντώνται εικονοστοιχεία άμμου, αφού στην κλάση αυτή συναντώνται περιοχές άμμου.

Οι κατηγορίες υδάτινων πόρων διαχωρίστηκαν πλήρως από τον αλγόριθμο ταξινόμησης, αφού επιτεύχθηκε ακρίβεια και αξιοπιστία για κάθε κλάση μεγαλύτερη και από 98%, χωρίς την ύπαρξη αξιόλογων εμπλοκών ανάμεσα στις διάφορες κατηγορίες.

Ως προς τις κλάσεις που ανήκουν στις καλλιέργειες διαπιστώνουμε ότι, όλες οι κλάσεις έχουν ταξινομηθεί επιτυχημένα από τον ταξινομητή, καθώς έχει πετύχει ακρίβειες παραγωγού και χρήστη ανώτερες του 70% σε κάθε περίπτωση.

Πιο συγκεκριμένα, η κατηγορία VNY, διαχωρίζεται άριστα καθώς η ακρίβεια παραγωγού και χρήστη ξεπερνά το 98% και το 95% αντίστοιχα, ενώ οι συγχύσεις που εμφανίζονται με τις άλλες κατηγορίες είναι αμελητέες.

Η κλάση FRT, εμφανίζει υψηλή ακρίβεια ανώτερη του 93% και λίγο χαμηλότερη αξιοπιστία της τάξης του 85%. Οι σημαντικότερες συγχύσεις που παρατηρούνται είναι αυτές με την κατηγορία OLG, καθώς παρατηρείται σφάλμα συμπερίληψης μεγαλύτερο από 4% και μικρότερες με τις κλάσεις SPSV και CNF με σφάλματα συμπερίληψης περί το 2%. Οι κλάσεις με τις οποίες εμφανίζονται συγχύσεις απαρτίζονται από δένδρα ή δενδρύλλια, παρόμοιου φυλλώματος με των οπωροφόρων δένδρων, με αποτέλεσμα να εμφανίζουν παρόμοια φασματικά χαρακτηριστικά, που οδηγούν σε σφάλματα στην ταξινόμηση. Όμως, τα σφάλματα αυτά είναι περιορισμένα και δεν οδηγούν σε μειωμένη ακρίβεια ταξινόμησης.

Η κατηγορία CRL εμφανίζει πολύ υψηλά ποσοστά ακρίβειας και αξιοπιστίας, μεγαλύτερα του 97%. Δεν εμφανίζει σημαντικές συγχύσεις με άλλες κατηγορίες, παρά μόνο παράλειψη του 2% των εικονοστοιχείων της τάξης και συμπερίληψής τους στην κατηγορία CTN. Οι δύο καλλιέργειες έχουν παρόμοιους κύκλους συγκομιδής, με αποτέλεσμα να υπάρχει μία δυσκολία στον ταξινομητή να διαχωρίσει τις καλλιέργειες αυτές. Η κατηγορία GRF, παρουσιάζει χαμηλά ποσοστά ακρίβειας και αξιοπιστίας (27,69% και 23,55% αντίστοιχως). Η συγκεκριμένη κλάση εμπλέκεται σε μικρό βαθμό με τις περισσότερες κλάσεις καλλιεργειών και συγκεκριμένα με τους αμπελώνες, τα οπωροφόρα δένδρα, τα δημητριακά, το βαμβάκι και τους ελαιώνες.

Η κατηγορία OLG, έχει διαχωριστεί πλήρως από τον ταξινομητή, αφού εμφανίζει ποσοστά ακρίβειας παραγωγού και χρήστη μεγαλύτερα από 96%. Ελέγχοντας τα σφάλματα παράλειψης δεν διαπιστώνουμε αξιοσημείωτες συγχύσεις. Η μόνη σύγχυση, που αν και μικρή (σφάλμα παράλειψης 2% και συμπερίληψης 1,6%), αξίζει να αναφερθεί είναι εκείνη με την κατηγορία SPSV. Οι δύο κατηγορίες, απαρτίζονται από γυμνό έδαφος και δενδρύλλια με σκούρο και σκληρό φύλλωμα, γεγονός που οδηγεί σε μικρά σφάλματα ταξινόμησης.

Η κατηγορία MAZ είναι μία κατηγορία καλλιεργειών την οποία ο ταξινομητής καταφέρνει να διαχωρίσει άριστα, καθώς πρόκειται για μια καλλιέργεια με πολύ

περισσότερη ανάγκη για νερό και αυξημένη υγρασία, που ανακλά αρκετά διαφορετικά σε σχέση με τις υπόλοιπες κατηγορίες.

Η κατηγορία VEG είναι μία κατηγορία καλλιεργειών, η οποία χαρακτηρίζεται από πολύ υψηλή ακρίβεια παραγωγής, ανώτερη του 98%, αλλά η ακρίβεια χρήστη είναι χαμηλότερη περί το 83%. Τα σφάλματα συμπερίληψης μαρτυρούν μικρή αδυναμία του ταξινομητή να τη διαχωρίσει από την καλλιέργεια των σιτηρών και του βαμβακιού, καθώς 5% των εικονοστοιχείων των κατηγοριών αυτών, συμπεριλαμβάνεται στην κατηγορία των λαχανικών. Τα μικρά αυτά σφάλματα του ταξινομητή οφείλονται στο γεγονός ότι οι κλάσεις αυτές εμφανίζουν κοινά φασματικά χαρακτηριστικά. Η κατηγορία λαχανικά, είναι μία ευρύτερη κατηγορία, που περιλαμβάνει τις εξής καλλιέργειες:

Πίνακας 5-2 Καλλιέργειες που συμπεριλαμβάνονται στην κλάση των VEG

ΚΑΛΛΙΕΡΓΕΙΕΣ ΛΑΧΑΝΙΚΩΝ
Λαχανικά Διάφορα
Τομάτες
Καρπούζι, πεπόνι
Φράουλες
Μπάμιες
Σπαράγγια

Οι καλλιέργειες αυτές εμφανίζουν ποικιλομορφία καθώς έχουν διαφορετικά ειδικά χαρακτηριστικά, διαφορετικές περιόδους σποράς και συγκομιδής. Οι διαφοροποιήσεις αυτές, δημιουργούν ένα φασματικό μοτίβο, το οποίο μπορεί σε ορισμένες περιπτώσεις να αντιστοιχηθεί με εκείνο των σιτηρών, αλλά και με εκείνο του βαμβακιού. Αποτέλεσμα είναι, ο ταξινομητής να οδηγείται σε κάποιες εσφαλμένες ταξινομήσεις ρίχει, λόγω αυτής της ταύτισης των φασματικών χαρακτηριστικών ορισμένων καλλιεργειών, που ανήκουν στην γενικότερη κατηγορία.

Η κατηγορία KWP, είναι εύκολα διαχωρίσιμη από το ανθρώπινο μάτι, καθώς είναι ένα φυτό δικοτυλήδονο και αναρριχόμενο, διαθέτει μεγάλο και στρογγυλωπό φύλλωμα και θυμίζει το αμπέλι. Η συγκομιδή τους γίνεται μετά τα μέσα του φθινοπώρου, ενώ πρόκειται για πολυετείς καλλιέργειες. Η φυσιολογία τους αυτή, τα διαφοροποιεί τόσο κατά την συλλογή δεδομένων εκπαίδευσης όσο και κατά τη διαδικασία ταξινόμησης τους, καθώς οι ακρίβειες παραγωγού και χρήστη ξεπερνούν το 96%. και δεν εμφανίζονται αξιόλογα σφάλματα παράλειψης ή και συμπερίληψης με άλλες κλάσεις. Η μόνη σύγχυση που μπορεί να παρατηρηθεί, αν και δεν είναι αξιοσημείωτη είναι εκείνη με την κλάση GRF.

Η καλλιέργεια TBC εμφανίζει υψηλή ακρίβεια παραγωγού, καθώς το ποσοστό αυτό ξεπερνά το 98%, όμως η ακρίβεια χρήστη είναι αρκετά μικρότερη καθώς δεν ξεπερνά το 70%. Το μειωμένο ποσοστό στην αξιοπιστία της συγκεκριμένης κλάσης, προκύπτει από την συμπερίληψη 16% και 12% εικονοστοιχείων των κατηγοριών CRL και CTN αντίστοιχα στην καλλιέργεια των καπνών. Τα σφάλματα αυτά του ταξινομητή

αποδίδονται στο γεγονός ότι η καλλιέργεια του καπνού εμφανίζει κάποια κοινά χαρακτηριστικά με την καλλιέργεια σιτηρών και βαμβακιού. Ο καπνός σπέρνεται την άνοιξη, από τον Απρίλη και μετά, όπως και το βαμβάκι, ενώ η συγκομιδή του γίνεται μέσα στο καλοκαίρι, όπως και η συγκομιδή των σιτηρών. Οι ανάγκες του σε υγρασία, ποικίλλουν, στις αρχές χρειάζεται περισσότερο νερό ενώ όταν αναπτύσσεται οι ανάγκες σε νερό είναι μικρότερες. Αποτέλεσμα είναι, να ανακλά ως ένα βαθμό παρόμοια με τις δύο αυτές κατηγορίες, και ο ταξινομητής να οδηγείται σε δυσκολία ταξινόμησης του συνόλου του ground truth στην σωστή κλάση.

Η κατηγορία RCF, ταξινομείται με μεγάλη ευκολία από τον ταξινομητή, μιας και τα ποσοστά ακρίβειας και αξιοπιστίας της κλάσης είναι ανώτερα από το 99 και το 98% αντίστοιχα. Συγχύσεις με τις υπόλοιπες κλάσεις δεν παρατηρούνται, γεγονός που οφείλεται στην διαφοροποίηση της συγκεκριμένης καλλιέργειας χάρη στα υψηλά ποσοστά υγρασίας, που την χαρακτηρίζουν και την κάνουν να ανακλά διαφοροποιημένα σε σχέση με τις άλλες κατηγορίες των καλλιεργειών.

Τέλος, η κατηγορία SLF, ταξινομείται και εκείνη εύκολα από τον ταξινομητή, καθώς χαρακτηρίζεται από πολύ υψηλή ακρίβεια παραγωγού και χρήστη, ενώ δεν εντοπίζονται αξιοσημείωτες συγχύσεις, με κάποια άλλη κατηγορία.

Συμπερασματικά, λοιπόν, ως προς τις κατηγορίες των καλλιεργειών, διαπιστώνεται ότι διαχωρίζονται αρκετά καλά από τον ταξινομητή, καθώς η ακρίβεια και η αξιοπιστία για κάθε κλάση ξεπερνά το 70%, με την πλειοψηφία των κατηγοριών (σ.σ. οι 11 από τις 13 κατηγορίες) να ξεπερνά το 95% .

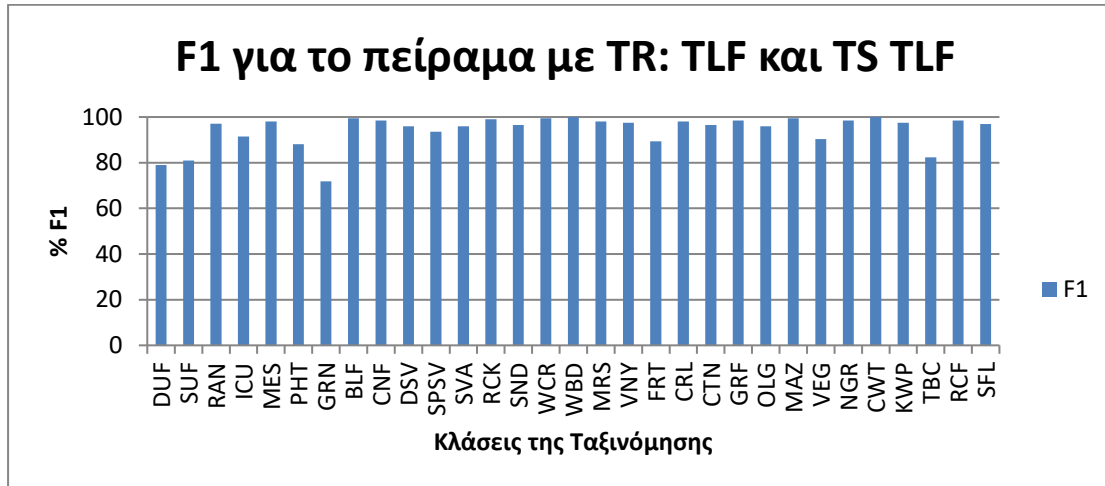
Πίνακας 5-3 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση στην εικόνα 35TLF

	DUF	SUF	RAN	ICU	MES	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MR	VNY	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL
DUF	78,3	11,4	0,73	4,65	0,21	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,06	0,24	0,00	0,00	0,00	0,23	0,00	0,02	0,07	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SUF	12,4	81,8	0,26	0,27	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,00	0,00	0,06	0,00	0,00	0,03	0,30	0,25	0,08	0,08	0,05	0,11	0,00	0,14	0,00	0,00	0,03	0,00	0,00	0,00
RAN	0,98	0,33	97,0	0,86	0,02	0,88	0,00	0,00	0,00	0,00	0,05	0,13	0,02	0,30	0,00	0,00	0,03	0,00	0,08	0,02	0,02	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
ICU	5,96	0,49	0,80	90,6	0,51	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,18	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MES	0,31	0,05	0,15	2,57	98,3	0,00	0,00	0,00	0,00	0,00	0,01	0,25	0,43	0,95	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
PHT	0,00	0,05	0,22	0,00	0,00	98,2	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,03	0,00	0,00	0,02	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,09	0,00	0,00	0,00
GR	0,31	0,11	0,04	0,05	0,02	0,00	100,	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,00	0,01	0,00	0,00	0,14	0,00	0,00	0,03	0,00	0,00	0,00
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,3	0,02	0,18	0,09	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,50	0,00	0,00	0,00	0,00	0,00
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,16	98,4	1,22	0,41	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,50	0,00	0,00	0,01	0,19	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,27	1,15	96,6	2,08	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SPS	0,00	0,11	0,11	0,00	0,00	0,00	0,00	0,00	0,08	1,58	93,7	0,38	0,04	0,00	0,00	0,00	0,03	0,03	0,08	0,00	0,00	0,00	2,13	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SVA	0,05	0,38	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,16	97,2	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
RCK	0,46	0,38	0,07	0,45	0,23	0,00	0,00	0,00	0,00	0,00	0,02	0,76	99,1	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SND	0,15	0,16	0,07	0,54	0,72	0,88	0,00	0,00	0,00	0,00	0,04	0,00	0,15	97,9	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,06	99,7	0,00	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WB	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,08	99,9	1,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MR	0,10	0,38	0,18	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,02	0,06	0,00	0,01	97,9	0,00	0,25	0,01	0,01	0,01	0,07	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00
VNY	0,00	0,55	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,03	98,5	0,25	0,05	0,11	0,01	0,16	0,03	0,00	0,00	0,00	0,03	0,00	0,00	0,00	
FRT	0,00	0,27	0,00	0,00	0,00	0,00	0,00	0,03	0,04	0,02	0,24	0,00	0,00	0,00	0,00	0,15	0,00	93,2	0,03	0,03	0,17	0,33	0,01	0,00	0,00	0,00	0,23	0,00	0,00	0,04	
CRL	0,21	0,55	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,12	0,25	0,00	0,00	0,00	0,03	0,17	0,67	97,3	1,39	0,09	0,10	0,29	0,00	0,00	0,00	0,00	0,17	0,00	0,23	
CTN	0,41	1,04	0,07	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,03	0,00	0,00	0,00	0,00	0,03	0,30	0,17	1,88	97,1	0,12	0,13	0,01	0,27	0,00	0,00	0,00	1,03	0,00	0,43	
GRF	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,07	0,38	0,00	0,00	0,00	0,00	0,07	0,67	0,07	0,03	97,9	0,03	0,04	0,27	0,00	0,00	0,72	0,00	0,37	0,58	
OLG	0,15	0,44	0,18	0,00	0,00	0,00	0,00	0,00	0,25	0,37	2,52	0,00	0,00	0,00	0,00	0,24	0,10	2,00	0,12	0,08	0,03	96,3	0,00	0,00	0,00	0,00	0,09	0,00	0,00	0,00	
MA	0,00	0,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,03	0,00	0,75	0,02	0,04	0,32	0,01	99,3	0,14	0,00	0,00	0,11	0,17	0,33	0,23	
VEG	0,00	0,71	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,00	0,00	0,10	0,17	0,10	0,22	0,09	0,01	0,05	98,9	0,00	0,00	0,00	0,52	0,00	0,16	
NG	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,4	0,00	0,00	0,00	0,00	0,00	
CW	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,15	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,	0,00	0,00	0,00	0,00	
KW	0,05	0,16	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	0,00	0,00	0,00	0,00	0,06	0,03	0,67	0,00	0,03	0,56	0,03	0,11	0,00	0,00	0,00	98,5	0,00	0,00	0,04	
TBC	0,05	0,22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,22	0,61	0,00	0,01	0,00	0,00	0,00	0,00	0,00	98,1	0,00	0,00	
RCF	0,00	0,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,00	0,00	0,10	0,08	0,00	0,01	0,16	0,01	0,03	0,14	0,00	0,00	0,09	0,00	99,2	0,00	
SFL	0,00	0,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,38	0,00	0,00	0,00	0,00	0,00	0,08	0,02	0,07	0,43	0,00	0,11	0,00	0,00	0,06	0,00	0,05	98,2		

Πίνακας 5-4 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση στην εικόνα 35TLF

	DUF	SUF	RAN	ICU	MES	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MR	VNY	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL
DUF	79,5	13,0	0,69	5,30	0,10	0,00	16,6	0,00	0,00	0,00	0,00	0,12	0,19	0,17	0,00	0,00	0,06	0,00	0,00	0,01	0,04	0,00	0,02	0,00	0,00	0,00	0,00	0,03	0,12	0,00	0,00
SUF	10,9	80,4	0,22	0,41	0,02	0,72	5,56	0,00	0,00	0,00	0,02	0,87	0,15	0,17	0,00	0,00	0,21	0,32	0,38	0,02	0,09	0,01	0,04	0,01	1,49	0,00	0,00	0,08	0,50	0,09	0,08
RAN	1,04	0,38	97,0	1,01	0,07	4,35	2,78	0,00	0,00	0,00	0,03	0,12	0,04	0,12	0,00	0,00	0,15	0,06	0,00	0,00	0,01	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
ICU	5,38	0,32	0,69	91,7	1,00	0,00	2,78	0,00	0,00	0,00	0,00	0,00	0,21	0,69	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MES	0,63	0,00	0,04	1,33	98,0	0,00	2,78	0,00	0,00	0,00	0,00	0,00	0,28	2,37	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
PHT	0,00	0,00	0,04	0,00	0,00	80,4	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
GR	0,00	0,00	0,00	0,00	0,00	0,00	55,5	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,7	0,06	0,30	0,01	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,61	0,00	0,01	0,02	0,01	0,00	0,00	1,95	0,00	0,00	0,00	0,00	0,00
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	99,3	3,31	0,42	0,00	0,00	0,00	0,07	0,00	0,03	0,00	1,92	0,00	0,00	0,00	0,87	0,00	0,00	0,49	0,00	0,00	0,00	0,00	0,00
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,16	0,42	95,1	2,86	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,31	0,00	0,01	0,01	0,44	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SPS	0,16	0,70	0,22	0,00	0,02	0,00	0,00	0,04	0,08	1,13	93,3	2,37	0,04	0,29	0,07	0,00	0,18	0,23	2,15	0,03	0,01	0,07	1,65	0,01	0,00	0,00	0,00	0,22	0,00	0,00	0,04
SVA	0,00	0,00	0,04	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,03	95,3	0,13	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,11	0,00	0,00	0,00	0,00	0,05	0,11
RCK	0,16	0,00	0,04	0,00	0,35	0,72	0,00	0,00	0,00	0,00	0,02	0,87	98,8	0,40	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SND	0,21	0,05	0,18	0,14	0,28	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	95,0	0,07	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,4	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WB	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	99,9	0,24	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MR	0,00	0,05	0,04	0,00	0,00	0,72	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,22	0,03	98,2	0,03	0,38	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00
VNY	0,37	0,48	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	96,4	0,00	0,01	0,04	0,02	0,02	0,00	0,34	0,00	0,00	0,03	0,00	0,14	0,00
FRT	0,00	0,16	0,04	0,00	0,00	0,00	2,78	0,00	0,01	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,09	0,10	85,8	0,02	0,01	0,07	0,13	0,03	0,23	0,00	0,00	0,22	0,00	0,05	0,04
CRL	0,47	2,04	0,26	0,00	0,03	7,25	2,78	0,00	0,00	0,00	0,00	0,00	0,04	0,23	0,00	0,00	0,18	0,78	1,15	98,9	3,93	0,31	0,31	0,03	5,15	0,00	0,00	0,03	12,2	0,05	0,38
CTN	0,84	0,91	0,15	0,09	0,00	2,90	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,09	0,78	0,54	0,67	95,5	0,07	0,10	0,03	5,49	0,00	0,00	0,20	16,2	0,09	0,57
GRF	0,00	0,27	0,00	0,00	0,00	0,00	2,78	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,03	1,38	0,02	0,06	98,7	0,02	0,11	1,14	0,00	0,00	1,68	0,00	0,78	1,74
OLG	0,31	1,07	0,33	0,00	0,02	0,00	0,00	0,00	0,05	0,16	3,23	0,12	0,02	0,29	0,00	0,00	0,39	0,94	4,45	0,04	0,11	0,06	96,3	0,01	0,23	0,00	0,17	0,25	0,05	0,00	
MA	0,00	0,00	0,00	0,00	0,00	0,72	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,29	0,23	0,20	0,01	0,12	0,00	99,7	1,72	0,00	0,92	0,00	0,41	1,25	
VEG	0,00	0,05	0,00	0,00	0,00	0,00	2,78	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,00	0,00	83,3	0,00	0,00	0,00	0,00	0,05	0,00	
NG	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	97,5	0,00	0,00	0,00	0,00	0,00	
CW	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,9	0,00	0,00	0,00	0,00	
KW	0,00	0,05	0,00	0,00	0,00	2,17	2,78	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,03	0,61	0,00	0,00	0,23	0,02	0,01	0,00	0,00	96,3	0,00	0,14	0,08	
TBC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,34	0,00	0,00	0,00	70,6	0,00	0,00	
RCF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	0,00	0,02	0,00	0,00	0,00	0,00	98,1	0,04	
SFL	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,01	0,05	0,14	0,00	0,02	0,46	0,00	0,03	0,00	0,00	95,7	

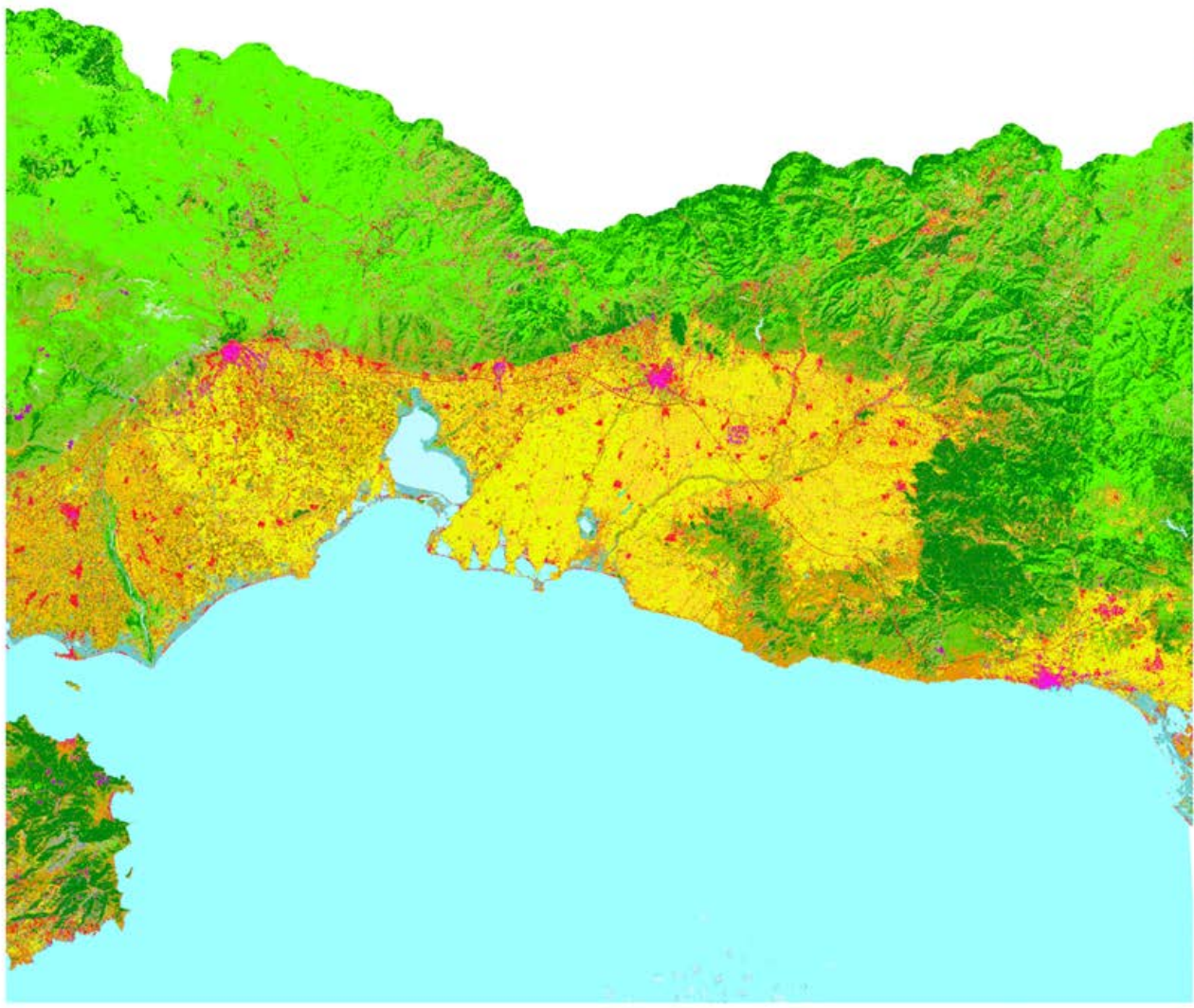
Στην συνέχεια παρατίθεται το γράφημα για τον δείκτη Fisher (F1) της ταξινόμησης, που αναλύθηκε παραπάνω. Στο συγκεκριμένο γράφημα εμφανίζεται το ποσοστό του F1 για κάθε κλάση της ταξινόμησης. Με το γράφημα αυτό, μπορούμε και εποπτικά να εντοπίσουμε τις κατηγορίες εκείνες, που απέδωσαν καλύτερα, αλλά και τις κατηγορίες αυτές, που εμφάνισαν συγχύσεις.



Γράφημα 5-3 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMF

Όπως διαπιστώνουμε, ο ταξινομητής έχει διαχωρίσει όλες τις κατηγορίες επιτυχώς, καθώς όλες οι κατηγορίες διακρίνονται από δείκτη F1 ανώτερο του 72%. Οι κατηγορίες όμως, οι οποίες διαχωρίστηκαν βέλτιστα, σύμφωνα με το δείκτη F1 είναι οι κατηγορίες των υδάτινων επιφανιών και συγκεκριμένα η WBD και η CWT, ενώ οι κατηγορίες με τα χαμηλότερα ποσοστά είναι η GRN, η DUF και η SUF. Η GRN είναι μία κατηγορία για την οποία τα δεδομένα εκπαίδευσης είναι περιορισμένα, γεγονός που οδηγεί σε δυσκολία εκπαίδευσης του ταξινομητή ως προς την αναγνώρισή της. Ακόμη, οι κατηγορίες DUF και SUF, είναι υποκατηγορίες του αστικού ιστού, ανακλούν παρόμοια στο Ηλεκτρομαγνητικό φάσμα, οδηγώντας έτσι σε μεταξύ τους συγχύσεις.

Τέλος, παρατίθεται ο χάρτης, που παρήχθη από το πείραμα αυτό.



Legend

- Dense Urban Fabric
- Sparse Urban Fabric
- Road and Rail Networks
- Industrial or Commercial Units
- Mineral Extraction Sites
- Photovoltaic Panels
- Greenhouses
- Broadleaved Forest
- Coniferous Forest
- Dense Sclerophyllous Vegetation
- Sparse Sclerophyllous Vegetation
- Sparsely Vegated Areas
- Bare Rocks
- Beaches, Dunes, Sands
- Water Courses
- Water Bodies
- Marshes
- Vineyards
- Fruit Trees and Berry Plantations
- Cereals
- Cotton
- Grass Fodders
- Olive Groves
- Maize
- Vegetables
- Natural Grassland
- Sea Coastal Water
- Kiwi Plants
- Tobacco
- Rice Fields
- Sunflowers

όγηση

Χάρτης 5.1 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG

Μελετώντας τον χάρτη, διαπιστώνεται ότι οι γενικές κατηγορίες της ταξινόμησης έχουν αποδοθεί σωστά και ότι παρέχει μία αξιόπιστη πρόβλεψη των χρήσεων/ καλύψεων γης και καλλιεργειών της περιοχής μελέτης. Παρατηρούμε, λοιπόν, ότι οι αστικές, οι δασικές περιοχές, οι υδάτινες επιφάνειες, καθώς και οι καλλιέργειες διαχωρίζονται σωστά από τον ταξινομητή.

Εάν παρατηρήσουμε τον χάρτη πιο λεπτομερώς και σε μεγαλύτερο zoom, θα διαπιστώσουμε ότι ορισμένες κλάσεις εμφανίζουν συγχύσεις με άλλες, επιβεβαιώνοντας τα αποτελέσματα του πίνακα σύγχυσης.

Πιο συγκεκριμένα, οι κατηγορίες DUF και SUF, εμφανίζουν μικρές συγχύσεις, καθώς εικονοστοιχεία τα οποία ανήκουν στην DUF ταξινομούνται στην SUF και το αντίστροφο. Ακόμη, εικονοστοιχεία, που ανήκουν στην DUF, αποδίδονται στην κατηγορία ICU, ενώ pixels της ICU αναγνωρίζονται ως pixel της DUF και της MES.



Εικόνα 5-1 Σύγχυση ανάμεσα στην DUF, την SUF και την ICU

Ακόμη, η κατηγορία PHT εμφανίζει μικρές συγχύσεις με πολλές κατηγορίες τεχνητών επιφανειών, καλλιεργειών και υδάτινων επιφανειών.



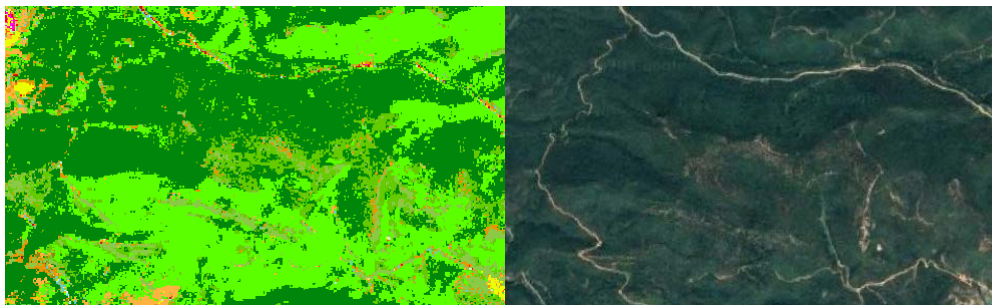
Εικόνα 5-2 Σύγχυση της PHT με πληθώρα κατηγοριών

Τέλος, η κατηγορία GRN, εμφανίζει συγχύσεις με τις υπόλοιπες κατηγορίες των τεχνητών επιφανειών.



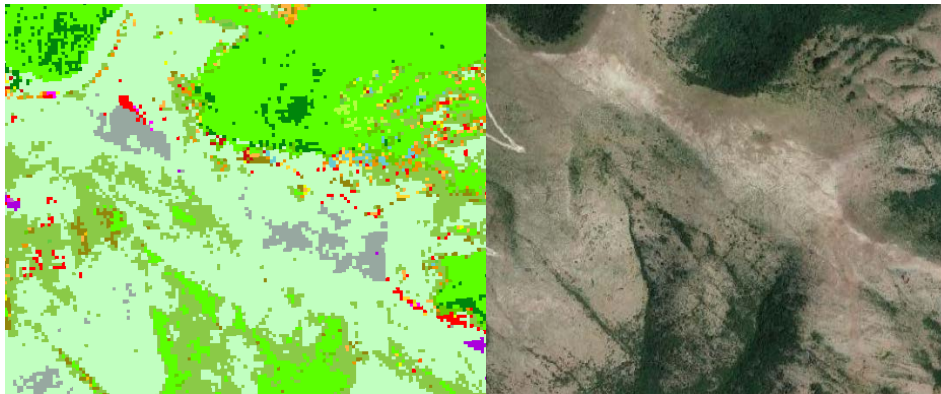
Εικόνα 5-3 Σύγκριση ανάμεσα στην GRN και στις κατηγορίες των τεχνητών επιφανειών

Οι κατηγορίες της φυσικής βλάστησης φαίνεται ότι αποδίδονται επιτυχώς στον παραγόμενο χάρτη καλύψεων γης. Τα εικονοστοιχεία του χάρτη απεικονίζουν τις σωστές κλάσεις, με αμελητέα σφάλματα.



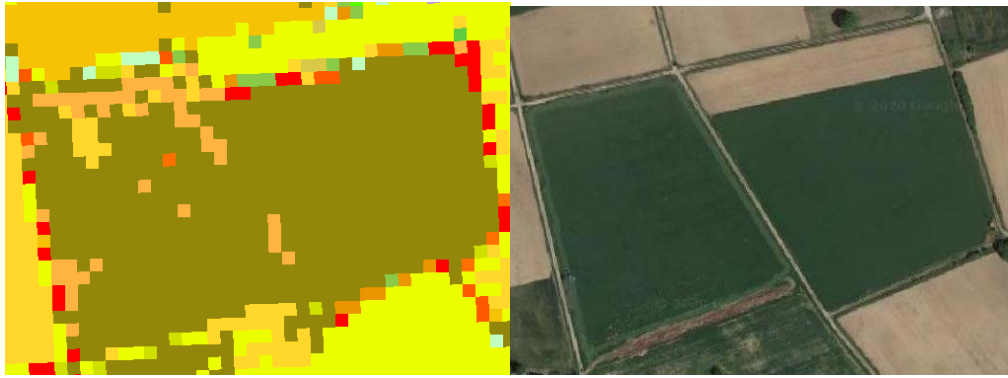
Εικόνα 5-4 Επιτυχής απόδοση των κατηγοριών φυσικής βλάστησης

Αντίστοιχα, οι κατηγορίες του γυμνού εδάφους προβλέπονται από τον ταξινομητή επιτυχώς.



Εικόνα 5-5 Επιτυχής πρόβλεψη των κατηγοριών SVA και RCK

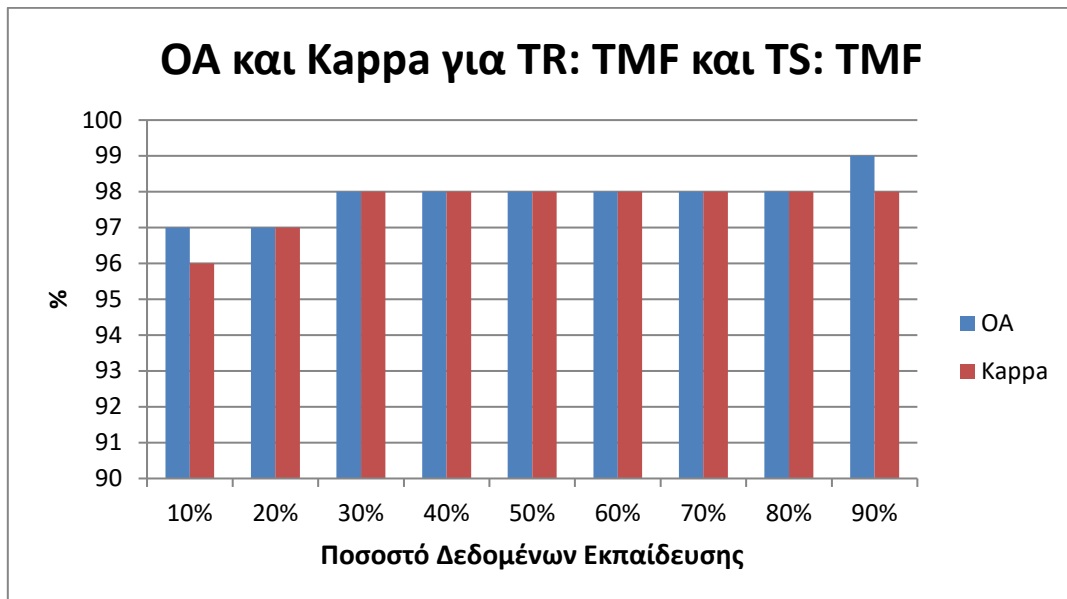
Τέλος, σχετικά με τις κατηγορίες καλλιεργειών, όλες οι κλάσεις αποδίδονται σωστά από τον ταξινομητή. Αμελητέες συγχύσεις εμφανίζονται στην κατηγορία TBC μα τις CTN, CRL.



Εικόνα 5-6 Αδυναμία απόδοσης της VNY. Συμπερίληψη στην BLF και στην CNF

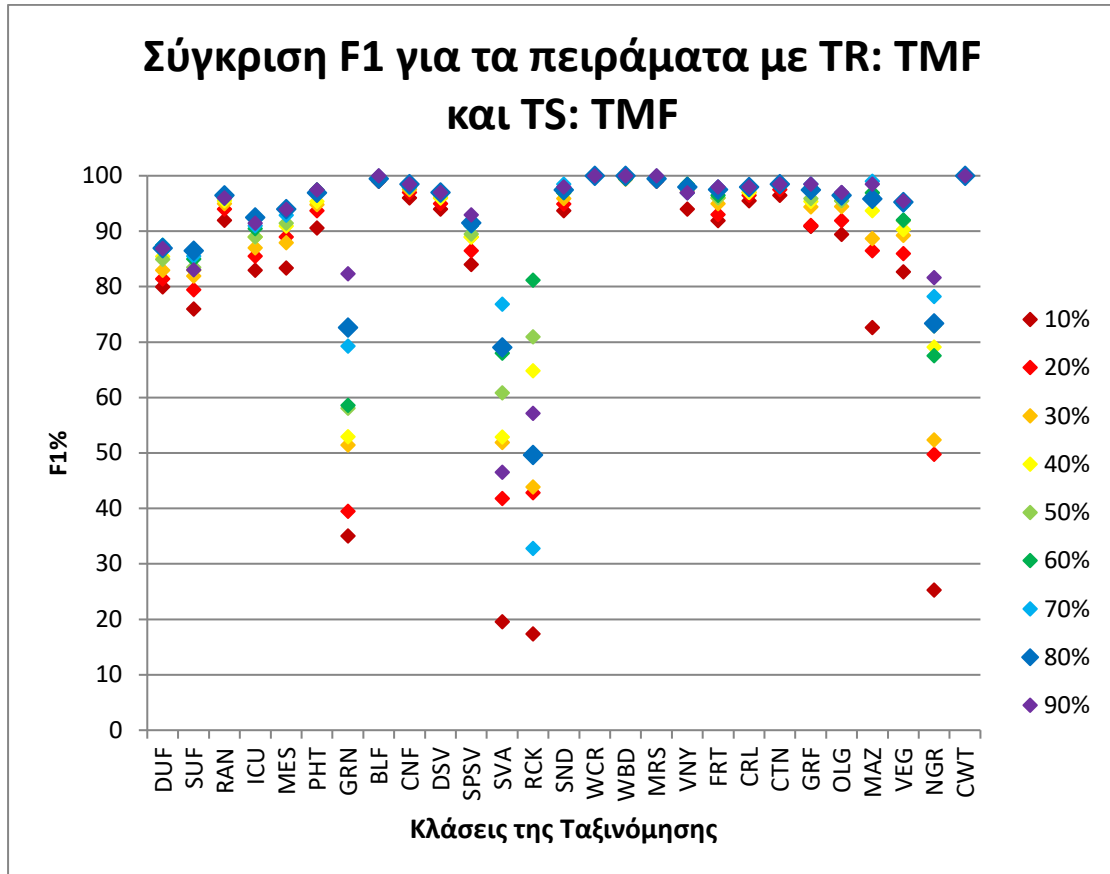
5.1.2 Δεύτερο πείραμα: ταξινόμηση στο tile 35TMF

Στο δεύτερο πείραμα που εφαρμόστηκε χρησιμοποιήθηκε το Tile 35TMF για να εκπαιδευτεί ο αλγόριθμος, ενώ ο έλεγχος της ταξινόμησης επιλέχθηκε να γίνει στο ίδιο tile. Χρησιμοποιήθηκαν και οι έξι επιλεγμένες εικόνες και από αυτές τα κανάλια 2,3,4,8,11,12, καθώς και οι δείκτες NDVI, MSAVI, NDWI, NDBI. Οι κλάσεις της ταξινόμησης ήταν 27 και όχι 31, όπως στα προηγούμενα πειράματα καθώς στο συγκεκριμένο tile, δεν απαντώνται δεδομένα εκπαίδευσης για τις κλάσεις KWP: ακτινίδια, TBC: καπνός, RCF: ορυζώνες και για την SFL: ηλιοτρόπια. Έτρεξαν 9 ταξινομήσεις, για εκπαίδευση του ταξινομητή με 10, 20, 30, 40, 50, 60, 70, 80 και 90% των δεδομένων αναφοράς. Τα αποτελέσματα που προέκυψαν από τις ταξινομήσεις δεν διέφεραν σημαντικά. Η συνολική ακρίβεια της ταξινόμησης και ο δείκτης Kappa, για τα εννέα πειράματα, παρουσιάζονται στο παρακάτω γράφημα.



Γράφημα 5-4 Συνολική ακρίβεια και δείκτης Kappa για την ταξινόμηση στην εικόνα 35TMF

Μελετώντας το γράφημα, διαπιστώνεται ότι και οι 9 ταξινομήσεις έχουν αποδώσει εξίσου καλά αποτελέσματα. Σημειώνεται ότι όσο περισσότερο ποσοστό των δεδομένων εκπαίδευσης χρησιμοποιούνται στην ταξινόμηση, τόσο υψηλότερης ακρίβειας αποτελέσματα αποδίδει ο αλγόριθμος, ενώ οι ταξινομήσεις στις οποίες χρησιμοποιήθηκε ποσοστό δεδομένων εκπαίδευσης από 30% και πάνω, αποδίδουν παρόμοια αποτελέσματα. Ελέγχοντας τα αποτελέσματα ακρίβειας και αξιοπιστίας της κάθε κλάσης της ταξινόμησης καθώς και το δείκτη F1, διαπιστώνεται ότι σχεδόν όλες οι κλάσεις αποδίδουν καλύτερα αποτελέσματα όσο αυξάνεται το ποσοστό των δεδομένων εκπαίδευσης που χρησιμοποιείται για εκπαίδευση. Όμως, τα αποτελέσματα της ακρίβειας και της αξιοπιστίας των κλάσεων δεν διαφέρουν σημαντικά στις ταξινομήσεις, που χρησιμοποιείται ποσοστό δεδομένων εκπαίδευσης από 50% και πάνω.



Γράφημα 5-5 Δείκτης F κάθε κλάσης για την ταξινόμηση στην εικόνα 35TMF

Επιλέγεται να παρατεθεί και αξιολογηθεί η ταξινόμηση με ποσοστό δεδομένων εκπαίδευσης ίσο με 50%, καθώς σημειώνει υψηλά ποσοστά στα μεγέθη αξιολόγησης της ταξινόμησης και των κλάσεων αυτής, τα οποία δεν μεταβάλλονται σημαντικά στις ταξινομήσεις με ποσοστό δεδομένων εκπαίδευσης πάνω από 50%.

Παρακάτω παρατίθεται ο πίνακας σύγχυσης της ταξινόμησης. Εν γένει η ταξινόμηση θεωρείται πετυχημένη, καθώς, η γενική ακρίβεια και ο δείκτης kappa ξεπερνούν το 98%.

Πίνακας 5-5 Πίνακας Σύγκρισης για ταξινόμηση στην εικόνα 35TMF

		ΚΛΑΣΕΙΣ ΤΟΥ ΑΠΟΤΕΛΕΣΜΑΤΟΣ ΤΗΣ ΤΑΞΙΝΟΜΗΣΗΣ																									sum	PA		
		DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPSV	SVA	RCK	SND	WCR	WBD	MRS	VNY	FRT	CRL	CTN	GRF	OLG	MAZ	VEG			NGR	CWT
ΚΛΑΣΕΙΣ ΤΟΥ ΧΑΡΤΗ ΤΟΥ GROUND TRUTH	DUF	1270	142	18	78	5	0	0	0	0	0	0	0	0	1	0	0	3	0	0	3	9	0	3	0	0	0	0	1532	0,83
	SUF	110	1509	7	20	2	0	0	0	0	0	8	0	0	0	0	0	11	0	2	12	115	0	5	0	0	0	0	1801	0,84
	RAN	14	16	3445	71	3	3	0	0	1	0	0	0	0	0	0	0	24	0	1	1	8	0	2	0	0	0	0	3589	0,96
	ICU	56	68	59	2133	38	2	2	0	0	0	0	0	0	4	0	0	1	0	0	1	1	0	0	0	0	0	2365	0,9	
	MES	1	4	11	61	706	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	785	0,9	
	PHT	1	3	8	4	1	642	0	0	0	0	0	0	0	0	0	0	0	1	0	0	4	11	0	0	0	0	675	0,95	
	GRN	7	1	16	31	2	0	47	0	0	0	0	0	0	0	0	0	3	0	0	1	5	0	0	0	0	0	113	0,42	
	BLF	0	0	0	0	0	0	0	43724	97	11	7	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	43841	1	
	CNF	0	0	0	0	0	0	0	239	20813	224	12	0	0	0	0	2	0	1	0	0	0	0	1	0	0	0	21292	0,98	
	DSV	0	0	0	0	0	0	0	9	246	13554	150	0	0	0	0	1	0	5	0	0	0	7	0	0	0	0	13972	0,97	
	SPSV	0	1	0	0	0	0	0	20	8	297	2855	0	0	0	0	8	0	1	0	2	0	35	0	0	0	0	3227	0,88	
	SVA	0	4	0	1	0	0	0	0	0	0	6	17	0	0	0	5	0	2	0	1	0	2	0	0	0	0	38	0,45	
	RCK	0	5	0	1	2	0	0	0	0	0	0	1	12	0	0	1	0	0	0	0	0	0	0	0	0	0	22	0,55	
	SND	3	1	7	20	3	0	0	0	0	0	0	0	0	751	0	0	0	0	0	0	0	0	0	0	0	0	785	0,96	
	WCR	0	0	1	0	0	0	0	0	1	0	0	0	0	0	10940	0	3	0	0	0	0	0	0	0	0	0	10945	1	
	WBD	0	0	1	0	0	0	0	0	0	0	0	0	0	0	15	3696	5	0	0	0	0	0	0	0	0	0	3717	0,99	
	MRS	0	7	9	1	0	0	0	1	0	0	0	0	0	1	6	0	23780	0	4	1	25	0	0	0	0	0	23835	1	
	VNY	0	2	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	472	3	0	0	0	0	0	0	0	480	0,98	
	FRT	0	3	1	0	0	0	25	6	12	6	0	0	0	0	0	49	2	3210	23	14	0	12	0	0	5	0	3368	0,95	
	CRL	0	6	0	0	0	0	0	0	0	2	0	0	0	0	15	0	12	4311	30	0	5	0	1	0	0	4382	0,98		
	CTN	1	33	2	0	1	0	0	0	0	2	0	0	0	0	0	38	3	17	48	12619	3	1	0	1	0	12769	0,99		
	GRF	0	2	0	0	0	1	4	0	0	1	0	0	0	0	0	3	2	0	16	2	438	1	0	0	0	470	0,93		
	OLG	0	16	1	0	0	0	0	0	0	24	73	0	0	0	0	2	3	22	0	4	0	2193	0	0	1	0	2339	0,94	
	MAZ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	8	0	0	140	0	150	0,93		
	VEG	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	46	38	0	0	0	506	0	594	0,85	
	NGR	0	0	0	0	0	0	22	0	2	7	0	0	0	0	0	0	0	0	15	0	0	1	3	0	0	76	126	0,6	
CWT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33442	33442	1		
sum	1463	1824	3586	2421	763	648	50	44044	21172	14124	3132	18	12	757	10961	3696	23955	482	3300	4467	12894	442	2271	140	508	82	33442	kappa: 0,98		
UA	0,87	0,83	0,96	0,88	0,93	0,99	0,94	0,99	0,98	0,96	0,91	0,94	1	0,99	1	1	0,99	0,98	0,97	0,97	0,98	0,99	0,97	1	1	0,93	1	OA: 0,98		

Μελετώντας επιμέρους την ακρίβεια και την αξιοπιστία ανά κατηγορία, διαπιστώνεται ότι όλες οι κατηγορίες έχουν ταξινομηθεί επιτυχώς μιας και τα ποσοστά τους ξεπερνούν το 83%. Εξαιρέση αποτελούν 4 κατηγορίες – οι GRN, SVA, RCK και NGR-όπου οι ακρίβειες παραγωγού χαρακτηρίζονται χαμηλές έως μέτριες, καθώς κυμαίνονται από 40 έως 60%.

Πιο συγκεκριμένα, εστιάζοντας στις κλάσεις των τεχνητών επιφανειών, διαπιστώνεται ότι οι κλάσεις DUF και SUF, εμφανίζουν μία μικρή σύγχυση μεταξύ τους, καθώς απαντώνται σφάλματα συμπερίληψης, αλλά και παράλειψης, που κυμαίνονται από 6 έως 9%. Επιπλέον η κατηγορία DUF, εμφανίζει μικρότερη σύγχυση και με την κατηγορία ICU, με σφάλματα παράλειψης και συμπερίληψης 5% και 4% αντίστοιχα. Οι συγχύσεις ανάμεσα στις κατηγορίες DUF και SUF είναι αναμενόμενες, καθώς πρόκειται για υποκατηγορίες της γενικής κατηγορίας: αστικός ιστός, παρουσιάζουν κοινά φασματικά χαρακτηριστικά μιας και είναι δύσκολο να διαχωριστούν πλήρως από τον αλγόριθμο.

Επιπλέον η σύγχυση της DUF με την ICU είναι επίσης ως ένα βαθμό αναμενόμενη, καθώς, τα κήρια που απαρτίζουν την κατηγορία ICU είναι δύσκολο να διαχωριστούν πλήρως από αυτά που απαρτίζουν την DUF.

Η κατηγορία RAN ταξινομείται με ευκολία από τον αλγόριθμο, καθώς χαρακτηρίζεται από πολύ υψηλά ποσοστά αξιοπιστίας και ακρίβειας – 96%- και συνεπώς δεν εμφανίζονται αξιοσημείωτες συγχύσεις με άλλες κλάσεις.

Ακόμη, η ICU εμφανίζει μικρές συγχύσεις (σφάλματα παράλειψης και συμπερίληψης από 1,6 έως 2,9%) σχεδόν με όλες τις κατηγορίες των τεχνητών επιφανειών και συγκεκριμένα με τις κλάσεις DUF, SUF, RAN, MES.

Η κατηγορία MES, χαρακτηρίζεται από πολύ υψηλή ακρίβεια (90%) και αξιοπιστία (93%) και η σημαντικότερη σύγχυση, που εμφανίζεται είναι με την κατηγορία ICU, με σφάλμα παράλειψης 7,77% και συμπερίληψης 4,98%.

Η κατηγορία PHT, χαρακτηρίζεται από πολύ υψηλή ακρίβεια παραγωγής (95%) και από πολύ υψηλή ακρίβεια χρήστη (99%). Δεν παρατηρούνται αξιοσημείωτα σφάλματα παράλειψης και συμπερίληψης. Αξίζει να σημειωθεί πως μικρή σύγχυση παρατηρείται με την κλάση CTN αφού 1,6% των εικονοστοιχείων της κλάσης PHT, παραλείπονται από την κατηγορία και αποδίδονται στην κλάση CTN. Αντιστοίχως, 1,2% των εικονοστοιχείων της κλάσης παραλείπονται και αποδίδονται στην κατηγορία RAN.

Τέλος η κλάση GRN, εμφανίζει χαμηλή ακρίβεια και υψηλή αξιοπιστία, καθώς 27% των pixel της ταξινομούνται στην ICU, 14% στην RAN, περίπου 6% στην DUF και περί το 4% στην CTN.

Οι συγχύσεις που παρουσιάζονται στην κατηγορία στην GRN είναι αποτέλεσμα του μικρού αριθμού δεδομένων εκπαίδευσης που παρέχονται στον ταξινομητή, καθώς δεν υπάρχει επαρκές δείγμα ώστε ο αλγόριθμος να εκπαιδευτεί κατάλληλα και να κατόπιν να διαχωρίσει αυτήν την κατηγορία.

Οι κατηγορίες της φυσικής βλάστησης ταξινομήθηκαν στο σύνολό τους επιτυχημένα από τον αλγόριθμο ταξινόμησης. Μικρές συγχύσεις παρουσιάζονται ανάμεσα στις

κλάσεις DSV και SPSV. Πιο συγκεκριμένα, η κλάση DSV, παρουσιάζει μικρά σφάλματα παράλειψης και συμπερίληψης με την κατηγορία SPSV. Οι κλάσεις αυτές, αποτελούν υποκατηγορίες της γενικότερης κλάσης σκληροφυλλική βλάστηση και η διαφοροποίησή τους έγκειται στην πυκνότητα της βλάστησης, γεγονός που καθιστά δύσκολο το διαχωρισμό τους από τον ταξινομητή Τέλος, μικρή σύγχυση, που διαφαίνεται από το σφάλμα συμπερίληψης 2%, εμφανίζεται στην κλάση OLG, με την κλάση SVA. Η μικρή αυτή σύγχυση, θεωρείται απόρροια της παρόμοιας ανάκλασης των δύο κατηγοριών στην περιοχή του πράσινου, του κόκκινου και του εγγύς υπέρυθρου.

Η κατηγορία NGR, χαρακτηρίζεται από υψηλή αξιοπιστία (93%), όχι όμως και από υψηλή ακρίβεια (60%), καθώς παρατηρούνται σφάλματα παράλειψης με ορισμένες κλάσεις. Η σημαντικότερη σύγχυση εμφανίζεται με την κλάση BLF (σφάλμα παράλειψης 17%). Μικρότερη σύγχυση εμφανίζεται με την κλάση FRT (σφάλμα παράλειψης 11%) και ακόμα μικρότερη με την SPSV (σφάλμα παράλειψης 5,55%).

Οι κατηγορίες του γυμνού εδάφους διαχωρίζονται εμφανίζουν διακυμάνσεις σε σχέση με την ακρίβεια και την αξιοπιστία, που τις χαρακτηρίζουν. Πιο συγκεκριμένα, η κατηγορία επιτυχώς, από τον ταξινομητή καθώς όλες οι κατηγορίες πετυχαίνουν ακρίβειες χρήστη και παραγωγού μεγαλύτερες του 95%. Σημαντικές συγχύσεις συνεπώς δεν παρατηρούνται. Αξιοσημείωτο είναι όμως το γεγονός, ότι στην κλάση SVA, παρατηρείται μικρή σύγχυση (σφάλμα συμπερίληψης περί τα 2%) με την κλάση SPSV. Η συμπερίληψη των εικονοστοιχείων της SPSV στην SVA είναι δικαιολογημένη, επειδή και οι δύο κατηγορίες απαρτίζονται από εικονοστοιχεία που περιέχουν γυμνό έδαφος και βλάστηση. Έτσι, τα φασματικά τους χαρακτηριστικά ως ένα βαθμό συγκλίνουν και δυσκολεύουν τον διαχωρισμό τους. Μία ακόμη μικρή αλλά αξιοσημείωτη σύγχυση είναι εκείνη της κατηγορία SND με την κατηγορία MES. Και η συγκεκριμένη σύγχυση έγκειται στο γεγονός ότι στα εικονοστοιχεία της MES απαντώνται εικονοστοιχεία άμμου.

Σε σχέση με τις κατηγορίες του γυμνού εδάφους, η κατηγορία που ταξινομήθηκε με μεγαλύτερη επιτυχία ήταν η κατηγορία SND, η οποία επέτυχε ποσοστά ακρίβειας και αξιοπιστίας 96% και 99% αντίστοιχα. Οι κατηγορίες SVA και RCK, χαρακτηρίζονται από μέτρια ακρίβεια (45% και 55%) και υψηλή αξιοπιστία (95% και 100%). Η κατηγορία SVA, εμφανίζει μεγάλα σφάλματα παράλειψης με τις κλάσεις SPSV (16%), MRS (13%) και SUF (10%) και μικρότερα με τις FRT (5%) και OLG (5%). Ως προς τα σφάλματα συμπερίληψης, το μοναδικό αξιοσημείωτο είναι εκείνο με την κατηγορία των απογυμνωμένων βράχων 5%.

Η κατηγορία RCK, εμφανίζει σφάλματα παράλειψης με πολυάριθμες κλάσεις. Το μεγαλύτερο σφάλμα εμφανίζεται ως προς την κατηγορία SUF (22%), μικρότερο ως προς την κατηγορία MES (9%) και ακόμη μικρότερο ως προς τις κατηγορίες MRS (5%), SVA (5%) και ICU (5%).

Η αποτυχία του αλγορίθμου να ταξινομήσει τις συγκεκριμένες κλάσεις είναι αποτέλεσμα του μικρού αριθμού δεδομένων εκπαίδευσης, καθώς οι συγκεκριμένες κατηγορίες δεν απαντώνται ευρέως στο συγκεκριμένο tile.

Οι κατηγορίες υδάτινων πόρων διαχωρίστηκαν πλήρως από τον αλγόριθμο ταξινόμησης, αφού επιτεύχθηκε ακρίβεια και αξιοπιστία για κάθε κλάση μεγαλύτερη και από 99%, χωρίς την ύπαρξη αξιόλογων εμπλοκών ανάμεσα στις διάφορες κατηγορίες.

Ως προς τις κλάσεις που ανήκουν στις καλλιέργειες διαπιστώνουμε ότι, όλες οι κλάσεις έχουν ταξινομηθεί επιτυχημένα από τον ταξινομητή, καθώς έχει πετύχει ακρίβειες παραγωγού και χρήστη ανώτερες του 70% σε κάθε περίπτωση.

Πιο συγκεκριμένα, η κατηγορία VNY, διαχωρίζεται άριστα καθώς η ακρίβεια παραγωγού και χρήστη ξεπερνά το 98%, ενώ οι συγχύσεις που εμφανίζονται με τις άλλες κατηγορίες είναι αμελητέες.

Η κλάση FRT, εμφανίζει υψηλή ακρίβεια 95% και αξιοπιστία (97%), ενώ και εκείνη δεν εμφανίζει αξιόλογες συγχύσεις με άλλες κλάσεις.

Η κατηγορία CRL εμφανίζει πολύ υψηλά ποσοστά ακρίβειας και αξιοπιστίας, μεγαλύτερα του 96%. Δεν εμφανίζει σημαντικές συγχύσεις με άλλες κατηγορίες, παρά μόνο σφάλμα συμπερίληψης της τάξης του 1% με την κατηγορία CTN και την κατηγορία VEG. Οι καλλιέργειες αυτές έχουν παρόμοια φασματικά χαρακτηριστικά και παρόμοιους κύκλους συγκομιδής, με αποτέλεσμα να υπάρχει μικρή δυσκολία στον ταξινομητή να διαχωρίσει τις καλλιέργειες αυτές.

Η κατηγορία GRF, παρουσιάζει υψηλά ποσοστά ακρίβειας και αξιοπιστίας (93% και 99% αντιστοίχως). Η συγκεκριμένη κλάση εμπλέκεται σε μικρό βαθμό με την κλάση CRL, καθώς εμφανίζει σφάλμα παράλειψης 3,4%.

Η κατηγορία OLG, έχει διαχωριστεί πλήρως από τον ταξινομητή, αφού εμφανίζει ποσοστά ακρίβειας παραγωγού και χρήστη μεγαλύτερα από 94%. Ελέγχοντας τα σφάλματα παράλειψης δεν διαπιστώνουμε αξιοσημείωτες συγχύσεις. Η μόνη σύγχυση, που αν και μικρή (σφάλμα παράλειψης 3,12% και συμπερίληψης 1,5%), αξίζει να αναφερθεί είναι εκείνη με την κατηγορία SPSV. Οι δύο κατηγορίες, απαρτίζονται από γυμνό έδαφος και δένδρۇλλια με σκούρο και σκληρό φύλλωμα, γεγονός που οδηγεί σε μικρά σφάλματα ταξινόμησης.

Η κατηγορία MAZ είναι μία κατηγορία καλλιεργειών την οποία ο ταξινομητής καταφέρνει να διαχωρίσει αρκετά καλά (ακρίβεια 94%, αξιοπιστία 100%). Η μόνη αξιοσημείωτη σύγχυση είναι με την κατηγορία CTN, καθώς παρατηρείται σφάλμα παράλειψης 5,33%.

Η κατηγορία VEG είναι μία κατηγορία καλλιεργειών, η οποία χαρακτηρίζεται από πολύ υψηλή αξιοπιστία ανώτερη, η οποία φτάνει το 99%, αλλά η ακρίβειά της είναι χαμηλότερη περί το 85%. Τα σφάλματα παράλειψης που παρατηρούνται είναι με την κατηγορία των σιτηρών(7,77%) και του βαμβακιού (6,39%). Τα μικρά αυτά σφάλματα του ταξινομητή οφείλονται στο γεγονός ότι οι κλάσεις αυτές εμφανίζουν κοινά φασματικά χαρακτηριστικά. Η κατηγορία λαχανικά, είναι μία ευρύτερη κατηγορία, που περιλαμβάνει πολλές και διαφορετικές καλλιέργειες, με διαφορετικά ειδικά χαρακτηριστικά, διαφορετικές περιόδους σποράς και συγκομιδής. Οι διαφοροποιήσεις αυτές, δημιουργούν ένα φασματικό μοτίβο, το οποίο μπορεί σε ορισμένες περιπτώσεις να αντιστοιχηθεί με εκείνο των σιτηρών, αλλά και με εκείνο του βαμβακιού. Αποτέλεσμα είναι, ο ταξινομητής να οδηγείται σε κάποιες εσφαλμένες ταξινομήσεις

rixel, λόγω αυτής της ταύτισης των φασματικών χαρακτηριστικών ορισμένων καλλιεργειών, που ανήκουν στην γενικότερη κατηγορία.

Συμπερασματικά, λοιπόν, ως προς τις κατηγορίες των καλλιεργειών, διαπιστώνεται ότι διαχωρίζονται αρκετά καλά από τον ταξινομητή, καθώς η ακρίβεια και η αξιοπιστία για κάθε κλάση ξεπερνά το 85%, με τις μισές (οι 4 από τις 8 κατηγορίες) να ξεπερνούν το 95%.

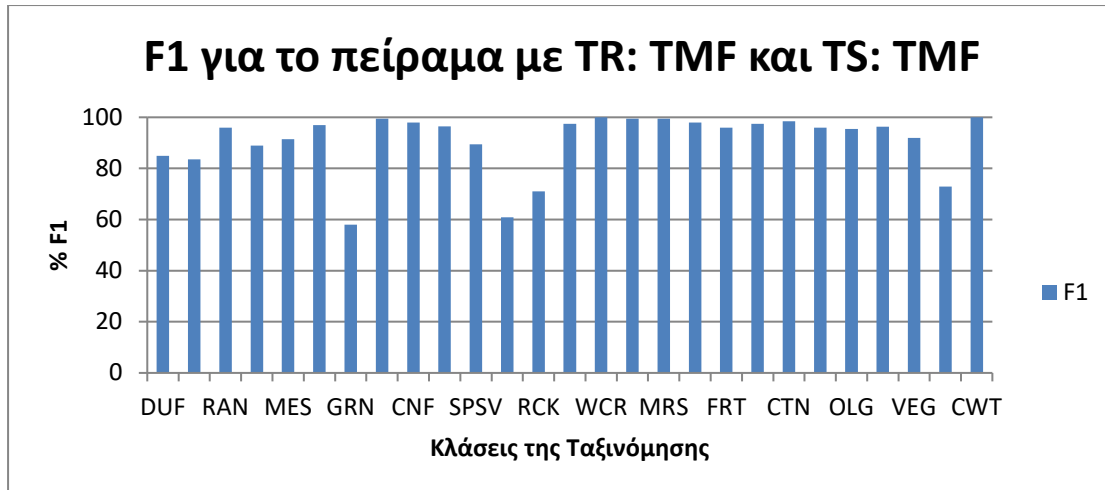
Πίνακας 5-6 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση στην εικόνα 35TMF

	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPSV	SVA	RCK	SND	WCR	WBD	MRS	VNY	FRT	CRL	CTN	GRF	OLG	MAZ	VEG	NGR	CWT			
DUF	82,90	6,11	0,39	2,37	0,13	0,15	6,19	0,00	0,00	0,00	0,00	0,00	0,00	0,38	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
SUF	9,27	83,79	0,45	2,88	0,51	0,44	0,88	0,00	0,00	0,00	0,03	10,53	22,73	0,13	0,00	0,00	0,03	0,42	0,09	0,14	0,26	0,43	0,68	0,00	0,17	0,00	0,00	0,00		
RAN	1,17	0,39	95,99	2,49	1,40	1,19	14,16	0,00	0,00	0,00	0,00	0,00	0,00	0,89	0,01	0,03	0,04	0,00	0,03	0,00	0,02	0,00	0,04	0,00	0,00	0,00	0,00	0,00		
ICU	5,09	1,11	1,98	90,19	7,77	0,59	27,43	0,00	0,00	0,00	0,00	2,63	4,55	2,55	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
MES	0,33	0,11	0,08	1,61	89,94	0,15	1,77	0,00	0,00	0,00	0,00	0,00	9,09	0,38	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
PHT	0,00	0,00	0,08	0,08	0,00	95,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,17	0,00	0,00		
GRN	0,00	0,00	0,00	0,08	0,00	0,00	41,59	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,21	0,00	0,00	0,00	0,00	0,00	0,00		
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,73	1,12	0,06	0,62	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,74	0,00	0,00	0,85	0,00	0,00	0,00	0,00	0,00	17,46	0,00	
CNF	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,22	97,75	1,76	0,25	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,18	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	1,05	97,01	9,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,36	0,00	0,00	0,00	1,03	0,00	0,00	0,00	1,59	0,00	0,00	
SPSV	0,00	0,44	0,00	0,00	0,00	0,00	0,00	0,02	0,06	1,07	88,47	15,79	0,00	0,00	0,00	0,00	0,00	0,63	0,18	0,05	0,02	0,21	3,12	0,00	0,00	0,00	5,56	0,00	0,00	
SVA	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	44,74	4,55	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
RCK	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	54,55	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SND	0,07	0,00	0,00	0,17	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	95,67	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WCR	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,95	0,40	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WBD	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,44	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MRS	0,20	0,61	0,67	0,04	0,00	0,15	2,65	0,00	0,01	0,01	0,25	13,16	4,55	0,00	0,03	0,13	99,77	0,00	1,45	0,34	0,30	0,64	0,09	0,00	0,00	0,00	0,00	0,00	0,00	
VNY	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	98,33	0,06	0,00	0,02	0,43	0,13	0,00	0,00	0,00	0,00	0,00	0,00	
FRT	0,00	0,11	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,03	5,26	0,00	0,00	0,00	0,00	0,02	0,63	95,31	0,27	0,13	0,00	0,94	1,33	0,34	11,90	0,00	0,00	0,00	
CRL	0,20	0,67	0,03	0,04	0,00	0,59	0,88	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,68	98,38	0,38	3,40	0,00	0,00	7,74	0,00	0,00	0,00	0,00	
CTN	0,59	6,39	0,22	0,04	0,25	1,63	4,42	0,00	0,00	0,00	0,06	2,63	0,00	0,00	0,00	0,00	0,10	0,00	0,42	0,68	98,83	0,43	0,17	5,33	6,40	0,00	0,00	0,00	0,00	
GRF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	93,19	0,00	0,00	0,00	0,79	0,00	0,00	0,00	
OLG	0,20	0,28	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,05	1,08	5,26	0,00	0,00	0,00	0,00	0,00	0,00	0,36	0,11	0,01	0,21	93,76	0,00	0,00	2,38	0,00	0,00	0,00	
MAZ	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	93,33	0,00	0,00	0,00	0,00	0,00	
VEG	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,01	0,00	0,00	0,00	85,19	0,00	0,00	0,00	0,00	
NGR	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15	0,00	0,00	0,00	0,04	0,00	0,00	60,32	0,00	0,00	0,00	
CWT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,00	0,00	

Πίνακας 5-7 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση στην εικόνα 35TMF

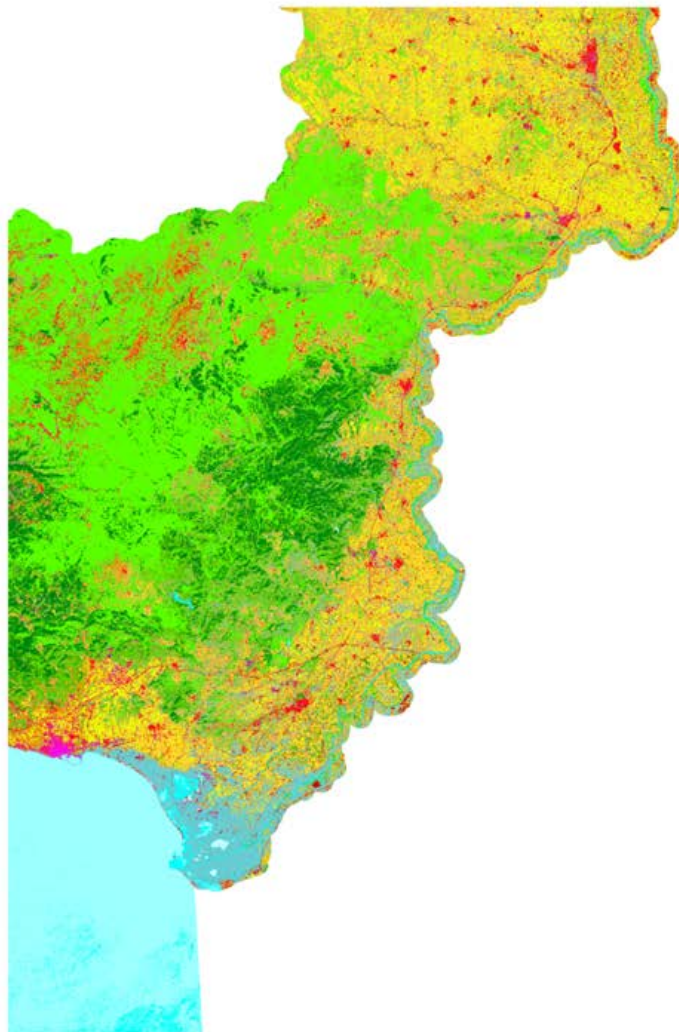
	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPSV	SVA	RCK	SND	WCR	WBD	MRS	VNY	FRT	CRL	CTN	GRF	OLG	MAZ	VEG	NGR	CWT
DUF	86,81	7,79	0,50	3,22	0,66	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,01	0,00	0,00	0,07	0,07	0,00	0,13	0,00	0,00	0,00	0,00
SUF	7,52	82,73	0,20	0,83	0,26	0,00	0,00	0,00	0,00	0,00	0,26	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,06	0,27	0,89	0,00	0,22	0,00	0,00	0,00	0,00
RAN	0,96	0,88	96,07	2,93	0,39	0,46	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00	0,03	0,02	0,06	0,00	0,09	0,00	0,00	0,00	0,00
ICU	3,83	3,73	1,65	88,10	4,98	0,31	4,00	0,00	0,00	0,00	0,00	0,00	0,00	0,53	0,00	0,00	0,00	0,00	0,00	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,00
MES	0,07	0,22	0,31	2,52	92,53	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00
PHT	0,07	0,16	0,22	0,17	0,13	99,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,09	0,09	0,00	0,00	0,00	0,00	0,00	0,00
GRN	0,48	0,05	0,45	1,28	0,26	0,00	94,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,02	0,04	0,00	0,00	0,00	0,00	0,00	0,00
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,27	0,46	0,08	0,22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,54	98,30	1,59	0,38	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,03	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	1,16	95,96	4,79	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15	0,00	0,00	0,00	0,31	0,00	0,00	0,00	0,00
SPSV	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,05	0,04	2,10	91,16	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,03	0,00	0,02	0,00	1,54	0,00	0,00	0,00	0,00
SVA	0,00	0,22	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,19	94,44	0,00	0,00	0,00	0,00	0,02	0,00	0,06	0,00	0,01	0,00	0,09	0,00	0,00	0,00	0,00
RCK	0,00	0,27	0,00	0,04	0,26	0,00	0,00	0,00	0,00	0,00	0,00	5,56	100,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SND	0,21	0,05	0,20	0,83	0,39	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,21	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WCR	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,81	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WBD	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,14	100,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MRS	0,00	0,38	0,25	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,05	0,00	99,27	0,00	0,12	0,02	0,19	0,00	0,00	0,00	0,00	0,00	0,00
VNY	0,00	0,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00	0,00	0,00	0,00	0,00	0,00	97,93	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
FRT	0,00	0,16	0,03	0,00	0,00	0,00	0,00	0,06	0,03	0,08	0,19	0,00	0,00	0,00	0,00	0,00	0,20	0,41	97,27	0,51	0,11	0,00	0,53	0,00	0,00	6,10	0,00
CRL	0,00	0,33	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,36	96,51	0,23	0,00	0,22	0,00	0,20	0,00	0,00
CTN	0,07	1,81	0,06	0,00	0,13	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00	0,16	0,62	0,52	1,07	97,87	0,68	0,04	0,00	0,20	0,00	0,00
GRF	0,00	0,11	0,00	0,00	0,00	0,00	2,00	0,01	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,01	0,41	0,00	0,36	0,02	99,10	0,04	0,00	0,00	0,00	0,00
OLG	0,00	0,88	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,17	2,33	0,00	0,00	0,00	0,00	0,00	0,01	0,62	0,67	0,00	0,03	0,00	96,57	0,00	0,00	1,22	0,00
MAZ	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,06	0,00	0,00	100,00	0,00	0,00	0,00
VEG	0,00	0,05	0,00	0,00	0,00	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	1,03	0,29	0,00	0,00	0,00	99,61	0,00	0,00
NGR	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,01	0,22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,45	0,00	0,00	0,23	0,13	0,00	0,00	92,68	0,00
CWT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,00

Στην συνέχεια παρατίθεται το γράφημα για τον δείκτη Fisher (F1) της βέλτιστης ταξινόμησης, που αναλύθηκε παραπάνω. Στο συγκεκριμένο γράφημα εμφανίζεται το ποσοστό του F1 για κάθε κλάση της ταξινόμησης. Με το γράφημα αυτό, μπορούμε και εποπτικά να εντοπίσουμε τις κατηγορίες εκείνες, που απέδωσαν καλύτερα, αλλά και τις κατηγορίες αυτές, που εμφάνισαν συγχύσεις.



Γράφημα 5-6 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMF

Όπως διαπιστώνεται, ο ταξινομητής έχει διαχωρίσει την πλειοψηφία των κατηγοριών επιτυχημένα, καθώς οι 25 από τις 27 κατηγορίες χαρακτηρίζονται από δείκτη F1 ανώτερο του 71%. Οι κατηγορίες όμως, οι οποίες διαχωρίστηκαν βέλτιστα, σύμφωνα με το δείκτη F1 είναι οι 4 κατηγορίες των υδάτινων επιφανιών και τα φυλλοβόλα δάση, ενώ οι κατηγορίες με τα χαμηλότερα ποσοστά είναι η GRN και η SVA. Οι κατηγορίες αυτές δεν διαθέτουν πολυπληθή δεδομένα εκπαίδευσης, γεγονός που οδηγεί σε δυσκολία της ταξινόμησής τους από τον αλγόριθμο RF. Τέλος, παρατίθεται ο χάρτης, που παρήχθη από το πείραμα αυτό.



Legend

- Dense Urban Fabric
- Sparse Urban Fabric
- Road and Rail Networks
- Industrial or Commercial Units
- Mineral Extraction Sites
- Photovoltaic Panels
- Greenhouses
- Broadleaved Forest
- Coniferous Forest
- Dense Sclerophyllous Vegetation
- Sparse Sclerophyllous Vegetation
- Sparsely Vegated Areas
- Bare Rocks
- Beaches, Dunes, Sands
- Water Courses
- Water Bodies
- Marshes
- Vineyards
- Fruit Trees and Berry Plantations
- Cereals
- Cotton
- Grass Fodders
- Olive Groves
- Maize
- Vegetables
- Natural Grassland
- Sea Coastal Water
- Kiwi Plants
- Tobacco
- Rice Fields
- Sunflowers

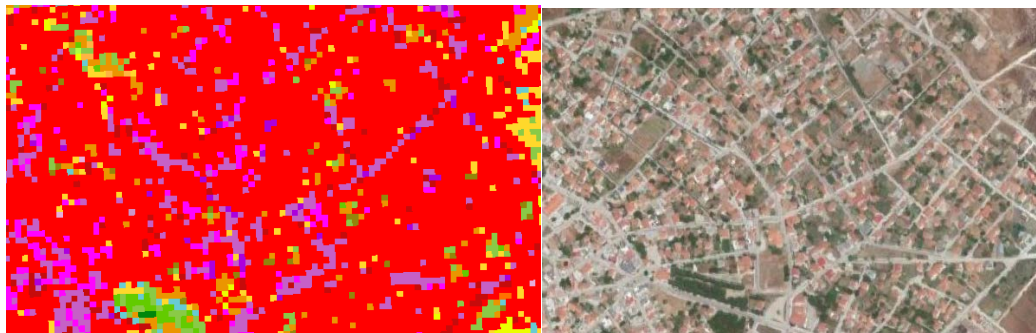
ματα και Αξιολόγηση

Χάρτης 5.2 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TLF και έλεγχο στην εικόνα 35TMF

Μελετώντας τον χάρτη, διαπιστώνεται ότι εν γένει οι αστικές, οι δασικές περιοχές, οι υδάτινες επιφάνειες, καθώς και οι καλλιέργειες έχουν αποδοθεί σωστά. Συνεπώς ο χάρτης παρέχει μία αξιόπιστη πρόβλεψη των χρήσεων/ καλύψεων γης και καλλιεργειών της περιοχής μελέτης

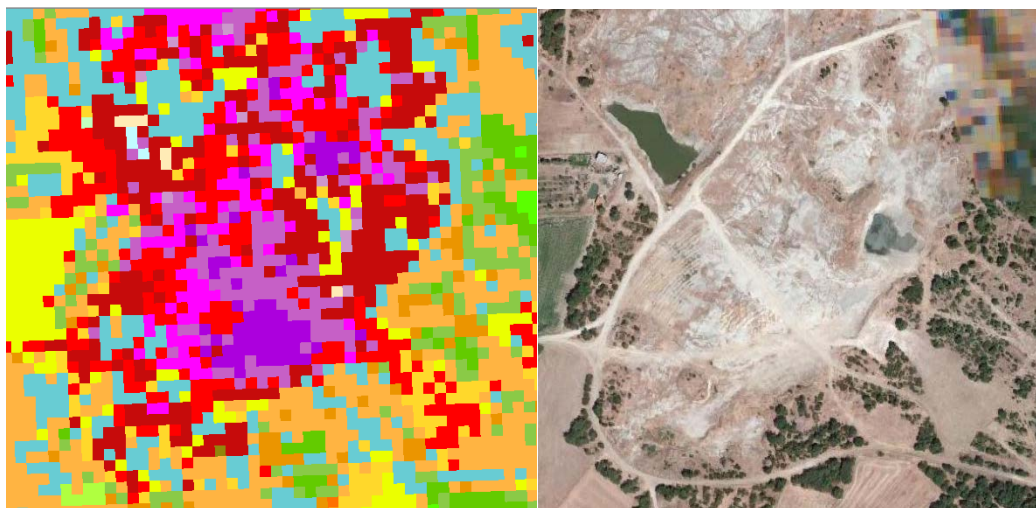
Εάν παρατηρήσουμε τον χάρτη πιο λεπτομερώς και σε μεγαλύτερο zoom, θα διαπιστώσουμε ότι ορισμένες κλάσεις εμφανίζουν συγχύσεις με άλλες, επιβεβαιώνοντας τα αποτελέσματα του πίνακα σύγχυσης. Αυτές οι κλάσεις, οι οποίες εμφανίζουν τα μεγαλύτερα προβλήματα είναι οι υποκατηγορίες των αστικών περιοχών.

Πιο συγκεκριμένα, οι κατηγορίες DUF και SUF, εμφανίζουν μικρές συγχύσεις, καθώς εικονοστοιχεία τα οποία ανήκουν στην SUF ταξινομούνται στην DUF και το αντίστροφο.



Εικόνα 5-7 Σύγκυση ανάμεσα στην DUF, την SUF και την ICU

Ακόμη, εικονοστοιχεία, που ανήκουν στην MES, αποδίδονται στις κλάσεις ICU, DUF και SUF.



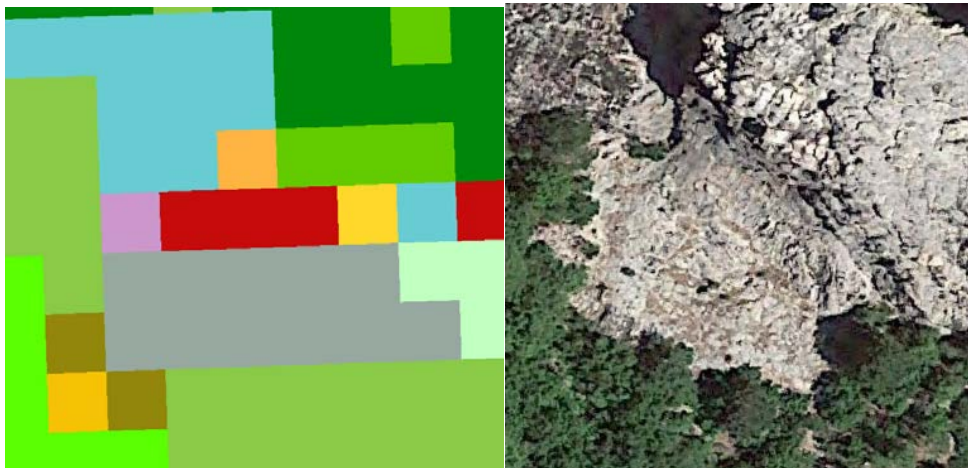
Εικόνα 5-8: Σύγκυση ανάμεσα στην MES, την DUF, την SUF και την ICU.

Τέλος, η κατηγορία GRN, εμφανίζει συγχύσεις με τις υπόλοιπες κατηγορίες των τεχνητών επιφανειών και ιδιαίτερος με τις κλάσεις MES και ICU.

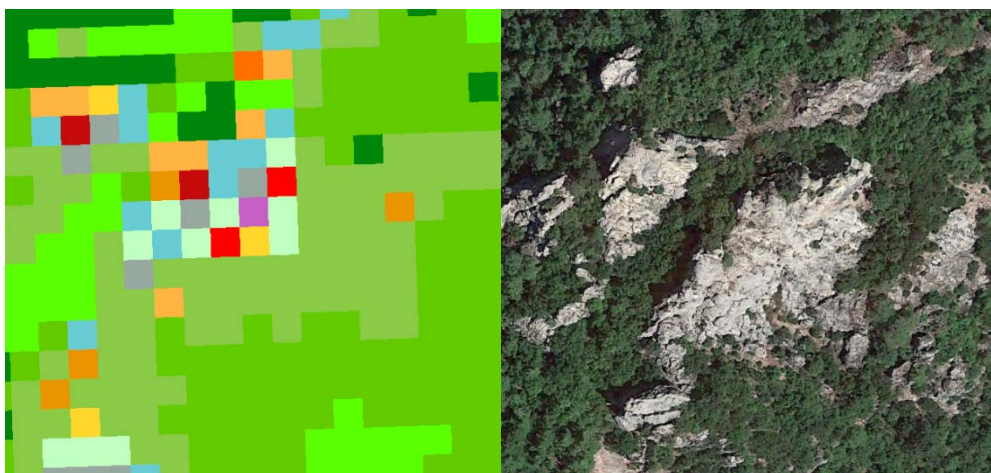


Εικόνα 5-9 Σύγκριση της GRN με την MES και την ICU.

Πέρα, από τις κατηγορίες των τεχνητών επιφανειών, δύο κλάσεις που εμφανίζουν σφάλματα στην απόδοσή τους είναι οι κατηγορίες του γυμνού εδάφους και συγκεκριμένα οι κατηγορίες SVA και RCK. Οι κλάσεις αυτές δεν διαθέτουν πολλά πολύγωνα εκπαίδευσης, με αποτέλεσμα ο ταξινομητής, να μη μπορεί να τις διαχωρίσει και κατά συνέπεια να τις αποδώσει αξιόπιστα, καθώς pixels της SVA αποδίδονται στην MRS και σε κατηγορίες των τεχνητών επιφανειών και pixels της RCK να αποδίδονται στην SVA.



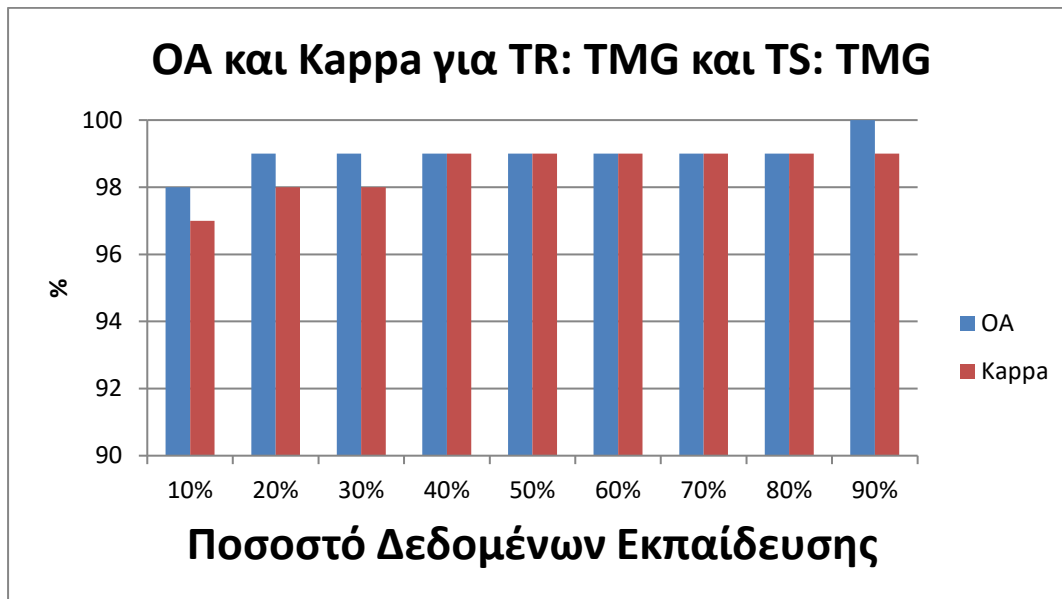
Εικόνα 5-10 Δυσκολία αναγνώρισης της SVA



Εικόνα 5-11 Δυσχέρεια του αλγορίθμου να αποδώσει την κατηγορία RCK

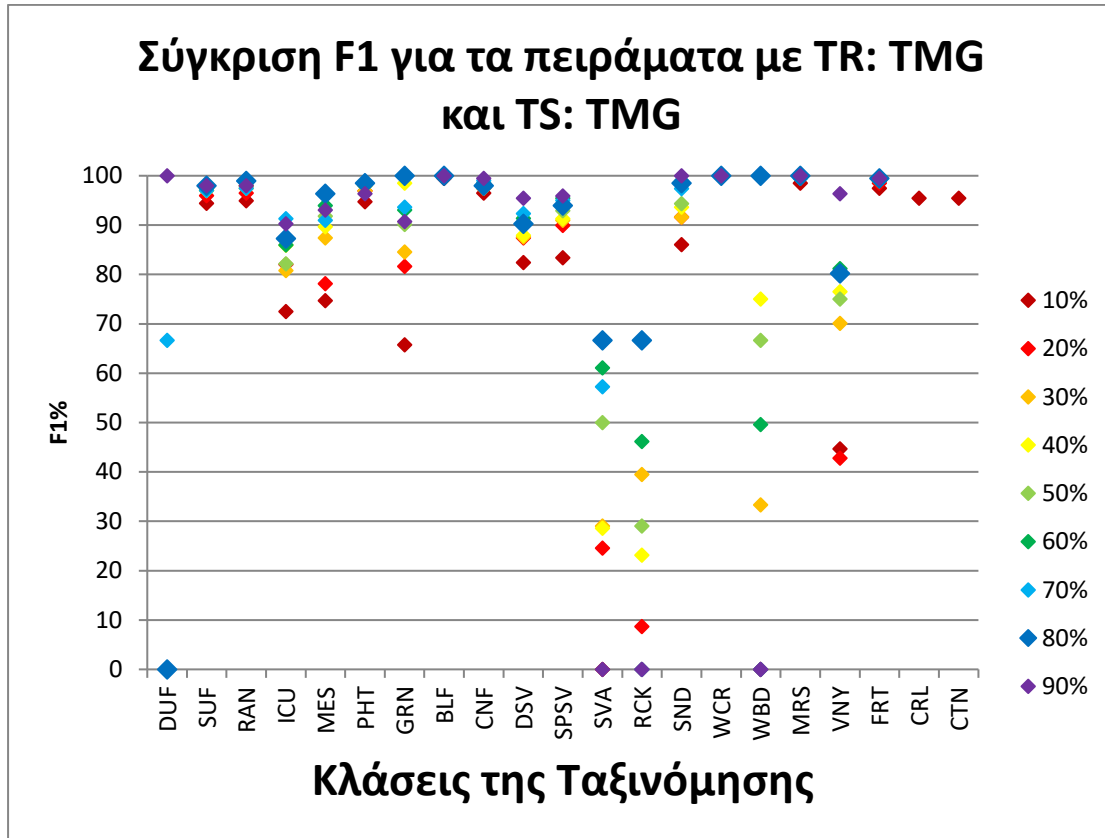
5.1.3 Τρίτο πείραμα: ταξινόμηση στο tile 35TMG

Στο τρίτο πείραμα που εφαρμόστηκε χρησιμοποιήθηκε το Tile 35TMG για να εκπαιδευτεί ο αλγόριθμος, ενώ ο έλεγχος της ταξινόμησης επιλέχθηκε να γίνει στο ίδιο tile. Χρησιμοποιήθηκαν και οι έξι επιλεγμένες εικόνες και από αυτές τα κανάλια 2,3,4,8,11,12, καθώς και οι δείκτες NDVI, MSAVI, NDWI, NDBI. Οι κλάσεις της ταξινόμησης ήταν 21 και όχι 31, καθώς στο συγκεκριμένο tile, δεν απαντώνται δεδομένα εκπαίδευσης για τις κλάσεις CRL, CTN, GRF, OLG, MAZ, VEG, NGR, CWT, KWP, TBC, RCF και για την SFL. Έτρεξαν 9 ταξινομήσεις, για εκπαίδευση του ταξινομητή με 10, 20, 30, 40, 50, 60, 70, 80 και 90% των δεδομένων αναφοράς. Τα αποτελέσματα που προέκυψαν από τις ταξινομήσεις δεν διέφεραν σημαντικά. Η συνολική ακρίβεια της ταξινόμησης και ο δείκτης Kappa, για τα εννέα πειράματα, παρουσιάζονται στο παρακάτω γράφημα.



Γράφημα 5-7 Συνολική ακρίβεια και δείκτης Kappa για την ταξινόμηση στην εικόνα 35TMG

Μελετώντας το γράφημα, διαπιστώνεται ότι και οι 9 ταξινομήσεις έχουν αποδώσει παρόμοιας ακρίβειας αποτελέσματα. Σημειώνεται ότι όσο περισσότερο ποσοστό των δεδομένων εκπαίδευσης χρησιμοποιούνται στην ταξινόμηση, τόσο υψηλότερης ακρίβειας αποτελέσματα αποδίδει ο αλγόριθμος, ενώ οι ταξινομήσεις στις οποίες χρησιμοποιήθηκε ποσοστό δεδομένων εκπαίδευσης από 40% και πάνω, αποδίδουν παρόμοια αποτελέσματα. Ελέγχοντας τα αποτελέσματα ακρίβειας και αξιοπιστίας της κάθε κλάσης της ταξινόμησης καθώς και το δείκτη F1, διαπιστώνεται ότι οι περισσότερες κλάσεις αποδίδουν καλύτερα αποτελέσματα όσο αυξάνεται το ποσοστό των δεδομένων εκπαίδευσης που χρησιμοποιείται για εκπαίδευση. Όμως, τα αποτελέσματα της ακρίβειας και της αξιοπιστίας των κλάσεων δεν διαφέρουν σημαντικά στις ταξινομήσεις, που χρησιμοποιείται ποσοστό δεδομένων εκπαίδευσης από 50% και πάνω. Επιλέγεται να παρατεθεί και αξιολογηθεί η ταξινόμηση με ποσοστό δεδομένων εκπαίδευσης ίσο με 50%, καθώς σημειώνει υψηλά ποσοστά στα μεγέθη αξιολόγησης της ταξινόμησης και των κλάσεων αυτής, τα οποία δεν μεταβάλλονται σημαντικά στις ταξινομήσεις με ποσοστό δεδομένων εκπαίδευσης πάνω από 50%.



Γράφημα 5-8 Δείκτης F1 για την ταξινόμηση στην εικόνα 35TMG

Παρακάτω παρατίθεται ο πίνακας σύγκρισης της ταξινόμησης. Εν γένει η ταξινόμηση θεωρείται πετυχημένη, καθώς, η γενική ακρίβεια και ο δείκτης kappa ξεπερνούν το 98%.

Πίνακας 5-8 Πίνακας Σύγκρισης για ταξινόμηση στην εικόνα 35TMG

		ΚΛΑΣΕΙΣ ΤΟΥ ΑΠΟΤΕΛΕΣΜΑΤΟΣ ΤΗΣ ΤΑΞΙΝΟΜΗΣΗΣ																				
		DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPSV	SVA	RCK	SND	WCR	WBD	MRS	VNY	FRT	sum	PA
ΚΛΑΣΕΙΣ ΤΟΥ ΧΑΡΤΗ ΤΟΥ GROUND TRUTH	DUF	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
	SUF	0	1822	11	6	2	1	0	0	0	0	0	0	0	2	0	0	0	0	11	1855	0,98
	RAN	0	8	1381	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1395	0,99
	ICU	0	19	2	81	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	105	0,77
	MES	0	3	4	2	65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	74	0,88
	PHT	0	11	2	0	0	337	0	0	0	0	0	0	0	0	0	0	0	0	0	350	0,96
	GRN	0	0	5	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	28	0,82
	BLF	0	0	0	0	0	0	0	2713	0	0	0	0	0	0	0	0	0	0	0	2714	
	CNF	0	0	0	0	0	0	0	4	8	2	0	0	0	0	0	0	0	0	2	6	1
	DSV	0	0	0	0	0	0	0	13	724	3	0	0	0	0	0	0	0	0	1	741	0,98
	SPSV	0	0	0	0	0	0	0	41	3	412	10	0	0	0	0	0	0	0	18	484	0,85
	SVA	0	0	0	0	0	0	0	17	0	8	344	0	0	0	0	0	0	0	8	377	0,91
	SVA	0	0	2	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	4	0,5
	RCK	0	0	7	0	0	0	0	0	0	0	0	0	2	2	0	0	0	0	1	12	0,17
	SND	0	7	7	0	0	0	0	0	0	0	0	0	0	0	149	0	0	0	0	163	0,91
	WCR	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	4110	0	0	0	4111	1
	WBD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	4	0,5
	MRS	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	286	0	0	287	1
	VNY	0	11	2	0	0	0	0	1	0	1	1	0	0	0	0	0	0	41	11	68	0,6
	FRT	0	2	0	0	0	0	0	20	1	0	7	0	0	0	0	0	0	0	0	5331	5361
sum	0	1886	1423	92	68	338	23	2722	6	737	426	362	4	2	152	4113	2	286	41	5387	a:	0,99
UA	0	0,97	0,97	0,88	0,96	1	1	1	1	0,98	0,97	0,95	0,5	1	0,98	1	1	1	1	0,99	OA:	0,99

Μελετώντας επιμέρους την ακρίβεια και την αξιοπιστία ανά κατηγορία, διαπιστώνεται ότι η πλειοψηφία των κλάσεων έχει ταξινομηθεί επιτυχώς, καθώς 14 από τις 19 κλάσεις χαρακτηρίζονται από ακρίβεια μεγαλύτερη του 77%. Όμως, υπάρχουν κατηγορίες οι οποίες εμφανίζουν μέτρια μέχρι και χαμηλά ποσοστά ακρίβειας και αξιοπιστίας, καθώς κυμαίνονται από 60 έως 0%. όπως είναι οι κατηγορίες DUF, SVA, RCK, WBD και VNY.

Πιο συγκεκριμένα, αν εστιάσουμε στις κλάσεις των τεχνητών επιφανειών, θα διαπιστώσουμε ότι η κλάση DUF, δεν διαχωρίστηκε από τον ταξινομητή. Αυτό οφείλεται στον μικρό αριθμό δεδομένων εκπαίδευσης που χρησιμοποιήθηκε. Η κατηγορία αυτή δεν απαντάται στο συγκεκριμένο tile και τα λιγοστά δεδομένα εκπαίδευσης δεν κατέστησαν ικανά να οδηγήσουν στο διαχωρισμό της από την κατηγορία SUF.

Οι κατηγορίες SUF και RAN, ταξινομούνται με ευκολία από τον αλγόριθμο, καθώς πετυχαίνει ποσοστά ακρίβειας παραγωγού και χρήστη ανώτερα του 96%. Συγχύσεις αξιόλογες δεν παρατηρούνται.

Η κατηγορία ICU, ταξινομείται ικανοποιητικά από τον ταξινομητή, καθώς εμφανίζει ακρίβεια παραγωγού ίση με 77% και ακρίβεια χρήστη ίση με 88%. Παρατηρώντας σφάλματα παράλειψης και συμπερίληψης, διαπιστώνουμε ότι τα μεγαλύτερα (18% και 6% αντίστοιχα) παρατηρούνται με την SUF. Μικρότερες συγχύσεις εμφανίζονται με τις κλάσεις RAN και MES. Η σύγχυση της DUF με την ICU είναι ως ένα βαθμό αναμενόμενη, καθώς, τα κτήρια που απαρτίζουν την κατηγορία ICU είναι δύσκολο να διαχωριστούν πλήρως από αυτά που απαρτίζουν την SUF.

Η κατηγορία MES, χαρακτηρίζεται από υψηλή ακρίβεια (88%) και από πολύ υψηλή αξιοπιστία (96%), ενώ οι μικρές συγχύσεις που κυμαίνονται από 1-5%, εμφανίζονται στις κατηγορίες ICU, RAN και SUF.

Η κατηγορία PHT, χαρακτηρίζεται από πολύ υψηλή ακρίβεια παραγωγής (96%) και χρήστη (99,7%). Δεν παρατηρούνται αξιοσημείωτα σφάλματα παράλειψης και συμπερίληψης. Αξίζει να σημειωθεί πως μικρή σύγχυση παρατηρείται με την κλάση SUF, αφού τα 3,14% των εικονοστοιχείων της κλάσης PHT, παραλείπονται από την κατηγορία και αποδίδονται στην κλάση SUF.

Τέλος η κλάση GRN, εμφανίζει υψηλή ακρίβεια (82%) και άριστη αξιοπιστία (100%). Από τα σφάλματα παράλειψης, διαπιστώνουμε σημαντική σύγχυση με την κλάση RAN, καθώς περίπου 17% των εικονοστοιχείων της κλάσης GRN, αποδίδονται στην κλάση RAN

Οι συγχύσεις που παρουσιάζονται στην κατηγορία στην GRN είναι αποτέλεσμα του μικρού αριθμού δεδομένων εκπαίδευσης που παρέχονται στον ταξινομητή, καθώς δεν υπάρχει επαρκές δείγμα ώστε ο αλγόριθμος να εκπαιδευτεί κατάλληλα και να κατόπιν να διαχωρίσει αυτήν την κατηγορία.

Οι κατηγορίες της φυσικής βλάστησης ταξινομήθηκαν στο σύνολό τους επιτυχημένα από τον αλγόριθμο ταξινόμησης, μιας και τα ποσοστά ακρίβειας χρήστη και παραγωγού ξεπερνούν σε κάθε περίπτωση το 85%.

Μικρή σύγχυση, εμφανίζεται ανάμεσα στις κλάσεις BLF και CNF, καθώς για την κλάση CNF, απαντάται σφάλμα παράλειψης 1,7% και σφάλμα συμπερίληψης 1% με την κλάση BLF.

Η κατηγορία , η οποία εμφάνισε τα μεγαλύτερα σφάλματα παράλειψης, ήταν η κατηγορία DSV, καθώς το 8% των εικονοστοιχείων του ground truth, που απαρτίζουν την κλάση DSV, ταξινομήθηκαν στην κλάση BLF. Μικρότερη σύγχυση, με σφάλμα παράλειψης 2% εντοπίζεται με την κατηγορία SPSV.

Αντίστοιχα σφάλματα παράλειψης, εντοπίζονται και στην κλάση SPSV, αφού 4% των εικονοστοιχείων του ground truth της αντιστοιχίζονται στην κλάση BLF και 2% στην κλάση DSV.

Η μικρή δυσκολία του ταξινομητή να διαχωρίσει πλήρως τις κλάσεις DSV και SPSV, είναι λογική και αναμενόμενη, καθώς πρόκειται για υποκατηγορίες, της κλάσης σκληροφυλλική βλάστηση. Όμως, οι συγχύσεις με των κλάσεων αυτών με την κατηγορία BLF δεν ήταν αναμενόμενη, καθώς, τα φυλλώματα τα δύο κατηγοριών, με αποτέλεσμα να εκπέμπουν αρκετά διαφορετικά στο φάσμα.

Σε σχέση με τις κατηγορίες του γυμνού εδάφους, η κατηγορία που ταξινομήθηκε με μεγαλύτερη επιτυχία ήταν η κατηγορία SND, η οποία επέτυχε ποσοστά ακρίβειας και αξιοπιστίας 91% και 98% αντίστοιχα.

Η κατηγορία SVA, χαρακτηρίζεται από μέτρια ακρίβεια και αξιοπιστία (50%), εμφανίζει μεγάλα σφάλματα παράλειψης και συμπερίληψης με τις κλάσεις RAN και RCK.

Η κατηγορία RCK, χαρακτηρίζεται από άριστη αξιοπιστία (100%) και από πολύ χαμηλή ακρίβεια παραγωγής (16%) καθώς εμφανίζει έντονα σφάλματα παράλειψης με τις κλάσεις RAN (50%), SVA (16%) και FRT (8%).

Η αποτυχία του αλγορίθμου να ταξινομήσει τις συγκεκριμένες κλάσεις, είναι αποτέλεσμα του μικρού αριθμού δεδομένων εκπαίδευσης, καθώς οι συγκεκριμένες κατηγορίες δεν απαντώνται ευρέως στο συγκεκριμένο tile.

Οι κατηγορίες υδάτινων πόρων διαχωρίστηκαν πλήρως από τον αλγόριθμο ταξινόμησης, αφού επιτεύχθηκε ακρίβεια και αξιοπιστία για κάθε κλάση μεγαλύτερη και από 99%, χωρίς την ύπαρξη αξιόλογων εμπλοκών ανάμεσα στις διάφορες κατηγορίες. Εξαίρεση αποτελεί η κατηγορία WBD, καθώς εμφανίζει μέτρια ακρίβεια παραγωγής (50%), λόγω της σύγχυσης της με την κλάση WCR. Η σύγχυση αυτή είναι αποτέλεσμα του μικρού αριθμού δεδομένων εκπαίδευσης για την κλάση των υδάτινων σωμάτων.

Τέλος, στο συγκεκριμένο tile, δεν υπάρχει μεγάλο πλήθος καλλιεργείων. Οι μόνες κατηγορίες, που απαντώνται είναι η κατηγορία VNY και η κατηγορία FRT.

Πιο συγκεκριμένα, η κατηγορία VNY, χαρακτηρίζεται από άριστη αξιοπιστία (100%) και από μέτρια προς καλή ακρίβεια (60%). Τα σημαντικότερα σφάλματα παράλειψης απαντώνται στις κλάσεις FRT και SUF, μία σύγχυση που δεν θεωρείται αναμενόμενη και οφείλεται στο λιγοστό ground truth.

Πίνακας 5-9 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση στην εικόνα 35TMG

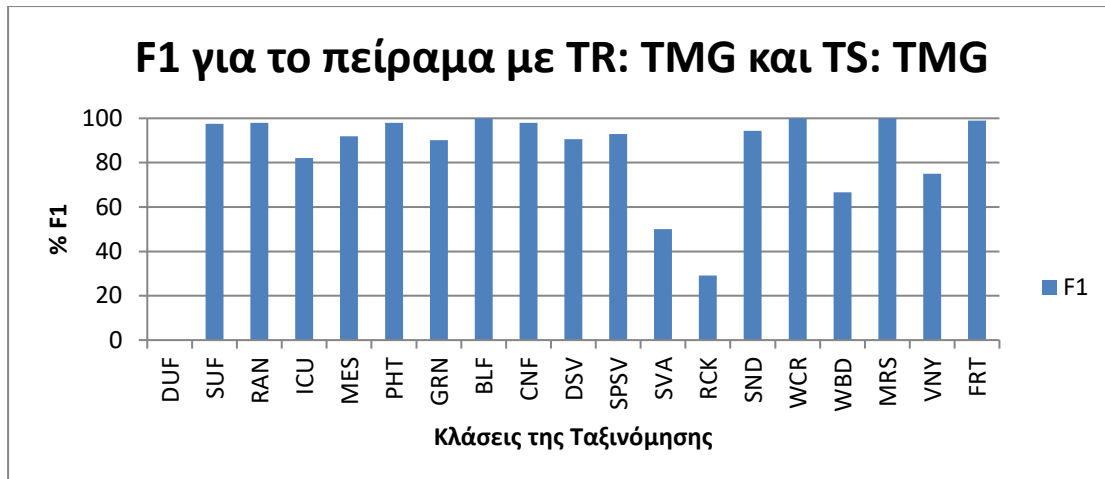
	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MRS	VNY	FRT	
DUF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SUF	100	98,2	0,57	18,1	4,05	3,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	4,29	0,00	0,00	0,00	16,1	0,04	
RAN	0,00	0,59	99,0	1,90	5,41	0,57	17,8	0,00	0,00	0,00	0,00	50,0	58,3	4,29	0,00	0,00	0,00	2,94	0,00	
ICU	0,00	0,32	0,14	77,1	2,70	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,35	0,00	0,00	
MES	0,00	0,11	0,00	0,95	87,8	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
PHT	0,00	0,05	0,00	0,00	0,00	96,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
GRN	0,00	0,00	0,00	0,00	0,00	0,00	82,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,9	1,75	8,47	4,51	0,00	0,00	0,00	0,00	0,00	0,00	1,47	0,37	
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	97,7	0,62	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,02	
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,40	85,1	2,12	0,00	0,00	0,00	0,00	0,00	0,00	1,47	0,00	
SPS	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2,07	91,2	0,00	0,00	0,00	0,00	0,00	0,00	1,47	0,13	
SVA	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	50,0	16,6	0,00	0,00	0,00	0,00	0,00	0,00	
RCK	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	16,6	0,00	0,00	0,00	0,00	0,00	0,00	
SND	0,00	0,11	0,00	0,95	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	91,4	0,00	0,00	0,00	0,00	0,00	
WC	0,00	0,00	0,00	0,95	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,9	50,0	0,00	0,00	0,00	
WB	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	50,0	0,00	0,00	0,00	
MRS	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,6	0,00	0,00	
VNY	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	60,2	0,00	
FRT	0,00	0,59	0,29	0,00	0,00	0,00	0,00	0,01	0,13	3,72	2,12	0,00	8,33	0,00	0,00	0,00	0,00	16,1	99,4	

Πίνακας 5-10 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση στην εικόνα 35TMG

	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MRS	VNY	FRT
DUF	0,00	0,16	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SUF	0,00	96,6	0,77	6,52	2,94	0,30	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,32	0,00	0,00	0,00	0,00	0,20
RAN	0,00	0,42	97,0	2,17	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07
ICU	0,00	1,01	0,14	88,0	1,47	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,66	0,02	0,00	0,00	0,00	0,00
MES	0,00	0,16	0,28	2,17	95,5	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
PHT	0,00	0,58	0,14	0,00	0,00	99,7	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
GRN	0,00	0,00	0,35	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,6	1,09	0,47	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	98,2	0,70	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15	0,41	96,7	2,76	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,33
SPS	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	1,88	95,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15
SVA	0,00	0,00	0,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	50,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00
RCK	0,00	0,00	0,49	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	50,0	100	0,00	0,00	0,00	0,00	0,00	0,02
SND	0,00	0,37	0,49	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	98,0	0,00	0,00	0,00	0,00	0,00
WC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,14	0,00	0,00	0,00	0,00	0,00	99,9	0,00	0,00	0,00	0,00
WB	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	100	0,00	0,00	0,00	0,00
MRS	0,00	0,00	0,00	1,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00
VNY	0,00	0,58	0,14	0,00	0,00	0,00	0,00	0,00	0,00	0,23	0,28	0,00	0,00	0,00	0,00	0,00	0,00	100	0,20
FRT	0,00	0,11	0,00	0,00	0,00	0,00	0,00	0,07	0,14	0,00	1,93	0,00	0,00	0,00	0,00	0,00	0,00	0,00	98,9

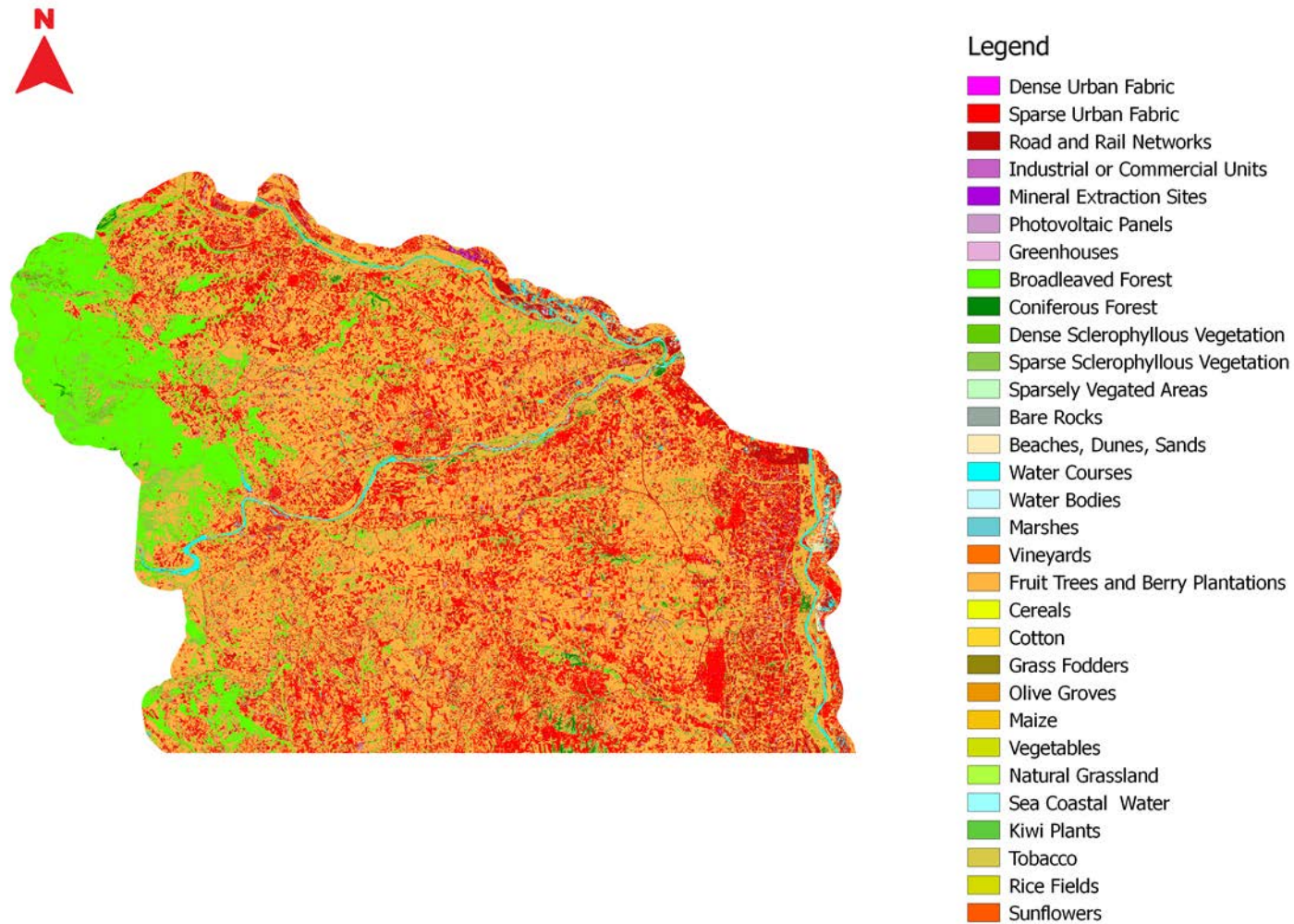
Η κλάση FRT, εμφανίζει υψηλή ακρίβεια και αξιοπιστία (99%), ενώ δεν εμφανίζει αξιολογες συγχύσεις με άλλες κλάσεις.

Στην συνέχεια παρατίθεται το γράφημα για τον δείκτη Fisher (F1) της βέλτιστης ταξινόμησης, που αναλύθηκε παραπάνω. Στο συγκεκριμένο γράφημα εμφανίζεται το ποσοστό του F1 για κάθε κλάση της ταξινόμησης. Με το γράφημα αυτό, μπορούμε και εποπτικά να εντοπίσουμε τις κατηγορίες εκείνες, που απέδωσαν καλύτερα, αλλά και τις κατηγορίες αυτές, που εμφάνισαν συγχύσεις.



Γράφημα 5-9 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMG

Διαπιστώνεται ότι ο ταξινομητής έχει διαχωρίσει την πλειοψηφία των κατηγοριών επιτυχημένα, καθώς οι 15 από τις 19 κατηγορίες χαρακτηρίζονται από δείκτη F1 ανώτερο του 75%. Οι κατηγορίες όμως, οι οποίες διαχωρίστηκαν βέλτιστα, σύμφωνα με το δείκτη F1 είναι οι κατηγορίες δασών BLF και CNF και οι καλλιέργειες οπωροφόρων δένδρων (FRT) ενώ οι κατηγορίες οι οποίες δεν ταξινομήθηκαν επιτυχώς σύμφωνα με τον F1 είναι η DUF, η SVA και RCK, ενώ μέτρια απόδοση είχε ο ταξινομητής και στο διαχωρισμό της WBD. Οι κατηγορίες αυτές διαθέτουν μικρό αριθμό δεδομένων εκπαίδευσης, καθώς δεν απαντώνται ευρέως στην περιοχή μελέτης, γεγονός που καθιστά προβληματικό ή και αδύνατο το διαχωρισμό και την αναγνώρισή τους από τον ταξινομητή. Τέλος, παρατίθεται ο χάρτης, που παρήχθη από το πείραμα αυτό.

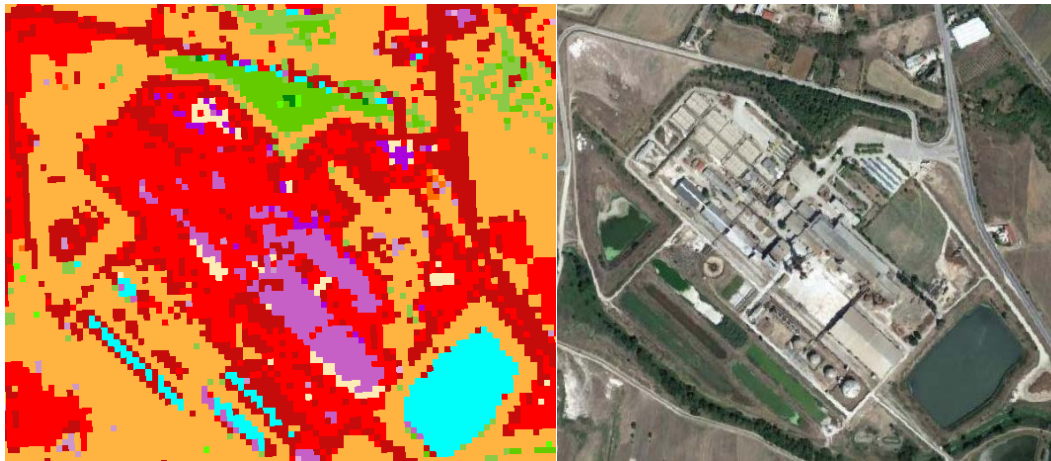


Χάρτης 5.3 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG

Μελετώντας τον χάρτη, διαπιστώνεται ότι εν γένει οι αστικές και οι δασικές περιοχές, έχουν αποδοθεί σωστά. Οι αγροτικές περιοχές, αναγνωρίζονται από τον ταξινομητή, καθώς, όμως το ground truth των καλλιεργειών στο συγκεκριμένο tile, διέθετε πολύγωνα μόνο για τις καλλιέργειες FRT και VNY (δεδομένα ΟΠΕΚΕΠΕ), ολόκληρες οι καλλιεργήσιμες περιοχές, έχουν αποδοθεί αναγκαστικά στις δύο κλάσεις, με την πλειοψηφία των εικονοστοιχείων να αναγνωρίζονται ως FRT. Συνεπώς ο χάρτης παρέχει μία αξιόπιστη πρόβλεψη των γενικών χρήσεων/ καλύψεων γης, που απαντώνται στην περιοχή αυτήν.

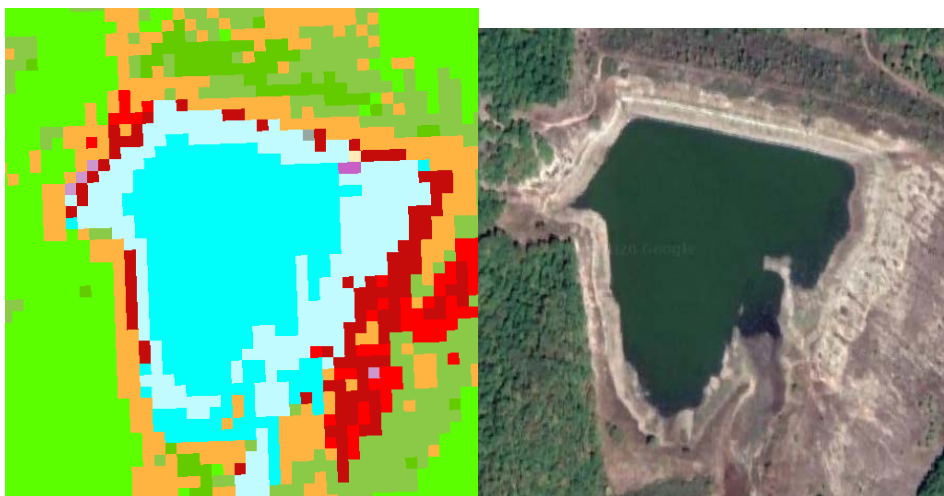
Εάν παρατηρήσουμε τον χάρτη πιο λεπτομερώς και σε μεγαλύτερο zoom, θα διαπιστώσουμε ότι ορισμένες κλάσεις εμφανίζουν συγχύσεις με άλλες, επιβεβαιώνοντας τα αποτελέσματα του πίνακα σύγχυσης.

Πιο συγκεκριμένα, η κατηγορία ICU εμφανίζει συγχύσεις με την SUF, καθώς εικονοστοιχεία τα οποία ανήκουν στην ICU αποδίδονται στην SUF.



Εικόνα 5-12 Σύγχυση ανάμεσα στην DUF, την SUF και την ICU

Επιπλέον, ενώ η WCR, αποδίδεται επιτυχημένα στον παραγόμενο χάρτη, η κατηγορία , δεν αποδίδεται εξίσου επιτυχημένα, αφού τα εικονοστοιχεία της αποδίδονται στην WCR.



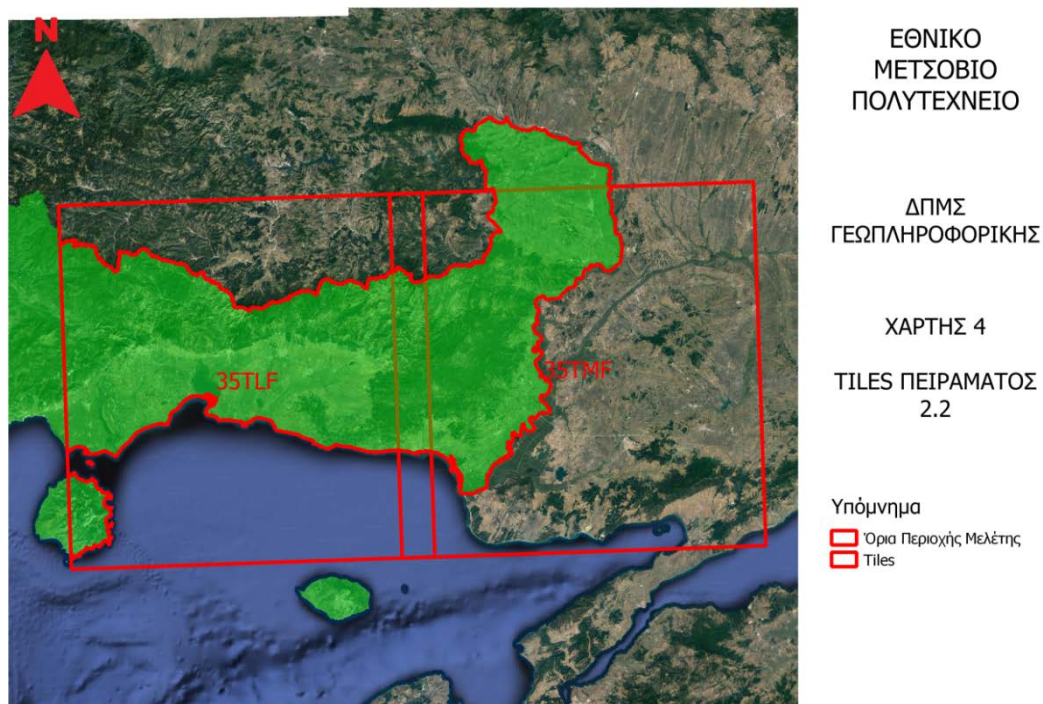
Εικόνα 5-13 Σύγχυση ανάμεσα WCR και στην WBD

5.2 Δεύτερη Κατηγορία Πειραμάτων- Ανά Δύο Tiles

Στην δεύτερη κατηγορία πειραμάτων, χρησιμοποιήθηκε διαφορετικό tile για την εκπαίδευση του ταξινομητή και διαφορετικό tile για τον έλεγχο της ταξινόμησης. Σε αυτού του τύπου τα πειράματα αναμένεται τόσο η γενική ακρίβεια, όσο και η ακρίβεια ορισμένων κλάσεων, ιδίως φυτοκάλυψης και καλλιέργειών να είναι κάτω του μετρίου. Αυτό συμβαίνει, διότι οι κατηγορίες φυτοκάλυψης και οι καλλιέργειες εμφανίζουν διαφοροποιημένα φασματικά χαρακτηριστικά από περιοχή σε περιοχή, καθώς ανά περιοχή εμφανίζονται διαφορετικές κλιματικές συνθήκες. Επιπλέον, η ενέργεια, που εκπέμπεται από τα αντικείμενα κάθε μίας κατηγορίας, διαφοροποιείται από tile σε tile. Αυτό συμβαίνει γιατί μεταβάλλονται οι ατμοσφαιρικές συνθήκες (π.χ. η συννεφιά, η υγρασία και η πυκνότητα του αέρα) -οι οποίες αποτελούν συνάρτηση των καιρικών συνθηκών και του γεωγραφικού πλάτους- και το ύψος του Ήλιου από τον ορίζοντα – που είναι συνάρτηση του γεωγραφικού πλάτους και του χρόνου λήψης των εικόνων [34]. Ο χρόνος λήψης των εικόνων των τριών tiles, δεν είναι ταυτίζεται για τις εικόνες των διαφορετικών ημερομηνιών, λόγω ύπαρξης νεφοκάλυψης στις εικόνες ίδιας ημερομηνίας (βλ. παράγραφο 3.2.1.). Συνεπώς, ο διαφορετικός χρόνος λήψης και το διαφορετικό γεωγραφικό πλάτος, οδηγούν σε διαφοροποιήσεις στα φασματικά χαρακτηριστικά των κλάσεων και συνεπώς σε μεταξύ τους σύγχυση.

5.2.1 Πρώτο πείραμα: εκπαίδευση στο tile 35TLF και έλεγχος στο Tile 35TMF

Στο πρώτο πείραμα που εφαρμόστηκε επιλέχθηκε να χρησιμοποιηθεί το Tile 35TLF για να εκπαιδευτεί ο αλγόριθμος, ενώ ο έλεγχος της ταξινόμησης επιλέχθηκε να γίνει στο Tile 35TMF. Χρησιμοποιήθηκαν και οι έξι επιλεγμένες εικόνες και από αυτές τα κανάλια 2,3,4,8,11,12, καθώς και οι δείκτες NDVI, MSAVI, NDWI, NDBI. Έτσι, η υλοποιήθηκε για υπερκύβους βάθους 10*6.



Χάρτης 5.4 Τα tiles, που χρησιμοποιήθηκαν από τον αλγόριθμο για την διεξαγωγή του πρώτου πειράματος.

Το πρόγραμμα έτρεξε και μέσα από ταξινόμηση RF, εξήχθησαν τα εξής αποτελέσματα.

Πίνακας 5-11 Πίνακας Σύγκρισης για ταξινόμηση με δεδομένα εκπαίδευσης από την εικόνα 35TML και έλεγχο στην εικόνα 35TMF

ΚΛΑΣΕΙΣ ΤΟΥ ΑΠΟΤΕΛΕΣΜΑΤΟΣ ΤΗΣ ΤΑΞΙΝΟΜΗΣΗΣ																																	
	DUF	SUF	RA	ICU	ME	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MR	VNY	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL	sum	PA
DU	219	229	49	412	38	0	0	0	0	0	0	0	4	24	0	0	0	0	0	79	23	3	3	0	0	0	0	0	0	0	0	306	0,72
SUF	725	191	10	33	0	0	0	0	0	0	13	14	2	5	0	0	0	15	9	536	210	12	72	2	14	0	0	1	6	0	3	360	0,53
RA	750	94	455	898	83	3	0	0	6	0	24	1	15	129	0	1	2	16	13	193	314	6	66	2	0	0	0	0	8	0	2	717	0,63
ICU	110	139	128	246	483	0	0	0	0	0	7	10	24	204	0	0	0	1	1	107	33	8	3	0	3	0	0	0	4	0	2	472	0,52
ME	78	17	7	250	918	0	0	0	0	0	4	3	37	111	0	0	0	1	0	92	49	0	1	0	0	0	0	0	2	0	0	157	0,58
PH	139	49	405	9	31	287	0	0	0	0	11	6	4	4	0	0	0	1	3	262	70	41	3	5	8	0	0	3	0	0	8	134	0,21
GR	41	9	17	66	28	0	10	0	0	0	3	0	1	8	0	0	0	4	0	21	9	0	3	0	1	0	0	0	5	0	0	226	0,04
BLF	0	0	1	0	0	0	1	621	248	119	28	0	0	0	0	1	4	0	18	0	0	33	6	287	0	81	0	35	0	3	0	876	0,71
CN	0	0	0	0	0	0	0	305	372	189	85	0	0	0	0	0	6	0	52	0	1	7	185	11	0	0	0	0	0	0	0	425	0,88
DS	0	0	0	0	0	0	0	362	111	144	151	0	0	0	0	0	3	1	5	0	0	9	414	8	0	1	0	42	0	1	0	279	0,52
SPS	0	8	1	0	0	1	0	225	311	841	389	0	0	0	0	0	24	18	14	7	2	28	105	5	0	0	0	17	0	0	0	645	0,6
SV	0	9	0	0	2	0	0	0	4	0	38	3	0	0	0	0	2	0	3	8	1	0	3	1	0	0	0	1	1	0	0	76	0,04
RCK	3	0	0	2	0	0	0	0	0	0	9	2	5	6	0	0	0	1	0	12	1	0	2	0	0	0	0	0	1	0	0	44	0,11
SN	29	10	0	149	295	0	0	0	0	0	0	0	64	999	0	0	0	1	0	9	11	0	2	0	0	0	0	0	0	0	0	156	0,64
WC	14	0	62	0	0	0	0	3	519	8	0	2	0	0	132	155	260	0	0	0	2	0	15	0	0	0	181	0	0	0	0	218	0,06
WB	0	0	6	0	0	0	0	0	1	0	0	0	0	0	242	262	39	0	0	0	0	0	0	0	0	0	452	0	0	0	0	743	0,35
MR	419	111	106	3	0	101	2	322	324	517	852	87	0	28	18	324	162	84	316	410	158	114	938	19	4	0	1	28	4	11	0	476	0,34
VN	2	79	0	0	0	0	0	0	0	1	23	1	0	0	0	0	2	400	6	82	236	4	109	0	9	0	0	3	1	0	1	959	0,42
FRT	1	17	14	0	0	4	0	511	338	142	176	27	0	0	0	1	143	92	165	141	927	185	393	640	23	0	0	8	0	3	23	673	0,25
CRL	96	73	157	0	0	24	0	0	0	0	31	15	0	0	0	0	0	4	100	741	497	131	73	65	12	0	0	2	50	6	12	876	0,85
CT	754	163	589	19	4	12	1	13	1	0	33	39	13	11	0	37	14	98	29	924	124	56	192	16	187	0	0	5	41	2	29	255	0,49
GR	0	63	2	0	0	2	1	32	0	1	68	0	0	0	0	0	2	111	18	132	66	260	83	8	21	0	0	47	0	8	14	939	0,28
OL	2	62	3	0	0	0	0	11	2	69	532	7	0	0	0	0	5	14	151	182	28	24	357	9	2	0	0	3	1	0	0	467	0,76
MA	0	4	11	0	0	0	0	0	0	0	0	0	0	0	0	1	8	5	0	71	140	12	1	43	2	0	0	0	0	1	299	0,14	
VE	19	26	3	0	0	0	0	6	0	1	2	2	0	0	0	12	2	4	11	452	445	156	13	8	13	0	0	0	3	8	2	118	0,01
NG	0	0	0	0	0	0	0	102	5	10	22	0	0	0	0	0	0	1	26	35	14	15	17	1	2	0	0	1	0	0	0	251	0
CW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	148	0	0	0	0	0	0	0	0	0	0	0	667	0	0	0	668	1
KW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TBC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RCF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SFL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
su	637	455	708	431	188	434	15	697	777	180	737	219	169	152	159	215	191	872	528	244	171	110	156	113	301	82	730	196	127	42	97	kap	0,59
UA	0,34	0,42	0,64	0,57	0,49	0,66	0,67	0,89	0,48	0,8	0,53	0,01	0,03	0,65	0,83	0,12	0,85	0,46	0,31	0,3	0,73	0,24	0,23	0,04	0,04	0	0,91	0	0	0	0	OA:	0,64

Όπως βλέπουμε στον παραπάνω πίνακα σύγχυσης η συνολική ακρίβεια είναι αρκετά ικανοποιητική καθώς ξεπερνά το 64%. Ακόμη, υπολογίστηκε ο δείκτης ακρίβειας KHAT (ή δείκτης συμφωνίας KAPPA), ο οποίος βρέθηκε ίσος με 59%. Συνεπώς, διαπιστώνουμε ότι η ταξινόμηση, σε ένα πείραμα στο οποίο τα δεδομένα εκπαίδευσης σε σχέση με τα δεδομένα ελέγχου προέρχονται από διαφορετικά tiles απέφερε θετικά αποτελέσματα.

Εκτός από τα μεγέθη που αφορούν την συνολική ακρίβεια της ταξινόμησης, σκόπιμο είναι να διαπιστώσουμε αν και κατά πόσο υπάρχουν σφάλματα ταξινόμησης αλλά και λανθασμένα ταξινομημένα εικονοστοιχεία για την κάθε κατηγορία ξεχωριστά. Μέσα από αυτόν τον έλεγχο θα καταφέρουμε να εντοπίσουμε κατηγορίες προβληματικές και να αντιληφθούμε λάθη και παραλείψεις τις διαδικασίας.

Πιο συγκεκριμένα, για τις κλάσεις που ανήκουν στην γενική κατηγορία των τεχνητών επιφανειών, βλέπουμε ότι υπάρχει μία μικρή δυσκολία του ταξινομητή να διαχωρίσει πλήρως τις κλάσεις DUF, SUF, RAN και ICU. Αυτή η δυσκολία ήταν αναμενόμενη καθώς, η κατηγορία DUF και SUF, είναι υποκατηγορίες του αστικού ιστού και άρα εμπρικλείουν ίδιου τύπου αντικείμενα που εκπέμπουν με παρόμοιο τρόπο στο ηλεκτρομαγνητικό φάσμα, ενώ η κατηγορία ICU, αποτελείται από περιοχές με τεχνητές επιφάνειες από τσιμέντο και άσφαλο, υλικά που εκπέμπουν παρόμοια με τις κατηγορίες DUF, SUF και RAN (Ασφαλτοστρωμένες περιοχές και δίκτυα μεταφοράς). Από τις συγκεκριμένες κατηγορίες παρατηρούμε ότι εκείνη με τα μικρότερα σφάλματα παράλειψης και συμπερίληψης είναι η RAN. Ενδιαφέρον παρουσιάζει το γεγονός ότι η DUF εμφανίζει ποσοστιαίο σφάλμα συμπερίληψης 11,83% με την κατηγορία CTN, ενώ στη SUF το σφάλμα συμπερίληψης αγγίζει το 36% σε σχέση με την κατηγορία CTN. Σε σχέση με τα σφάλματα παράλειψης της SUF εντοπίζεται ότι περί το 15% της κατηγορίας παραλείφθηκε από τον αλγόριθμο και ταξινομήθηκε στην CRL. Ακόμη αξιο αναφοράς κρίνεται το γεγονός ότι ο αλγόριθμος παρέλειψε 15% εικονοστοιχεία της κατηγορίας RAN και τα ταξινόμησε στην κατηγορία MRS.

Σε σχέση με την κατηγορία MES, διαπιστώνεται ότι η κατηγορία MES μπλέκεται έντονα με την κατηγορία ICU, καθώς το ποσοστό συμπερίληψής τους είναι αντίστοιχα 25,7% και 20,8%. Ακόμη η κατηγορία MES, εμπλέκεται με την κατηγορία SND με ποσοστό συμπερίληψης 15,7%. Αυτές οι συγχύσεις είναι αναμενόμενες, καθώς οι κατηγορίες ICU, MES και SND, έχουν παρόμοια φασματικά χαρακτηριστικά.

Τέλος, η κατηγορία PHT, εμφανίζει έντονο σφάλμα παράλειψης σε σχέση με την κατηγορίες RAN (30%) και μικρότερο με την κατηγορία DUF (10,30%), ενώ αξιοσημείωτο είναι το σφάλμα παράλειψης ίσο με 19,4%, που εμφανίζει σε σχέση με την κατηγορία CRL. Η κατηγορία GRN ταξινομείται ικανοποιητικά από τον αλγόριθμο, όμως εμφανίζει μεγάλο σφάλμα παράλειψης με την κατηγορία ICU (29,20%), την κατηγορία DUF (18%) και με την κατηγορία MES (12%). Οι κατηγορίες PHT και GRN, διαθέτον πενιχρό GT καθώς δεν απαντώνται ευρέως στην περιοχή μελέτης. Συνεπώς τα αποτελέσματα του ταξινομητή κρίνονται ιδιαίτερα ικανοποιητικά, συνυπολογίζοντας τα λιγοστά δεδομένα εκπαίδευσης.

Εν γένει η ακρίβεια παραγωγής για όλες τις κλάσεις των τεχνητών επιφανειών είναι ικανοποιητική καθώς κυμαίνεται από 52-72%. Αντίθετα η αξιοπιστία- ακρίβεια χρήστη

εμφανίζει χαμηλότερα ποσοστά, τα οποία κυμαίνονται από 34-89%. Οι κλάσεις με την μικρότερη αξιοπιστία είναι η DUF με ποσοστό 34% και η SUF με ποσοστό 42%.

Οι κατηγορίες φυσικής βλάστησης εμφανίζουν υψηλά ποσοστά ακρίβειας και αξιοπιστίας. Πιο συγκεκριμένα, η κατηγορία BLF, ταξινομείται επιτυχημένα, καθώς χαρακτηρίζεται από ακρίβεια που αγγίζει το 71% και αξιοπιστία ανώτερη του 89%, ενώ η μόνη σημαντική σύγχυση που παρουσιάζει είναι με την κατηγορία CNF, καθώς το 28% των pixel του ground truth δεν ταξινομούνται στην συγκεκριμένη κατηγορία αλλά στην CNF. Η συγκεκριμένη σύγχυση είναι εν μέρει αναμενόμενη, ιδίως αν συνυπολογίσουμε τους παράγοντες τόπου και τις μεταβολές στα χαρακτηριστικά της φυτοκάλυψης από περιοχή σε περιοχή.

Η κατηγορία CNF, εμφανίζει υψηλή ακρίβεια παραγωγής 87,56%. Όμως η ακρίβεια χρήστη είναι κάτω του μετρίου με σκορ ίσο με 47,96%, γεγονός που προκύπτει από το έντονο σφάλμα συμπερίληψης που εμφανίζεται με την κατηγορία BLF (περίπου 32%) και από το μικρότερο σφάλμα συμπερίληψης που εμφανίζει με την κατηγορία DSV. Η σύγχυση με την συγκεκριμένη κατηγορία είναι αναμενόμενη καθώς οι δύο κατηγορίες είναι παρόμοιας φύσης, και η δυσκολία να διακριθούν τόσο από τον ταξινομητή όσο και από τον χρήστη κατά τη συλλογή του ground truth είναι ιδιαίτερα αυξημένη. Αντίστοιχα η κατηγορία DSV χαρακτηρίζεται από μέτρια ακρίβεια παραγωγής (52,64%), καθώς το σφάλμα παράλειψης ως προς την κατηγορία CNF είναι ιδιαίτερα αυξημένο (39,89%), ενώ η ακρίβεια χρήστη είναι ιδιαίτερα ικανοποιητική (80%), καθώς δεν εμφανίζονται σημαντικά σφάλματα συμπερίληψης πέρα από 10,5% των pixel της κατηγορίας CNF που συμπεριλαμβάνονται στην DSV.

Η κατηγορία SPSV, ταξινομείται ικανοποιητικά, με ποσοστό ακρίβειας παραγωγής ίσο με 60,40% και ακρίβεια χρήστη ίσο με 52,83%. Τα σφάλματα που εντοπίζονται είναι ως προς την κλάση DSV (σφάλμα παράλειψης: 13% και συμπερίληψης: 10%), μία σύγχυση αναμενόμενη καθώς και οι δύο κλάσεις αποτελούν υποκατηγορίες της κλάσης σκληροφυλλικής βλάστησης, αλλά και ως προς τις κλάσεις OLG (σφάλμα παράλειψης 16% και συμπερίληψης 7%) και MRS (ποσοστό συμπερίληψης 11%). Τέλος η κατηγορία NGR δεν έχει διαχωριστεί από την κατηγορία BLF, καθώς 40% του πλήθους των εικονοστοιχείων της ταξινομήθηκε εσφαλμένα σ' αυτήν, ενώ το σφάλμα συμπερίληψης ως προς την κατηγορία αυτή ξεπέρασε το 98%.

Σε σχέση με τις κατηγορίες του γυμνού εδάφους, η μόνη κατηγορία που ταξινομήθηκε επιτυχώς ήταν η SND, η οποία επέτυχε ποσοστά ακρίβειας και αξιοπιστίας 63,67% και 65,33% αντίστοιχα. Οι συγχύσεις που εμφανίζει είναι με τις κλάσεις RAN, ICU, MES με ποσοστά παράλειψης και συμπερίληψης έως 25% με την κλάση ICU. Οι συγχύσεις αυτές είναι αναμενόμενες, καθώς εικονοστοιχεία άμμου απαντώνται στις συγκεκριμένες κλάσεις.

Ως προς τις κατηγορίες SVA και RCK, ο ταξινομητής δεν κατάφερε να πετύχει υψηλά ποσοστά ακρίβειας και αξιοπιστίας. (1-11%). Η αποτυχία του αλγορίθμου να ταξινομήσει τις συγκεκριμένες κλάσεις είναι αποτέλεσμα του μικρού αριθμού δεδομένων εκπαίδευσης, καθώς οι συγκεκριμένες κατηγορίες δεν απαντώνται ευρέως στα συγκεκριμένα tiles.

Στην γενική κατηγορία των υδάτινων πόρων διαπιστώνεται ότι υπάρχει μεγάλη διαφοροποίηση ως προς τα ποσοστά ακρίβειας και αξιοπιστίας κάθε κλάσης. Η ακρίβεια παραγωγής για τις τρεις κλάσεις WCR, WBD και MRS κυμαίνονται χαμηλά αφού είναι ίσα με 6%, 35% και 34% αντιστοίχως. Αντίθετα, η αξιοπιστία εμφανίζει πολύ υψηλότερα ποσοστά για τις κλάσεις WCR και MRS (83 και 85% αντιστοίχως), όχι όμως και για την κατηγορία WBD που η αξιοπιστία της είναι ιδιαίτερα χαμηλή (12%). Πιο συγκεκριμένα, η έντονη σύγχυση που παρατηρείται είναι ανάμεσα στις κλάσεις WCR και WBD, καθώς 71% των ρixelς της κατηγορίας WCR ταξινομούνται στην WBD και το αντίστροφο. Η κατηγορία MRS εμφανίζει μικρές αλλά ποικίλες συγχύσεις με τις κλάσεις των δασών, του νερού και ορισμένων καλλιεργειών (Φρούτα, Δημητριακά, Βαμβάκι). Το υψηλότερο σφάλμα που παρουσιάζει είναι σφάλμα συμπερίληψης ίσο με 14% με την κατηγορία WCR, μία σύγχυση αναμενόμενη καθώς ρixelς νερού απαντώνται στα δεδομένα εκπαίδευσης της MRS κατηγορίας. Αντίθετα, η κατηγορία CWT, είναι η κατηγορία εκείνη που ο ταξινομητής πέτυχε να διαχωρίσει άριστα, καθώς η ακρίβεια και η αξιοπιστία ξεπέρασαν το 90%.

Ως προς τις κλάσεις που ανήκουν στις καλλιέργειες διαπιστώνουμε ότι, υπάρχει δυσκολία του ταξινομητή να διαχωρίσει τις κλάσεις. Πιο συγκεκριμένα, η κατηγορία VNY, διαθέτει σχετικά μικρό αριθμό ρixelς ground truth και παρόλο που πρόκειται για μία κατηγορία εύκολα διαχωρίσιμη από τις υπόλοιπες, ο ταξινομητής πετυχαίνει ακρίβεια και αξιοπιστία κατηγορίας κάτω του μετρίου (41,71% και 45,87%). Παρουσιάζονται συγχύσεις μικρότερες με την κατηγορία MRS και μεγαλύτερες με διάφορες κατηγορίες καλλιεργειών, όπως την κατηγορία CRL (σφάλμα παράλειψης 8,5%), την κατηγορία OLG (σφάλμα παράλειψης 11,3%) την κατηγορία GRF (σφάλμα συμπερίληψης 12,8%) αλλά και την κατηγορία CTN, στην οποία το σφάλμα συμπερίληψης είναι ίσο με 11,2% όμως το σφάλμα παράλειψης ξεπερνά το 24%.

Η κλάση FRT, εμφανίζει χαμηλά ποσοστά ακρίβειας (24,56%) και αξιοπιστίας (31,33%), καθώς ο ταξινομητής δεν καταφέρνει να τη διαχωρίσει επαρκώς από τις υπόλοιπες κλάσεις των καλλιεργειών και ιδίως από την CRL, την CTN και την MAZ. Αξιοσημείωτο είναι το γεγονός ότι, η συγκεκριμένη κατηγορία εμφανίζει μεγάλο σφάλμα συμπερίληψης ως προς την κατηγορία των MRS, καθώς περί το 60% των ρixel της κατηγορίας MRS ταξινομείται λανθασμένα στην FRT.

Η κατηγορία CRL εμφανίζει πολύ υψηλό ποσοστό ακρίβειας ταξινόμησης. Δεν εμφανίζει όμως και μεγάλο ποσοστό αξιοπιστίας αξιοπιστίας, καθώς αυτό περιορίζεται στο 30%. Αυτό οφείλεται στο υψηλό σφάλμα συμπερίληψης (37,8%) ως προς την κατηγορία CTN. Ακόμα, αξιοσημείωτο είναι το γεγονός ότι και αυτή η κατηγορία, εμφανίζει σημαντική σύγχυση με την κατηγορία MRS, καθώς 16,78% του ground truth της κατηγορίας MRS ταξινομείται εσφαλμένως στη κατηγορία CRL.

Η κατηγορία CTN χαρακτηρίζεται από μέτρια ακρίβεια παραγωγής (48,78%), εμφανίζει όμως υψηλό ποσοστό ακρίβειας χρήστη (72,78%). Η κυριότερη σύγχυση που παρατηρείται είναι με την κατηγορία CTN, καθώς 36,2% των εικονοστοιχείων της κατηγορίας CTN παραλείπονται και ταξινομούνται στην κατηγορία CRL. Άλλη σημαντική σύγχυση δεν παρατηρείται. Αξίζει όμως να σημειωθεί ότι 9% των εικονοστοιχείων της κατηγορίας MRS, συμπεριλαμβάνονται στην κλάση CTN.

Η κατηγορία GRF, παρουσιάζει χαμηλά ποσοστά ακρίβειας και αξιοπιστίας (27,69% και 23,55% αντιστοίχως). Η συγκεκριμένη κλάση εμπλέκεται σε μικρό βαθμό με τις περισσότερες κλάσεις καλλιεργειών και συγκεκριμένα με τους αμπελώνες, τα οπωροφόρα δένδρα, τα δημητριακά, το βαμβάκι και τους ελαιώνες.

Η κατηγορία OLG, εμφανίζει υψηλό ποσοστό ακρίβειας παραγωγής, το οποίο ισούται με 76,33%. Ελέγχοντας τα σφάλματα παράλειψης διαπιστώνουμε ότι πάνω από 11% του GT της κατηγορίας αυτής ταξινομείται εσφαλμένα στην κατηγορία SPSV. Αυτή η σύγχυση είναι ως έναν βαθμό αναμενόμενη καθώς οι συγκριμένες κλάσεις εμφανίζουν παρόμοια φασματικά χαρακτηριστικά. Ελέγχοντας την αξιοπιστία της κλάσης OLG και τα σφάλματα συμπερίληψης, διαπιστώνουμε ότι η αξιοπιστία της κατηγορίας αυτής είναι ιδιαίτερα χαμηλή (22,78%), καθώς περίπου 60% που ground truth της κλάσης MRS έχουν εσφαλμένα ταξινομηθεί ως OLG.

Η κατηγορία MAZ είναι μία κατηγορία καλλιεργειών την οποία ο ταξινομητής δεν καταφέρνει να διαχωρίσει. Τα ποσοστά ακρίβειας παραγωγού και χρήστη είναι ιδιαίτερα χαμηλά (14,38% και 3,80% αντιστοίχως), ενώ τα πολύ υψηλά σφάλματα συμπερίληψης μαρτυρούν έντονη σύγχυση με τις παραπλήσιες φασματικά κατηγορίες του βαμβακιού και των δημητριακών.

Η κατηγορία VEG είναι άλλη μία κατηγορία καλλιεργειών, που ο ταξινομητής απέτυχε να ταξινομήσει. Πιο συγκεκριμένα, τα ποσοστά ακρίβειας και αξιοπιστίας είναι ιδιαίτερα χαμηλά (1% και 4,3% αντιστοίχως), ενώ κοιτώντας τα σφάλματα παράλειψης και συμπερίληψης διαπιστώνουμε ότι τα pixels της κατηγορίας αυτής εμπλέκονται έντονα με αυτά της κατηγορίας CTN και CRL. Αξιοσημείωτο είναι το γεγονός ότι τα εικονοστοιχεία του ground truth δεν ήταν πολλαπλά.

Συμπερασματικά λοιπόν ως προς τις κατηγορίες των καλλιεργειών, διαπιστώνουμε ότι εμφανίζονται συγχύσεις μεταξύ τους με κυριότερες εκείνες ανάμεσα στις μονοετείς καλλιέργειες. Οι συγχύσεις αυτές είναι αναμενόμενες, καθώς πρόκειται για κατηγορίες με παρόμοια φασματικά χαρακτηριστικά. Εκτός αυτού, η διαδικασία εκπαίδευσης του ταξινομητή σε ένα tile και ελέγχου σε άλλο, είναι πρακτική που μπορεί να προκαλέσει σύγχυση στην ταξινόμηση των κλάσεων, λόγω μεταβολών των χαρακτηριστικών των κατηγοριών από περιοχή σε περιοχή, ιδίως σε κλάσεις καλλιεργειών, τα χαρακτηριστικά των οποίων και η χρονική μεταβολή τους εξαρτάται σε μεγάλο βαθμό από τις κλιματικές συνθήκες της περιοχής μελέτης. Αξιοσημείωτο είναι το γεγονός ότι η κατηγορία βάλτη-έλη εμφανίζει συγχύσεις με τις κατηγορίες των καλλιεργειών.

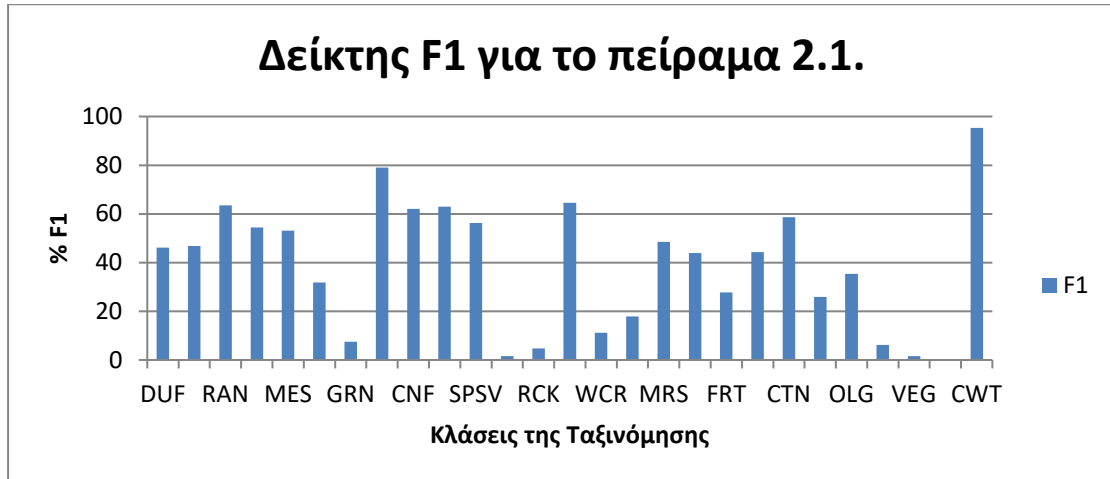
Πίνακας 5-12 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF

	DUF	SUF	RAN	ICU	MES	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MR	VNY	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL
DUF	71,7	20,1	10,4	23,3	4,97	10,3	18,1	0,00	0,00	0,00	0,00	0,00	6,82	1,85	0,06	0,00	0,88	0,21	0,01	1,10	2,95	0,00	0,04	0,00	1,60	0,00	0,00	0,00	0,00	0,00	0,00
SUF	7,48	53,2	1,31	2,94	1,08	3,63	3,98	0,00	0,00	0,00	0,12	11,8	0,00	0,64	0,00	0,00	0,23	8,24	0,25	0,83	6,41	6,71	1,33	1,34	2,19	0,00	0,00	1,00	6,00	0,00	3,00
RAN	1,60	0,28	63,4	2,71	0,45	30,0	7,52	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,28	0,08	2,24	0,00	0,21	1,79	2,31	0,21	0,06	3,68	0,25	0,00	0,00	0,00	8,00	0,00	2,00
ICU	13,4	0,92	12,5	52,2	15,9	0,67	29,2	0,00	0,00	0,00	0,00	0,00	4,55	9,50	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	4,00	0,00	2,00
MES	1,24	0,00	1,16	10,2	58,4	2,30	12,3	0,00	0,00	0,00	0,00	2,63	0,00	18,8	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2,00	0,00	0,00
PHT	0,00	0,00	0,04	0,00	0,00	21,2	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,21	0,00	0,06	0,27	0,05	0,21	0,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	8,00
GR	0,00	0,00	0,00	0,00	0,00	0,00	4,42	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,00	0,00	0,00	0,00	0,00	5,00	0,00	0,00
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	70,9	7,18	1,30	3,49	0,00	0,00	0,00	0,01	0,00	6,77	0,00	7,59	0,00	0,05	3,41	0,24	0,00	0,51	40,6	0,00	35,0	0,00	3,00	0,00
CNF	0,00	0,00	0,08	0,00	0,00	0,00	0,00	28,3	87,5	39,8	4,82	5,26	0,00	0,00	2,37	0,01	6,80	0,00	5,02	0,00	0,00	0,00	0,04	0,00	0,00	1,99	0,00	0,00	0,00	0,00	0,00
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,14	4,44	51,6	13,0	0,00	0,00	0,00	0,04	0,00	1,08	0,10	2,11	0,00	0,00	0,11	1,47	0,00	0,08	3,98	0,00	42,0	0,00	1,00	0,00
SPS	0,00	0,36	0,33	0,15	0,25	0,82	1,33	0,03	0,20	5,44	60,4	50,0	20,4	0,00	0,00	0,00	1,79	2,40	2,61	0,35	0,13	7,24	11,3	0,00	0,17	8,76	0,00	17,0	0,00	0,00	0,00
SVA	0,00	0,39	0,01	0,21	0,19	0,44	0,00	0,00	0,00	0,00	0,00	3,95	4,55	0,00	0,01	0,00	0,18	0,10	0,40	0,17	0,15	0,00	0,15	0,00	0,17	0,00	0,00	1,00	1,00	0,00	0,00
RCK	0,13	0,06	0,21	0,51	2,36	0,30	0,44	0,00	0,00	0,00	0,00	0,00	11,3	4,08	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00
SND	0,78	0,14	1,80	4,31	7,07	0,30	3,54	0,00	0,00	0,00	0,00	0,00	13,6	63,6	0,00	0,00	0,06	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	6,04	3,26	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00
WB	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	11,9	35,3	6,80	0,01	0,00	0,14	0,00	0,00	0,00	0,33	1,01	0,00	0,22	0,00	0,00	0,00	0,00
MR	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,37	2,63	0,00	0,00	0,00	0,52	34,0	0,21	2,12	0,00	0,05	0,21	0,11	2,68	0,17	0,00	0,00	28,0	4,00	11,0	0,00
VNY	0,00	0,42	0,22	0,02	0,06	0,07	1,77	0,00	0,00	0,00	0,28	0,00	2,27	0,06	0,00	0,00	0,18	41,7	1,37	0,05	0,38	11,8	0,30	1,67	0,34	0,40	0,00	3,00	1,00	0,00	1,00
FRT	0,00	0,25	0,18	0,02	0,00	0,22	0,00	0,02	0,12	0,02	0,22	3,95	0,00	0,00	0,00	0,00	6,64	0,63	24,5	1,14	0,11	1,92	3,23	0,00	0,93	10,3	0,00	8,00	0,00	3,00	23,0
CRL	2,58	14,8	2,69	2,26	5,86	19,4	9,29	0,00	0,00	0,00	0,11	10,5	27,2	0,57	0,00	0,00	8,61	8,55	20,9	84,6	36,2	14,0	3,89	23,7	38,0	13,9	0,00	2,00	50,0	6,00	12,0
CTN	0,75	5,83	4,38	0,70	3,12	5,19	3,98	0,00	0,00	0,00	0,03	1,32	2,27	0,70	0,01	0,00	3,31	24,6	13,7	5,67	48,7	7,03	0,60	46,8	37,4	5,58	0,00	5,00	41,0	2,00	29,0
GRF	0,10	0,33	0,08	0,17	0,00	3,04	0,00	0,04	0,02	0,03	0,43	0,00	0,00	0,00	0,00	0,00	0,24	0,42	2,75	1,49	0,22	27,6	0,51	4,01	13,1	5,98	0,00	47,0	0,00	8,00	14,0
OLG	0,10	2,00	0,92	0,06	0,06	0,22	1,33	0,01	0,43	1,48	16,3	3,95	4,55	0,13	0,07	0,00	19,6	11,3	5,83	0,83	0,75	8,84	76,3	0,33	1,09	6,77	0,00	3,00	1,00	0,00	0,00
MA	0,00	0,06	0,03	0,00	0,00	0,37	0,00	0,33	0,03	0,03	0,08	1,32	0,00	0,00	0,00	0,00	0,04	0,00	9,50	0,74	0,06	0,85	0,19	14,3	0,67	0,40	0,00	0,00	0,00	0,00	1,00
VEG	0,00	0,39	0,00	0,06	0,00	0,59	0,44	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,94	0,34	0,14	0,73	2,24	0,04	0,67	1,09	0,80	0,00	3,00	8,00	2,00	
NG	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	
CW	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	8,28	60,8	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,7	0,00	0,00	0,00	0,00
KW	0,00	0,03	0,00	0,00	0,00	0,22	0,00	0,04	0,00	0,15	0,26	1,32	0,00	0,00	0,00	0,00	0,06	0,31	0,12	0,02	0,02	5,01	0,06	0,00	0,00	0,40	0,00	0,00	0,00	0,00	
TBC	0,00	0,17	0,11	0,08	0,13	0,00	2,21	0,00	0,00	0,00	0,00	1,32	2,27	0,00	0,00	0,00	0,01	0,10	0,00	0,57	0,16	0,00	0,02	0,00	0,25	0,00	0,00	0,00	0,00	0,00	
RCF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,04	0,07	0,01	0,85	0,00	0,00	0,67	0,00	0,00	0,00	0,00	0,00	
SFL	0,00	0,08	0,03	0,04	0,00	0,59	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,34	0,14	0,11	1,49	0,00	0,33	0,17	0,00	0,00	0,00	0,00	0,00	

Πίνακας 5-13 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF

	DUF	SUF	RAN	ICU	MES	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MR	VNY	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL	
DUF	34,5	5,03	0,69	9,56	2,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2,37	1,57	0,00	0,00	0,00	0,00	0,00	0,32	0,13	0,27	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SUF	11,3	42,1	0,14	0,77	0,00	0,00	0,00	0,00	0,00	0,00	0,18	6,39	1,18	0,33	0,00	0,00	0,00	1,72	0,17	2,19	1,23	1,09	0,46	0,18	4,65	0,00	0,00	0,51	4,72	0,00	3,09	
RAN	11,7	2,06	64,2	20,8	4,41	0,69	0,00	0,00	0,01	0,00	0,33	0,46	8,88	8,44	0,00	0,00	0,01	1,83	0,25	0,79	1,83	0,54	0,42	0,18	0,00	0,00	0,00	0,00	6,30	0,00	2,06	
ICU	17,3	3,05	1,81	57,2	25,6	0,00	0,00	0,00	0,00	0,00	0,09	4,57	14,2	13,3	0,00	0,00	0,00	0,11	0,02	0,44	0,19	0,72	0,02	0,00	1,00	0,00	0,00	0,00	3,15	0,00	2,06	
MES	1,22	0,37	0,10	5,80	48,7	0,00	0,00	0,00	0,00	0,00	0,05	1,37	21,8	7,26	0,00	0,00	0,00	0,11	0,00	0,38	0,29	0,00	0,01	0,00	0,00	0,00	0,00	0,00	1,57	0,00	0,00	
PHT	2,18	1,08	5,72	0,21	1,65	66,1	0,00	0,00	0,00	0,00	0,15	2,74	2,37	0,26	0,00	0,00	0,00	0,11	0,06	1,07	0,41	3,71	0,02	0,44	2,66	0,00	0,00	1,53	0,00	0,00	8,25	
GR	0,64	0,20	0,24	1,53	1,49	0,00	66,6	0,00	0,00	0,00	0,04	0,00	0,59	0,52	0,00	0,00	0,00	0,46	0,00	0,09	0,05	0,00	0,02	0,00	0,33	0,00	0,00	0,00	3,94	0,00	0,00	
BLF	0,00	0,00	0,01	0,00	0,00	0,00	6,67	89,1	32,0	0,66	0,38	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,34	0,00	0,00	2,99	0,04	25,4	0,00	98,7	0,00	17,8	0,00	7,14	0,00	
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	4,38	47,9	10,4	1,15	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,98	0,00	0,01	0,63	1,18	0,97	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	14,3	80,0	20,5	0,00	0,00	0,00	0,00	0,00	0,02	0,11	0,09	0,00	0,00	0,82	2,64	0,71	0,00	1,22	0,00	21,4	0,00	2,38	0,00	
SPS	0,00	0,18	0,01	0,00	0,00	0,23	0,00	0,32	0,40	4,66	52,8	0,00	0,00	0,00	0,00	0,00	0,13	2,06	0,27	0,03	0,01	2,54	6,72	0,44	0,00	0,00	0,00	8,67	0,00	0,00	0,00	
SVA	0,00	0,20	0,00	0,00	0,11	0,00	0,00	0,00	0,01	0,00	0,52	1,37	0,00	0,00	0,00	0,00	0,01	0,00	0,06	0,03	0,01	0,00	0,02	0,09	0,00	0,00	0,00	0,51	0,79	0,00	0,00	
RCK	0,05	0,00	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,12	0,91	2,96	0,39	0,00	0,00	0,00	0,11	0,00	0,05	0,01	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,79	0,00	0,00	
SND	0,45	0,22	0,00	3,46	15,6	0,00	0,00	0,00	0,00	0,00	0,00	0,00	37,8	65,3	0,00	0,00	0,00	0,11	0,00	0,04	0,06	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
WC	0,22	0,00	0,88	0,00	0,00	0,00	0,00	0,00	0,67	0,04	0,00	0,91	0,00	0,00	83,0	71,9	13,6	0,00	0,00	0,00	0,01	0,00	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
WB	0,00	0,00	0,08	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	15,2	12,1	0,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	6,19	0,00	0,00	0,00	
MR	6,57	2,44	15,0	0,07	0,00	23,2	13,3	4,63	4,17	2,87	11,5	39,7	0,00	1,83	1,13	15,0	85,0	9,63	59,9	16,7	9,23	10,3	59,9	1,68	1,33	0,00	0,00	14,2	3,15	26,1	0,00	
VNY	0,03	1,73	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,31	0,46	0,00	0,00	0,00	0,00	0,01	45,8	0,11	0,34	1,38	0,36	0,70	0,00	2,99	0,00	0,00	1,53	0,79	0,00	1,03	
FRT	0,02	0,37	0,20	0,00	0,00	0,92	0,00	0,73	0,43	0,79	2,39	12,3	0,00	0,00	0,00	0,00	0,75	10,5	31,3	5,78	5,42	16,7	2,51	56,6	7,64	0,00	0,00	4,08	0,00	7,14	23,7	
CRL	1,51	1,60	2,22	0,00	0,00	5,53	0,00	0,00	0,00	0,00	0,42	6,85	0,00	0,00	0,00	0,00	0,00	0,46	1,89	30,3	2,90	11,8	0,47	5,75	3,99	0,00	0,00	1,02	39,3	14,2	12,3	
CTN	11,8	35,9	8,31	0,44	0,21	2,76	6,67	0,02	0,00	0,00	0,45	17,8	7,69	0,72	0,00	0,17	0,07	11,2	0,55	37,8	72,7	5,07	1,23	1,42	62,1	0,00	0,00	2,55	32,2	4,76	29,9	
GRF	0,00	1,38	0,03	0,00	0,00	0,46	6,67	0,05	0,00	0,01	0,92	0,00	0,00	0,00	0,00	0,00	0,01	12,7	0,34	0,54	0,39	23,5	0,53	0,71	6,98	0,00	0,00	23,9	0,00	19,0	14,4	
OLG	0,03	1,36	0,04	0,00	0,00	0,00	0,00	0,02	0,00	0,38	7,21	3,20	0,00	0,00	0,00	0,00	0,03	1,61	2,86	0,74	0,16	2,17	22,7	0,80	0,66	0,00	0,00	1,53	0,79	0,00	0,00	
MA	0,00	0,09	0,16	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,57	0,00	0,29	0,82	1,09	0,01	3,81	0,66	0,00	0,00	0,00	0,00	0,00	1,03	
VEG	0,30	0,57	0,04	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,03	0,91	0,00	0,00	0,00	0,06	0,01	0,46	0,21	1,85	2,60	14,1	0,08	0,71	4,32	0,00	0,00	0,00	2,36	19,0	2,06	
NG	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15	0,01	0,06	0,30	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,49	0,14	0,08	1,36	0,11	0,09	0,66	0,00	0,00	0,51	0,00	0,00	0,00	
CW	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,57	0,69	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	91,3	0,00	0,00	0,00	0,00	
KW	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
TBC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
RCF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
SFL	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	

Στην συνέχεια παρατίθεται το γράφημα για τον δείκτη F (F1). Στο συγκεκριμένο γράφημα εμφανίζεται το ποσοστό του F1 για κάθε κλάση της ταξινόμησης. Με το γράφημα αυτό, μπορούμε και εποπτικά να εντοπίσουμε τις κατηγορίες εκείνες, που απέδωσαν καλύτερα, αλλά και τις κατηγορίες αυτές, που εμφάνισαν συγχύσεις.

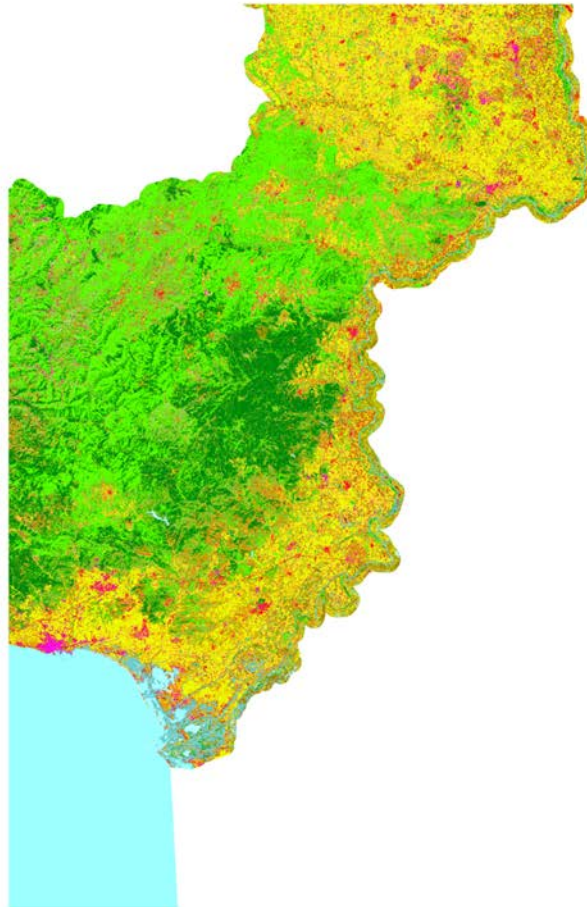


Γράφημα 5-10 Δείκτης F1 για την ταξινόμηση με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMF.

Όπως διαπιστώνουμε, η κατηγορία εκείνη με τον μέγιστο δείκτη F1, είναι η κατηγορία των παράκτιων υδάτων (CWT). Ακολουθεί η κατηγορία BLF και έπειτα οι κατηγορίες CNF, DSV και SPSV. Αντίθετα, οι κατηγορίες εκείνες, που ο ταξινομητής δυσκολεύτηκε να διαχωρίσει ήταν οι κατηγορίες εκείνες, που δεν απαντώνται ευρέως στο συγκεκριμένο tile και ο αλγόριθμος δυσκολεύεται να αναγνωρίσει. Αυτές είναι οι κατηγορίες GRN, SVA, RCK, MAZ, VEG και NGR.

Τα αποτελέσματα του δείκτη F1, επιβεβαιώνουν τις παρατηρήσεις, που προέκυψαν από τη μελέτη του πίνακα σύγχυσης.

Τέλος, παρατίθεται ο χάρτης, που παρήχθη από το πείραμα αυτό.



Legend

-  Dense Urban Fabric
-  Sparse Urban Fabric
-  Road and Rail Networks
-  Industrial or Commercial Units
-  Mineral Extraction Sites
-  Photovoltaic Panels
-  Greenhouses
-  Broadleaved Forest
-  Coniferous Forest
-  Dense Sclerophyllous Vegetation
-  Sparse Sclerophyllous Vegetation
-  Sparsely Vegated Areas
-  Bare Rocks
-  Beaches, Dunes, Sands
-  Water Courses
-  Water Bodies
-  Marshes
-  Vineyards
-  Fruit Trees and Berry Plantations
-  Cereals
-  Cotton
-  Grass Fodders
-  Olive Groves
-  Maize
-  Vegetables
-  Natural Grassland
-  Sea Coastal Water
-  Kiwi Plants
-  Tobacco
-  Rice Fields
-  Sunflowers

Χάρτης 5.5 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMF

Μελετώντας τον χάρτη, διαπιστώνουμε ότι σε γενικές γραμμές δίνει μία αξιόπιστη πρόβλεψη των χρήσεων/ καλύψεων γης και καλλιεργειών της περιοχής μελέτης. Οι καλύψεις γης, που επικρατούν στην περιοχή, προβλέπονται σωστά από τον αλγόριθμο RF, καθώς οι εδαφοκαλύψεις μεγάλου εμβαδού αναγνωρίζονται σωστά από τον ταξινομητή. Παρατηρούμε, λοιπόν, ότι οι αστικές, οι δασικές περιοχές, τα νερά, αλλά και οι διάφορες καλλιέργειες διαχωρίζονται αξιόπιστα από τον ταξινομητή.

Εάν παρατηρήσουμε τον χάρτη πιο λεπτομερώς και σε μεγαλύτερο zoom, θα διαπιστώσουμε ότι εμφανίζονται αρκετές συγχύσεις.

Για την γενική κατηγορία τεχνητών επιφανειών ,εμφανίζονται συγχύσεις ανάμεσα στις υποκατηγορίες της. Όπως είχε ήδη διαπιστωθεί κατά την ανάλυση του πίνακα σύγχυσης, η κατηγορία DUF, παρουσιάζει έντονη σύγχυση με την κατηγορία SUF, αλλά και με την κατηγορία ICU. Αυτό διαπιστώνεται και από την παρατήρηση του εικονιστικού αποτελέσματος, καθώς, σε περιοχές που θεωρούνται αμιγώς πυκνός αστικός ιστός, ή αμιγώς αραιός αστικός ιστός, οι δύο εδαφοκαλύψεις συνυπάρχουν



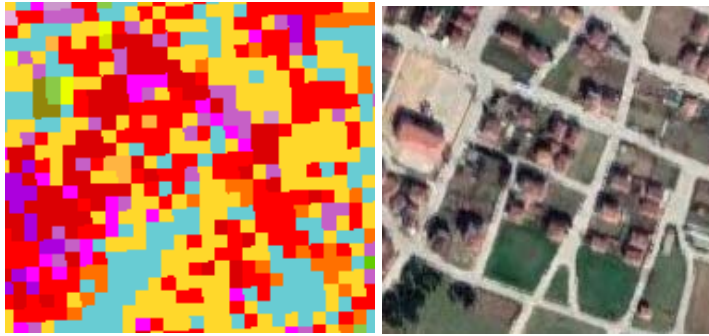
Εικόνα 5-14 Σύγχυση ανάμεσα στην DUF και την SUF.

Επιπλέον στις περιοχές οι οποίες θεωρούνται αμιγώς πυκνός αστικός ιστός ή αραιός αστικός ιστός, απαντώνται εικονοστοιχεία τα οποία ο ταξινομητής έχει αποδώσει στην κλάση ICU, στην κλάση MES αλλά και στην κλάση MRS. Οι συγχύσεις αυτές διαφαίνονταν από τα σφάλματα συμπερίληψης και επιβεβαιώνονται στο εικονιστικό αποτέλεσμα.



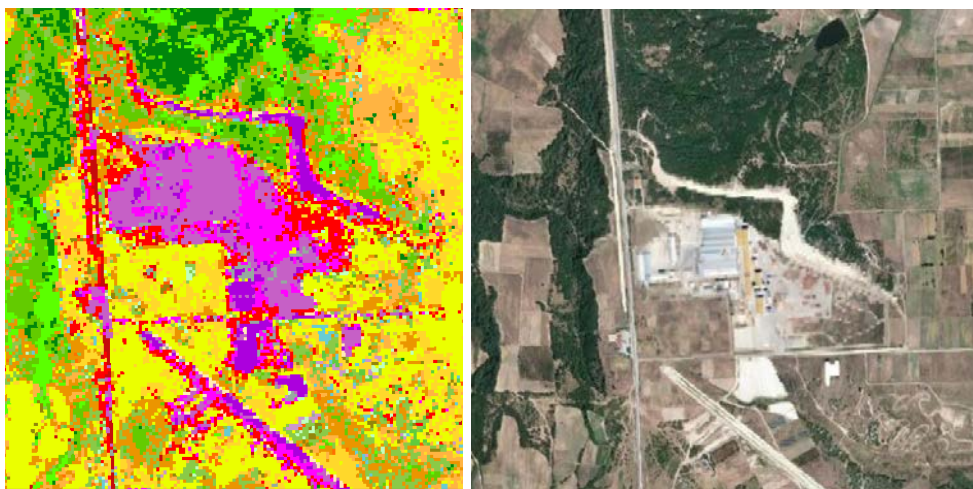
Εικόνα 5-15 Σύγχυση ανάμεσα στις DUF, SUF, RAN, ICU, MES, MRS

Ακόμη, εικονοστοιχεία τα οποία ανήκουν στην κλάση SUF, προβλέπονται ως εικονοστοιχεία της CTN, γεγονός, που διαφαίνεται και από τον πίνακα σύγχυσης.



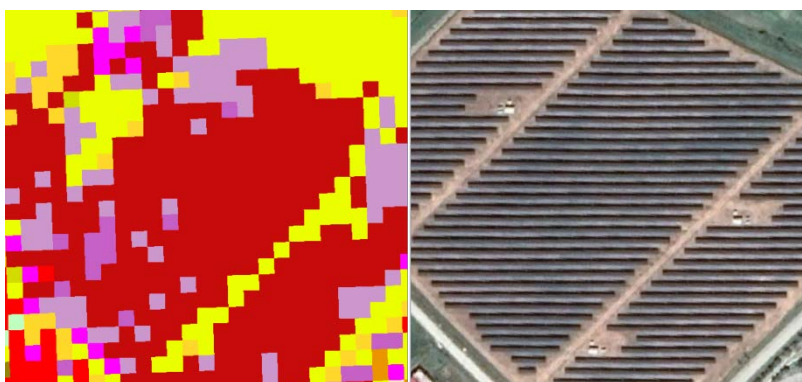
Εικόνα 5-16 Σύγκριση ανάμεσα στις DUF, SUF και CTN.

Οι κλάσεις ICU, MES και RAN διαχωρίζονται αρκετά ικανοποιητικά, όπως φαίνεται και στην παρακάτω εικόνα.



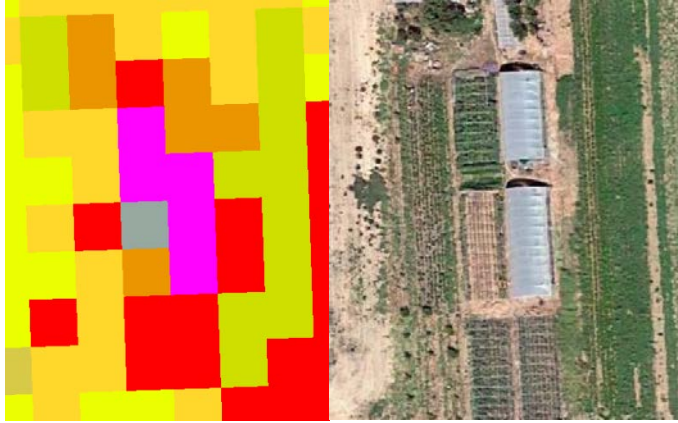
Εικόνα 5-17 Επιτυχής Ταξινόμηση ICU και MES.

Η κατηγορία PHT εμφανίζει έντονες συγχύσεις με την κατηγορία RAN, καθώς εικονοστοιχεία της παραλείπονται και ταξινομούνται ως RAN.



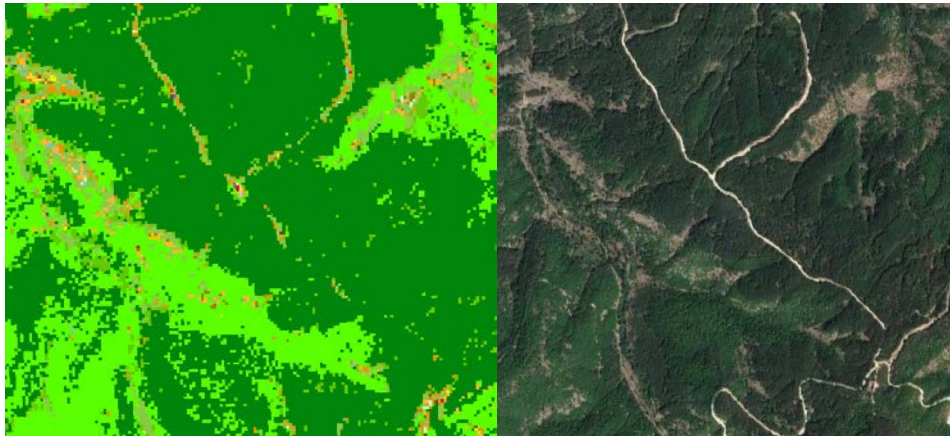
Εικόνα 5-18 Συγχύσεις ανάμεσα στην κλάση PHT και στην RAN

Τέλος, η κλάση GRN δεν προβλέπεται σωστά από τον αλγόριθμο, καθώς τα εικονοστοιχεία της αποδίδονται σε άλλες κλάσεις των τεχνητών επιφανειών.



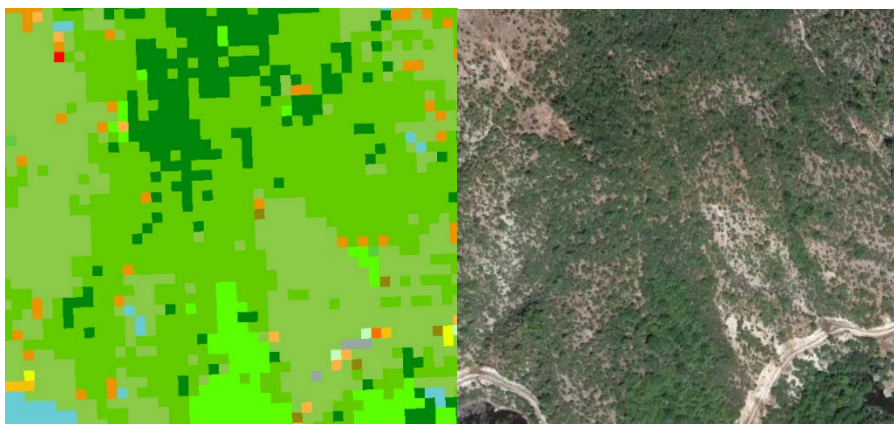
Εικόνα 6.6: Αδυναμία πρόβλεψης της GRN.

Οι κατηγορίες της φυσικής βλάστησης φαίνεται ότι προβλέπονται σε μεγάλο βαθμό από τον αλγόριθμο. Συγχύσεις, εμφανίζονται ανάμεσα στις κλάσεις BLF και CNF.



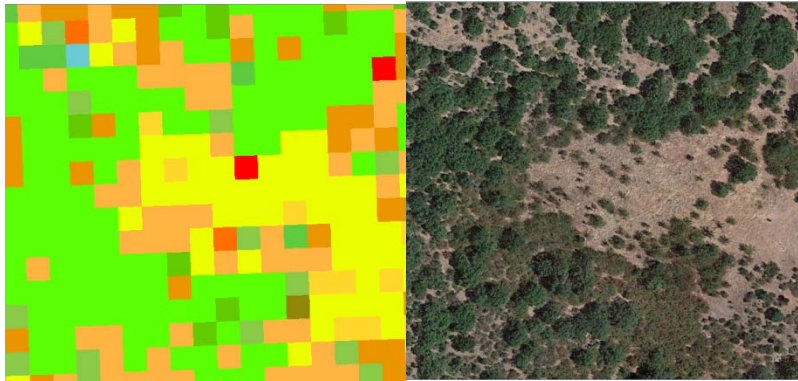
Εικόνα 6.7: Μικρή σύγχυση ανάμεσα στην BLF και την CNF.

Ακόμη, επιβεβαιώνοντας τα αποτελέσματα του πίνακα σύγχυσης, διαπιστώνουμε μικρές συγχύσεις ανάμεσα στις κλάσεις CNF και DSV, αλλά και στις κλάσεις DSV και SPSV.



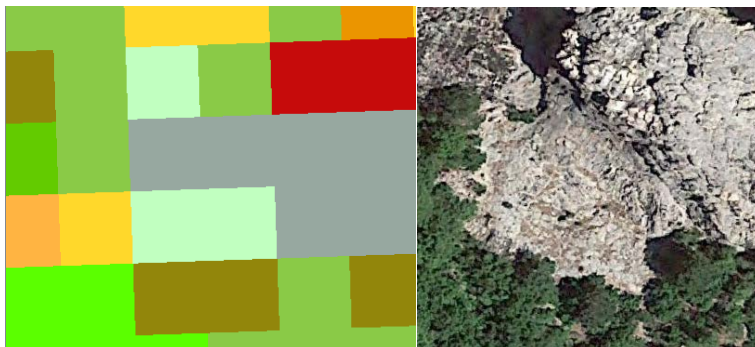
Εικόνα 6.8: Μικρή σύγχυση ανάμεσα στην DSV, SPSV, CNF.

Η κατηγορία εκείνη η οποία δυσκολεύει τον ταξινομητή στην αναγνώριση και πρόβλεψή της είναι εκείνη των φυσικών βοσκότοπων.



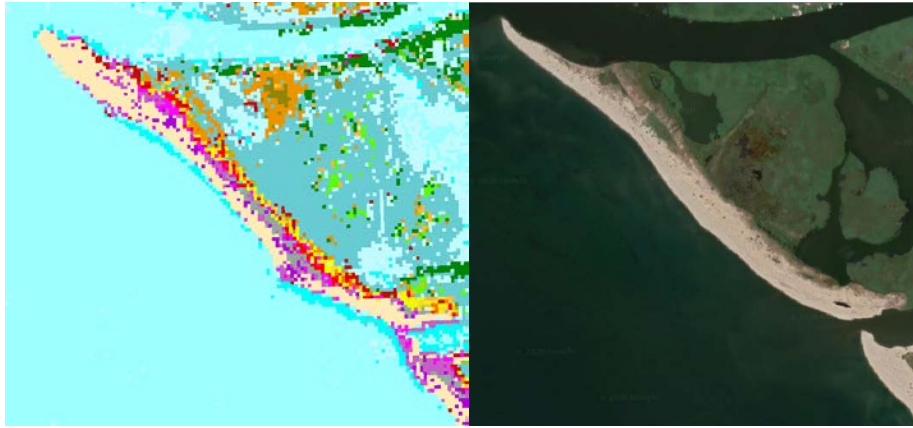
Εικόνα 6.9: Δυσκολία του ταξινομητή να προβλέψει την κατηγορία NGR.

Ως προς τις κλάσεις του γυμνού εδάφους, διαπιστώνουμε ότι οι κλάσεις RCK και SVA, δεν μπορούν να προβλεφθούν επιτυχώς στον χάρτη, καθώς τα εικονοστοιχεία τους ταξινομούνται σε πλήθος διάφορων άλλων κλάσεων. Η αδυναμία του ταξινομητή να διαχωρίσει τις συγκεκριμένες κλάσεις είναι απόρροια του μικρού αριθμού των δεδομένων εκπαίδευσης, καθώς αυτές οι κλάσεις δεν απαντώνται στα συγκεκριμένα layers.



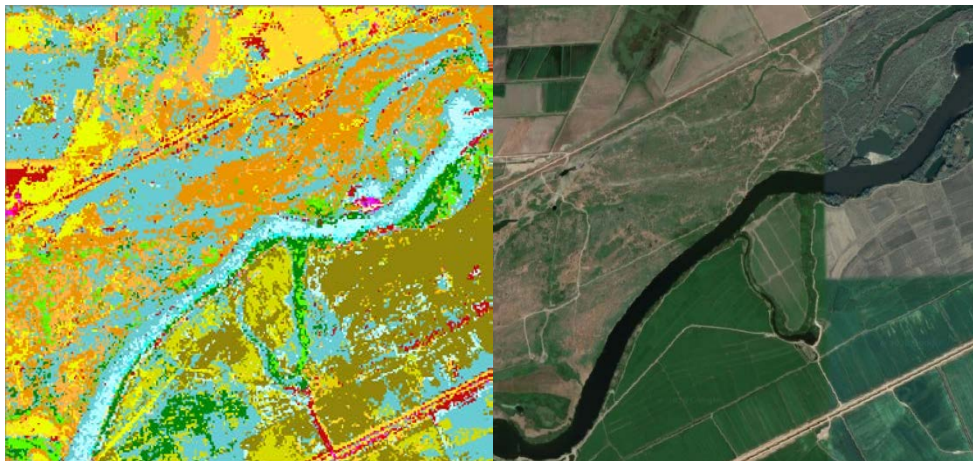
Εικόνα 6.10: Σύγχυση ανάμεσα στις κλάσεις RCK, SVA, SPSV και RAN.

Από την άλλη μεριά η κλάση SND, καταφέρνει να προβλεφθεί ικανοποιητικά από τον ταξινομητή, με μικρές συγχύσεις με την κλάση MES, σύγχυση που διαφαίνεται και από τον πίνακα σύγχυσης.

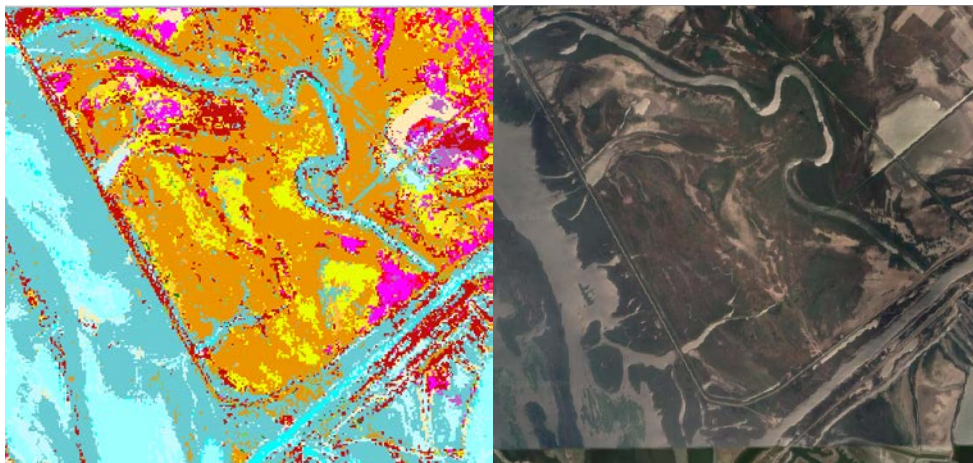


Εικόνα 5-19 Σύγκριση της κλάσης SND με την RCK και την MES

Ως προς τις υδάτινες επιφάνειες, διαπιστώνουμε ότι, ενώ η κατηγορία CWT αποδίδονται εν γένει σωστά στον χάρτη, οι υπόλοιπες κλάσεις, εμπλέκονται μεταξύ τους και το αποτέλεσμα του χάρτη δεν ανταποκρίνεται στην πραγματικότητα, καθώς εικονοστοιχεία της WCR αποδίδονται ως WBD ή MRS, εικονοστοιχεία WBD αποδίδονται ως CWT και ως MRS και εικονοστοιχεία που ανήκουν στην MRS, προβλέπονται ως καλλιέργειες (ελιές και δημητριακά) και δάση.

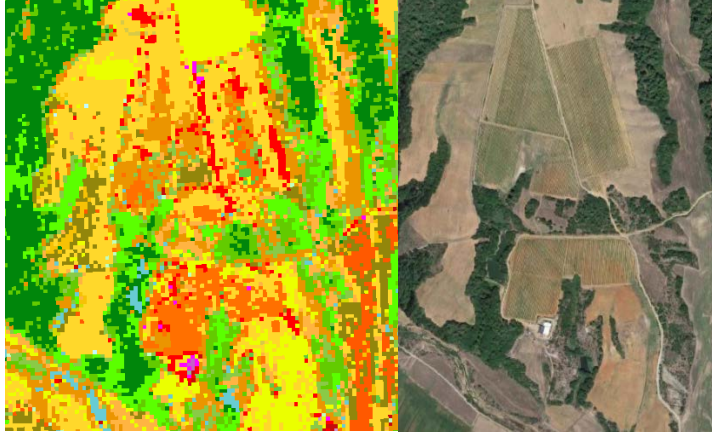


Εικόνα 5-20 Δυσκολία αναγνώρισης της WCR. Λάθος πρόβλεψη ως WBD και MRS



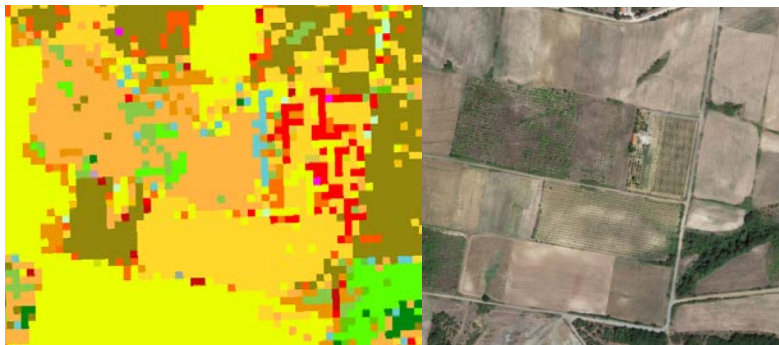
Εικόνα 5-21 Δυσκολία αναγνώρισης της MRS. Λάθος πρόβλεψη ως OLG και CRL

Σε σχέση με τις κλάσεις των καλλιεργειών εντοπίζεται πληθώρα εμπλοκών και λανθασμένων προβλέψεων, ανάμεσα στις υποκατηγορίες των καλλιεργειών αλλά και με υδάτινες επιφάνειες και αστικές περιοχές. Πιο συγκεκριμένα, η κατηγορία VNY, εμπλέκεται με άλλες καλλιέργειες και συγκεκριμένα με τους Ελαιώνες, τα βαμβάκι τις χορτονομές και τα δημητριακά.

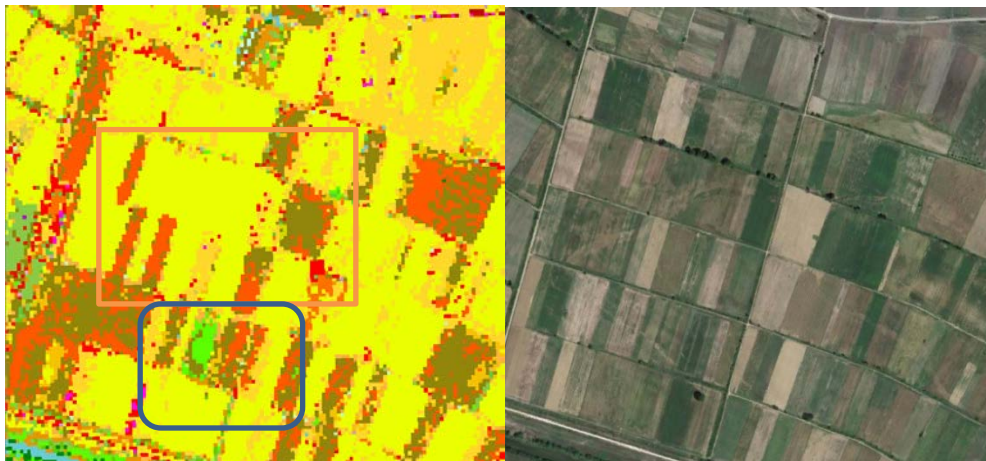


Εικόνα 5-22 Σύγκριση VNY με κλάσεις καλλιεργειών

Η κατηγορία των οπωροφόρων δένδρων, εμφανίζει συγχύσεις με τις καλλιέργειες βαμβακιού. Το ίδιο συμβαίνει και με τα δημητριακά, εικονοστοιχεία των οποίων αποδίδονται στο βαμβάκι.

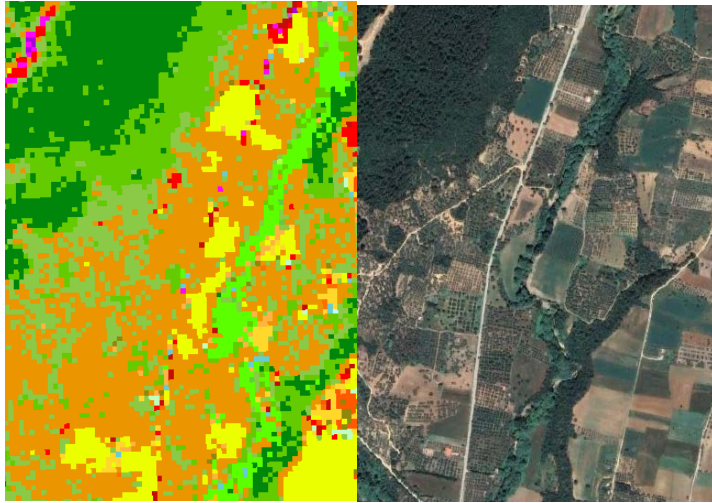


Εικόνα 5-23 Σύγκριση FRT με CTN



Εικόνα 5-24 Σύγκριση CRL με CTN και GRF με BLF και SPSV

Τα εικονοστοιχεία, που αντιπροσωπεύουν ελαιώνες, ταξινομούνται επιτυχώς, με μικρό αριθμό pixels να καταλήγει στην κλάση SPSV.



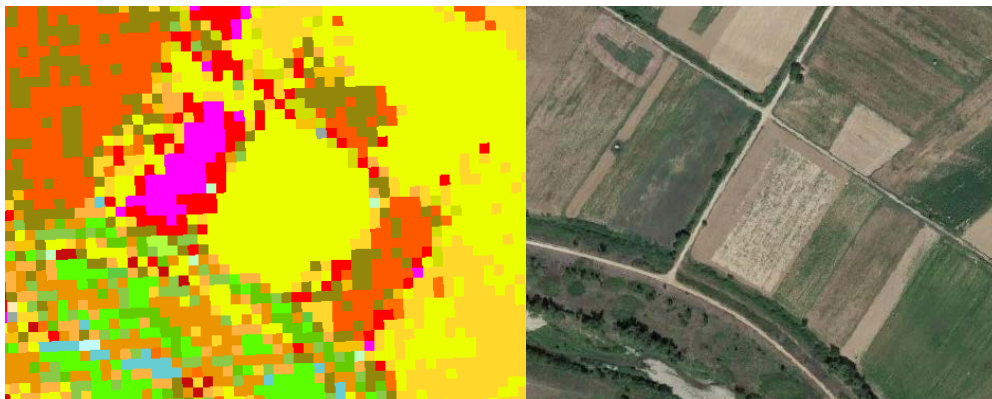
Εικόνα 5-25 Επιτυχής προσδιορισμός των pixels OLG στην εικόνα

Τα εικονοστοιχεία που αποτελούν χωράφια αραβόσιπου, προσδιορίζονται ως βαμβάκι.



Εικόνα 5-26 Εικονοστοιχεία της κλάσης GRF, ταξινομούνται ως CTN

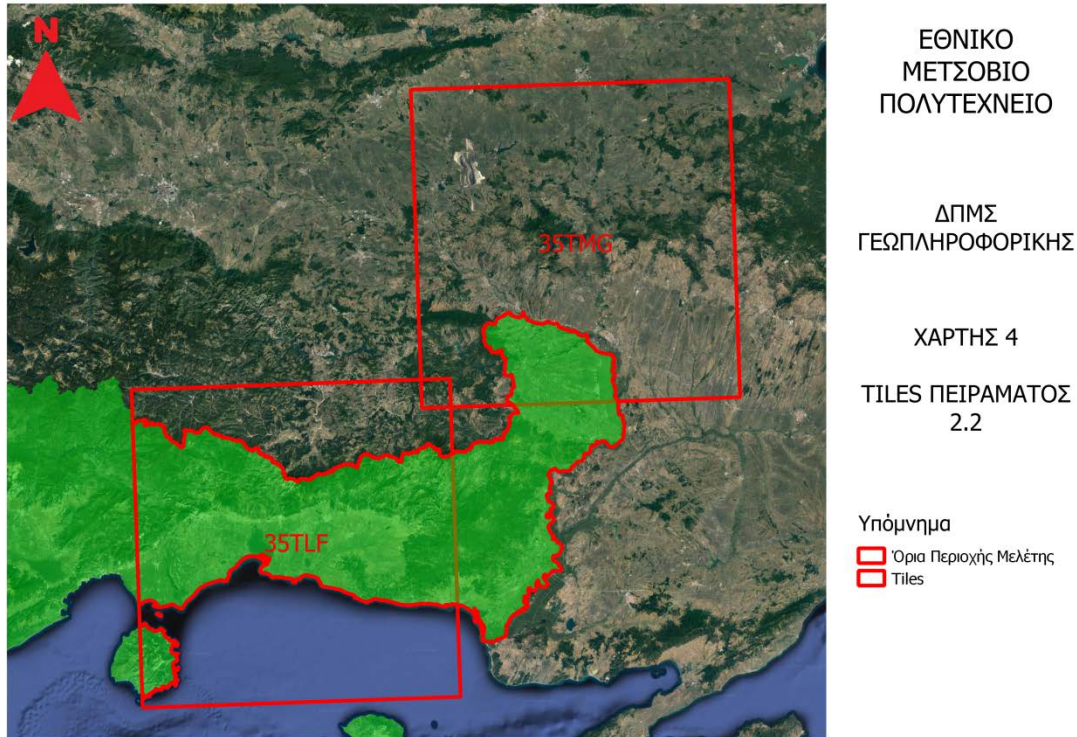
Τέλος, η κλάση VEG, δεν αποδίδεται σωστά στον χάρτη, καθώς εικονοστοιχεία της ταξινομούνται σε άλλες καλλιέργειες, πχ. στα δημητριακά.



Εικόνα 5-27 Απόδοση της VEG ως CRL

5.2.2 Δεύτερο πείραμα: εκπαίδευση στο tile 35TLF και έλεγχος στο tile 35TMF

Στο πρώτο πείραμα που εφαρμόστηκε επιλέχθηκε να χρησιμοποιηθεί το Tile 35TLF για να εκπαιδευτεί ο αλγόριθμος, ενώ ο έλεγχος της ταξινόμησης επιλέχθηκε να γίνει στο Tile 35TMG. Χρησιμοποιήθηκαν και οι έξι επιλεγμένες εικόνες και από αυτές τα κανάλια 2,3,4,8,11,12, καθώς και οι δείκτες NDVI, MSAVI, NDWI, NDBI. Έτσι, η ταξινόμηση υλοποιήθηκε για υπερκύβους βάρους 10*6.



Χάρτης 5.6 Τα tiles, που χρησιμοποιήθηκαν από τον αλγόριθμο για την διεξαγωγή του πρώτου πειράματος

Το πρόγραμμα έτρεξε και μέσα από ταξινόμηση RF, εξήχθησαν τα εξής αποτελέσματα.

Πίνακας 5-14 Πίνακας Σύγκρισης για ταξινόμηση με δεδομένα εκπαίδευσης από την εικόνα 35TLF και έλεγχο στην εικόνα 35TMG

		ΚΛΑΣΕΙΣ ΤΟΥ ΑΠΟΤΕΛΕΣΜΑΤΟΣ ΤΗΣ ΤΑΞΙΝΟΜΗΣΗΣ																																	
		DUF	SUF	RA	ICU	ME	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MR	VNY	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL	sum	PA	
ΚΛΑΣΕΙΣ ΤΟΥ ΧΑΡΤΗ ΤΟΥ GROUND TRUTH	DUF	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0,83	
	SUF	882	1394	20	39	3	2	0	0	0	0	22	19	0	25	0	0	2	31	46	728	339	41	71	16	9	0	0	1	7	0	12	3709	0,38	
	RA	262	114	1307	152	29	39	0	0	0	0	95	8	6	266	25	7	203	3	18	36	62	22	107	10	1	0	0	0	2	16	2790	0,47		
	ICU	57	2	4	84	35	0	0	0	0	0	0	0	0	0	24	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	210	0,4	
	ME	64	1	35	31	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	147	0,1	
	PHT	18	4	182	1	0	356	0	0	0	0	0	0	0	0	0	0	0	0	0	13	67	26	21	2	1	0	0	0	0	0	0	1	700	0,51
	GR	14	1	1	4	6	0	13	0	0	0	0	0	1	0	2	0	0	0	1	0	8	4	0	1	0	0	0	0	0	0	0	0	56	0,23
	BLF	0	0	0	0	0	0	0	5303	1234	1	1	0	0	0	0	0	0	0	0	15	0	0	0	0	1	0	0	0	0	0	0	0	5429	0,98
	CNF	0	0	0	0	0	0	0	651	820	1	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	2	0	0	0	1481	0,55	
	DSV	0	0	0	0	0	0	0	389	388	157	27	0	0	0	0	0	0	0	2	0	0	0	3	0	0	2	0	0	0	0	0	0	968	0,16
	SPS	0	0	0	0	0	0	1	405	29	146	38	1	0	0	0	0	17	0	17	4	1	1	28	65	1	0	0	0	0	0	0	754	0,05	
	SVA	0	0	1	0	0	0	0	0	0	0	3	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0
	RCK	4	4	9	0	0	0	0	0	0	0	4	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0
	SND	73	18	37	61	6	0	0	0	0	0	0	0	0	0	63	0	60	2	0	0	2	2	0	2	0	0	0	0	0	0	0	326	0,19	
	WC	0	0	1	0	0	0	0	18	42	8	0	0	0	0	0	12	4210	536	0	0	0	0	0	0	0	0	0	3394	0	0	0	0	8221	0
	WB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0
	MR	88	12	0	431	41	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	574	0
	VNY	1	39	1	0	0	0	0	44	0	0	4	2	0	0	0	0	0	0	2	8	18	7	3	3	2	0	0	0	0	0	1	135	0	
	FRT	1	35	0	1	0	0	0	5093	529	125	260	88	0	1	0	0	151	3	1930	177	52	155	144	1943	2	8	0	7	1	14	2	1072	0,18	
	CRL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CTN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
GRF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
OLG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
VEG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
NG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
CW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
KW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
TBC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
RCF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
SFL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
sum	1469	1625	1598	804	135	397	14	5963	3042	438	455	125	7	386	37	4277	920	38	2043	1036	504	247	361	2046	15	10	3394	10	8	16	32	kapp	0,44		
UA	0	0,86	0,82	0,1	0,11	0,9	0,93	0,89	0,27	0,36	0,08	0	0	0,16	0,32	0	0	0	0,94	0	0	0	0	0	0	0	0	0	0	0	0	0	0,7	0,7	

Όπως βλέπουμε στον παραπάνω πίνακα σύγκυσης η συνολική ακρίβεια είναι αρκετά ικανοποιητική καθώς ισούται με 64%. Ακόμη, υπολογίστηκε ο δείκτης ακρίβειας KHAM (ή δείκτης συμφωνίας KAPPA), ο οποίος βρέθηκε ίσος με 59%. Συνεπώς, διαπιστώνουμε ότι η ταξινόμηση, σε ένα πείραμα στο οποίο τα δεδομένα εκπαίδευσης σε σχέση με τα δεδομένα ελέγχου προέρχονται από διαφορετικά tiles απέφερε θετικά αποτελέσματα.

Εκτός από τα μεγέθη που αφορούν την συνολική ακρίβεια της ταξινόμησης, σκόπιμο είναι να διαπιστώσουμε αν και κατά πόσο υπάρχουν σφάλματα ταξινόμησης αλλά και λανθασμένα ταξινομημένα εικονοστοιχεία για την κάθε κατηγορία ξεχωριστά. Μέσα από αυτόν τον έλεγχο θα καταφέρουμε να εντοπίσουμε κατηγορίες προβληματικές και να αντιληφθούμε λάθη και παραλείψεις τις διαδικασίας.

Πιο συγκεκριμένα, για τις κλάσεις που ανήκουν στην γενική κατηγορία των τεχνητών επιφανειών, βλέπουμε ο ταξινομητής εμφανίζει δυσκολία στο διαχωρισμό τους. Πιο συγκεκριμένα, οι κατηγορίες DUF και SUF, εμπλέκονται καθώς η DUF εμφανίζει σημαντικό σφάλμα συμπερίληψης που φτάνει το 60% με την SUF. Αυτό το σφάλμα επιφέρει χαμηλή ακρίβεια χρήστη και ακρίβεια παραγωγής για τις κλάσεις DUF και SUF αντίστοιχα. Επιπλέον, οι δύο αυτές κλάσεις εμφανίζουν κατά αντιστοιχία υψηλή ακρίβεια παραγωγής και χρήστη αντίστοιχα, καθώς αυτές οι τιμές ξεπερνάν το 80%. Η σύγκυση αυτή είναι δικαιολογημένη καθώς οι δύο αυτές κατηγορίες αποτελούν υποκατηγορίες του αστικού ιστού.

Για την κατηγορία RAN, ο ταξινομητής επιτυγχάνει σχετικά ικανοποιητικά αποτελέσματα ταξινόμησης, αφού η ακρίβεια παραγωγού δεν ξεπερνά το 47%, ενώ η ακρίβεια χρήστη είναι αρκετά υψηλή με το ποσοστό της να αγγίζει το 82%. Ως προς τα σφάλματα παράλειψης και συμπερίληψης, δεν προκύπτει σύγκυση με συγκεκριμένες κλάσεις.

Για τις κατηγορίες ICU και MES, διαπιστώνουμε ότι ο ταξινομητής δυσκολεύεται να τις διαχωρίσει, αφού οι ακρίβειες παραγωγής είναι ίσες με 40 και 10% αντίστοιχα, ενώ οι ακρίβειες χρήστη είναι και για τις δύο κατηγορίες κοντά στο 10%. Από τα σφάλματα παράλειψης και συμπερίληψης διαπιστώνουμε ότι οι συγκεκριμένες κλάσεις εμπλέκονται αναμεταξύ τους αλλά και με τις κλάσεις DUF, RAN. Αυτές οι συγχύσεις είναι αναμενόμενες λόγω της παραπλήσιας φύσης των συγκεκριμένων κατηγοριών. Δυσκολία ήταν αναμενόμενη καθώς, η κατηγορία DUF και SUF, είναι υποκατηγορίες του αστικού ιστού και άρα εμπρικλείουν ίδιου τύπου αντικείμενα που εκπέμπουν με παρόμοιο τρόπο στο ηλεκτρομαγνητικό φάσμα, ενώ η κατηγορία ICU, αποτελείται από περιοχές με τεχνητές επιφάνειες από τσιμέντο και άσφαλτο, υλικά που εκπέμπουν παρόμοια με τις κατηγορίες DUF, SUF και RAN.

Η κατηγορία PHT, χαρακτηρίζεται από σχετικά υψηλές ακρίβειες παραγωγής και χρήστη, αφού η ακρίβεια παραγωγής καταφέρνει να αγγίξει το 51%, ενώ η ακρίβεια χρήστη ξεπερνά το 89%. Η μόνη σύγκυση που διαφαίνεται από τα σφάλματα παράλειψης (ίσο με 26%) και συμπερίληψης (ίσο με 9%), είναι με την κατηγορία των οδικών δικτύων μεταφοράς.

Τέλος, η κατηγορία GRN χαρακτηρίζεται από χαμηλή ακρίβεια παραγωγής περίπου ίση με το 23%, παρουσιάζει όμως υψηλή αξιοπιστία η οποία φτάνει το 93%.

Ελέγχοντας τα σφάλματα παράλειψης, διαπιστώνεται ότι 25% των εικονοστοιχείων των δεδομένων εκπαίδευσης της συγκεκριμένης κλάσης ταξινομείται στην DUF, 7% στην ISU, 10% στην MES, καθώς και 14% στην CRL, μία σύγκυση μη αναμενόμενη, λόγω του γεγονότος ότι οι δύο αυτές κλάσεις ανακλούν με διαφορετικό τρόπο.

Στις κατηγορίες φυσικής βλάστησης τα ποσοστά ακρίβειας και αξιοπιστίας ποικίλουν. Πιο συγκεκριμένα, η κατηγορία BLF, ταξινομείται πολύ επιτυχημένα, αφού η ακρίβεια παραγωγής ξεπερνά το 97% και η ακρίβεια χρήστη το 89%. χαρακτηρίζεται από ακρίβεια που αγγίζει το 71% και αξιοπιστία ανώτερη του 89%, ενώ δεν εμφανίζει αξιόλογες συγχύσεις.

Η κατηγορία CNF, δεν ταξινομείται εξίσου επιτυχημένα μιας και η αξιοπιστία είναι χαμηλή της τάξης του 26%. Η ακρίβεια παραγωγής από την άλλη πλευρά είναι ανώτερη καθώς ξεπερνά το 55%. Ως προς τα σφάλματα παράλειψης και συμπερίληψης, διαπιστώνεται ότι η σημαντικότερη σύγκυση που εμφανίζεται είναι ανάμεσα στα φυλλοβόλα και στα κωνοφόρα δάση, καθώς πάνω από 40% των εικονοστοιχείων των δεδομένων εκπαίδευσης της CNF παραλείπονται και αποδίδονται στην BLF και το 44% των εικονοστοιχείων ταξινομούνται στον χάρτη ως BLF, ενώ χαρακτηρίζουν περιοχές CNF. Ακόμα σημαντικές συγχύσεις, έτσι όπως προκύπτουν από τα σφάλματα συμπερίληψης, εντοπίζουμε με τις κατηγορίες DSV και FRT. Η σύγκυση με την κατηγορία DSV είναι αναμενόμενη καθώς οι δύο κατηγορίες είναι παρόμοιας φύσης, και η δυσκολία να διακριθούν τόσο από τον ταξινομητή όσο και από τον χρήστη κατά τη συλλογή του ground truth είναι ιδιαίτερα αυξημένη. Η σύγκυση με την κατηγορία FRT δεν είναι εξίσου αναμενόμενη, καθώς τα οπωροφόρα δένδρα παρουσιάζουν διαφοροποιήσεις στα φασματικά τους χαρακτηριστικά, διαφοροποιήσεις που όμως δεν είναι ικανές να οδηγήσουν σε πλήρως πετυχημένη διάκριση από τον αλγόριθμο ταξινόμησης. Η σύγκυση αυτή, μπορεί να δικαιολογηθεί και λόγω των διαφορετικών ειδών οπωροφόρων δένδρων που καλλιεργούνται στις δύο περιοχές, οδηγώντας σε μεταβολή της φασματικής υπογραφής της κλάσης από περιοχή σε περιοχή και πιθανότητα παράλειψης των συγκεκριμένων εικονοστοιχείων από την κατηγορία FRT, και συμπερίληψής τους στην κατηγορία CNF.

Η κατηγορία DSV χαρακτηρίζεται από χαμηλή ακρίβεια παραγωγής (16%), καθώς το σφάλμα παράλειψης ως προς την κατηγορία CNF και την BLF είναι ανώτερο του 40%. Η ακρίβεια χρήστη είναι μέτρια προς χαμηλή (36%), καθώς εμφανίζεται σημαντικό σφάλματα συμπερίληψης (33%) με την κατηγορία SPSV, σύγκυση αναμενόμενη καθώς οι δύο κατηγορίες αποτελούν υποκατηγορίες της γενικής κλάσης: Σκληροφυλική Βλάστηση. Ακόμα αξιοσημείωτο είναι το σφάλμα συμπερίληψης (28,58%) που εμφανίζει η DSV με τη FRT.

Η κατηγορία SPSV, δεν ταξινομείται ικανοποιητικά απ' τον αλγόριθμο, αφού τα ποσοστά ακρίβειας παραγωγής χρήστη δεν ξεπερνούν το 5% και το 8% αντιστοίχως. Οι σημαντικότερες συγχύσεις της κατηγορίας αυτήν είναι κυρίως με τις κλάσεις FRT, BLF και λιγότερο σημαντικές με την DSV και την RAN.

Σε σχέση με τις κατηγορίες του γυμνού εδάφους, και οι τρεις κατηγορίες διέθεταν πενιχρό ground truth. Έτσι, δεν ήταν δυνατή η ταξινόμησή τους από τον ταξινομητή. Η κατηγορία SND, ήταν η μόνη που ο ταξινομητής κατάφερε σε μικρό βαθμό να διαχωρίσει πετυχαίνοντας ακρίβεια παραγωγής 19,3% και χρήστη 16,3%. Συγχύσεις

εμφανίζονται με όλες σχεδόν τις κλάσεις των αστικών περιοχών (DUF, SUF, RAN, & ICU). Οι συγχύσεις αυτές είναι αναμενόμενες, καθώς εικονοστοιχεία άμμου απαντώνται στις συγκεκριμένες κλάσεις.

Ως προς τις κατηγορίες SVA και RCK, ο ταξινομητής δεν κατάφερε να τις ταξινομήσει. Η αποτυχία του αλγορίθμου να ταξινομήσει τις συγκεκριμένες κλάσεις είναι αποτέλεσμα του μικρού αριθμού δεδομένων εκπαίδευσης, καθώς οι συγκεκριμένες κατηγορίες δεν απαντώνται ευρέως στα συγκεκριμένα tiles.

Αντίστοιχα, η γενική κατηγορία των υδάτινων πόρων δεν ήταν δυνατό να ταξινομηθεί επιτυχημένα, καθώς λίμνες, ποτάμια και παράκτια ύδατα δεν απαντώνται στο tile της εικόνας 35TMG. Η κατηγορία MRS αν και διέθετε κάποιον αριθμό πολυγώνων εκπαίδευσης, δεν διαχωρίστηκε από τον ταξινομητή, αφού, 75% των pixels αποδόθηκαν στην ICU, ενώ ανήκουν στην MRS, 58% που αναπαριστούν περιοχές SND και 22% που αναπαριστούν περιοχές WCR, συμπεριλήφθηκαν στην SND.

Ως προς τις κλάσεις που ανήκουν στις καλλιέργειες διαπιστώνεται ότι, στην εικόνα 35TMG δύο κατηγορίες απαντώνται, η VNY και η FRT. Η κατηγορία VNY δεν ταξινομείται επιτυχημένα, καθώς τα δεδομένα είναι πενιχρά, αφού τα εικονοστοιχεία ταξινόμησης και ελέγχου δεν ξεπερνάν τα 135. Για την κατηγορία FRT, βλέπουμε ότι η αξιοπιστία της είναι ιδιαίτερα υψηλή αφού ξεπερνά το 94%, όχι όμως και η ακρίβεια παραγωγού, που είναι ίση με 18%.

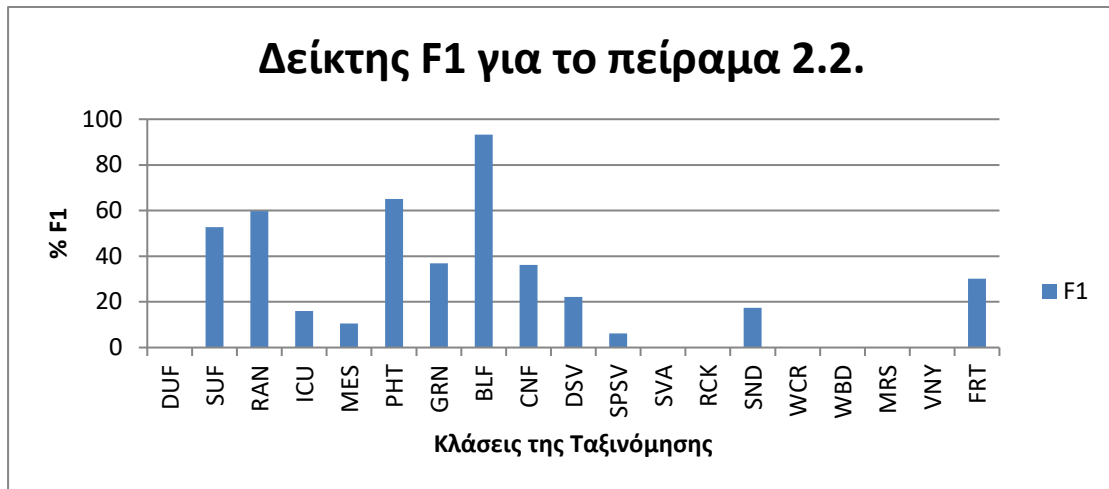
Πίνακας 5-15 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMG

	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPSV	SVA	RCK	SND	WCR	WBD	MRS	VNY	FRT
DUF	83,33	23,78	9,39	27,14	43,54	2,57	25,00	0,00	0,00	0,00	0,00	0,00	17,39	22,39	0,00	0,00	15,33	0,74	0,01
SUF	16,67	37,58	4,09	0,95	0,68	0,57	1,79	0,00	0,00	0,00	0,00	0,00	17,39	5,52	0,00	0,00	2,09	28,89	0,33
RAN	0,00	0,54	46,85	1,90	23,81	26,00	1,79	0,00	0,00	0,00	0,00	12,50	39,13	11,35	0,01	0,00	0,00	0,74	0,00
ICU	0,00	1,05	5,45	40,00	21,09	0,14	7,14	0,00	0,00	0,00	0,00	0,00	0,00	18,71	0,00	0,00	75,09	0,00	0,01
MES	0,00	0,08	1,04	16,67	10,20	0,00	10,71	0,00	0,00	0,00	0,00	0,00	0,00	1,84	0,00	0,00	7,14	0,00	0,00
PHT	0,00	0,05	1,40	0,00	0,00	50,86	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
GRN	0,00	0,00	0,00	0,00	0,00	0,00	23,21	0,00	0,00	0,00	0,13	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	97,69	43,96	40,19	53,71	0,00	0,00	0,00	0,22	0,00	0,00	32,59	47,50
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2,27	55,37	40,08	3,85	0,00	0,00	0,00	0,51	0,00	0,00	0,00	4,93
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	16,22	19,36	0,00	0,00	0,00	0,10	0,00	0,00	0,00	1,17
SPSV	0,00	0,59	3,41	0,00	0,00	0,14	0,00	0,00	0,00	2,79	5,04	37,50	17,39	0,00	0,00	0,00	0,00	2,96	2,42
SVA	0,00	0,51	0,29	0,00	0,00	0,71	1,79	0,00	0,00	0,00	0,13	0,00	4,35	0,00	0,00	0,00	0,00	1,48	0,82
RCK	0,00	0,00	0,22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	12,50	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SND	0,00	0,67	9,53	11,43	0,00	0,14	3,57	0,00	0,00	0,00	0,00	37,50	4,35	19,33	0,00	0,00	0,00	0,00	0,01
WCR	0,00	0,00	0,90	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15	0,00	0,00	0,00	0,00
WBD	0,00	0,00	0,25	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	18,40	51,21	0,00	0,00	0,00	0,00
MRS	0,00	0,05	7,28	0,00	0,00	0,14	0,00	0,00	0,00	0,00	2,25	0,00	0,00	0,61	6,52	100,0	0,17	0,00	1,41
VNY	0,00	0,84	0,11	0,00	0,00	0,00	1,79	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03
FRT	0,00	1,24	0,65	0,00	0,00	1,86	0,00	0,03	0,00	0,21	2,25	0,00	0,00	0,00	0,00	0,00	0,00	1,48	18,00
CRL	0,00	19,63	1,29	1,90	0,68	9,57	14,29	0,00	0,00	0,00	0,53	0,00	0,00	0,61	0,00	0,00	0,17	5,93	1,65
CTN	0,00	9,14	2,22	0,00	0,00	3,71	7,14	0,00	0,00	0,00	0,13	0,00	0,00	0,61	0,00	0,00	0,00	13,33	0,48
GRF	0,00	1,11	0,79	0,00	0,00	3,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,00	0,00	0,00	0,00	5,19	1,45
OLG	0,00	1,91	3,84	0,00	0,00	0,29	1,79	0,00	0,00	0,31	3,71	0,00	0,00	0,61	0,00	0,00	0,00	2,22	1,34
MAZ	0,00	0,43	0,36	0,00	0,00	0,14	0,00	0,00	0,47	0,00	8,62	0,00	0,00	0,00	0,00	0,00	0,00	2,22	18,12
VEG	0,00	0,24	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,00	0,00	0,00	0,00	1,48	0,02
NGR	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,21	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07
CWT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	41,28	0,00	0,00	0,00	0,00
KWP	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07
TBC	0,00	0,19	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
RCF	0,00	0,00	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,13
SFL	0,00	0,32	0,57	0,00	0,00	0,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,74	0,02

Πίνακας 5-16 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMG

	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPSV	SVA	RCK	SND	WCR	WBD	MRS	VNY	FRT	CRL	CTN	GRF	OLG	MAZ	VEG	NGR	CWT	KWP	TBC	RCF	SFL	
DUF	0,34	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SUF	60,04	85,78	1,25	4,85	2,22	0,50	0,00	0,00	0,00	0,00	4,84	15,20	0,00	6,48	0,00	0,00	0,22	81,58	2,25	70,27	67,26	16,60	19,67	0,78	60,00	0,00	0,00	10,00	87,50	0,00	37,50	
RAN	17,84	7,02	81,79	18,91	21,48	9,82	0,00	0,00	0,00	0,00	20,88	6,40	85,71	68,91	67,57	0,16	22,07	7,89	0,88	3,47	12,30	8,91	29,64	0,49	6,67	0,00	0,00	0,00	0,00	12,50	50,00	
ICU	3,88	0,12	0,25	10,45	25,93	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	6,22	0,00	0,00	0,00	0,00	0,00	0,39	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
MES	4,36	0,06	2,19	3,86	11,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
PHT	1,23	0,25	11,39	0,12	0,00	89,67	0,00	0,00	0,00	0,00	0,22	4,00	0,00	0,26	0,00	0,00	0,11	0,00	0,64	6,47	5,16	8,50	0,55	0,05	0,00	0,00	0,00	0,00	0,00	0,00	3,13	
GRN	0,95	0,06	0,06	0,50	4,44	0,00	92,86	0,00	0,00	0,00	0,00	0,80	0,00	0,52	0,00	0,00	0,00	2,63	0,00	0,77	0,79	0,00	0,28	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	88,93	40,57	0,23	0,22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,73	0,00	0,00	0,00	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,09	26,96	0,23	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,34	0,00	0,00	0,00	20,00	0,00	0,00	0,00	
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,65	12,75	35,84	5,93	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00	0,00	0,00	0,83	0,00	0,00	20,00	0,00	0,00	0,00	0,00	0,00	
SPSV	0,00	0,00	0,00	0,00	0,00	0,00	7,14	0,68	0,95	33,33	8,35	0,80	0,00	0,00	0,00	0,00	1,85	0,00	0,83	0,39	0,20	0,40	7,76	3,18	6,67	0,00	0,00	0,00	0,00	0,00	0,00	
SVA	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,66	0,00	14,29	0,78	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
RCK	0,27	0,25	0,56	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,88	0,80	0,00	0,26	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
SND	4,97	1,11	2,32	7,59	4,44	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	16,32	0,00	1,40	0,22	0,00	0,00	0,19	0,40	0,00	0,55	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
WCR	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,03	1,38	1,83	0,00	0,00	0,00	0,00	32,43	98,43	58,26	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,00	0,00	0,00	0,00	0,00	
WBD	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,76	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
MRS	5,99	0,74	0,00	53,61	30,37	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,00	0,00	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
VNY	0,07	2,40	0,06	0,00	0,00	0,00	0,00	0,07	0,00	0,00	0,88	1,60	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,77	3,57	2,83	0,83	0,15	13,33	0,00	0,00	0,00	0,00	0,00	3,13	
FRT	0,07	2,15	0,00	0,12	0,00	0,00	0,00	8,54	17,39	28,54	57,14	70,40	0,00	0,26	0,00	0,00	16,41	7,89	94,47	17,08	10,32	62,75	39,89	94,97	13,33	80,00	0,00	70,00	12,50	87,50	6,25	

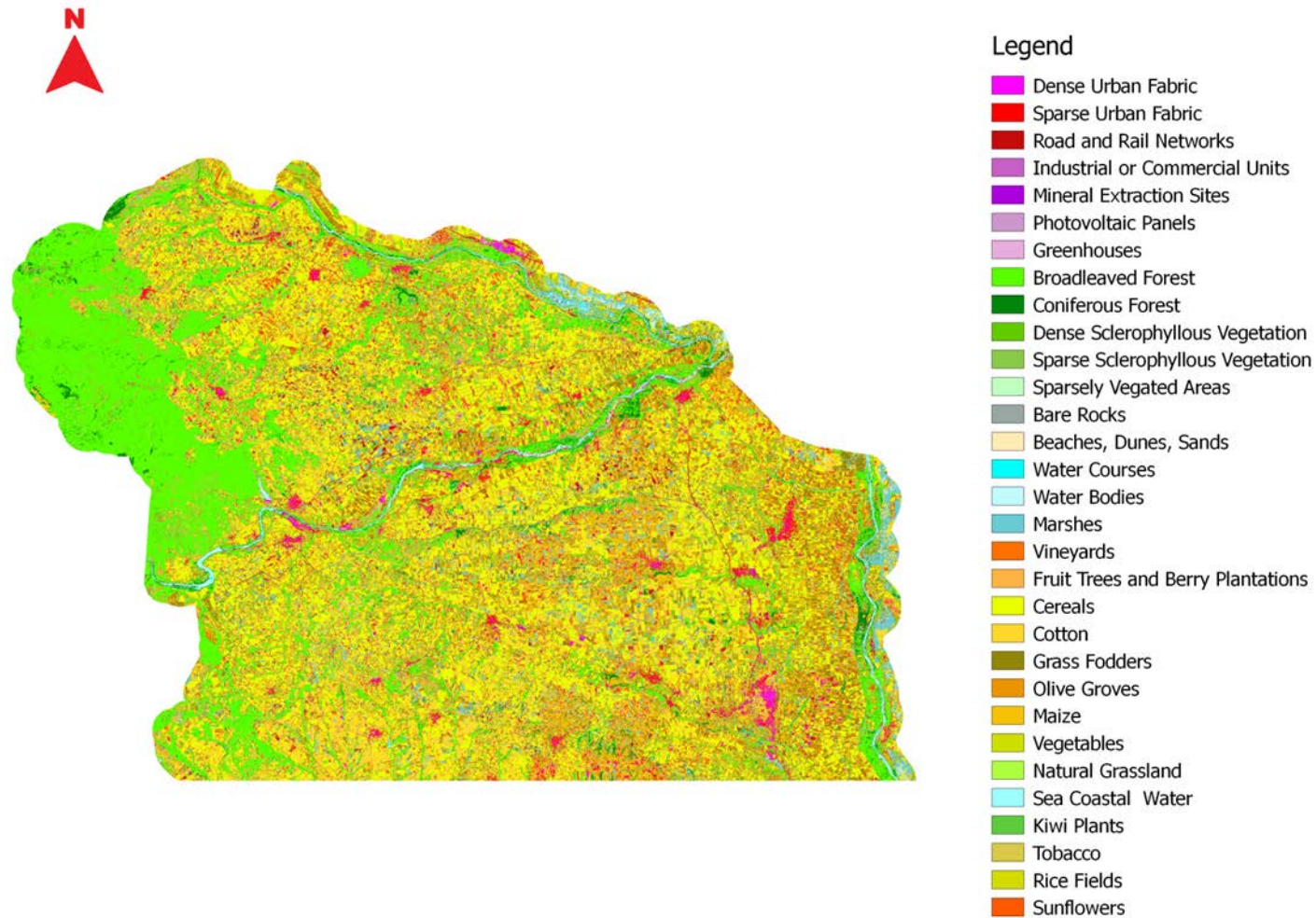
Στην συνέχεια παρατίθεται το γράφημα για τον δείκτη F (F1). Στο συγκεκριμένο γράφημα εμφανίζεται το ποσοστό του F1 για κάθε κλάση της ταξινόμησης. Με το γράφημα αυτό, μπορούμε και εποπτικά να εντοπίσουμε τις κατηγορίες εκείνες, που απέδωσαν καλύτερα, αλλά και τις κατηγορίες αυτές, που εμφάνισαν συγχύσεις.



Γράφημα 5-11 Δείκτης F1 για την ταξινόμηση με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG

Όπως διαπιστώνουμε, η κατηγορία εκείνη με τον μέγιστο δείκτη F1, είναι η κατηγορία BLF και έπειτα η κατηγορία PHT. Ακολουθούν οι κατηγορίες CNF, DSV και SPSV. Αντίθετα, οι κατηγορίες εκείνες, που ο ταξινομητής δυσκολεύτηκε να διαχωρίσει ήταν οι κατηγορίες εκείνες, που δεν απαντώνται ευρέως στο συγκεκριμένο tile και ο αλγόριθμος δυσκολεύεται να αναγνωρίσει. Αυτές είναι οι κατηγορίες SVA, RCK, WCR, WBD, MRS και VNY. Τα αποτελέσματα του δείκτη F1, επιβεβαιώνουν τις παρατηρήσεις, που προέκυψαν από τη μελέτη του πίνακα σύγχυσης.

Τέλος, παρατίθεται ο χάρτης, που παρήχθη από το πείραμα αυτό.

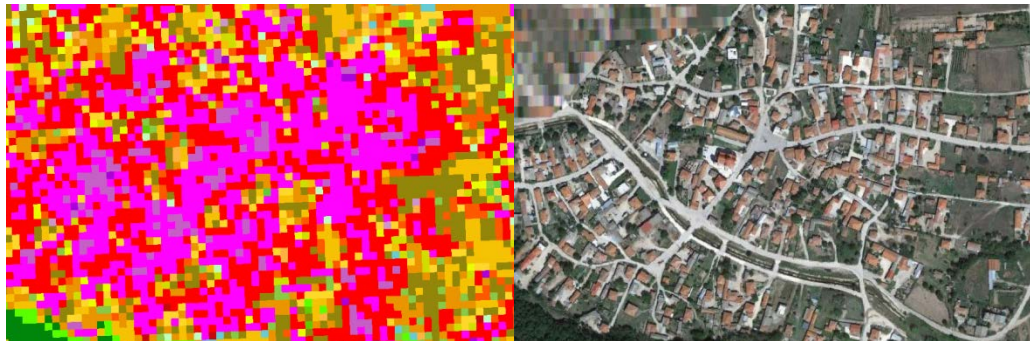


Χάρτης 5.7 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TLF και έλεγχο στην εικόνα 35TMG

Μελετώντας τον χάρτη, διαπιστώνουμε ότι σε γενικές γραμμές δίνει μία αξιόπιστη πρόβλεψη των χρήσεων/ καλύψεων γης και καλλιεργειών της περιοχής μελέτης. Οι καλύψεις γης, που επικρατούν στην περιοχή, προβλέπονται σχετικά επιτυχημένα, από τον αλγόριθμο RF, καθώς οι βασικές καλύψεις της περιοχής μελέτης προβλέπονται σωστά από τον ταξινομητή. Παρατηρούμε, λοιπόν, ότι οι αστικές, οι δασικές περιοχές, αλλά και οι καλλιεργήσιμες περιοχές διαχωρίζονται εν γένει σωστά από τον ταξινομητή.

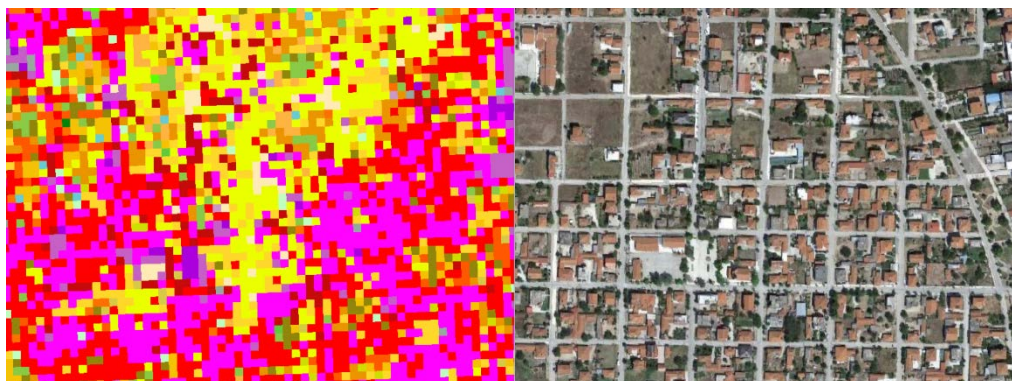
Εάν παρατηρήσουμε τον χάρτη πιο λεπτομερώς και σε μεγαλύτερο zoom, θα διαπιστώσουμε ότι εμφανίζονται αρκετές συγχύσεις.

Για την γενική κατηγορία τεχνητών επιφανειών ,εμφανίζονται συγχύσεις ανάμεσα στις υποκατηγορίες της. Όπως είχε ήδη διαπιστωθεί κατά την ανάλυση του πίνακα σύγχυσης, η κατηγορία DUF, εμφανίζει έντονο σφάλμα συμπερίληψης με την κατηγορία SUF, καθώς πάν από το 60% των εικονοστοιχείων, τα οποία ταξινομούνται στη συγκεκριμένη κλάση δεν ανήκουν σε εκείνη αλλά στη SUF. Αυτή η παρατήρηση, η οποία προκύπτει από τη μελέτη του πίνακα σύγχυσης, είναι εμφανής και στον παραγόμενο χάρτη, καθώς εικονοστοιχεία της SUF, απεικονίζονται ως DUF



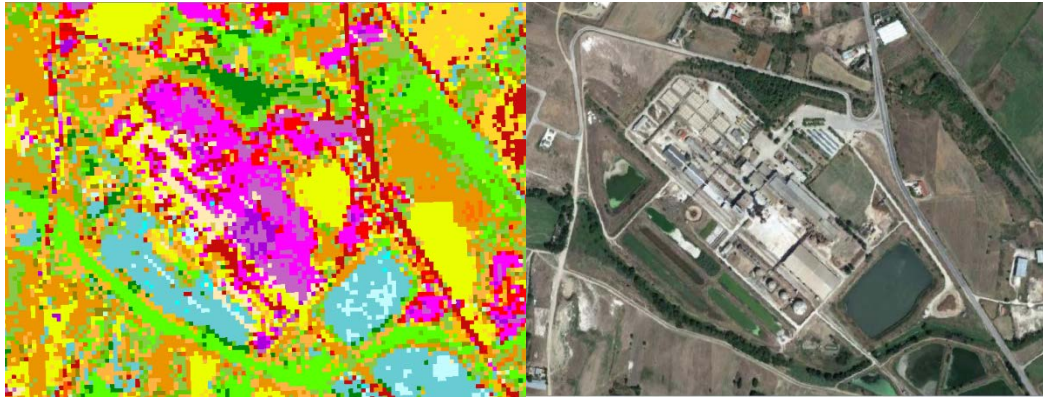
Εικόνα 5-28 Σύγχυση ανάμεσα στην DUF και την SUF

Ακόμη, εικονοστοιχεία τα οποία ανήκουν στην κατηγορία SUF, παραλείπονται από αυτήν την κλάση και αποδίδονται στην λάση CRL. Η συγκεκριμένη σύγχυση εντοπίζεται από τα σφάλματα παράλειψης και επιβεβαιώνεται στον παραγόμενο χάρτη.



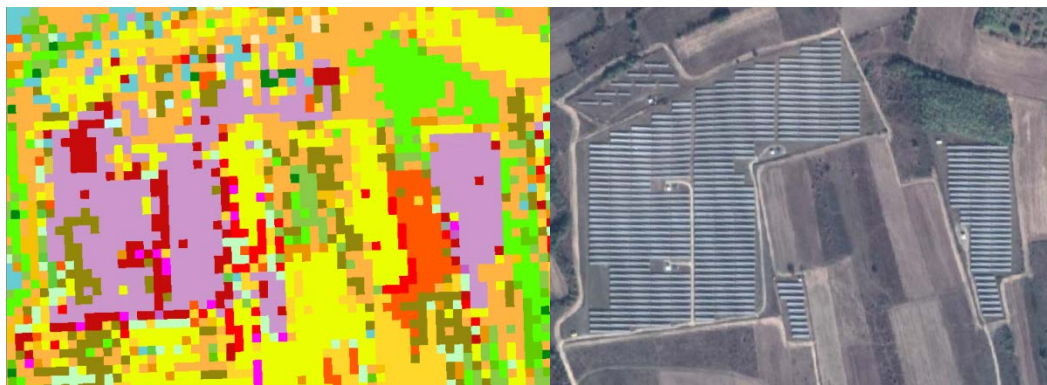
Εικόνα 5-29 Σύγχυση ανάμεσα στην DUF και την CRL

Επιπλέον, η κατηγορία RAN, προβλέπεται αρκετά επιτυχημένα από τον ταξινομητή. Όχι, όμως και οι κλάσεις ICU και MES, οι οποίες εμφανίζουν σύγχυση τόσο μεταξύ τους όσο και με τις υπόλοιπες κατηγορίες των τεχνητών επιφανειών.



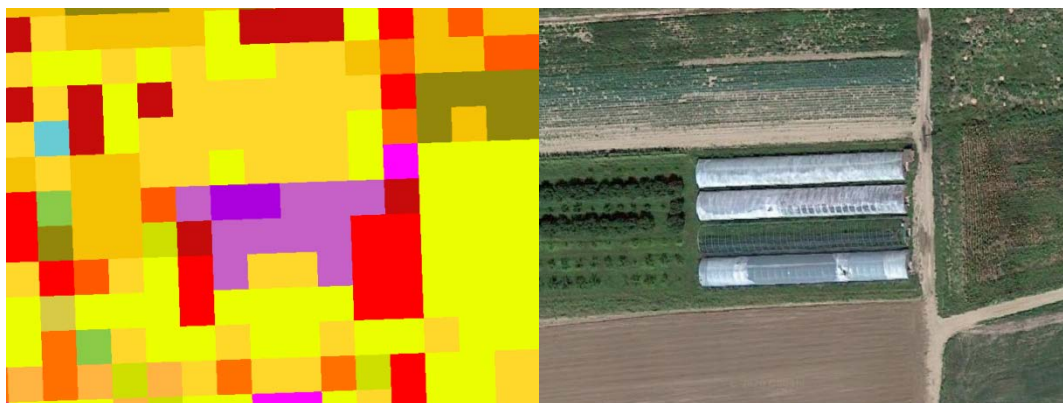
Εικόνα 5-30 Σύγκριση της ICU με DUF και την MES. Επιτυχής Ταξινόμηση της RAN

Η κατηγορία ΡΗΤ αποδίδεται επιτυχώς από τον ταξινομητή. Μικρός αριθμός των εικονοστοιχείων της και εμφανίζει συγχύσεις με την κατηγορία RAN, καθώς εικονοστοιχεία της παραλείπονται και ταξινομούνται ως RAN.



Εικόνα 5-31 Επιτυχημένη απόδοση των ΡΗΤ

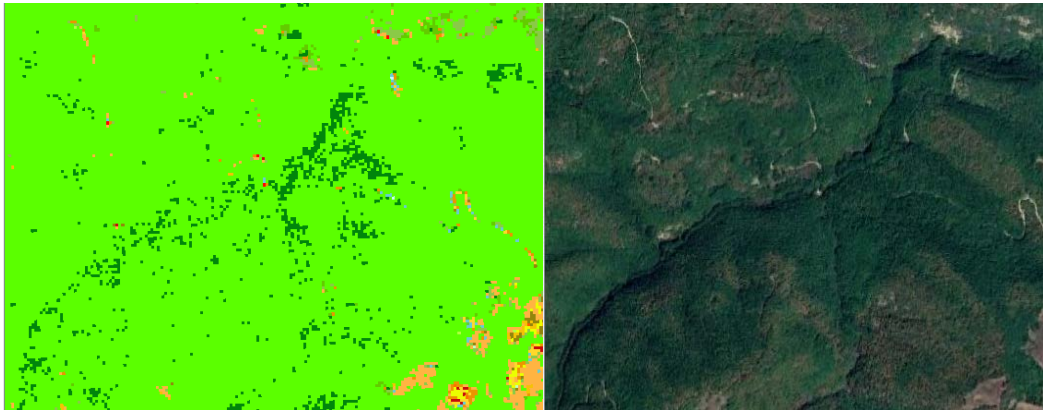
Τέλος, η κλάση GRN δεν προβλέπεται σωστά από τον αλγόριθμο, καθώς τα εικονοστοιχεία της αποδίδονται σε άλλες κλάσεις των τεχνητών επιφανειών, αλλά και σε κατηγορίες καλλιεργειών.



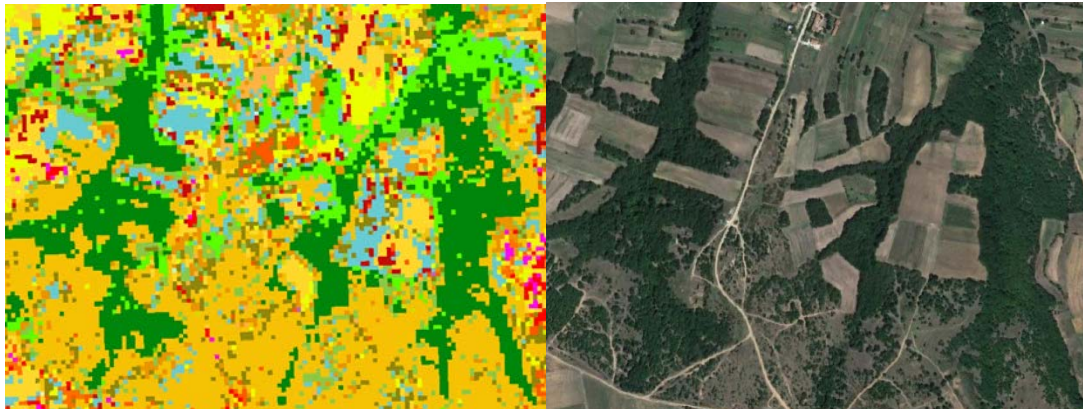
Εικόνα 5-32 Αδυναμία πρόβλεψης της GRN

Οι κατηγορίες της φυσικής βλάστησης φαίνεται ότι προβλέπονται σε μεγάλο βαθμό από τον αλγόριθμο. Τα εικονοστοιχεία του χάρτη απεικονίζουν τις σωστές κλάσεις, με

λίγα pixels των κατηγοριών αυτών να αποδίδονται σε λανθασμένη κατηγορία. Οι συγχύσεις αυτές εμφανίζονται κατά βάση ανάμεσα στις κλάσεις φυσικής βλάστησης.

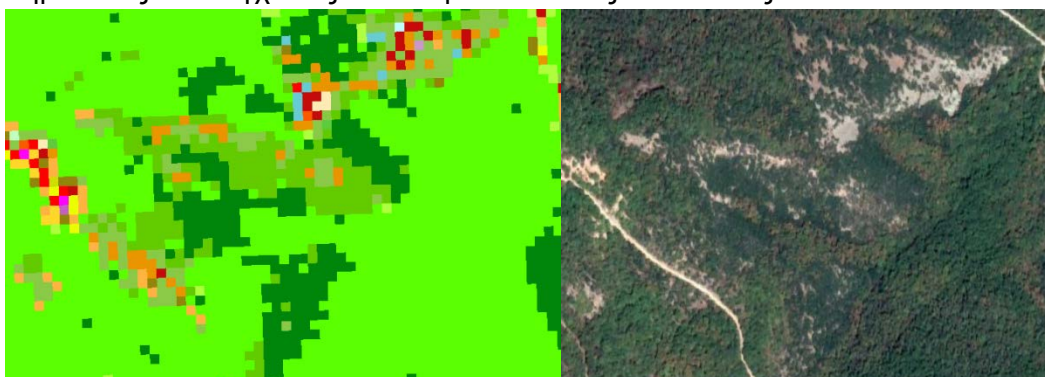


Εικόνα 5-33 Μικρή σύγχυση ανάμεσα στην BLF και την CNF



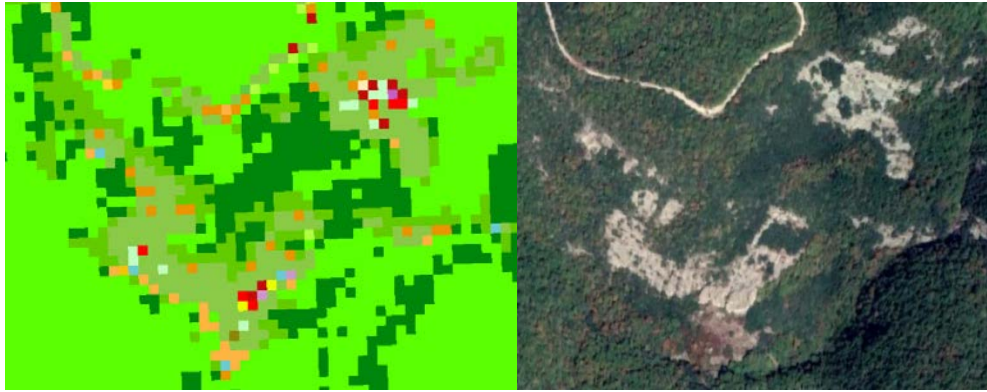
Εικόνα 5-34 Επιτυχημένη ταξινόμηση της κατηγορίας CNF

Ακόμη, επιβεβαιώνοντας τα αποτελέσματα του πίνακα σύγχυσης, διαπιστώνουμε σημαντικές συγχύσεις ανάμεσα στις κλάσεις CNF και DSV.

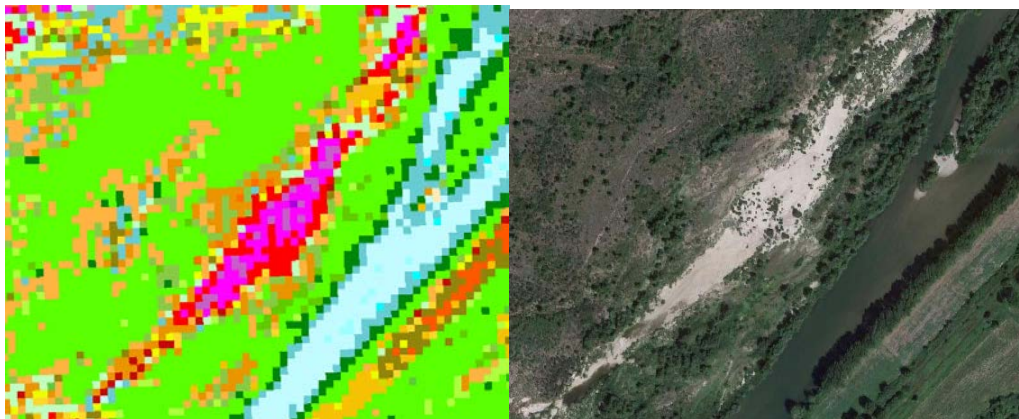


Εικόνα 5-35 Σύγχυση ανάμεσα στην DSV και στην CNF

Ως προς τις κλάσεις του γυμνού εδάφους, διαπιστώνουμε ότι δεν μπορούν να προβλεφθούν επιτυχώς στον χάρτη, καθώς τα εικονοστοιχεία τους ταξινομούνται σε πλήθος διάφορων άλλων κλάσεων. Η αδυναμία του ταξινομητή να διαχωρίσει τις συγκεκριμένες κλάσεις είναι απόρροια του μικρού αριθμού των δεδομένων εκπαίδευσης, καθώς αυτές οι κλάσεις δεν απαντώνται στα συγκεκριμένα layers.



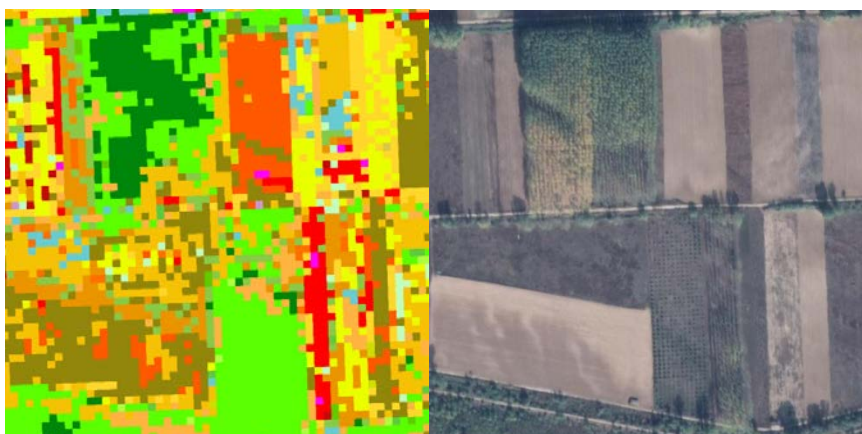
Εικόνα 5-36 Λανθασμένη απόδοση της RCK



Εικόνα 5-37 Λανθασμένη απόδοση της SND ως DUF και SUF. Λανθασμένη απόδοση της WCR ως WBD

Αντίστοιχα με τις κατηγορίες του γυμνού εδάφους και οι κατηγορίες των υδάτινων επιφανειών, δεν αποδίδονται επιτυχημένα από τον ταξινομητή, καθώς εμπλέκονται μεταξύ τους. Μεγάλη δυσκολία εμφανίζεται στην πρόβλεψη των υπορρευμάτων, καθώς τα pixels τους αποδίδονται στην κατηγορία των υδάτινων επιφανειών.

Τέλος, σχετικά με τις κατηγορίες καλλιεργειών, που απαντώνται στο συγκεκριμένο tile, επιβεβαιώνεται η αδυναμία του ταξινομητή να τις διαχωρίσει και συνεπώς να τις αποδώσει στον παραγόμενο χάρτη. Τα εικονοστοιχεία της κλάσης VNY αποδίδονται στην κλάση SUF.



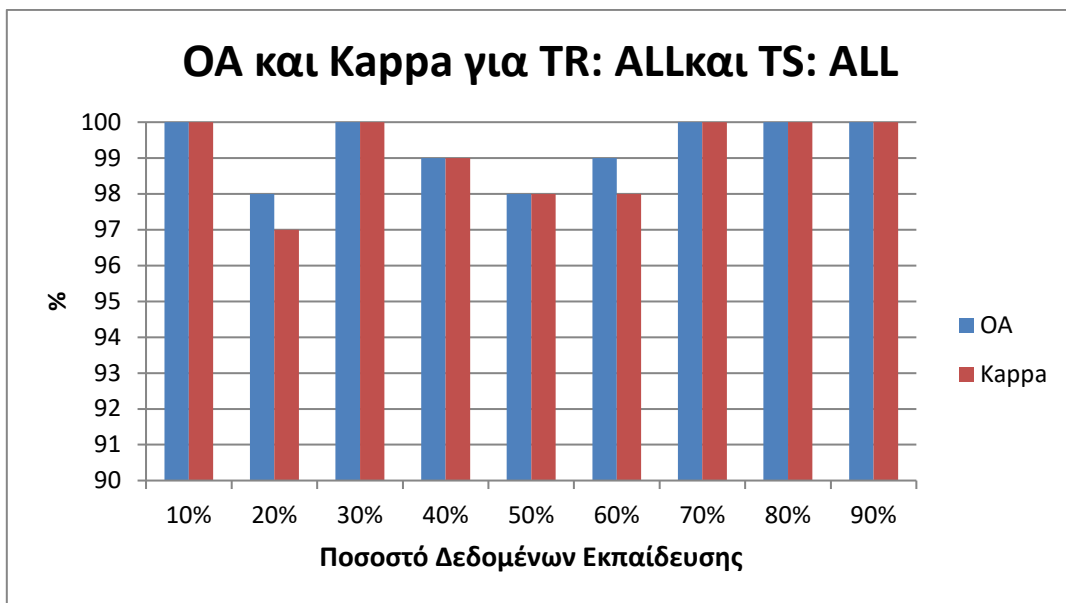
Εικόνα 5-38 Αδυναμία απόδοσης της VNY. Συμπερίληψη στην BLF και στην CNF

5.3 Τρίτη Κατηγορία- Ανά Τρία Tiles

Στην τρίτη κατηγορία πειραμάτων, χρησιμοποιήθηκαν όλα τα tiles τόσο στην εκπαίδευση, όσο και στον έλεγχο της ταξινόμησης. Εφαρμόστηκαν 9 πειράματα με διαφορετικό ποσοστό του ground truth, που κυμαίνονταν από 10 έως 90%, το οποίο χρησιμοποιήθηκε στην ταξινόμηση. Σκοπός ήταν η διερεύνηση του βέλτιστου ποσοστού δεδομένων εκπαίδευσης για την συγκεκριμένη ταξινόμηση.

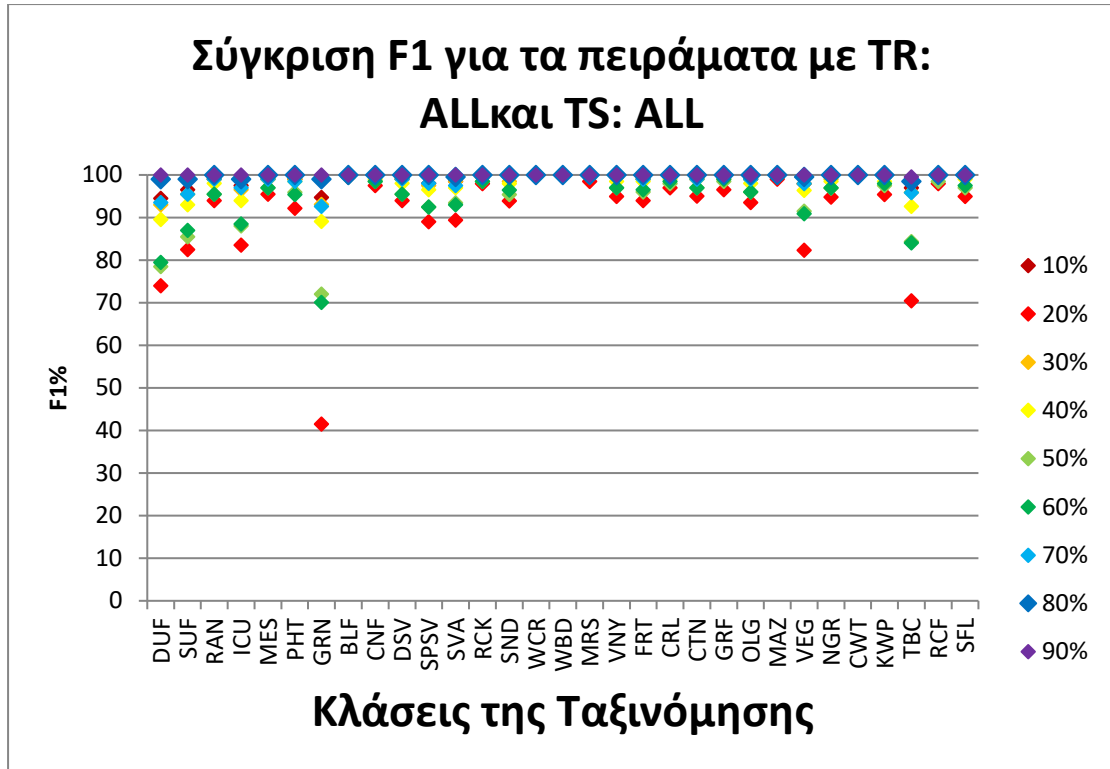
Χρησιμοποιήθηκαν και οι έξι επιλεγμένες εικόνες και από αυτές τα κανάλια 2,3,4,8,11,12, καθώς και οι δείκτες NDVI, MSAVI, NDWI, NDBI.

Τα αποτελέσματα που προέκυψαν από τις ταξινομήσεις δεν διέφεραν σημαντικά.. Η συνολική ακρίβεια της ταξινόμησης και ο δείκτης Kappa, για τα εννέα πειράματα, παρουσιάζονται στο παρακάτω γράφημα.



Γράφημα 5-12 Συνολική ακρίβεια και δείκτης Kappa για την ταξινόμηση με εκπαίδευση και έλεγχο σε όλα τα tile

Ελέγχοντας τα μεγέθη αξιολόγησης της ταξινόμησης, καθώς και το δείκτη F1 για την κάθε κλάση διαπιστώσαμε ότι και οι 9 ταξινομήσεις έχουν αποδώσει εξίσου καλά αποτελέσματα. Όμως, εκείνες που απέδωσαν τα καλύτερα ήταν αυτές για τις οποίες χρησιμοποιήθηκε ποσοστό του ground truth στην εκπαίδευση του αλγορίθμου ίσο με 10, με 30 και με ποσοστό ανώτερο του 70%. Επιλέγεται να παρατεθεί η ταξινόμηση στην οποία αξιοποιήθηκε το 70% του ground truth για εκπαίδευση του αλγορίθμου, επειδή τα μεγέθη αξιολόγησης της ακρίβειας της ταξινόμησης, δεν μεταβάλλονται αξιολογικά, χρησιμοποιώντας πέραν του 70% του ground truth για εκπαίδευση.



Γράφημα 5-13 Δείκτης F1 κάθε κλάσης για την ταξινόμηση που συμμετέχουν και τα 3 tiles

Παρακάτω παρατίθεται ο πίνακας σύγκρισης της ταξινόμησης. Η ταξινόμηση θεωρείται πετυχημένη, καθώς, η γενική ακρίβεια και ο δείκτης kappa αγγίζουν το 100%, ενώ ο δείκτης F1 για κάθε κλάση αγγίζει το 93%.

Πίνακας 5-17 Πίνακας Σύγκρισης για ταξινόμηση με δεδομένα εκπαίδευσης από την όλα τα tiles και έλεγχο σε όλα τα tiles

		ΚΛΑΣΕΙΣ ΤΟΥ ΑΠΟΤΕΛΕΣΜΑΤΟΣ ΤΗΣ ΤΑΞΙΝΟΜΗΣΗΣ																														sum	PA				
		DUF	SUF	RA	ICU	ME	PHT	GR	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MR	VNY	FRT	CRL	CTN	GRF	OLG	MA	VEG	NG	CW	KW	TBC	RCF	SFL					
ΚΛΑΣΕΙΣ ΤΟΥ ΧΑΡΤΗ ΤΟΥ GROUND IRLISH	DUF	1758	655	36	269	16	0	4	0	0	0	0	0	0	16	0	0	8	0	0	8	24	0	8	0	0	0	0	0	0	0	0	0	0	1863	0,94	
	SUF	541	2881	47	68	8	4	0	0	0	0	35	0	0	0	0	0	4	12	36	36	158	16	28	0	0	0	0	0	0	0	0	0	0	2981	0,97	
	RA	66	64	4132	144	11	0	0	0	0	0	8	0	0	8	0	0	19	0	8	12	20	0	12	0	0	0	0	0	0	0	0	0	0	4170	0,99	
	ICU	277	95	173	2452	40	0	0	0	0	0	0	0	0	3	12	0	0	0	0	0	4	16	0	0	0	0	0	0	0	0	0	0	0	2514	0,98	
	ME	20	8	48	109	3532	0	0	0	0	0	0	0	0	16	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3554	0,99	
	PHT	0	16	24	4	0	6174	0	0	0	0	4	0	0	0	0	0	4	0	12	28	16	0	0	0	0	0	0	0	0	0	0	0	0	6282	0,98	
	GR	20	20	20	28	4	0	865	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	961	0,9	
	BLF	0	0	0	0	0	0	0	5108	251	64	8	0	0	0	0	0	0	0	19	0	0	0	0	0	0	12	0	0	0	0	0	0	0	5112	1	
	CNF	0	0	0	0	0	0	0	379	4519	638	96	0	0	0	0	0	0	0	0	12	0	0	0	16	0	0	0	0	0	0	0	0	0	0	4530	1
	DSV	0	0	0	0	0	0	0	279	1251	1924	662	0	0	0	0	0	0	0	0	8	0	0	0	40	0	0	0	0	0	0	0	0	0	0	1946	0,99
	SPS	4	16	0	0	0	0	0	91	96	920	8103	4	0	0	0	0	43	0	12	0	0	0	520	0	0	0	0	0	0	0	0	0	0	0	8273	0,98
	SVA	0	12	4	0	0	0	0	0	0	0	41	4479	8	0	0	0	4	0	4	0	4	0	4	0	4	0	0	0	0	0	0	0	0	4560	0,98	
	RCK	4	12	34	4	64	0	0	0	0	0	8	16	2522	7	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	2537	0,99	
	SND	24	16	15	48	69	0	0	0	0	0	20	4	0	1424	4	0	4	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	1445	0,99	
	WC	0	0	4	0	0	0	0	0	0	0	4	0	0	0	0	8848	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8850	1	
	WB	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	12	6908	68	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	6909	1	
	MR	4	20	8	0	0	0	0	0	0	0	4	4	4	0	0	10	1481	4	20	0	20	4	16	0	0	0	0	0	0	0	0	0	0	1482	1	
	VNY	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0	16	1949	8	20	64	0	27	8	0	0	0	0	0	0	0	0	0	0	1967	0,99	
	FRT	0	27	4	0	0	0	0	77	52	19	51	0	0	0	0	44	0	5365	68	55	18	112	0	0	0	0	0	0	0	0	0	0	0	5418	0,99	
	CRL	4	15	8	0	0	0	0	0	0	0	12	0	0	0	0	27	12	36	2647	395	12	44	91	0	0	0	0	4	8	0	0	0	0	2653	1	
	CTN	0	110	0	0	0	0	0	4	0	0	16	0	0	0	0	48	16	24	895	1854	0	11	0	0	0	0	0	0	3	0	4	1866	0,99			
	GRF	0	8	0	0	0	0	0	8	0	0	12	0	0	0	0	4	0	16	55	20	6007	4	12	0	0	0	8	0	16	4	6024	1				
	OLG	0	44	12	0	0	0	0	24	177	127	453	0	0	0	0	63	8	36	48	44	12	1076	0	0	0	0	8	4	0	0	0	1087	0,99			
	MA	0	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	12	43	16	0	1662	0	0	0	4	4	0	4	1663	1			
VEG	4	28	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	8	0	87	142	8	8	8	7617	0	0	8	8	0	8	7930	0,96				
NG	0	0	0	0	0	0	0	54	16	0	0	0	0	0	0	0	0	0	36	0	0	4	4	0	0	9433	0	0	0	0	0	0	9547	0,99			
CW	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	7186	0	0	0	0	0	7186	1			
KW	0	8	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	4	8	4	4	47	8	12	0	0	0	1917	0	0	0	0	1926	0,99			
TBC	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	78	109	4	4	0	0	0	0	0	4150	0	0	0	4357	0,95		
RCF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4	4	7	8	0	0	0	0	0	0	0	0	1169	0	1171	1			
SFL	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	23	4	16	0	19	0	0	0	0	0	0	1421	1430	0,99				
sum	1855	3006	4177	2520	3553	6178	869	5118	4537	1942	8247	4507	2525	1431	8850	6908	1485	1956	5396	2661	1866	6024	1085	1664	7617	9445	7186	1919	4177	1171	1423	kapp	1				
UA	0,95	0,96	0,99	0,97	0,99	1	1	1	1	0,99	0,98	0,99	1	1	1	1	1	1	1	0,99	0,99	0,99	1	0,99	1	1	1	1	1	0,99	1	1	OA:	1			

Όπως βλέπουμε στον παραπάνω πίνακα σύγχυσης η ταξινόμηση παρουσιάζει άριστα ποσοστά συνολικής ακρίβειας και δείκτη kappa. Συνεπώς, θεωρείται πως οι κλάσεις της ταξινόμησης, διαχωρίστηκαν επαρκώς από τον ταξινομητή.

Εκτός από τα μεγέθη που αφορούν την συνολική ακρίβεια της ταξινόμησης, σκόπιμο είναι να διαπιστωθεί αν και κατά πόσο υπάρχουν σφάλματα ταξινόμησης αλλά και λανθασμένα ταξινομημένα εικονοστοιχεία για την κάθε κατηγορία ξεχωριστά.

Πιο συγκεκριμένα, για τις κλάσεις που ανήκουν στην γενική κατηγορία των τεχνητών επιφανειών, διαπιστώνεται ότι όλες οι κατηγορίες ταξινομήθηκαν με ευκολία, αφού τα ποσοστά ακρίβειας και αξιοπιστίας του ξεπερνούν το 90% σε κάθε περίπτωση. Μικρή δυσκολία του ταξινομητή εμφανίζεται στο να διαχωρίσει πλήρως τις κλάσεις DUF, SUF, και ICU. Τα σφάλματα παράλειψης και συμπερίληψης σε αυτές τις κλάσεις κυμαίνονται από το 1,5% έως το 3,6%, με τη σημαντικότερη σύγχυση να εμφανίζεται ανάμεσα στις κλάσεις DUF και SUF. Αυτή η δυσκολία ήταν αναμενόμενη καθώς, η κατηγορία DUF και SUF, είναι υποκατηγορίες του αστικού ιστού και άρα εμπλεκούνται ίδιου τύπου αντικείμενα που εκπέμπουν με παρόμοιο τρόπο στο ηλεκτρομαγνητικό φάσμα.

Η κατηγορία ICU, εμπλέκεται σε μικρότερο βαθμό με την κατηγορία DUF. Και αυτή η σύγχυση είναι αναμενόμενη, αφού η κατηγορία απαρτίζεται από περιοχές με τεχνητές επιφάνειες τσιμέντου και ασφάλτου, υλικά που εκπέμπουν παρόμοια με τις κατηγορίες DUF, SUF.

Οι κατηγορίες RAN, MES, PHT, χαρακτηρίζονται από πολύ υψηλά ποσοστά ακρίβειας χρήστη και παραγωγού, ενώ δεν απαντώνται συγχύσεις με κάποια κατηγορία.

Τέλος, η κατηγορία GRN ταξινομείται ικανοποιητικά από τον αλγόριθμο, όμως εμφανίζει μικρά σφάλμα παράλειψης (2% έως 3%) με τις κατηγορίες ICU, DUF, SUF, RAN.

Οι κατηγορίες φυσικής βλάστησης εμφανίζουν και εκείνες πολύ υψηλά ποσοστά ακρίβειας και αξιοπιστίας. Οι συγχύσεις είναι πολύ μικρές, καθώς τα σφάλματα παράλειψης και συμπερίληψης δεν υπερβαίνουν τα 1%. Αξίζει, όμως να σημειωθεί ότι οι μικρές συγχύσεις που παρατηρούνται είναι ανάμεσα στην κλάση DSV με την κλάση CNF και την κλάση SPSV, και της SPSV με την DSV και την OLG. Οι συγχύσεις αυτές, έχουν παρατηρηθεί σε όλα τα πειράματα, τα οποία υλοποιήθηκαν, και όπως έχει ήδη αναφερθεί είναι προϊόν των παρόμοιων φασματικών χαρακτηριστικών αυτών των κλάσεων.

Ακόμη, η κατηγορία NGR, χαρακτηρίζεται από πολύ υψηλά ποσοστά ακρίβειας και αξιοπιστίας, αξίζει όμως να αναφερθεί η μικρή σύγχυση που εμφανίζει με την κλάση BLF, μία σύγχυση, που έχει παρατηρηθεί εν γένει σε όλες τις ταξινομήσεις που έχουν λάβει χώρα.

Σε σχέση με τις κατηγορίες του γυμνού εδάφους, ο ταξινομητής πέτυχε σε όλες πολύ υψηλά ποσοστά ακρίβειας παραγωγού και χρήστη. Αν εξετάσουμε τα μικρά σφάλματα παράλειψης και συμπερίληψης, καταλήγουμε στο γεγονός ότι η κατηγορία SVA, εμπλέκεται σε μικρό βαθμό με την κατηγορία SPSV και με την κατηγορία RCK. Αυτά τα σφάλματα έχουν παρατηρηθεί σε όλες τις ταξινομήσεις και απορρέουν από την παρόμοια φασματική υπογραφή αυτών των κλάσεων, μιας και τα δομικά τους υλικά

συμμετέχουν σε όλες αυτές τις κλάσεις και εμπεριέχονται σε όλων τα εικονοστοιχεία του ground truth.

Στην γενική κατηγορία των υδάτινων πόρων διαπιστώνεται ότι όλες οι τάξεις ταξινομήθηκαν επιτυχώς και δεν παρατηρείται κάποια αξιολογη σύγχυση.

Ως προς τις κλάσεις που ανήκουν στις καλλιέργειες διαπιστώνουμε ότι, ο ταξινομητής πετυχαίνει να διαχωρίσει πλήρως τις επιμέρους τάξεις. Αν και αξιολογες συγχύσεις δεν παρατηρούνται αξίζει να αναφερθούμε σε ορισμένες.

Οι κλάσεις CRL και CTN εμφανίζουν μικρές συγχύσεις μεταξύ τους.

Ορισμένα pixels της κλάσης OLG αποδίδονται στην κλάση SPSV.

Η κλάση των VEG μπλέκεται σε μικρό βαθμό με τις κλάσεις CRL και CTN.

Τέλος, η κατηγορία TBC, μπλέκεται σε μικρό βαθμό με τις κλάσεις CRL και CTN.

Αυτές οι συγχύσεις, είναι αναμενόμενες και προκύπτουν λόγω των παρόμοιων φασματικών χαρακτηριστικών των κλάσεων αυτών.

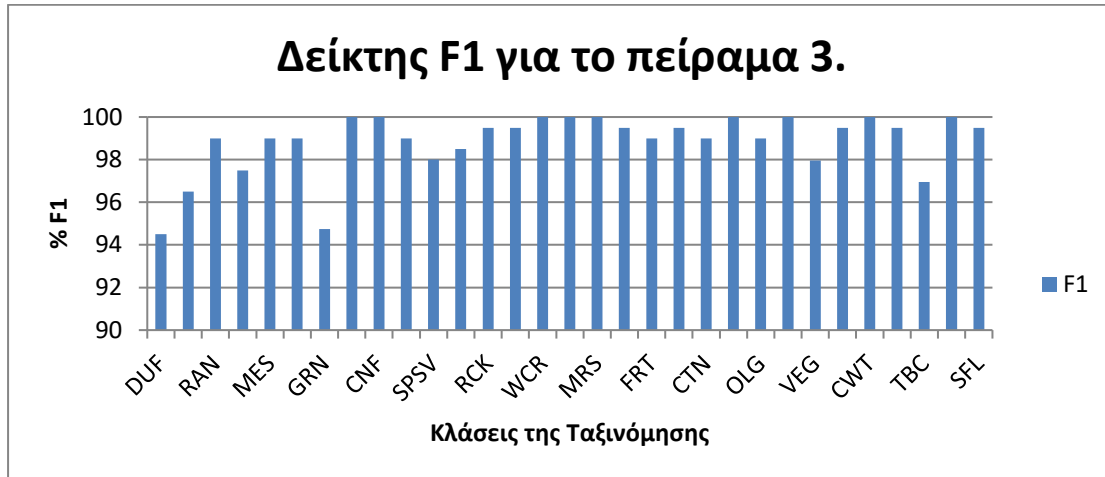
Πίνακας 5-18 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF

	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MRS	VNY	FRT	CRL	CTN	GRF	OLG	MAZ	VEG	NGR	CWT	KW	TBC	RCF	SFL
DUF	94,4	1,81	0,16	1,10	0,06	0,00	2,08	0,00	0,00	0,00	0,00	0,00	0,02	0,17	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00
SUF	3,52	96,6	0,15	0,38	0,02	0,25	2,08	0,00	0,00	0,00	0,02	0,26	0,05	0,11	0,00	0,00	0,01	0,20	0,05	0,01	0,06	0,01	0,04	0,00	0,35	0,00	0,00	0,04	0,28	0,00	0,10
RAN	0,19	0,16	99,1	0,69	0,14	0,38	2,08	0,00	0,00	0,00	0,00	0,09	0,13	0,10	0,00	0,00	0,01	0,00	0,01	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
ICU	1,44	0,23	0,35	97,5	0,31	0,06	2,91	0,00	0,00	0,00	0,00	0,00	0,02	0,33	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MES	0,09	0,03	0,03	0,16	99,3	0,00	0,42	0,00	0,00	0,00	0,00	0,00	0,25	0,48	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
PHT	0,00	0,01	0,00	0,00	0,00	98,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
GRN	0,02	0,00	0,00	0,00	0,00	0,00	90,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,9	0,08	0,14	0,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,00	0,00	0,57	0,00	0,00	0,00	0,00	0,00
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	99,7	0,64	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00	0,00	0,00	0,16	0,00	0,00	0,17	0,00	0,00	0,00	0,00	0,00
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,14	98,8	1,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SPS	0,00	0,12	0,02	0,00	0,00	0,06	0,42	0,00	0,02	0,34	97,9	0,90	0,03	0,14	0,00	0,00	0,00	0,00	0,09	0,00	0,01	0,02	0,42	0,00	0,05	0,00	0,00	0,02	0,00	0,00	0,00
SVA	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	98,2	0,06	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
RCK	0,00	0,00	0,00	0,01	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,18	99,4	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SND	0,09	0,00	0,02	0,05	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	98,5	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	99,9	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WB	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	99,9	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
MRS	0,04	0,01	0,05	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,05	0,09	0,00	0,03	0,01	0,01	99,9	0,08	0,08	0,01	0,03	0,01	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
VNY	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,0	0,00	0,00	0,01	0,00	0,01	0,00	0,10	0,00	0,00	0,02	0,00	0,03	0,00
FRT	0,00	0,12	0,02	0,00	0,00	0,19	0,00	0,00	0,00	0,01	0,09	0,00	0,00	0,00	0,00	0,00	0,01	0,04	99,0	0,01	0,01	0,03	0,03	0,00	0,00	0,38	0,00	0,04	0,00	0,03	0,06
CRL	0,04	0,12	0,03	0,02	0,00	0,45	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,10	0,13	99,7	0,48	0,09	0,04	0,01	1,10	0,00	0,00	0,02	1,79	0,03	0,16
CTN	0,13	0,53	0,05	0,06	0,00	0,25	0,00	0,00	0,00	0,00	0,00	0,09	0,00	0,00	0,00	0,00	0,01	0,33	0,10	0,15	99,3	0,03	0,04	0,03	1,79	0,00	0,00	0,02	2,50	0,06	0,03
GRF	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	99,7	0,01	0,01	0,10	0,04	0,00	0,24	0,09	0,07	0,11	
OLG	0,04	0,09	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,63	0,09	0,00	0,00	0,00	0,00	0,01	0,14	0,21	0,02	0,01	0,01	99,0	0,00	0,10	0,04	0,00	0,04	0,09	0,00	0,00
MAZ	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,03	0,00	0,02	0,00	99,9	0,10	0,00	0,06	0,00	0,00	0,13	
VEG	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	96,0	0,00	0,00	0,00	0,00	0,00	
NGR	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	98,8	0,00	0,00	0,00	0,00	
CWT	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,	0,00	0,00	0,00	0,00	
KW	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,00	99,4	0,00	0,00	0,00	
TBC	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00	0,00	95,2	0,00	0,00	
RCF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,10	0,00	0,00	0,00	99,7	0,00	
SFL	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,4	

Πίνακας 5-19 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF

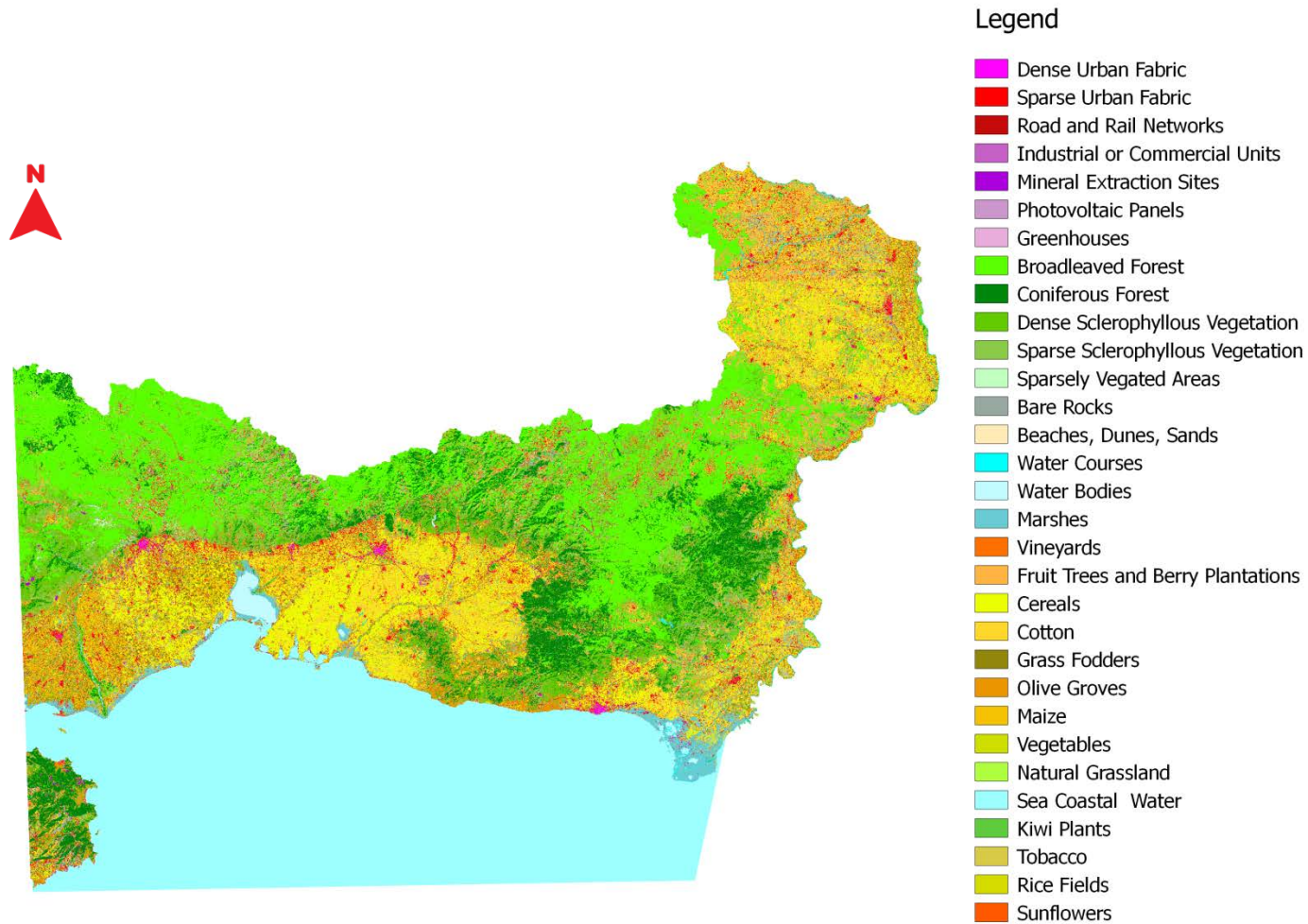
	DUF	SUF	RAN	ICU	MES	PHT	GRN	BLF	CNF	DSV	SPS	SVA	RCK	SND	WC	WB	MRS	VNY	FRT	CRL	CTN	GRF	OLG	MAZ	VEG	NGR	CWT	KW	TBC	RCF	SFL
DUF	94,7	2,18	0,09	1,07	0,05	0,00	0,46	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,00	0,00	0,01	0,00	0,00	0,00	0,01	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SUF	2,92	95,8	0,11	0,27	0,02	0,06	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,07	0,01	0,08	0,03	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
RAN	0,36	0,21	98,9	0,57	0,03	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,06	0,00	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
ICU	1,49	0,32	0,41	97,3	0,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,08	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
MES	0,11	0,03	0,11	0,43	99,4	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
PHT	0,00	0,05	0,06	0,02	0,00	99,9	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
GRN	0,11	0,07	0,05	0,11	0,01	0,00	99,5	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
BLF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,8	0,06	0,03	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,00	0,00	
CNF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	99,5	0,33	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
DSV	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,28	99,0	0,80	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
SPS	0,02	0,05	0,00	0,00	0,00	0,00	0,00	0,02	0,02	0,47	98,2	0,09	0,00	0,00	0,00	0,00	0,03	0,00	0,02	0,00	0,00	0,00	0,48	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
SVA	0,00	0,04	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	99,3	0,03	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
RCK	0,02	0,04	0,08	0,02	0,18	0,00	0,00	0,00	0,00	0,01	0,36	99,8	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
SND	0,13	0,05	0,04	0,19	0,19	0,00	0,00	0,00	0,00	0,00	0,02	0,09	0,00	99,5	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
WC	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,9	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
WB	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	100,	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
MRS	0,02	0,07	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,09	0,02	0,00	0,00	0,00	99,7	0,02	0,04	0,00	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
VNY	0,00	0,13	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	99,6	0,01	0,01	0,03	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
FRT	0,00	0,09	0,01	0,00	0,00	0,00	0,00	0,02	0,01	0,01	0,06	0,00	0,00	0,00	0,00	0,00	0,03	0,00	99,4	0,03	0,03	0,03	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
CRL	0,02	0,05	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,02	0,06	0,07	99,4	0,21	0,02	0,04	0,05	0,00	0,00	0,00	0,02	0,19	0,00	
CTN	0,00	0,37	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,03	0,08	0,04	0,34	99,3	0,00	0,01	0,00	0,00	0,00	0,00	0,07	0,00	0,03	
GRF	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,02	0,01	99,7	0,00	0,01	0,00	0,00	0,00	0,04	0,00	0,14	0,03	
OLG	0,00	0,15	0,03	0,00	0,00	0,00	0,00	0,00	0,04	0,07	0,55	0,00	0,00	0,00	0,00	0,04	0,04	0,07	0,02	0,02	0,02	99,1	0,00	0,00	0,00	0,00	0,04	0,10	0,00	0,00	
MAZ	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,02	0,03	0,00	99,9	0,00	0,00	0,00	0,02	0,10	0,00	0,03	
VEG	0,02	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,03	0,08	0,01	0,01	0,00	100,	0,00	0,00	0,19	0,07	0,00		
NGR	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	0,00	0,00	0,01	0,00	0,00	0,00	99,8	0,00	0,00	0,00	0,00	0,00	
CWT	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,	0,00	0,00	0,00	0,00	0,00	
KW	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,01	0,00	0,00	0,08	0,01	0,01	0,00	0,00	99,8	0,00	0,00	0,00	0,00	
TBC	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,06	0,01	0,00	0,00	0,00	0,00	0,00	99,3	0,00	0,00	0,00	
RCF	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	99,8	0,00	0,00	
SFL	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01	0,00	0,03	0,00	0,01	0,00	0,00	0,00	0,00	0,00	99,9	

Στην συνέχεια παρατίθεται το γράφημα για τον δείκτη Fisher (F1) της βέλτιστης ταξινόμησης, που αναλύθηκε παραπάνω. Στο συγκεκριμένο γράφημα εμφανίζεται το ποσοστό του F1 για κάθε κλάση της ταξινόμησης. Με το γράφημα αυτό, μπορούμε και εποπτικά να εντοπίσουμε τις κατηγορίες εκείνες, που απέδωσαν καλύτερα, αλλά και τις κατηγορίες αυτές, που εμφάνισαν συγχύσεις.



Γράφημα 5-14 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMF

Όπως διαπιστώνεται, ο ταξινομητής έχει διαχωρίσει όλες τις κατηγορίες επιτυχημένα, καθώς ο δείκτης F1 είναι ανώτερος του 94%, για κάθε κλάση. Οι κατηγορίες όμως, οι οποίες διαχωρίστηκαν βέλτιστα, σύμφωνα με το δείκτη F1 είναι οι κατηγορίες των υδάτινων επιφανειών, των δασών και των περισσότερων καλλιεργειών, ενώ οι κατηγορίες οι οποίες εμφάνισαν χαμηλότερο δείκτη Fisher, ήταν εκείνων των τεχνητών επιφανειών και ιδίως η DUF και η GRN. Τέλος, παρατίθεται ο χάρτης, που παρήχθη από το πείραμα αυτό.

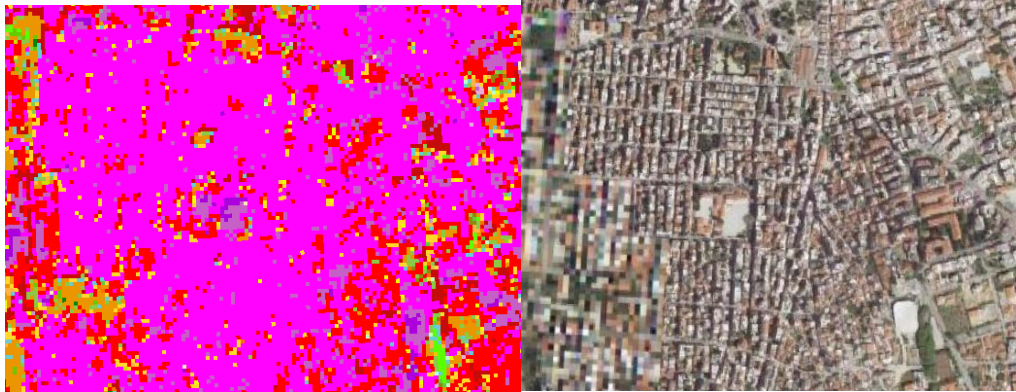


Χάρτης 5.8 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG

Μελετώντας τον χάρτη, διαπιστώνεται ότι οι γενικές κατηγορίες (τεχνητές επιφάνειες, φυσική βλάστηση, γυμνό έδαφος, καλλιέργειες, υδάτινες επιφάνειες) έχουν αποδοθεί σωστά.

Εάν παρατηρήσουμε τον χάρτη λεπτομερώς θα διαπιστώσουμε ορισμένα σφάλματα του ταξινομητή στην απόδοση των κλάσεων. Τα περισσότερα εντοπίζονται στις κατηγορίες των τεχνητών επιφανειών, ενώ μικρές συγχύσεις εμφανίζονται ανάμεσα στις κλάσεις των καλλιεργειών.

Πιο συγκεκριμένα, η κατηγορία DUF, εμφανίζει μικρές συγχύσεις με την SUF.



Εικόνα 5-39 Μικρή σύγχυση της DUF με την SUF

Επιπροσθέτως, η ICU εμφανίζει μικρές συγχύσεις με την DUF, αλλά και με την MES.



Εικόνα 5-40 Σύγχυση ανάμεσα στην ICU και τις DUF- MES

Ως προς τις κλάσεις των καλλιεργειών, εντοπίζονται μικρές συγχύσεις, με σημαντικότερες την απόδοση εικονοστοιχείων της VEG στην CTN και στην CRL και την απόδοση εικονοστοιχείων της κλάσης TBC στις κλάσεις CTN, CRL.



Εικόνα 5-41 Μικρή σύγχυση ανάμεσα στην VEG και στην CTN

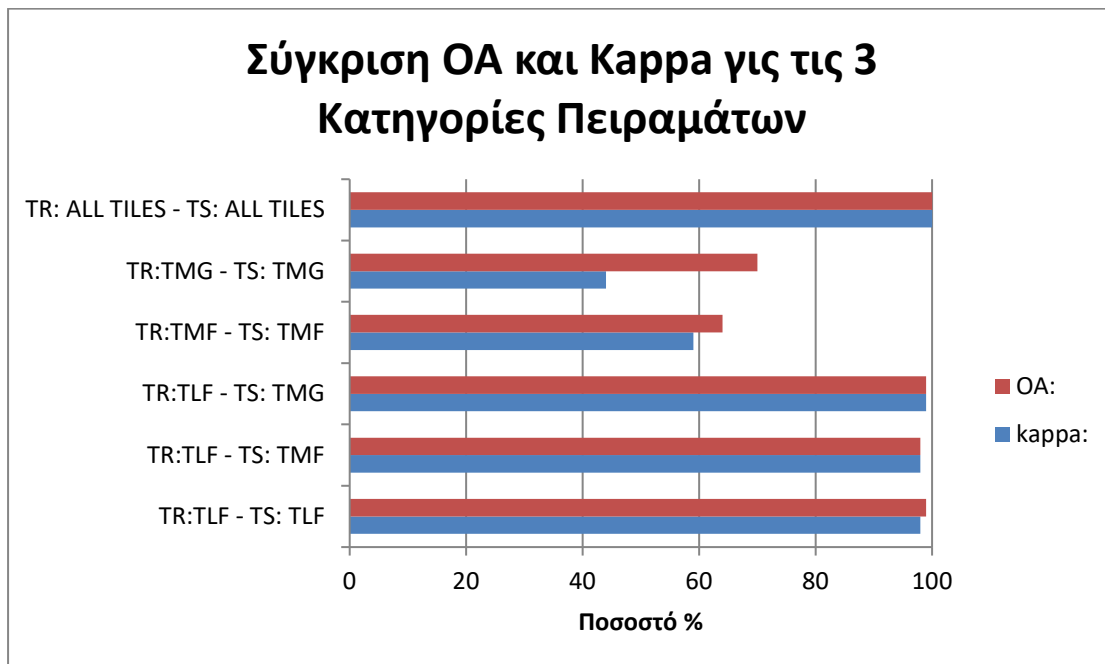


Εικόνα 5-42 Μικρή σύγχυση ανάμεσα στην TBC και στην CTN

5.4 Συγκριτική Αξιολόγηση των Πειραμάτων

Στην συγκεκριμένη παράγραφο, συγκρίνονται τα αποτελέσματα των τριών ομάδων πειραμάτων. Καθώς, στην πρώτη και στην τρίτη κατηγορία πειραμάτων, υλοποιήθηκαν πολλαπλά πειράματα, σε μια προσπάθεια διερεύνησης του βέλτιστου ποσοστού δεδομένων εκπαίδευσης, επιλέγεται να συγκριθούν τα βέλτιστα πειράματα των δύο αυτών κατηγοριών, Μέσα από αυτήν την σύγκριση εξάγονται σημαντικές παρατηρήσεις.

Αρχικά συγκρίνονται τα ποσοστά της συνολικής ακρίβειας και του δείκτη Kappa των 6 πειραμάτων, που αναλύθηκαν στην προηγούμενη παράγραφο.



Γράφημα 5-15 Σύγκριση της συνολικής ακρίβειας και του δείκτη Kappa των 6 αναλυθέντων πειραμάτων

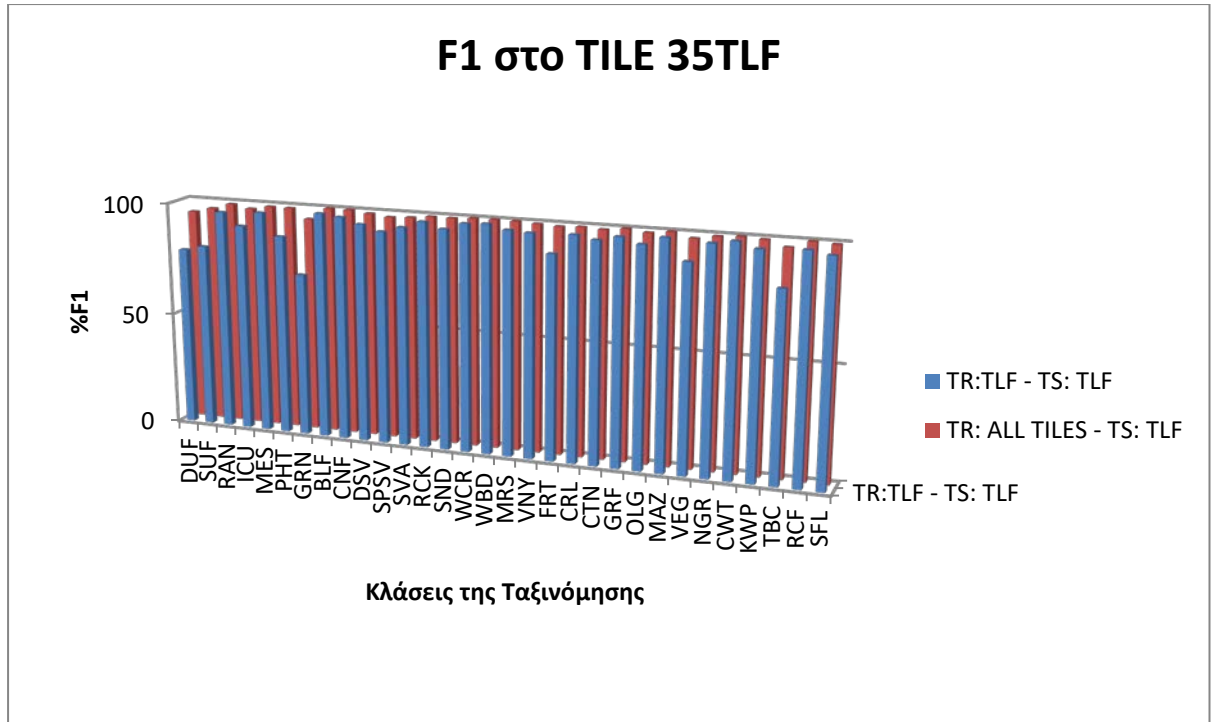
Διαπιστώνεται ότι τα βέλτιστα αποτελέσματα επιτυγχάνονται όταν ο ταξινομητής εκπαιδεύεται και ελέγχεται χρησιμοποιώντας τα δεδομένα εκπαίδευσης των τριών tiles, ενώ τα χειρότερα αποτελέσματα προκύπτουν όταν ο αλγόριθμος εκπαιδεύεται και ελέγχεται με ground truth διαφορετικών tiles.

Στην συνέχεια, κρίνεται σκόπιμο να ελεγχθεί, πως επηρεάζεται η κάθε κλάση ανά διαφορετικό πείραμα. Η σύγκριση αυτή έχει νόημα να εφαρμοσθεί ανά tile, επειδή πρώτον δεν απαντώνται όλες οι κλάσεις σε όλα τα tile και δεύτερον γιατί τα tiles 35TMG και 35TMF, συμμετέχουν και στις 3 ομάδες πειραμάτων, καθιστώντας έτσι δυνατή την σύγκριση των αποτελεσμάτων των τριών κατηγοριών.

Για να επιτευχθεί αυτή η σύγκριση, επιλέγεται να χρησιμοποιηθεί ο δείκτης F1.

Στην συνέχεια, παρατίθενται τα γραφήματα, που παρουσιάζουν τον δείκτη F1 για την κάθε κλάση σε κάθε ένα tile.

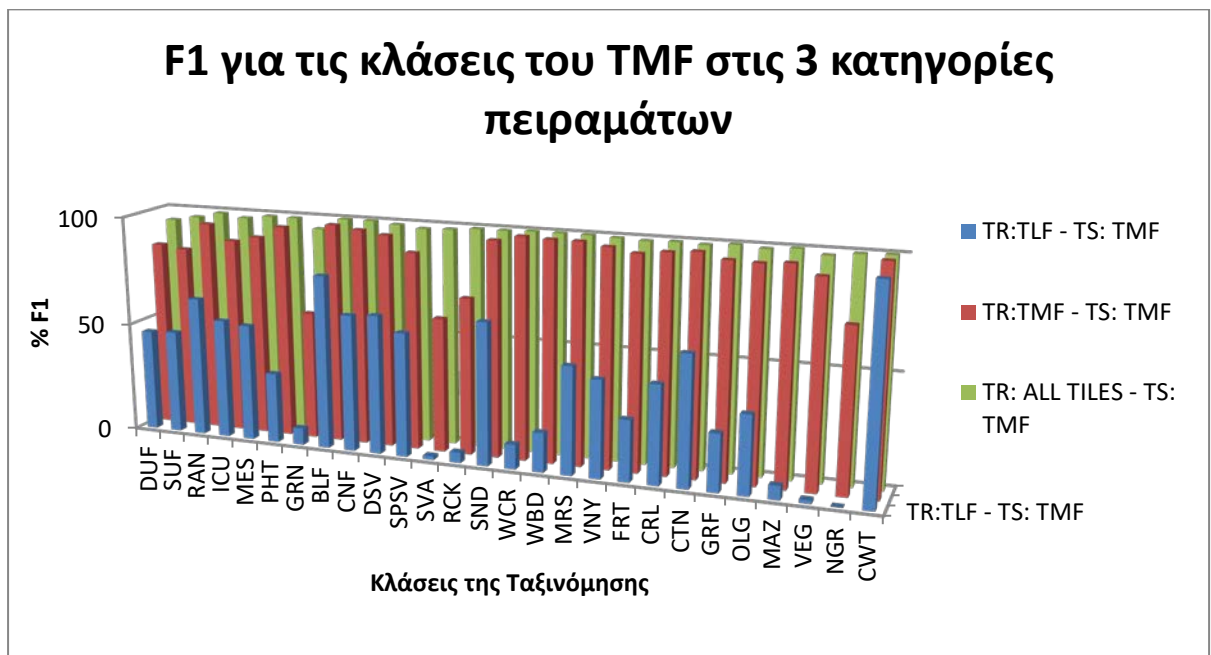
Για το tile 35TLF:



Γράφημα 5-16 Σύγκριση του δείκτη F1 για κάθε κλάση για τις κατηγορίες πειραμάτων 1 και 3 για το tile 35TLF

Το συγκεκριμένο tile αξιοποιήθηκε για έλεγχο στις δύο κατηγορίες πειραμάτων, την πρώτη και την τρίτη. Διαπιστώνεται ότι όλες οι κλάσεις πετυχαίνουν καλύτερα αποτελέσματα, όταν ο αλγόριθμος εκπαιδεύεται με δεδομένα εκπαίδευσης από όλα τα tiles.

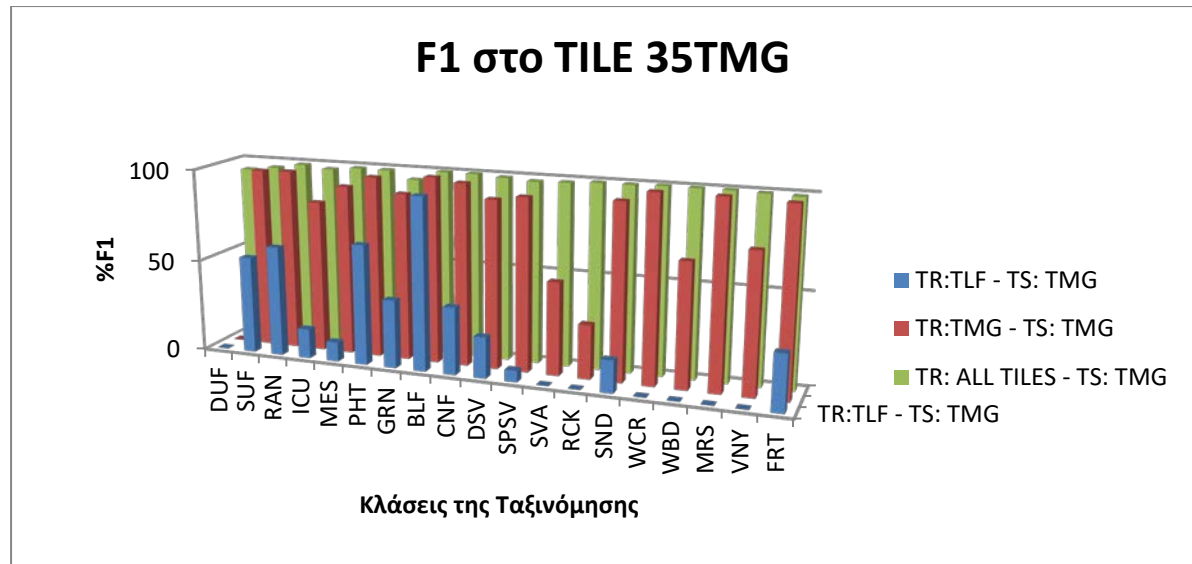
Για το tile 35TMF:



Γράφημα 5-17 Σύγκριση του δείκτη F1 για κάθε κλάση για τις τρεις κατηγορίες πειραμάτων για το tile 35TMF

Το συγκεκριμένο tile συμμετείχε και στις 3 κατηγορίες πειραμάτων. Όπως διαπιστώνουμε από το γράφημα του δείκτη F1, όλες οι κλάσεις απέδωσαν τα χαμηλότερα αποτελέσματα, όταν ο αλγόριθμος εκπαιδεύτηκε και ελέγχθηκε με δεδομένα από δύο διαφορετικά tiles. Όταν τα δεδομένα εκπαίδευσης προήλθαν από το ίδιο tile, οι αποδόσεις ήταν ιδιαίτερα υψηλές. Βέλτιστες, όμως ήταν όταν αξιοποιήθηκαν δεδομένα εκπαίδευσης και από τα τρία tiles.

Για το tile 35TMF:



Γράφημα 5-18 Σύγκριση του δείκτη F1 για κάθε κλάση για τις τρεις κατηγορίες πειραμάτων για το tile 35TMG

Αντίστοιχα ήταν τα αποτελέσματα των συγκρίσεων του F1 δείκτη και για το tile 35TMG, καθώς τα χειρότερα αποτελέσματα προέκυψαν όταν τα δεδομένα εκπαίδευσης και ελέγχου προέρχονταν από διαφορετικά tiles και τα καλύτερα όταν τα δεδομένα εκπαίδευσης και ελέγχου προέρχονταν κι από τα τρία tiles. Αξιοσημείωτο είναι το γεγονός, ότι κλάσεις οι οποίες στην πρώτη και στη δεύτερη κατηγορία πειραμάτων, απέδωσαν χαμηλά έως μέτρια αποτελέσματα, στην τρίτη κατηγορία πειραμάτων αύξησαν θεαματικά το δείκτη Fisher.

Εάν συγκρίνουμε τους παραγόμενους χάρτες, των τριών πειραματικών κατηγοριών, θα διαπιστώσουμε και πάλι, ότι η πρώτη κατηγορία πειραμάτων, αποδίδει αξιόπιστο εικονιστικό αποτέλεσμα, καθώς τα λάθη της χαρτογραφικής απόδοσης είναι περιορισμένα. Αντίθετα, ο χάρτης που εξάγεται από τα πειράματα της δεύτερης κατηγορίας, είναι αξιόπιστος ως προς τις γενικές κατηγορίες καλύψεων γης, εμφανίζει όμως σημαντικές συγχύσεις, ιδίως σε τάξεις με χαμηλό αριθμό δεδομένων εκπαίδευσης. Τέλος, ο χάρτης που παράγεται από το πείραμα της τρίτης κατηγορίας πειραμάτων, είναι ο πιο ακριβής, αφού η πλειοψηφία των κλάσεων αποδίδεται σωστά και οι συγχύσεις, που εμφανίζονται είναι λιγοστές και άνευ σημασίας.

Αξίζει να σημειωθεί ότι ο χάρτης που παράχθηκε για το tile 35TMG, από το πείραμα της πρώτης κατηγορίας, με εκπαίδευση και έλεγχο του ταξινομητή με ground truth από το tile αυτό, ενώ έχει αποδώσει σωστά τις περισσότερες κατηγορίες, δεν έχει αποδώσει σωστά τις κατηγορίες βλάστησης, σε αντίθεση με τους χάρτες της κατηγορίας 2 και 3.

Αυτό συμβαίνει, καθώς στο tile αυτό δεν διατίθεται ground truth πολλαπλών κατηγοριών βλάστησης, με αποτέλεσμα, ο ταξινομητής να αποδίδει εικονοστοιχεία κατηγοριών βλάστησης για τα οποία δεν έχει εκπαιδευτεί σε λανθασμένες κατηγορίες. Αντίθετα, όταν αξιοποιείται ground truth και από άλλα tiles, ο ταξινομητής εκπαιδεύεται κατάλληλα και αποδίδει σωστότερα τις κατηγορίες καλλιεργειών.

Πίνακας 5-20: Διαφοροποίηση στον παραγόμενο χάρτη του tile 35TMG.

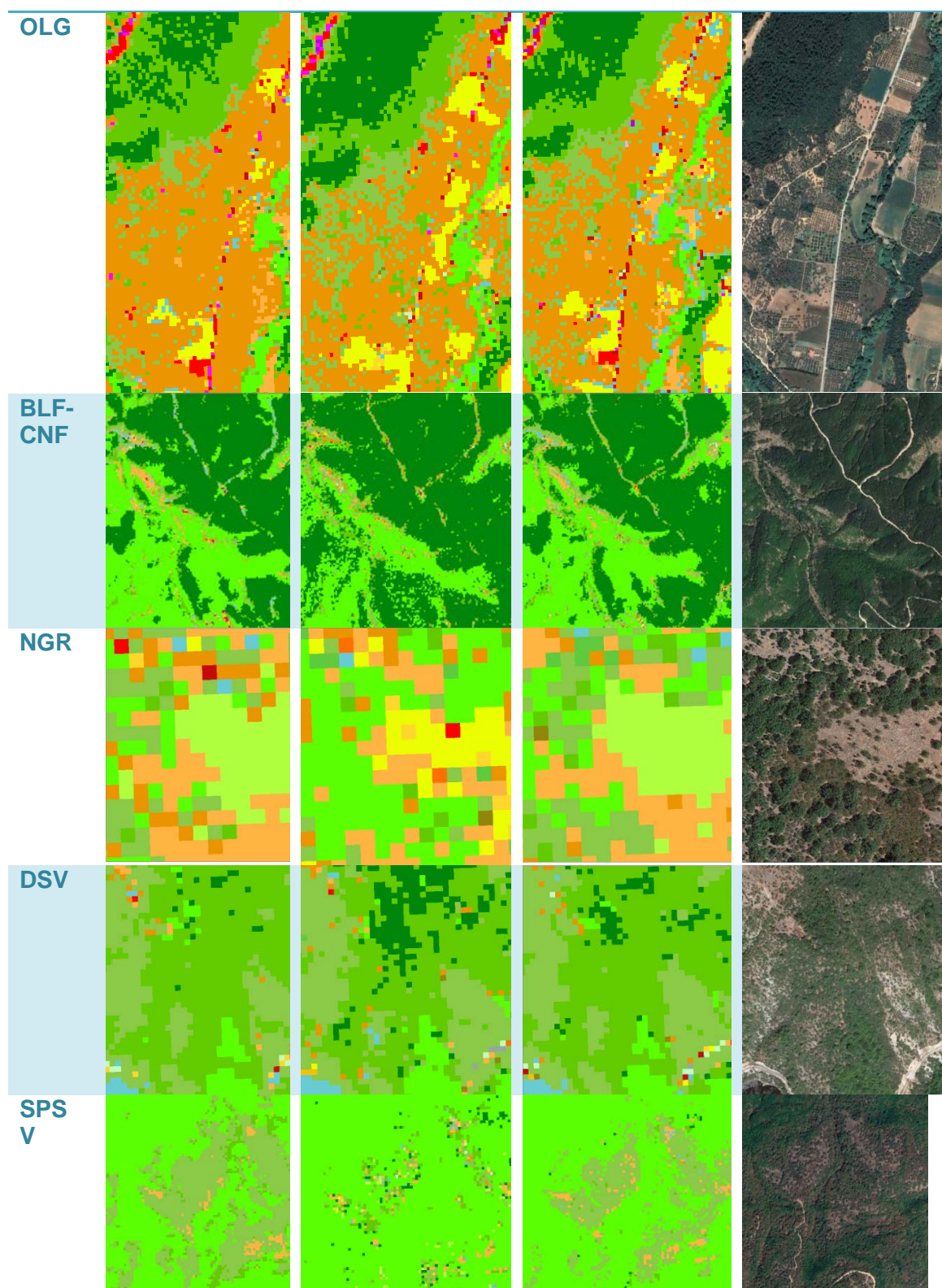


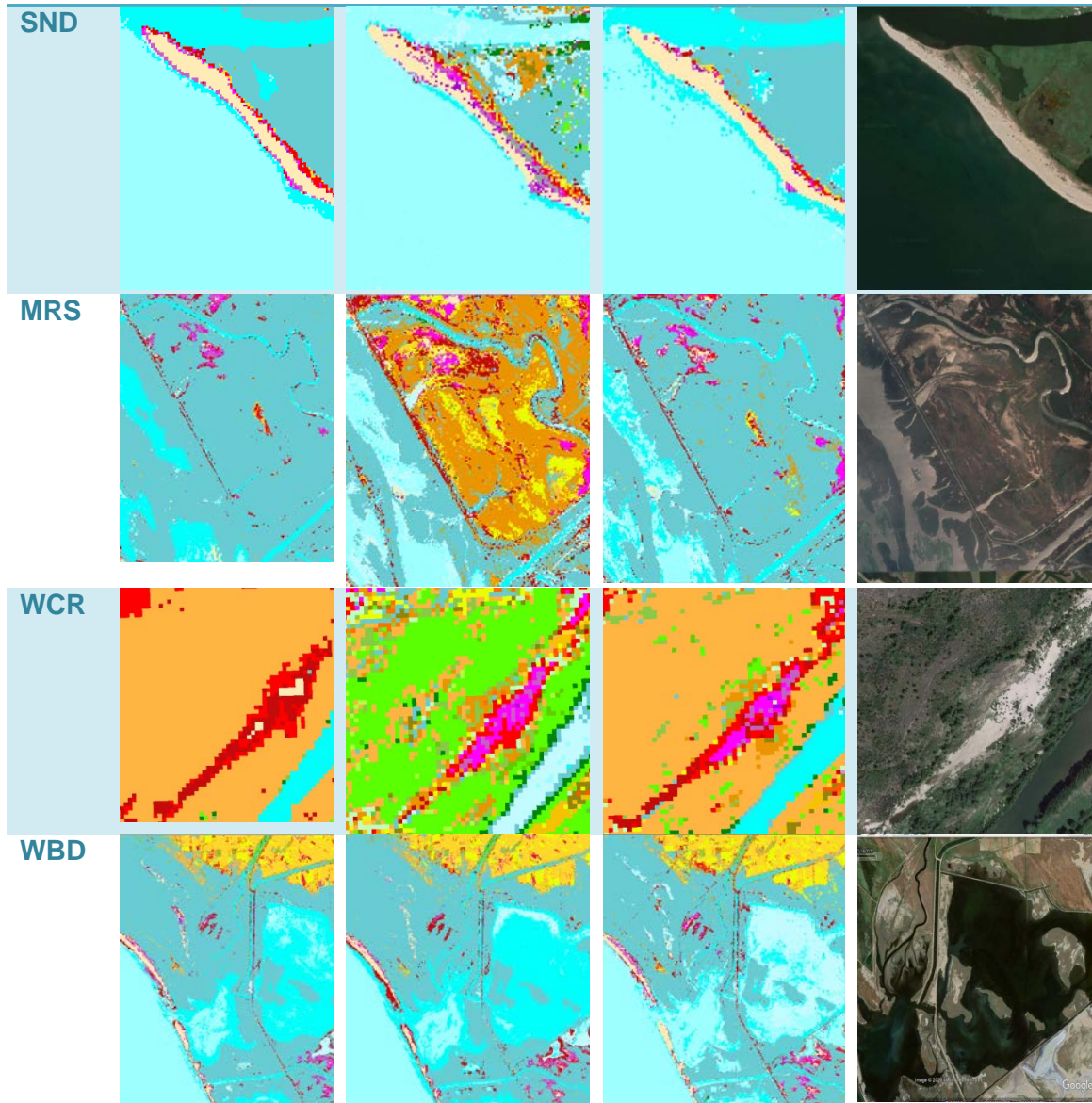
Παρακάτω παρατίθεται ένας πίνακας στον οποίο παρουσιάζονται διάφορες κλάσεις της ταξινόμησης, έτσι όπως απεικονίζονται στους παραγόμενους χάρτες.

Πίνακας 5-21 Σύγκριση των κλάσεων όπως αποδίδονται στους χάρτες των τριών κατηγοριών

Κλάση	1 ^η Κατηγορία	2 ^η Κατηγορία	3 ^η Κατηγορία	4 ^η Κατηγορία
DUF				
SUF				
ICU				
MES				
PHT				
GRN				







Ελέγχοντας τις διαφοροποιήσεις στους παραγόμενους χάρτες των τριών κατηγοριών πειραμάτων, διαπιστώνουμε ότι τόσο η πρώτη όσο και η τρίτη πειραματική κατηγορία, αποδίδουν αξιόπιστα χαρτογραφικά αποτελέσματα. Οι παραγόμενοι χάρτες είναι παρόμοιοι, χωρίς αξιολογες διαφοροποιήσεις. Ορισμένες κλάσεις αποδίδονται καλύτερα στους χάρτες της πρώτης κατηγορίας (DUF, SUF, GRN, FRT, DSV), άλλες αποδίδονται σχεδόν πανομοιότυπα στους χάρτες των δύο κατηγοριών πειραμάτων (PHT, CRL, VNY, BLF, CNF), όμως οι περισσότερες κλάσεις αποδίδονται καλύτερα στον χάρτη της τρίτης κατηγορίας πειραμάτων (ICU, MES, NGR, SPSV, SND, MRS, WCR, WBD).

Συνεπώς, ως βέλτιστος χάρτης κρίνεται εκείνος ο οποίος προέρχεται από την τρίτη ομάδα πειραμάτων, επιβεβαιώνοντας και τις παρατηρήσεις από τα μεγέθη αξιολόγησης της ακρίβειας της ταξινόμησης (OA, Kappa, F1).

6

ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ

6.1 Συμπεράσματα

Στην παρούσα μελέτη, στοχεύθηκε η παραγωγή χάρτη καλύψεων γης και καλλιεργειών για εκτενή περιοχή της Ανατολικής Μακεδονίας και Θράκης, που απεικονίζεται σε 3 Sentinel-2 tiles. Οι κλάσεις της ταξινόμησης βασίστηκαν στη δομή του επιπέδου L3 της ονοματολογίας του CORINE, ενώ προστέθηκαν και περαιτέρω κλάσεις σύμφωνα με τα δεδομένα καλλιεργειών, που προέρχονται από τον ΟΠΕΚΕΠΕ. Τα πολυφασματικά δεδομένα που αξιοποιήθηκαν προέρχονταν από 6 διαφορετικές ημερομηνίες για το έτος 2016, του δορυφόρου Sentinel-2.

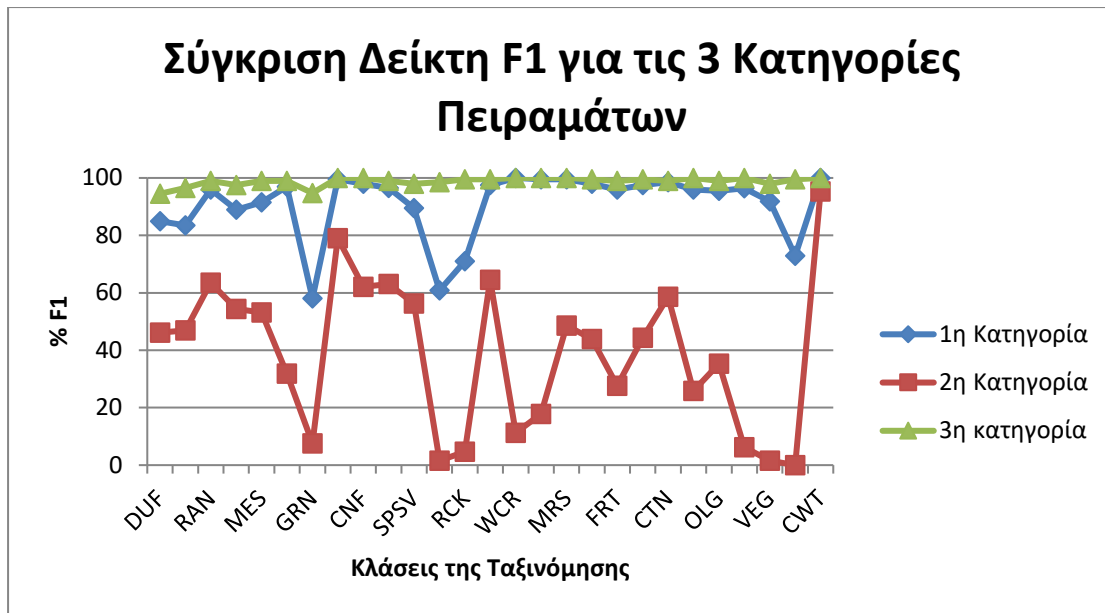
Η μεθοδολογία ταξινόμησης βασίστηκε σε φασματικά χαρακτηριστικά και τον αλγόριθμο Random Forest. Οι ταξινομήσεις που υλοποιήθηκαν, ομαδοποιήθηκαν σε 3 κατηγορίες. Η πρώτη ομαδοποίηση υλοποιήθηκε ανά tile, δηλαδή περιελάμβανε εκπαίδευση και έλεγχο στο ίδιο μοντέλο και περιοχή, εφαρμόζοντας 9 ταξινομήσεις ανά tile, με διαφορετικό ποσοστό δεδομένων εκπαίδευσης κάθε φορά (10-90%). Η δεύτερη κατηγορία πειραμάτων, περιλάμβανε εκπαίδευση και έλεγχο σε διαφορετικά μοντέλα και η τρίτη κατηγορία περιλάμβανε εκπαίδευση και έλεγχο σε ένα κοινό μοντέλο, προερχόμενο από τα τρία tiles, εφαρμόζοντας 9 ταξινομήσεις για 5 διαφορετικά ποσοστά εκπαίδευσης του αλγορίθμου (10-90%). Μετά το πέρας των πειραμάτων, έγιναν οι κατάλληλες συγκρίσεις των αποτελεσμάτων.

Σύμφωνα με τους πίνακες σύγχυσης των ταξινομήσεων, οι οποίοι αναλύθηκαν διεξοδικά στο κεφάλαιο 5, η συνολική ακρίβεια και ο δείκτης Kappa, συγκέντρωσαν τα παρακάτω ποσοστά:

Πίνακας 6-1 Διακυμάνσεις της Συνολικής Ακρίβειας και του Δείκτη Kappa

Κατηγορία Πειραμάτων	Συνολική Ακρίβεια	Δείκτης Kappa
Πρώτη- Ανά Tile	97-100%	96-99%
Δεύτερη- Ανά Δύο Tiles	64-70%	44-59%
Τρίτη- Ανά Τρία Tiles	98-100%	97-100%

Παρακάτω παρατίθεται γράφημα με την μεταβολή του δείκτη F1 για τις κλάσεις της ταξινόμησης ανά κατηγορία πειράματος.



Γράφημα 6-1 Σύγκριση των κλάσεων όπως αποδίδονται στους χάρτες των τριών κατηγοριών

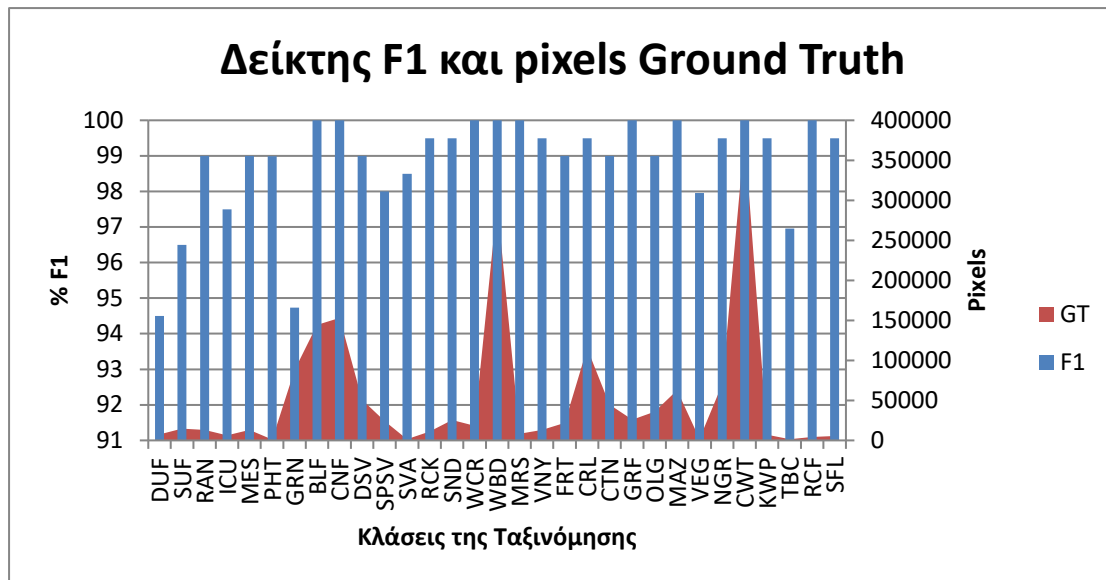
Αναλύοντας τα πειράματα των τριών κατηγοριών προέκυψαν κάποιες αξιόλογες παρατηρήσεις, που οδήγησαν σε ορισμένα συμπεράσματα. Οι παρατηρήσεις αυτές αφορούν τόσο τις μεμονωμένες κατηγορίες πειραμάτων, όσο και την συνολική διαδικασία περαίωσης της συγκεκριμένης μεταπτυχιακής εργασίας. Συνυπολογίζοντας, λοιπόν, την συνολική ακρίβεια, το δείκτη Karra, και το δείκτη F1 για τις διάφορες κλάσεις, καταλήγουμε στα εξής συμπεράσματα:

- Η πρώτη κατηγορία πειραμάτων, στην οποία η εκπαίδευση και ο έλεγχος της ταξινόμησης γίνεται σε ένα μεμονωμένο μοντέλο και tile, αποδίδει υψηλής ακρίβειας αποτελέσματα.
- Η δεύτερη κατηγορία πειραμάτων, στην οποία εκπαίδευση και έλεγχος γίνεται σε διαφορετικά μοντέλα αποδίδει μέτρια αποτελέσματα. Πιο συγκεκριμένα, κλάσεις, που διαθέτουν πολυπληθή δεδομένα εκπαίδευσης ταξινομούνται αξιόπιστα από τον αλγόριθμο ταξινόμησης, ενώ κλάσεις οι οποίες δεν διαθέτουν πολυπληθή, δεδομένα εκπαίδευσης, δεν ταξινομούνται επιτυχώς από τον αλγόριθμο.
- Η τρίτη κατηγορία πειραμάτων, η οποία αξιοποιεί δεδομένα εκπαίδευσης και ελέγχου και από τα 3 tiles της ταξινόμησης, κατασκευάζοντας ένα κοινό μοντέλο, οδηγεί στα βέλτιστα αποτελέσματα.

Σχετικά με τις κλάσεις και την ατομική ακρίβεια και αξιοπιστία της κάθε μίας, διαπιστώνεται ότι οι κλάσεις που αναγνωρίζονται με μεγαλύτερη ευκολία κι επιτυχία από τον ταξινομητή είναι εκείνες που αφορούν εδαφοκαλύψεις ενός είδους αντικειμένων (π.χ. μεμονωμένες φυτοκαλύψεις και καλλιέργειες, υδάτινες επιφάνειες κ.α.), ενώ δυσκολίες παρουσιάζονται στην αναγνώριση κλάσεων που εμπερικλείουν διαφορετικά αντικείμενα (π.χ. αστικός ιστός, βιομηχανικές περιοχές), που προκύπτουν ως υποκατηγορίες γενικών εδαφοκαλύψεων (π.χ. πυκνός και αραιός ιστός, πυκνή και αραιή σκληροφυλλική βλάστηση), κλάσεις, που εμφανίζουν παρόμοια φασματικά χαρακτηριστικά (π.χ. βαμβάκι, δημητριακά, καπνός), καθώς και κλάσεις, που

διαθέτουν μη επαρκή ή χαμηλής ποιότητας δεδομένα εκπαίδευσης (π.χ. θερμοκήπια, αραιή βλάστηση, απογυμνωμένα βράχια).

Παρακάτω παρατίθεται ένα γράφημα, στο οποίο παρουσιάζεται ο δείκτης F1 και το πλήθος του ground truth, για την καλύτερη ταξινόμηση από την τρίτη ομάδα πειραμάτων, ώστε να γίνει αντιληπτή η συσχέτιση ανάμεσα στο πλήθος των δεδομένων εκπαίδευσης και της απόδοσης της κάθε κλάσης.



Γράφημα 6-2 Συσχέτιση δείκτη F1 και εικονοστοιχείων ground truth

Ακόμη, διαπιστώνεται ότι κλάσεις, οι οποίες δεν διέθεταν πολυπληθές ground truth, χρησιμοποιώντας κοινό μοντέλο κι από τα τρία tiles, αυξάνουν την ατομική τους ακρίβεια και αξιοπιστία και συνεπώς το δείκτη F1, έως και 70 ποσοστιαίες μονάδες, καθιστώντας έτσι δυνατό τον διαχωρισμό τους (π.χ. RCK, SVA).

Συμπερασματικά, διαπιστώθηκε ότι

- Ο ταξινομητής στην πλειονότητα των περιπτώσεων απέδωσε τα βέλτιστα αποτελέσματα για τις κατηγορίες με πολυπληθές ground truth.
- Η επιλογή ως προς το διαχωρισμό του ποσοστού ground truth για εκπαίδευση και έλεγχο, υπαγορεύεται κάθε φορά από το εκάστοτε πείραμα. Σύμφωνα με τα πειράματα της πρώτης κατηγορίας πειραμάτων, αυξάνοντας το ποσοστό του ground truth, που χρησιμοποιείται για την εκπαίδευση του αλγορίθμου, καταφέρνουμε να επιτύχουμε υψηλότερα ποσοστά ακρίβειας. Επιπλέον, μετά το ποσοστό του 50%, τα αποτελέσματα της ταξινόμησης δεν εμφανίζουν αξιοσημείωτες μεταβολές. Αντίθετα, στην τρίτη κατηγορία πειραμάτων, παρατηρείται διακύμανση των αποτελεσμάτων σε σχέση με το ποσοστό του ground truth, που χρησιμοποιείται για εκπαίδευση του αλγορίθμου, καθώς τα βέλτιστα αποτελέσματα προέκυψαν, χρησιμοποιώντας το ποσοστό για εκπαίδευση του αλγορίθμου ίσο με 10%, ίσο με 30% και ανώτερο του 70%. Βάσει της συνολικής αξιολόγησης όλων των πειραμάτων, κρίνεται ότι τα υψηλά ποσοστά για τα δεδομένα εκπαίδευσης ίσα με 10%, οφείλονται σε τυχαία επιλογή αντιπροσωπευτικού δείγματος των κατηγοριών.

- Οι κλάσεις που διαχωρίζονται με δυσκολία από τον ταξινομητή είναι εκείνες που εμπρικλείουν διαφορετικά αντικείμενα, που προκύπτουν ως υποκατηγορίες γενικών εδαφοκαλύψεων ή που διαθέτουν μη επαρκή ή χαμηλής ποιότητας δεδομένα εκπαίδευσης.
- Η ακρίβεια των κλάσεων που διαθέτουν μικρό αριθμό πολύγωνων εκπαίδευσης βελτιώνεται άρδην με την χρήση μοντέλου προερχόμενο από πολλαπλά tiles.
- Η δημιουργία καλύψεων γης και καλλιεργειών με εκπαίδευση και έλεγχο σε διαφορετικά tiles, είναι εφικτή, καθώς αποδίδει ανώτερα του μετρίου αποτελέσματα. Για βελτίωση των αποτελεσμάτων, συστήνεται η χρήση περισσότερων κοινών ημερομηνιών/εικόνων ή/και τεχνικές για την παραγωγή σύνθετων κοινών ημερομηνιών/εικόνων και όταν αυτό δεν είναι εφικτό σύμπτυξη των κλάσεων, καθώς οι γενικές κατηγορίες των καλύψεων γης, αποδίδονται επιτυχώς σε ένα τέτοιο πείραμα.

6.2 Προοπτικές

Η ενασχόληση με το συγκεκριμένο πλαίσιο εργασίας, ανέδειξε επιπλέον προοπτικές εφαρμογής του. Κρίνεται λοιπόν σκόπιμη η περαιτέρω διερεύνηση και πειραματισμός προς την κατεύθυνση:

- ✓ Σύγκρισης της απόδοσης και άλλων αλγορίθμων μηχανικής μάθησης ή και βαθιάς μάθησης.
- ✓ Σύγκρισης της ποσοτικής αξιολόγησης των πειραμάτων με χωρισμό των δεδομένων εκπαίδευσης και ελέγχου σε επίπεδο πολυγώνου και όχι εικονοστοιχείου.
- ✓ Μελέτη απόδοσης της μεθοδολογίας με χρήση περισσότερων εικόνων και ημερομηνιών και από άλλους δέκτες ή/και χρονιές.
- ✓ Μελέτη απόδοσης της μεθοδολογίας με χρήση περισσότερων χαρακτηριστικών ταξινόμησης όπως επιπρόσθετα φασματικά αλλά και χαρακτηριστικά υψής όπως και με χρήση κάποιου ψηφιακού μοντέλου εδάφους.
- ✓ Σύγκριση της απόδοσης σε ταξινόμηση σε διαχρονικά (temporal) χαρακτηριστικά με ταυτόχρονη μείωση της διάστασης των δεδομένων
- ✓ Εφαρμογή επεξεργασίας post-classification για ενοποίηση περιοχών με κοινή κατηγορία και εξάλειψη του θορύβου τύπου αλατοπίπερο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] A. Stoian, V. Poulain, J. Inglada, V. Poughon, and D. Derksen, “Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems,” Jun. 2019, doi: 10.20944/preprints201906.0270.v1.
- [2] C. Karakizi, K. Karantzalos, M. Vakalopoulou, and G. Antoniou, “Detailed Land Cover Mapping from Multitemporal Landsat-8 Data of Different Cloud Cover,” *Remote Sensing*, vol. 10, no. 8, p. 1214, Aug. 2018, doi: 10.3390/rs10081214.
- [3] C. Marais Sicre *et al.*, “Early Detection of Summer Crops Using High Spatial Resolution Optical Image Time Series,” *Remote Sensing*, vol. 8, no. 7, p. 591, Jul. 2016, doi: 10.3390/rs8070591.
- [4] J. Inglada *et al.*, “Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery,” *Remote Sensing*, vol. 7, no. 9, pp. 12356–12379, Sep. 2015, doi: 10.3390/rs70912356.
- [5] E. Adam, O. Mutanga, J. Odindi, and E. M. Abdel-Rahman, “Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers,” *International Journal of Remote Sensing*, vol. 35, no. 10, pp. 3440–3458, May 2014, doi: 10.1080/01431161.2014.903435.
- [6] Δ. Ρόκος, *Φωτοερμηνεία-Τηλεπισκόπηση*. Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο, 1987.
- [7] Δ. Αργιαλάς, *Φωτοερμηνεία- Τηλεπισκόπηση*. Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο, 1999.
- [8] “Spectrum | physics,” *Encyclopedia Britannica*. [Online]. Available: <https://www.britannica.com/science/spectrum>. [Accessed: 16-Feb-2020].
- [9] Κ. Περάκης, Α. Μωυσιάδης, and Ι. Φαρασλής, *Η τηλεπισκόπηση σε 13 ενότητες*. 2015.
- [10] “Copernicus: Sentinel-2 - Satellite Missions - eoPortal Directory.” [Online]. Available: <https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-2>. [Accessed: 16-Feb-2020].
- [11] Δ. Αργιαλάς, *Ψηφιακή Τηλεπισκόπηση*. Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο, 1999.
- [12] K. Karantzalos, D. Bliziotis, and A. Karmas, “A Scalable Geospatial Web Service for Near Real-Time, High-Resolution Land Cover Mapping,” *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 8, no. 10, pp. 4665–4674, Oct. 2015, doi: 10.1109/JSTARS.2015.2461556.
- [13] C. Karakizi, G. Antoniou, and K. Karantzalos, “Towards Joint Land Cover and Crop Type Mapping with Numerous Classes,” 2018, pp. 2980–2983, doi: 10.1109/IGARSS.2018.8517473.
- [14] A. Bannari, D. Morin, F. Bonn, and A. R. Huete, “A review of vegetation indices,” *Remote Sensing Reviews*, vol. 13, no. 1–2, pp. 95–120, Aug. 1995, doi: 10.1080/02757259509532298.
- [15] J. Qi, A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian, “A modified soil adjusted vegetation index,” *Remote Sensing of Environment*, vol. 48, no. 2, pp. 119–126, May 1994, doi: 10.1016/0034-4257(94)90134-1.
- [16] H. Xu, “Extraction of Urban Built-up Land Features from Landsat Imagery Using a Thematicoriented Index Combination Technique,” *photogramm eng remote sensing*, vol. 73, no. 12, pp. 1381–1391, Dec. 2007, doi: 10.14358/PERS.73.12.1381.

- [17] J. Al-doski, S. B. Mansor, and H. Z. M. Shafri, "Image Classification in Remote Sensing," vol. 3, p. 8, 2013.
- [18] "Al-doski et al. - 2013 - Image Classification in Remote Sensing.pdf." .
- [19] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [20] T. Kavzoglu, "Chapter 33 - Object-Oriented Random Forest for High Resolution Land Cover Mapping Using Quickbird-2 Imagery," in *Handbook of Neural Computation*, P. Samui, S. Sekhar, and V. E. Balas, Eds. Academic Press, 2017, pp. 607–619.
- [21] L. Rokach, "Ensemble-based classifiers," *Artif Intell Rev*, vol. 33, no. 1, pp. 1–39, Feb. 2010, doi: 10.1007/s10462-009-9124-7.
- [22] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, 2nd ed. New York: Springer-Verlag, 2009.
- [23] A. Criminisi, E. Konukoglu, and J. Shotton, "Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning," Oct. 2011.
- [24] E. Soares, K. Oliveira-Esquerre, A. Aguiar, G. L. Botelho, and A. Kiperstok, "Development of a model to identify combined use in residential water end use events," in *Computer Aided Chemical Engineering*, 2018, pp. 1951–1956.
- [25] S. R. Joelsson, J. Benediktsson, and J. Sveinsson, "Random forest classifiers for hyperspectral data," 2005, p. 4 pp., doi: 10.1109/IGARSS.2005.1526129.
- [26] Y. Xie, Z. Sha, and M. Yu, "Remote sensing imagery in vegetation mapping: a review," *J Plant Ecol*, vol. 1, no. 1, pp. 9–23, Mar. 2008, doi: 10.1093/jpe/rtm005.
- [27] S. Visa, B. Ramsay, A. Ralescu, and E. VanDerKnaap, "Confusion Matrix-Based Feature Selection," *All Faculty Articles*, pp. 120–127, Jan. 2011.
- [28] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sensing of Environment*, vol. 62, no. 1, pp. 77–89, Oct. 1997, doi: 10.1016/S0034-4257(97)00083-7.
- [29] "Περιφέρεια Κεντρικής Μακεδονίας," *Βικιπαίδεια*. 25-Dec-2019.
- [30] "Κλίμα | Ανατολική Μακεδονία - Θράκη." [Online]. Available: <https://www.emtgreece.com/el/climate>. [Accessed: 16-Feb-2020].
- [31] Γ. Αντωνίου, "Χαρτογράφηση κάλυψης γης με χρήση δεδομένων sentinel-2 από πολλαπλές ημερομηνίες μέσω μηχανών διανυσματικής υποστήριξης," Feb. 2018, doi: <http://dx.doi.org/10.26240/heal.ntua.14942>.
- [32] "CORINE Land Cover," *European Environment Agency*. [Online]. Available: <https://www.eea.europa.eu/publications/COR0-landcover>. [Accessed: 16-Feb-2020].
- [33] Κ. Τσιώκου, "Χαρτογράφηση Καλλιεργειών και Κάλυψης Γης από Διαχρονικά Δεδομένα Sentinel-2 για τα έτη 2016 και 2017.," May 2019, doi: <http://dx.doi.org/10.26240/heal.ntua.9412>.
- [34] Ι. Παρχαρίδης, *Αρχές δορυφορικής τηλεπισκόπησης*. 2015.

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 2-1 Το ηλεκτρομαγνητικό φάσμα. Πηγή: Περάκης, 2016[9].....	6
Εικόνα 2-2 Πορεία Επιβλεπόμενης Ταξινόμησης (Πηγή: Al-doski, 2013, p.p. 143)[17]	18
Εικόνα 2-3 Πορεία Μη Επιβλεπόμενης Ταξινόμησης (Πηγή: Al-doski, 2013, p.p. 143)[17].....	19
Εικόνα 2-4 Σχηματική απεικόνιση του «random forest» αλγορίθμου, βασισμένο στην μέθοδο «bagging».....	22
Εικόνα 4-1 Περιοχές Ground Truth σε απεικόνιση raster εικόνας.....	44
Εικόνα 5-1 Σύγκυση ανάμεσα στην DUF, την SUF και την ICU	60
Εικόνα 5-2 Σύγκυση της PHT με πληθώρα κατηγοριών	60
Εικόνα 5-3 Σύγκυση ανάμεσα στην GRN και στις κατηγορίες των τεχνητών επιφανειών	61
Εικόνα 5-4 Επιτυχής απόδοση των κατηγοριών φυσικής βλάστησης	61
Εικόνα 5-5 Επιτυχής πρόβλεψη των κατηγοριών SVA και RCK	61
Εικόνα 5-6 Αδυναμία απόδοσης της VNY. Συμπερίληψη στην BLF και στην CNF... ..	62
Εικόνα 5-7 Σύγκυση ανάμεσα στην DUF, την SUF και την ICU	74
Εικόνα 5-8: Σύγκυση ανάμεσα στην MES, την DUF, την SUF και την ICU.....	74
Εικόνα 5-9 Σύγκυση της GRN με την MES και την ICU.	75
Εικόνα 5-10 Δυσκολία αναγνώρισης της SVA	75
Εικόνα 5-11 Δυσχέρεια του αλγορίθμου να αποδώσει την κατηγορία RCK.....	75
Εικόνα 5-12 Σύγκυση ανάμεσα στην DUF, την SUF και την ICU	84
Εικόνα 5-13 Σύγκυση ανάμεσα WCR και στην WBD	84
Εικόνα 5-14 Σύγκυση ανάμεσα στην DUF και την SUF.	96
Εικόνα 5-15 Σύγκυση ανάμεσα στις DUF, SUF, RAN, ICU, MES, MRS.....	96
Εικόνα 5-16 Σύγκυση ανάμεσα στις DUF, SUF και CTN.....	97
Εικόνα 5-17 Επιτυχής Ταξινόμηση ICU και MES.....	97
Εικόνα 5-18 Συγγύσεις ανάμεσα στην κλάση PHT και στην RAN	97
Εικόνα 5-19 Σύγκυση της κλάσης SND με την RCK και την MES	100
Εικόνα 5-20 Δυσκολία αναγνώρισης της WCR. Λάθος πρόβλεψη ως WBD και MRS	100
Εικόνα 5-21 Δυσκολία αναγνώρισης της MRS. Λάθος πρόβλεψη ως OLG και CRL	100
Εικόνα 5-22 Σύγκυση VNY με κλάσεις καλλιεργειών	101
Εικόνα 5-23 Σύγκυση FRT με CTN.....	101
Εικόνα 5-24 Σύγκυση CRL με CTN και GRF με BLF και SPSV	101
Εικόνα 5-25 Επιτυχής προσδιορισμός των pixel OLG στην εικόνα.....	102
Εικόνα 5-26 Εικονοστοιχεία της κλάσης GRF, ταξινομούνται ως CTN	102
Εικόνα 5-27 Απόδοση της VEG ως CRL	102
Εικόνα 5-28 Σύγκυση ανάμεσα στην DUF και την SUF	112
Εικόνα 5-29 Σύγκυση ανάμεσα στην DUF και την CRL	112
Εικόνα 5-30 Σύγκυση της ICU με DUF και την MES. Επιτυχής Ταξινόμηση της RAN	113
Εικόνα 5-31 Επιτυχημένη απόδοση των PHT.....	113
Εικόνα 5-32 Αδυναμία πρόβλεψης της GRN	113
Εικόνα 5-33 Μικρή σύγκυση ανάμεσα στην BLF και την CNF.....	114

Εικόνα 5-34 Επιτυχημένη ταξινόμηση της κατηγορίας CNF	114
Εικόνα 5-35 Σύγκριση ανάμεσα στην DSV και στην CNF	114
Εικόνα 5-36 Λανθασμένη απόδοση της RCK	115
Εικόνα 5-37 Λανθασμένη απόδοση της SND ως DUF και SUF. Λανθασμένη απόδοση της WCR ως WBD.....	115
Εικόνα 5-38 Αδυναμία απόδοσης της VNY. Συμπερίληψη στην BLF και στην CNF.....	115
Εικόνα 5-39 Μικρή σύγκριση της DUF με την SUF.....	125
Εικόνα 5-40 Σύγκριση ανάμεσα στην ICU και τις DUF- MES.....	125
Εικόνα 5-41 Μικρή σύγκριση ανάμεσα στην VEG και στην CTN	126
Εικόνα 5-42 Μικρή σύγκριση ανάμεσα στην TBC και στην CTN.....	126

ΚΑΤΑΛΟΓΟΣ ΧΑΡΤΩΝ

Χάρτης 5.1 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG	59
Χάρτης 5.2 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TLF και έλεγχο στην εικόνα 35TMF	73
Χάρτης 5.3 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG	83
Χάρτης 5.4 Τα tiles, που χρησιμοποιήθηκαν από τον αλγόριθμο για την διεξαγωγή του πρώτου πειράματος,	85
Χάρτης 5.5 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMF	95
Χάρτης 5.6 Τα tiles, που χρησιμοποιήθηκαν από τον αλγόριθμο για την διεξαγωγή του πρώτου πειράματος	103
Χάρτης 5.7 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TLF και έλεγχο στην εικόνα 35TMG	111
Χάρτης 5.8 Χάρτης Καλύψεων Γης και Καλλιεργειών της ταξινόμησης με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG	124

ΚΑΤΑΛΟΓΟΣ ΓΡΑΦΗΜΑΤΩΝ

Γράφημα 5-1 Συνολική ακρίβεια και δείκτης Kappa για την ταξινόμηση στην εικόνα 35TLF	49
Γράφημα 5-2 Δείκτης F1 κάθε κλάσης για την ταξινόμηση στην εικόνα 35TLF.	50
Γράφημα 5-3 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMF.....	58
Γράφημα 5-4 Συνολική ακρίβεια και δείκτης Kappa για την ταξινόμηση στην εικόνα 35TMF.....	63
Γράφημα 5-5 Δείκτης F κάθε κλάσης για την ταξινόμηση στην εικόνα 35TMF	64
Γράφημα 5-6 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMF.....	72
Γράφημα 5-7 Συνολική ακρίβεια και δείκτης Kappa για την ταξινόμηση στην εικόνα 35TMG.....	76
Γράφημα 5-8 Δείκτης F1 για την ταξινόμηση στην εικόνα 35TMG.....	77
Γράφημα 5-9 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMG.....	82

Γράφημα 5-10 Δείκτης F1 για την ταξινόμηση με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMF.....	94
Γράφημα 5-11 Δείκτης F1 για την ταξινόμηση με εκπαίδευση στην εικόνα 35TF και έλεγχο στην εικόνα 35TMG	110
Γράφημα 5-12 Συνολική ακρίβεια και δείκτης Kappa για την ταξινόμηση με εκπαίδευση και έλεγχο σε όλα τα tile	116
Γράφημα 5-13 Δείκτης F1 κάθε κλάσης για την ταξινόμηση που συμμετέχουν και τα 3 tiles	117
Γράφημα 5-14 Δείκτης F1 για την ταξινόμηση με εκπαίδευση και έλεγχο στην εικόνα 35TMF.....	123
Γράφημα 5-15 Σύγκριση της συνολικής ακρίβειας και του δείκτη Kappa των 6 αναλυθέντων πειραμάτων	127
Γράφημα 5-16 Σύγκριση του δείκτη F1 για κάθε κλάση για τις κατηγορίες πειραμάτων 1 και 3 για το tile 35TLF.....	128
Γράφημα 5-17 Σύγκριση του δείκτη F1 για κάθε κλάση για τις τρεις κατηγορίες πειραμάτων για το tile 35TMF	128
Γράφημα 5-18 Σύγκριση του δείκτη F1 για κάθε κλάση για τις τρεις κατηγορίες πειραμάτων για το tile 35TMG.....	129
Γράφημα 6-1 Σύγκριση των κλάσεων όπως αποδίδονται στους χάρτες των τριών κατηγοριών	136
Γράφημα 6-2 Συσχέτιση δείκτη F1 και εικονοστοιχείων ground truth	137

ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 4-1 Διάγραμμα Ροής του Προγράμματος Ταξινόμησης, στην Python3...	46
--	----

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 2-1 Κανάλια Sentinel-2	11
Πίνακας 2-2 Συσχέτιση των Τιμών του Δείκτη NDVI και των Καλύψεων Γης	14
Πίνακας 2-3 Συσχέτιση των Τιμών του Δείκτη MSAVI και των Καλύψεων Γης.....	15
Πίνακας 2-4 Μέθοδοι Ταξινόμησης, Παραδείγματα και Χαρακτηριστικά.....	17
Πίνακας 3-1 Εικόνες Ταξινόμησης.....	31
Πίνακας 3-2 Ελληνική Ονοματολογία CORINE LAND COVER. Πηγή: http://portal.survey.ntua.gr [32].....	34
Πίνακας 3-3 Κλάσεις της Ταξινόμησης	36
Πίνακας 3-4 Πλήθος Εικονοστοιχίων Δεδομένων Εκπαίδευσης ανά Κατηγορία Ταξινόμησης.....	39
Πίνακας 5-1 Πίνακας Σύγχυσης για ταξινόμηση στην εικόνα 35TLF.	51
Πίνακας 5-2 Καλλιέργειες που συμπεριλαμβάνονται στην κλάση των VEG	54
Πίνακας 5-3 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση στην εικόνα 35TLF	56
Πίνακας 5-4 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση στην εικόνα 35TLF	57
Πίνακας 5-5 Πίνακας Σύγχυσης για ταξινόμηση στην εικόνα 35TMF	65

Πίνακας 5-6 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση στην εικόνα 35TMF.....	70
Πίνακας 5-7 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση στην εικόνα 35TMF.....	71
Πίνακας 5-8 Πίνακας Σύγκυσης για ταξινόμηση στην εικόνα 35TMG.....	78
Πίνακας 5-9 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση στην εικόνα 35TMG.....	81
Πίνακας 5-10 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση στην εικόνα 35TMG.....	81
Πίνακας 5-11 Πίνακας Σύγκυσης για ταξινόμηση με δεδομένα εκπαίδευσης από την εικόνα 35TML και έλεγχο στην εικόνα 35TMF.....	87
Πίνακας 5-12 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF	92
Πίνακας 5-13 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF	93
Πίνακας 5-14 Πίνακας Σύγκυσης για ταξινόμηση με δεδομένα εκπαίδευσης από την εικόνα 35TLF και έλεγχο στην εικόνα 35TMG.....	104
Πίνακας 5-15 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMG.....	108
Πίνακας 5-16 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMG.....	109
Πίνακας 5-17 Πίνακας Σύγκυσης για ταξινόμηση με δεδομένα εκπαίδευσης από την όλα τα tiles και έλεγχο σε όλα τα tiles	118
Πίνακας 5-18 Ποσοστιαία Σφάλματα Παράλειψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF	121
Πίνακας 5-19 Ποσοστιαία Σφάλματα Συμπερίληψης για την ταξινόμηση με εκπαίδευση στην εικόνα 35TML και έλεγχο στην εικόνα 35TMF	122
Πίνακας 5-20: Διαφοροποίηση στον παραγόμενο χάρτη του tile 35TMG.....	130
Πίνακας 5-21 Σύγκριση των κλάσεων όπως αποδίδονται στους χάρτες των τριών κατηγοριών	131
Πίνακας 6-1 Διακυμάνσεις της Συνολικής Ακρίβειας και του Δείκτη Karra	135