

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ**



**ΤΜΗΜΑ ΓΕΩΓΡΑΦΙΑΣ**

**Π.Μ.Σ**

**“ΓΕΩΓΡΑΦΙΑ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΗ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ”**

**Μεταπτυχιακή Διπλωματική Διατριβή:  
“Παρακολούθηση αλλαγών εδαφοκάλυψης στην  
Λέσβο τα τελευταία 30 χρόνια με σύζευξη  
εικονοληπτικών Radar και οπτικών δεκτών”**

**ΒΕΛΛΙΟΣ ΓΕΩΡΓΙΟΣ**

**A.M.:1632017007**

**Επιβλέπων Καθηγητής: Κουκούλας Σωτήριος**

**Μυτιλήνη, Οκτώβριος 2019**

**UNIVERSITY OF AEGEAN**



**DEPARTMENT OF GEOGRAPHY**

**P.S.P**

**“GEOGRAPHY AND APPLIED GEO-INFORMATICS”**

**Master Thesis:**

**“Monitoring landcover changes in Lesvos the last 30  
years, using image fusion of radar and optical  
satellites”**

**VELLIOS GEORGIOS**

**A.M.:1632017007**

**Supervisor: Koukoulas Sotirios**

**Mytilene, September 2019**

**Τριμελής Επιτροπή**

**Κουκούλας Σωτήριος**

**Κίζος Αθανάσιος**

**Σουλακέλλης Νικόλαος**

---

---

---

Ο άνθρωπος μοιάζει με κλάσμα  
όπου ο αριθμητής είναι ο πραγματικός εαυτός του  
και ο παρονομαστής η ιδέα που έχει για τον εαυτό του.  
Όσο μεγαλύτερος ο παρονομαστής, τόσο μικρότερη η αξία του κλάσματος.  
Και όσο ο παρονομαστής διογκώνεται προς το άπειρο,  
τόσο το κλάσμα τείνει προς το μηδέν.

Leo Tolstoy (1828-1910)

## Ευχαριστίες

Για την διεκπεραίωση της παρούσας διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω τον κ. Κουκούλα για την εμπιστοσύνη που μου έδειξε, καθώς και τα μέλη της τριμελούς επιτροπής για τις παρατηρήσεις τους. Επίσης θα ήθελα να ευχαριστήσω τον κ. Καβουρδάκη για την βοήθειά του στο επίπεδο του προγραμματισμού στην γλώσσα R, αλλά και όποτε τον χρειάστηκα. Ακόμη θα ήθελα να ευχαριστήσω την οικογένειά μου, που μου στάθηκε δίπλα μου τόσο κατά την διάρκεια της συγγραφής της, όσο και κατά την διάρκεια των σπουδών μου. Τέλος θέλω να ευχαριστήσω την Λ.Μ. και τον Γ.Κ. που ήταν δίπλα μου και με στήριξαν.

## Περίληψη

Σκοπός της παρούσας διπλωματικής μεταπτυχιακής διατριβής είναι η σύζευξη οπτικών δεκτών και εικονοληπτικών Radar, με σκοπό τον εντοπισμό αλλαγών εδαφοκάλυψης στην Λέσβο τα τελευταία 30 χρόνια (1993-2018). Για την επίτευξη του σκοπού της διπλωματικής, χρησιμοποιήθηκαν δεδομένα από οπτικούς δέκτες Landsat, δεδομένα από τα εικονοληπτικά Radar ERS-1 και Sentinel-1, λόγοι καναλιών των οπτικών δεκτών, δείκτες βλάστησης καθώς και το οδικό δίκτυο της περιοχής μελέτης. Η ταξινόμηση έγινε με τον ταξινομητή Random Forest και ο εντοπισμός αλλαγών εδαφοκάλυψης με την μέθοδο διασταύρωσης πινάκων και εικόνων. Ακόμη έγινε προσπάθεια για σύζευξη των δορυφορικών δεδομένων. Τέλος η ακρίβεια των ταξινομήσεων είχε εύρος από 81,89% έως 88,63%, ενώ οι χάρτες εντοπισμού αλλαγών είχαν ακρίβεια 72,58% και 76,86%.

Λέξεις Κλειδιά: Radar, σύζευξη, Random Forest Ταξινομητής, Λέσβος,

Εντοπισμός Αλλαγών

**Abstract**

The purpose of this thesis is to fuse optical and imaging radars satellites with the aim of detecting land cover changes in Lesvos over a period of the last 30 years (1993-2018). To achieve this goal, data from Landsat optical receivers, imaging Radar ERS-1 and Sentinel-1, band ratios, vegetation indices, and the study area road network has been used. The classification has been done with the Random Forest classifier and the detection of land cover changes using the intersection of tables and images. Efforts have also been made to fuse satellite data in order to achieve accuracy of the rankings ranged from 81.89% to 88.63%, while change detection maps had accuracy of 72.58% and 76.86%.

Key Words: Radar, Fusion, Random Forest Classifier, Lesvos, Change Detection

## Περιεχόμενα

### ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

|  |    |
|--|----|
| 1. Δορυφορικοί αισθητήρες, είδη δορυφόρων. ....          | 2  |
| 1.1. Παθητικοί αισθητήρες. ....                          | 6  |
| 1.2. Ενεργητικοί αισθητήρες. ....                        | 7  |
| 2. Μέθοδοι ταξινόμησης δορυφορικών δεδομένων. ....       | 12 |
| 2.1. Ταξινομητές Δέντρων Απόφασης. ....                  | 13 |
| 2.2. Ταξινομητής Random Forest. ....                     | 17 |
| 3. Μέθοδοι εντοπισμού αλλαγών χρήσεων/καλύψεων γης. .... | 23 |
| 4. Σύνοψη. ....  | 32 |

### ΚΕΦΑΛΑΙΟ 2: Περιοχή Μελέτης

|                       |    |
|-----------------------|----|
| Περιοχή Μελέτης. .... | 35 |
|-----------------------|----|

### ΚΕΦΑΛΑΙΟ 3: Δεδομένα

|                                   |    |
|-----------------------------------|----|
| 1. Συλλογή δεδομένων. ....        | 37 |
| 1.1. Διανυσματικά δεδομένα. ....  | 37 |
| 1.2. Δορυφορικά δεδομένα. ....    | 38 |
| 1.3. Μετεωρολογικά δεδομένα. .... | 41 |

### ΚΕΦΑΛΑΙΟ 4: Μεθοδολογία

|  |    |
|--|----|
| 1. Προεπεξεργασία δεδομένων. ....                        | 44 |
| 2. Δημιουργία δειγματοληπτικών σημείων. ....             | 47 |
| 3. σύζευξη εικονοληπτικών Radar και οπτικών δεκτών. .... | 50 |
| 4. Ταξινόμηση με τον αλγόριθμο RandomForest. ....        | 51 |
| 5. Αξιολόγηση ακρίβειας ταξινόμησης. ....                | 52 |
| 6. Εντοπισμός αλλαγών εδαφοκάλυψης. ....                 | 53 |

### ΚΕΦΑΛΑΙΟ 5: Αποτελέσματα

|                    |    |
|--------------------|----|
| Αποτελέσματα. .... | 54 |
|--------------------|----|

### ΚΕΦΑΛΑΙΟ 6: Συζήτηση

|                |    |
|----------------|----|
| Συζήτηση. .... | 66 |
|----------------|----|



**ΚΕΦΑΛΑΙΟ 7: Συμπεράσματα**

|                          |           |
|--------------------------|-----------|
| <b>Συμπεράσματα.....</b> | <b>66</b> |
|--------------------------|-----------|

**ΚΕΦΑΛΑΙΟ 8: Βιβλιογραφία**

|  |           |
|--|-----------|
| <b>1. Βιβλία.....</b>                    | <b>71</b> |
| <b>2. Επιστημονική βιβλιογραφία.....</b> | <b>71</b> |
| <b>3. Διαδικτυακές πηγές.....</b>        | <b>75</b> |

## ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

Τα τηλεπισκοπικά δεδομένα, προσφέρουν σημαντικές πληροφορίες για την επιφάνεια της γης, με αποτέλεσμα πολλοί ερευνητές να τα χρησιμοποιούν για τις εφαρμογές τους. Λόγω των χαρακτηριστικών των δορυφορικών εικόνων, καταγραφή ηλεκτρομαγνητικής ακτινοβολίας από την επιφάνεια της γης, μπορούν να χρησιμοποιηθούν για περιβαλλοντικές, δημογραφικές και κοινωνικοοικονομικές εφαρμογές, σε διάφορες χρονικές και χωρικές κλίμακες, χωρίς να είναι υποχρεωτική η παρουσία των ερευνητών στο πεδίο, αλλά και έχοντας εικόνα από περιοχές που είναι δυσπρόσιτες.

Μια εφαρμογή είναι και η παρακολούθηση, στιγμιαία ή και διαχρονικά, της χρήσης ή της κάλυψης γης. Σύμφωνα με τους Campbell και Wynne (2011), η χρήση γης αναφέρεται στον τρόπο που χρησιμοποιείται η γη από τον άνθρωπο, με έμφαση στο ρόλο της γης στις οικονομικές δραστηριότητες, παραδείγματος χάριν μια βιομηχανική περιοχή, ένα δάσος για παραγωγή ξυλείας. Η κάλυψη γης αφορά τα ορατά χαρακτηριστικά της επιφάνειας της γης, δηλαδή στην κατηγοριοποίηση των διαφόρων φυσικών ή ανθρωπογενών στοιχείων του εδάφους, παραδείγματος χάριν αστική περιοχή, δασικές εκτάσεις.

Οι ίδιοι, Campbell και Wynne (2011), υποστηρίζουν ότι πρέπει να εξετάζονται μαζί η κάλυψη και η χρήση γης μιας περιοχής αναγνωρίζοντας όμως την διάκριση μεταξύ των δύο τους.

Στην παρούσα μεταπτυχιακή διπλωματική διατριβή, έγινε μια προσπάθεια εντοπισμού αλλαγών των καλύψεων γης της Λέσβου, από το 1993 έως το 2018, με χρήση τηλεπισκοπικών δεδομένων εικονοληπτικών Radar και οπτικών δεκτών καθώς και σύζευξη αυτών.

## 1. Δορυφορικοί αισθητήρες, είδη δορυφόρων

Με τον όρο τηλεπισκόπηση εννοούμε την λήψη πληροφοριών για ένα αντικείμενο, μια περιοχή ή ένα φαινόμενο, με την χρήση ανιχνευτικών συσκευών, οι οποίες δεν βρίσκονται σε επαφή με το αντικείμενο/φαινόμενο/περιοχή παρακολούθησης. Οι πληροφορίες αυτές είναι η μέτρηση και αποτύπωση της ηλεκτρομαγνητικής ακτινοβολίας που ανακλάται ή εκπέμπεται από το υπό παρακολούθηση “αντικείμενο”. Όταν η ανίχνευση και καταγραφή των πληροφοριών αυτών, γίνεται από ανιχνευτές που φέρονται από δορυφόρους, τότε μιλάμε για δορυφορική τηλεπισκόπηση.

Τα ηλεκτρομαγνητικά κύματα δημιουργούνται από περιοδικές κινήσεις φορτισμένων σωματιδίων, οι οποίες δημιουργούν περιοδικές διαταραχές στο ηλεκτρικό τους πεδίο, που με την σειρά τους δημιουργούν ένα μαγνητικό πεδίο κάθετο στο ηλεκτρικό. Τα δύο πεδία αυτά είναι αλληλένδετα και πάντα κάθετα μεταξύ τους (Καρτάλης και Φείδας (2015), σελ.1-5).

Συνεπώς η ηλεκτρομαγνητική ακτινοβολία είναι ένα ηλεκτρομαγνητικό κύμα. Η ηλεκτρομαγνητική ακτινοβολία διαιρείται σε περιοχές με βάση το μήκος του κύματος. Το ηλεκτρομαγνητικό φάσμα έχει εύρος από 0.1  $\mu\text{m}$  έως και 100m και χωρίζεται σε περιοχές. Η ορατή περιοχή του φάσματος έχει εύρος από 0.38 $\mu\text{m}$  έως 0.72 $\mu\text{m}$ , τα όρια δηλαδή της ανθρώπινης όρασης. Η υπέρυθρη περιοχή έχει μήκη κύματος μεγαλύτερα του ορατού, 0.7 $\mu\text{m}$  έως 1000 $\mu\text{m}$ , ενώ η μικροκυματική ακτινοβολία περιλαμβάνει μήκη κύματος από 1mm έως 1m, και έχει την ιδιότητα να διαπερνά νέφη και να εισχωρεί στην γήινη επιφάνεια, δεν είναι αντιληπτή από το ανθρώπινο μάτι, ενώ οι αισθητήρες που χρησιμοποιούν την μικροκυματική ακτινοβολία μπορούν να καταγράψουν επιφανειακές ιδιότητες του περιβάλλοντος οι οποίες σχετίζονται με την τοπογραφία, την μορφολογία, την τραχύτητα του

εδάφους, την περιεκτικότητα σε νερό στην βλάστηση και στο έδαφος, καθώς και στην διηλεκτρική συμπεριφορά. Ακόμη η μικροκυματική ακτινοβολία αποτελείται από τις ζώνες P, L, S, C, K, Q, V και W. Πριν την ορατή περιοχή είναι η υπεριώδης περιοχή (Καρτάλης και Φείδας (2015), σελ.1-5; Παρχαρίδης (2015),σελ. 21-22).

Οι δορυφορικοί αισθητήρες χωρίζονται σε 2 μεγάλες κατηγορίες ανάλογα με την πηγή ακτινοβολίας που χρησιμοποιούν για την καταγραφή της ηλεκτρομαγνητικής ακτινοβολίας που ανακλάται ή εκπέμπεται από τον στόχο. Η πρώτη κατηγορία είναι οι παθητικοί αισθητήρες και είναι αυτοί που καταγράφουν την εκπεμπόμενη ή ανακλώμενη ακτινοβολία από τον στόχο, και καταγράφουν από την υπεριώδη περιοχή μέχρι την υπέρυθρη περιοχή του φάσματος της ηλεκτρομαγνητικής ακτινοβολίας. Η δεύτερη κατηγορία είναι οι ενεργητικοί αισθητήρες, όπου χρησιμοποιούν δικιά τους πηγή ακτινοβολίας, όπως οι δορυφόροι Radar, την οποία κατευθύνουν προς τον υπομελέτη στόχο και στην συνέχεια καταγράφουν την επιστροφή της ακτινοβολίας (Καρτάλης και Φείδας (2015), σελ.45-46).

Στους αισθητήρες Radar, πολύ σημαντικό ρόλο έχει η πόλωση της ακτινοβολίας καθώς και η γωνία πρόσπτωσης. Ως πόλωση της μικροκυματικής ακτινοβολίας ορίζεται η κατεύθυνση της έντασης του ηλεκτρικού πεδίου της ακτινοβολίας σε σχέση με το οριζόντιο επίπεδο. Η πόλωση είναι μια παράμετρος από την οποία εξαρτάται το ποσοστό της οπισθοσκεδαζόμενης ακτινοβολίας (ανακλώμενη). Η πόλωση είναι είτε οριζόντια, όταν το διάνυσμα της έντασης του ηλεκτρικού πεδίου είναι κάθετο στην επιφάνεια της γης είτε κάθετη όταν το διάνυσμα είναι παράλληλο στην επιφάνεια της γης. Όμως πρέπει να σημειωθεί πως η πόλωση της ακτίνας που ταξιδεύει προς την γη, δεν είναι ούτε οριζόντια, αλλά ούτε κάθετη, αν και μπορεί να αναλυθεί σε οριζόντια και κάθετη πόλωση. Κατά την αλληλεπίδραση της ακτινοβολίας με το επίπεδο, δύναται

να αλλάξει η πόλωση της ακτινοβολίας. Αυτό γίνεται λόγω της τραχύτητας του εδάφους (Richards και Jia (2006), σελ. 9; Richards (2009), σελ. 28-29)

Για την δημιουργία της χωρικής ανάλυσης των Radar, χρησιμοποιούνται διαφορετικές αρχές στην παράλληλη κατεύθυνση της κίνησης της πλατφόρμας, αζιμούθιο, και διαφορετικές στην κατεύθυνση κάθετη της κίνησης της πλατφόρμας (εύρος). Για την ανάλυση στο εύρος χρησιμοποιείται ο χρόνος που απαιτείται για να επιστρέψει ο παλμός στην κεραία του Radar, οπότε και υπολογίζεται η απόσταση του στόχου από το Radar. Γι' αυτό τον λόγο η πλατφόρμα δεν κοιτάει στο ναδίρ, αλλά κάθετα και προς τα πλάγια. Επίσης το πλάτος της λωρίδας που καταγράφει το Radar, έχει να κάνει με τις ιδιότητες της κεραίας.

Για την ανάλυση στο αζιμούθιο χρησιμοποιείται η αρχή Doppler. Καθώς η πλατφόρμα μετακινείται, αλλάζει η συχνότητα του σήματος που δέχεται πίσω (όπως όταν έχουμε έναν σταθερό παρατηρητή, και ένα ασθενοφόρο τον πλησιάζει με σταθερή ταχύτητα), με αποτέλεσμα να υπολογίζεται η χωρική ανάλυση στο αζιμούθιο.

Οι αισθητήρες Radar, χωρίζονται σε RAR (Real Aperture Radar) και σε SAR (Synthetic Aperture Radar) και η διαφορά τους είναι ότι ενώ και οι δύο αισθητήρες έχουν σταθερό μήκος κεραίας, οι αισθητήρες SAR, χρησιμοποιούν την κίνηση της πλατφόρμας για να αυξήσουν την ανάλυση στο αζιμούθιο, ενώ η ανάλυση εξαρτάται από το μέγεθος της κεραίας, που μπορεί να είναι πολύ μικρό (Richards (2009), σελ 62).

Ένα χαρακτηριστικό των εικόνων των Radar, είναι ο συντελεστής οπισθοσκέδασης, δηλαδή η ένταση του οπισθοσκεδαζόμενου σήματος και ο οποίος μετριέται σε decibel (dB). Όσο μεγαλύτερος είναι ο συντελεστής οπισθοσκέδασης, τόσο πιο τραχιά είναι μια επιφάνεια.

Ένα ακόμη χαρακτηριστικό, είναι η υφή. Δημιουργούνται προβλήματα παρόμοια με αυτά των οπτικών δεκτών γνωστά και ως φαινόμενο του

αλατοπίπερου (“salt and pepper effect”), που στα Radar ονομάζονται στίγματα (“speckle effect”). Τα στίγματα προκύπτουν από την σκέδαση της ακτίνας και σε άλλες επιφάνειες ύστερα από την πρόσπτωση στον στόχο, οι οποίες βρίσκονται εντός της περιοχής μελέτης. Αυτό προκαλεί φαινόμενα αφαιρετικής ή προσθετικής συμβολής μέσα σε κάθε εικονοστοιχείο, με αποτέλεσμα κάποιες περιοχές να φαίνονται πιο σκούρες από ότι θα έπρεπε, ή πιο ανοικτόχρωμες. Αυτό διορθώνεται με χρήση φίλτρων (speckle filters).

Επίσης ένα χαρακτηριστικό των εικόνων, είναι το σχήμα και το μέγεθος. Η γεωμετρία των Radar, δημιουργεί γεωμετρικές παραμορφώσεις στην εικόνα. Την παραμόρφωση της κεκλιμένης απόστασης (slant range distortion), η οποία οφείλεται στο ότι το Radar μετρά την απόσταση σε στόχους που βρίσκονται κάθετα στην τροχιά του και όχι οριζόντια, με αποτέλεσμα τα μεγέθη και τα σχήματα να αλλοιώνονται όσο πλησιάζουμε το ίχνος του δορυφόρου, εγγύς περιοχή.

Δύο παραμορφώσεις που σχετίζονται με το ανάγλυφο, είναι το φαινόμενο της σμίκρυνσης (foreshortening) και το φαινόμενο της αναστροφής (layover). Το φαινόμενο της σμίκρυνσης δημιουργείται, όταν η δέσμη της ακτινοβολίας του Radar, φτάσει πρώτα στην βάση μιας πλαγιάς, η οποία έχει προσανατολισμό προς τον αισθητήρα, πριν φτάσει στην κορυφή. Η ένταση του φαινομένου εξαρτάται από την κλίση της πλαγιάς και την γωνία πρόσπτωσης της ακτινοβολίας. Το φαινόμενο της αναστροφής, συμβαίνει όταν ο παλμός φτάσει πρώτα στην κορυφή μιας πλαγιάς, η οποία έχει προσανατολισμό προς τον αισθητήρα, πριν φτάσει στην βάση της πλαγιάς, δηλαδή θα καταγραφεί πρώτα η κορυφή, η οποία θα φαίνεται μετατοπισμένη κατά το Radar, και μετά η βάση, με αποτέλεσμα η πλαγιά να φαίνεται αντεστραμμένη (Καρτάλης και Φείδας (2015), σελ.138-147; Richards (2009), σελ.111-112).

Γενικά, οι δορυφορικοί αισθητήρες έχουν κάποια χαρακτηριστικά που

είναι σημαντικά ώστε να επιλεγούν ανάλογα τον στόχο ή το υπομελέτη φαινόμενο. Ο συνδυασμός της χωρικής, της φασματικής, της ραδιομετρικής και της χρονικής διακριτικής ικανότητας του κάθε αισθητήρα είναι αυτός που μας κατευθύνει στην κατάλληλη επιλογή αισθητήρα (Καρτάλης και Φείδας (2015), σελ.50-53).

Οι δορυφόροι κατηγοριοποιούνται ως χαμηλής ή υψηλής τροχιάς, ανάλογα το ύψος περιστροφής γύρω από την γη, και ως ισημερινής, πολικής και πλάγιας ή σχεδόν πολικής τροχιάς, ανάλογα με την τροχιά που κινούνται (Καρτάλης και Φείδας (2015), σελ.60-64).

### **1.1. Παθητικοί αισθητήρες**

Ο Landsat 5 αποτελεί την πέμπτη αποστολή του προγράμματος Landsat, που ξεκίνησε το 1972 και μετράει 8 εκτοξεύσεις. Από εκείνη την εποχή οι δορυφόροι παρέχουν μια τεράστια πληθώρα πληροφοριών, μέσω του τεράστιου αριθμού πολυφασματικών εικόνων. Η συνεργασία της NASA με την Αμερικανική Γεωλογική Υπηρεσία (USGS), έκανε το πρόγραμμα Landsat το μακροβιότερο πρόγραμμα συλλογής μέτριας διακριτικής ικανότητας τηλεπισκοπικών δεδομένων. Οι εικόνες παρέχουν πληροφορίες σε όσους ασχολούνται με την γεωργία, τη γεωλογία, τη δασοκομία, την εκπαίδευση, τις φυσικές καταστροφές, το χωροταξικό σχεδιασμό και την έρευνα αλλαγών σε παγκόσμιο επίπεδο. Ο Landsat 5 εκτοξεύθηκε στις 1/3/1984 και σταμάτησε την αποστολή δεδομένων τον Ιανουάριο του 2013. Είναι ο δορυφόρος με την μεγαλύτερη διάρκεια παροχής δορυφορικών δεδομένων και βρισκόταν σε ύψος 705 χιλιομέτρων και ήταν ηλιοσύγχρονης τροχιάς. Είχε δύο αισθητήρες, ένα πολυφασματικό (MSS) και έναν θεματικό (TM) (30 μέτρα χωρική διακριτική ικανότητα για το ορατό, μέσο και κοντινό υπέρυθρο, και 120 μέτρα για το θερμικό κανάλι).

Ακολουθούσε μια σειρά σταθερών εδαφικών σημείων (path) που ορίζονταν από το δεύτερο Παγκόσμιο Σύστημα Αναφοράς (WRS- 2), όπως όλοι οι δορυφόροι Landsat.

Ο Landsat 8 αποτελεί την όγδοη αποστολή του προγράμματος Landsat. Βρίσκεται, και αυτός σε ύψος 705 χιλιομέτρων και έχει ηλιοσύγχρονη τροχιά. Ο Landsat 8 εκτοξεύθηκε στις 11 Φεβρουαρίου του 2013. Ο Landsat 8 φέρει 2 εργαλεία σάρωσης. Το σύστημα OLI περιλαμβάνει πολύ πιο αναπτυγμένη τεχνολογία σε σύγκριση με το Enhanced Thematic Mapper (ETM+) του Landsat 7. Συλλέγει δεδομένα σε εννιά φασματικά κανάλια με χωρική ικανότητα 30 μέτρα για όλα τα κανάλια εκτός από το παγχρωματικό κανάλι που έχει χωρική ικανότητα 15 μέτρα. Το εύρος των καναλιών του αισθητήρα σχεδιάστηκε έτσι ώστε να αποφεύγονται τα φαινόμενα ατμοσφαιρικών απορροφήσεων. Το παγχρωματικό κανάλι του OLI είναι επίσης πιο ανεπτυγμένο από αυτό του ETM+ με σκοπό τον καλύτερο διαχωρισμό της βλάστησης από το έδαφος μιας εικόνας. Υπάρχουν επίσης και δύο καινούργια κανάλια το Coastal/Aerosol (κανάλι 1), σημαντικό για παράκτιες παρατηρήσεις και το Cirrus (κανάλι 9) με στόχο την ανίχνευση των λεπτών σύννεφων. Σε αντίθεση με τον Landsat 7, ο Landsat 8 έχει δύο θερμικά υπέρυθρα κανάλια (κανάλια 10, 11). Τα δύο θερμικά κανάλια ακολουθούν την ίδια τεχνολογία με το σύστημα OLI και σκοπός τους είναι να μετράνε τη θερμική ενέργεια που εκπέμπεται από την επιφάνεια της γης. Είναι ευαίσθητα στη θερμοκρασία της επιφάνειας του εδάφους. Ο αισθητήρας, λόγω των διπλών θερμικών καναλιών, υπερέχει του ενός που έχει ο ETM+ αν και έχουν χωρική διακριτική ικανότητα 100 μέτρα (USGS).

## 1.2. Ενεργητικοί αισθητήρες

Από τον δορυφόρο ERS παρέχονται 3 τύποι δεδομένων μέσω του



αποθετηρίου της ESA (European Space Agency) και της πλατφόρμας EOli (Earth Observation Link). Οι 3 τύποι δεδομένων είναι τα αρχικά μη επεξεργασμένα δεδομένα (Raw Data) και δύο τύποι δεδομένων που παράγονται όταν ο αισθητήρας είναι σε Image mode. Οι δύο τύποι δεδομένων που παράγονται σε Image Mode είναι τα Single Look Complex (SLC) και Precision Image (PI).

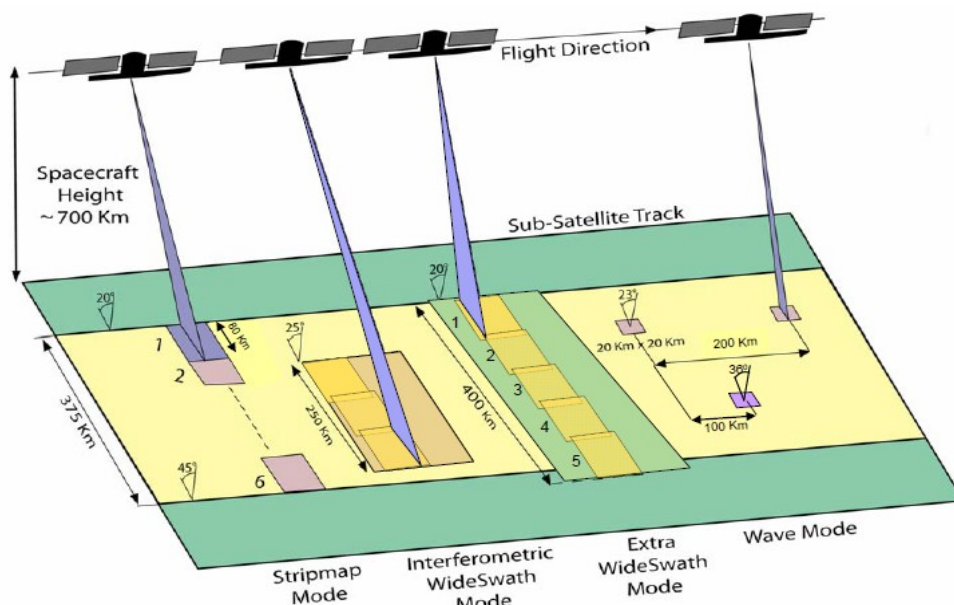
Τα SLC, παράγονται από Level 0 δεδομένα (Raw Data) και διατηρούν την φάση. Χρησιμοποιούνται κυρίως για εκτίμηση της ποιότητας, βαθμονόμηση και συμβολομετρικές (interferometry) εφαρμογές. Ακόμη έχει πραγματοποιηθεί ο ελάχιστος αριθμός διορθώσεων και παρεμβολών, ενώ οι παράμετροι βαθμονόμησης παρέχονται στα μεταδεδομένα των προϊόντων. Τέλος παρέχονται με χωρική διακριτική ικανότητα 8 μέτρων ως προς το range και 4 μέτρα ως προς το αζιμούθιο (μεταβάλλεται ελαφρώς ανάλογα με την συχνότητα επανάληψης των παλμών) (Earth ESA, 10/06/2019).

Τα PI, είναι προϊόντα πολλών όψεων (Multilook), με μειωμένο το φαινόμενο των κηλιδώσεων (speckle effect). Ο συγκεκριμένος τύπος αρχείων είναι πιο κατάλληλος για εφαρμογές τηλεπισκόπησης, αλλά και για σκοπούς βαθμονόμησης. Τα δεδομένα αυτά είναι βαθμονομημένα και διορθωμένα για της παραμορφώσεις των σφαλμάτων που προέρχονται από την κεραία, όπως η απώλεια της εξάπλωσης λόγω της εμβέλειας. Επίσης δεν έχουν υποστεί διορθώσεις για τα φαινόμενα που προκαλούνται από την τοπογραφία και επηρεάζουν την ραδιομετρία της εικόνας. Ακόμη τα δεδομένα δεν είναι γεω-κωδικοποιημένα και δεν έχουν αφαιρεθεί παραμορφώσεις της τοπογραφίας, όπως είναι η σμίκρυνση (foreshortening) και η πτύχωση (layover). Τέλος παρέχονται με χωρική διακριτική ικανότητα 12,5 μέτρων ως προς range και 12,5 μέτρων ως προς αζιμούθιο, με χωρική διακριτική ικανότητα 25 μέτρα, σύμφωνα με τους Amarsaikhani et al (2012). (Earth ESA, 10/06/2019)

Το πρόγραμμα Sentinel 1, είναι το πρώτο δορυφορικό πρόγραμμα που σχεδιάστηκε από την ESA, ως μέρος του προγράμματος Copernicus. Αποτελείται από δύο δίδυμους δορυφόρους τον Sentinel-1A, που εκτοξεύθηκε στις 03/04/2014 και τον Sentinel-1B, που εκτοξεύθηκε στις 25/05/2016. Δημιουργήθηκαν με σκοπό να είναι η συνέχεια των δορυφόρων ERS και Envisat.

Είναι εφοδιασμένο με το όργανο radar συνθετικού ανοίγματος (SAR), το οποίο λειτουργεί στην ζώνη μικροκυματικής C, με κάθετη και οριζόντια πόλωση, ενώ παρέχει δεδομένα σε μονή (HH, VV), δηλαδή στέλνει και δέχεται παλμούς με ίδια πόλωση, και διπλή πόλωση (HH+HV, VV+VH), δηλαδή στέλνει παλμούς σε μία πόλωση (H ή V) και δέχεται με δύο πολώσεις (HV ή VH).

Επίσης το όργανο SAR, είναι ικανό να λειτουργεί σε κύκλο 25 λεπτών ανά τροχιά, στις λειτουργίες λήψεις Stripmap Mode (SM), Interferometric



Εικόνα 1: Σχεδιαστική απεικόνιση τρόπων λειτουργίας αισθητήρα Sentinel 1

πηγή: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/sar-instrument/acquisition-modes>

Wide Swath Mode (IW), Extra Wide Swath Mode (EW) και Wave Mode (WV). Επιπροσθέτως στις λειτουργίες λήψεις SM, IW και EW modes, παρέχονται δεδομένα και μονής πόλωσης, αλλά και διπλής πόλωσης, σε αντίθεση με την λειτουργία λήψης WV που παρέχονται εικόνες μόνο μονής πόλωσης.

Η λειτουργία λήψης Stripmap Mode (SM), είναι μια τυπική λειτουργία των SAR, όπου η λωρίδα γης φωτίζεται με συνεχή ακολουθία παλμών (6 παλμοί), ενώ η ακτίνα από την κεραία “δείχνει” σε μια σταθερή γωνία ως προς το αζιμούθιο και σε μια σχεδόν σταθερή γωνία εκτός ναδίρ.

Η λειτουργία λήψης Interferometric Wide Swath Mode (IW), λαμβάνει δεδομένα από μια ευρύτερη λωρίδα, όπου είναι αποτέλεσμα 3 υπο-λωρίδων, χρησιμοποιώντας την τεχνική απεικόνισης TOPSAR. Η συγκεκριμένη τεχνική είναι μια μορφή της απεικόνισης ScanSAR (η ακτίνα της κεραία γυρίζει κυκλικά αναμεταξύ των υπο-λωρίδων), όπου για κάθε παλμό, ο οποίος είναι συγχρονισμένος σε κάθε πέρασμα από την ίδια περιοχή, η ακτίνα γυρίζει ηλεκτρονικά από πίσω προς τα εμπρός στην κατεύθυνση του αζιμούθιου. Με αυτό τον τρόπο επιτυγχάνεται μεγαλύτερη ποιότητα στην εικόνα.

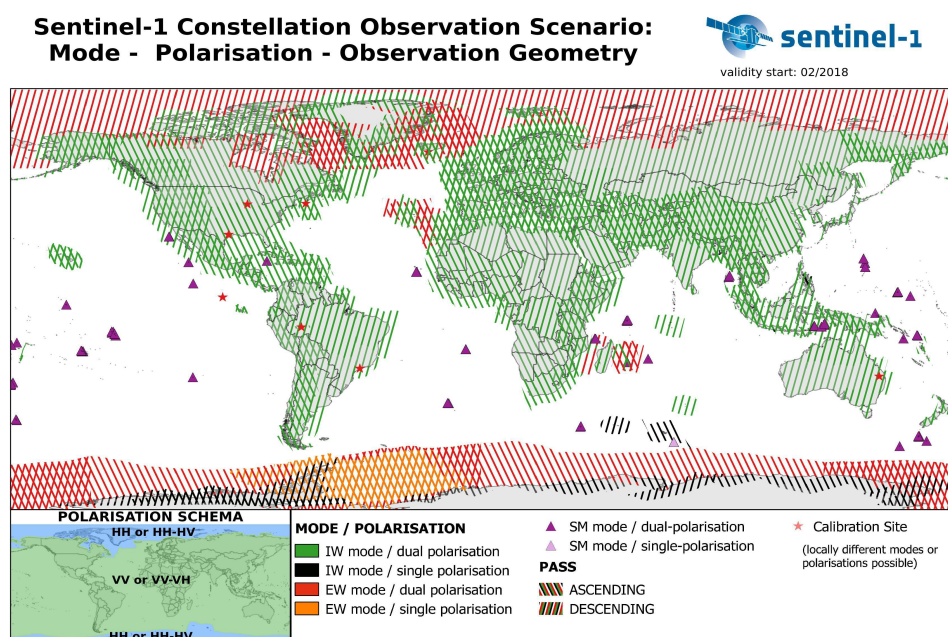
Η λειτουργία λήψης Extra-Wide Swath Mode (EW), χρησιμοποιεί και αυτή την τεχνική απεικόνισης TOPSAR. Επίσης παρέχει μια πολύ μεγάλη λωρίδα κάλυψης, η οποία δημιουργείται από 5 υπο-λωρίδες, εις βάρος μιας περαιτέρω μείωσης της ανάλυσης.

Η λειτουργία λήψης Wave Mode (WV), αποκτά μικρές σκηνές Stripmap, γνωστές ως βινιέτες, που βρίσκονται σε τακτά χωρικά διαστήματα των 100 χιλιομέτρων κατά μήκος της διαδρομής, ανάμεσα στο κέντρο των διαδοχικών βινιετών, όπως ο ERS και ο ENVISAT στην απεικόνιση ASAR Wave. Με αυτό τον τρόπο και την δημιουργία υπό-δειγματοληψίας, παρέχεται η δυνατότητα δημιουργία χαμηλού όγκου δεδομένων. Επίσης ο

τρόπος με τον οποίο καταγράφονται οι βινιέτες ονομάζεται “άλμα βατράχου” (“leap-frog”), και ουσιαστικά η πρώτη βινιέτα αποκτάται σε μικρή γωνία πρόσπτωσης σε κοντινή απόσταση ενώ η επόμενη σε μεγαλύτερη απόσταση και με μεγάλη γωνία πρόσπτωσης. Τέλος η συγκεκριμένη λειτουργία λήψης, σχεδιάστηκε συγκεκριμένα για εφαρμογές σχετικά με τους ωκεανούς.

Πάνω από το έδαφος και τις παράκτιες περιοχές ο δορυφόρος λειτουργεί σε IW mode, ενώ πάνω από τους ωκεανούς σε WV mode και πάνω από τις θάλασσες και τους πόλους, λειτουργεί σε IW και EW mode.

Η πόλωση για της περιοχές που βρίσκονται σε πολικό περιβάλλον, και σε περιοχές πάγου και θάλασσας, είναι είτε HH-HV είτε σε HH, ενώ στις υπόλοιπες περιοχές είναι σε VV-VH ή σε VV, εκτός από την Βαλτική Θάλασσα, όπου παρατηρείται και με HH-HV.



Εικόνα 2: Περιοχές κάλυψης Sentinel 1, ανάλογα τον τρόπο λειτουργίας του αισθητήρα και την πόλωση

πηγή: <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/observation-scenario>

Τα δεδομένα που αποκτούνται από τον Sentinel-1A/B από τις

προαναφερθέντες λειτουργίες λήψης, κατηγοριοποιούνται ως δεδομένα επιπέδου 1 (Level 1 Products) και επιπέδου 2 (Level 2 Products). Τα δεδομένα επιπέδου 1 χωρίζονται σε Single-Look Complex (SLC) και σε Ground Range Detected (GRD), ενώ ανάλογα με την ανάλυση κατηγοριοποιούνται σε πλήρη ανάλυση (Full Resolution - FR), σε υψηλή ανάλυση (High Resolution - HR) και σε μεσαία ανάλυση (Medium Resolution - MR). Τα δεδομένα επιπέδου 2 είναι για εφαρμογές σχετικά με τους ωκεανούς και χωρίζονται σε Ocean SWell spectra (OSW), Ocean Wind field (OWI) και Radial Surface VeLocity (RVL) (Sentinel missions, 10/06/2019).

## **2. Μέθοδοι ταξινόμησης δορυφορικών δεδομένων**

Οι χρήσεις αλλά και οι καλύψεις γης μιας οποιασδήποτε περιοχής, είναι από τις θεμελιώδεις μεταβλητές όπου επηρεάζουν το φυσικό καθώς και το ανθρωπογενές περιβάλλον.

Με την χρήση των δορυφορικών δεδομένων, γίνεται δυνατή η δημιουργία θεματικών χαρτών, για την αποτύπωση της εδαφοκάλυψης, καθώς αναπαριστούν την αποτύπωση της γήινης επιφάνειας, με συγκεκριμένη χωρική αλλά και χρονική κλίμακα. Για να δημιουργηθούν αυτοί οι θεματικοί χάρτες, συνήθως, τα δορυφορικά δεδομένα ταξινομούνται με χρήση διάφορων τεχνικών (Foody, 2002).

Μια πληθώρα μεθόδων ταξινόμησης έχουν δημιουργηθεί και χρησιμοποιηθεί για την δημιουργία θεματικών χαρτών εδαφοκάλυψης με την χρήση δορυφορικών δεδομένων. Ταξινόμηση στην τηλεπισκόπηση, είναι η μέθοδος όπου κάθε εικονοστοιχείο μιας δορυφορικής εικόνας, αποκτά μια “ταμπέλα” με βάση τα χαρακτηριστικά του. Αυτά τα χαρακτηριστικά του μπορεί να είναι είτε οι φασματικές του αποκρίσεις σε

διάφορα μήκη κύματος (οπτικοί δέκτες), είτε η υφή του (εικονοληπτικά Radar) (Pal και Mather, 2001). Οι μέθοδοι αυτές, κατηγοριοποιούνται σε αυτές της μη επιβλεπόμενης ταξινόμησης, όπως ο αλγόριθμος K-means, της επιβλεπόμενης ταξινόμησης, όπως ο αλγόριθμος Maximum Likelihood (μέγιστης πιθανοφάνειας), και των αλγορίθμων εκπαίδευσης, όπως τα τεχνητά νευρωνικά δίκτυα (artificial neural networks), τα δέντρα απόφασης (Decision Trees), οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines) και οι ταξινομητές συνόλων (Ensemble of Classifiers) (Rodriguez-Galiano et al, 2012).

Στην παρούσα εργασία έγινε χρήση του ταξινομητή Random Forest, ο οποίος ταξινομητής συνόλων και είναι ένας συνδυασμός αλγορίθμου εκπαίδευσης μηχανής (Machine Learning Algorithm) με ταξινομητή δέντρων απόφασης. Στις παρακάτω 2 υποενότητες γίνεται μία εισαγωγή στον αλγόριθμο Random Forest για την κατανόηση της λειτουργίας του. Αρχικά θα αναφερθούμε στους ταξινομητές δέντρων απόφασης και στην συνέχεια στον αλγόριθμο RF.

## **2.1. Ταξινομητές Δέντρων Απόφασης**

Οι ταξινομητές δέντρων απόφασης (από εδώ και στο εξής θα αναφέρονται ως DTC, Decision Trees Classifier) χρησιμοποιούνται ευρεία στην τηλεπισκόπηση. Η κύρια ικανότητα των δέντρων απόφασης (από εδώ και στο εξής θα αναφέρονται DT, Decision Trees), είναι να “σπάνε” μία σύνθετη διαδικασία απόφασης σε άλλες απλούστερες, παρέχοντας έτσι μια λύση που είναι ευκολότερη να ερμηνευθεί. Ανήκουν στην κατηγορία των αλγορίθμων εκπαίδευσης, όπου χρησιμοποιούν τυχαίο δείγμα των δεδομένων, αντιπροσωπευτικό του πληθυσμού και συνήθως είναι τα 2/3 του συνόλου των δεδομένων, ως δεδομένα εκπαίδευσης και το υπόλοιπο 1/3

των δεδομένων ως δεδομένα ελέγχου, ενώ ανήκει και στην κατηγορία ταξινομητών επιπέδου εικονοστοιχείου. Σκοπός των DTC, είναι με βάση τα χαρακτηριστικά ενός συνόλου αντικειμένων, να υπολογιστεί η κλάση που ανήκει το κάθε αντικείμενο έπειτα από την εκπαίδευση του αλγορίθμου.

Τα δέντρα αποτελούνται από κόμβους και από πλευρές που ενώνουν τους κόμβους. Σε κάθε κόμβο εισέρχεται μία πλευρά εκτός από τον αρχικό κόμβο που ονομάζεται ρίζα ή τερματικός κόμβος (root node), ο οποίος περιέχει όλες τις κατηγορίες στις οποίες μπορεί να ταξινομηθεί ένα δείγμα. Κάθε πλευρά ενώνει 2 κόμβους, έστω ο “Z” κόμβος ενώνεται με τον “Ω”, ο “Z” ονομάζεται πρόγονος/πατέρας του “Ω”, ενώ ο “Ω” απόγονος/γιος του “Z”. Επίσης κάθε κόμβος που δεν έχει απόγονο/γιο, ονομάζεται τερματικός κόμβος ή φύλλο. Ακόμη κάθε κόμβος έχει έναν πρόγονο, εκτός από τον αρχικό κόμβο ή αλλιώς ρίζα, και 2 ή περισσότερους απογόνους, εκτός από τον τερματικό κόμβο ή αλλιώς φύλλο. Τέλος, κάθε κόμβος έχει ακριβώς μία εισερχόμενη πλευρά σε αυτόν.

Δυαδικό δέντρο (binary tree) ονομάζεται κάθε δέντρο όπου κάθε κόμβος διακρίνεται σε δεξιός ή αριστερός γιος και κάθε κόμβος έχει μόνο 1 αριστερό ή δεξιό γιο. Ενώ αν δύο εσωτερικοί κόμβοι περιέχουν τουλάχιστον 1 κοινή κλάση, τότε λέμε πως έχουν κατηγορίες επικάλυψης. Μέσο πλάτος ενός δέντρου, ονομάζεται ο μέσος όρος κόμβων σε ένα Level ή Layer του δέντρου. Ενώ μέσο βάθος ονομάζεται ο μέσος όρος των Level σε ένα δέντρο. Το μέσο πλάτος, αντανακλά στο σχετικό βάρος της ακρίβειας του ταξινομητή ενώ το μέσο βάθος στο σχετικό βάρος που δίνεται στην αποδοτικότητα του δέντρου. Σε κάθε Level του δέντρου, η απόφαση διαχωρίζεται σε έναν κόμβο χαρακτηριστικών που παρουσιάζεται ως τρίγωνο, σύμφωνα με τον Slagle, και τους κόμβους των αποφάσεων, οι οποίοι δείχνουν τα πιθανά μονοπάτια της ταξινόμησης στο επόμενο Level. (Slagle, 1971). Ως ύψος ενός κόμβου ορίζεται το μέγιστο μήκος

μονοπατιού από τον κόμβο “Z” μέχρι το φύλλο του δέντρου. (Hauska και Swain, 1975)

Ένα καλό δέντρο θα πρέπει να έχει ίδιο πλάτος και βάθος, αν και δεν είναι πάντα εφικτό, θα πρέπει να μην έχει τερματικό/φύλλο, στο πρώτο Level (ή Layer) και να έχει όσο το δυνατόν περισσότερους κόμβους στο 2ο Level. Τέλος, η μορφή του δέντρου, σχετίζεται με την διαχωριστικότητα εσωτερικά των κλάσεων και το χρησιμοποιούμενο μέτρο απόστασης.

Οι DTC, όπως προαναφέρθηκε, διαιρούν τα δεδομένα εκπαίδευσης σε υποσύνολα, όπου το καθένα αντιπροσωπεύει μια κλάση. Τα αποτελέσματα αυτής της διαδικασίας είναι συνήθως μεγάλα και πολύπλοκα δέντρα. Για να μειωθεί το συγκεκριμένο πρόβλημα προτείνεται να γίνει “κλάδεμα” του δέντρου για να μειωθούν τα σφάλματα ταξινόμησης (Pal και Mather, 2001), καθώς αυτοί οι κόμβοι που θα αφαιρεθούν, μπορούν να συνδεθούν με θόρυβο των δεδομένων, αφού πρώτα γίνει η εκπαίδευση του αλγορίθμου. Με αυτό τον τρόπο επιτυγχάνεται η βελτίωση της προβλεπτικής ισχύς του μοντέλου, όμως το “κλάδεμα” είναι κάπως υποκειμενική διαδικασία και δίνει το ελάχιστο σφάλμα για ένα σύνολο δεδομένων ελέγχου (Hornig, 2010).

Για τον υπολογισμό των σφαλμάτων ταξινόμησης, ύστερα από την χρήση των DT, χρησιμοποιείται ο λόγος των δειγμάτων που ταξινομήθηκαν λάθος, προς τον αριθμό των δειγμάτων ελέγχου που χρησιμοποιήθηκαν συνολικά. Μια πιο σύνθετη μέθοδος εκτίμησης του ποσοστού των σφαλμάτων ταξινόμησης είναι η μέθοδος K-fold cross-validation, όπου το πλήθος των δεδομένων  $X$ , χωρίζονται σε  $N$  σχεδόν ίσα μέρη ( $X_1, X_2, \dots, X_N$ ), με  $N=10$  συνήθως. Έπειτα χρησιμοποιεί τα  $X-X_N$  ως δεδομένα εκπαίδευσης και τα  $X_N$  ως δεδομένα ελέγχου. Αυτό επαναλαμβάνεται για  $n$  φορές και τελικά προκύπτει ο μέσος όρος σφαλμάτων ταξινόμησης για το



N.

Τα πλεονεκτήματα των DTC είναι ότι μπορεί ο εκπαιδευτής του αλγορίθμου να τερματίσει οποιονδήποτε κόμβο αν δεν τον ενδιαφέρει η ταξινόμηση που προκύπτει από αυτόν. Επίσης περίπλοκες αποφάσεις, μπορούν να παρθούν με την ένωση απλούστερων αποφάσεων (μικρότερα δέντρα) και την ένωσή τους σε διάφορα επίπεδα του δέντρου.

Επιπροσθέτως, με την χρήση των DTC, επιτυγχάνονται γρήγορες ταξινομήσεις μόλις αναπτυχθούν οι κανόνες ταξινόμησης, καθώς είναι και εύκολη η παρουσίαση και ερμηνεία των κανόνων, όταν χρησιμοποιούνται οι DTC. Επειδή είναι μη παραμετρικός αλγόριθμος, είναι εύκολο να ενσωματώσει ένα εύρος αριθμητικών ή κατηγορικών δεδομένων (Hornig, 2010).

Ακόμη ένα πλεονέκτημα των DTC είναι ότι κάθε δείγμα εξετάζεται για ένα υποσύνολο των κλάσεων, δηλαδή παρέχεται η ευελιξία στον αναλυτή να επιλέξει διαφορετικά υποσύνολα χαρακτηριστικών του δείγματος σε διαφορετικούς μη τερματικούς κόμβους, επιτυγχάνοντας έτσι την διάκριση των κλάσεων εντός του κόμβου με καλύτερο τρόπο. Αντίθετα, άλλοι ταξινομητές χρησιμοποιούν το σύνολο των χαρακτηριστικών του δείγματος, για να διακριθούν όλες οι κλάσεις.

Τέλος άλλο ένα πλεονέκτημα των DTC, είναι ότι μπορεί σε πολυμεταβλητή ανάλυση με μεγάλο πλήθος χαρακτηριστικών, σε ένα μικρό δείγμα δεδομένων εκπαίδευσης, να αποφευχθεί το πρόβλημα της πολυδιάστασης των δεδομένων, χρησιμοποιώντας μικρότερο πλήθος χαρακτηριστικών για κάθε μη τερματικό κόμβο και χωρίς σημαντική μείωση την επιδόσεων του ταξινομητή.

Τα DT όμως, παρουσιάζουν και μειονεκτήματα. Ένα μειονέκτημα είναι όταν υπάρχει επικάλυψη, και ειδικότερα όταν ο αριθμός των κλάσεων είναι μεγάλος, τότε μπορεί ο αριθμός των φύλλων, να είναι αρκετά μεγαλύτερος,

ακόμη και από των αριθμών των κλάσεων, αυξάνοντας έτσι και τον χρόνο αναζήτησης, αλλά και τις απαιτήσεις για χώρο μνήμης.

Άλλο ένα μειονέκτημα, είναι ότι τα σφάλματα μπορεί να συσσωρεύονται από το ένα επίπεδο στο άλλο σε περιπτώσεις μεγάλων δέντρων, ενώ δεν μπορεί ταυτόχρονα να βελτιστοποιείται η ακρίβεια και η αποτελεσματικότητα.

Τέλος ένα ακόμη μειονέκτημα είναι ότι η επίδοση ενός ταξινομητή DT, εξαρτάται σημαντικά από το πόσο καλά έχει σχεδιαστεί το δέντρο, και το αν οι μεταβλητές που εξετάζονται είναι ανεξάρτητες μεταξύ τους και τα χαρακτηριστικά που χρησιμοποιούνται για την εκπαίδευση του αλγορίθμου, πρέπει να είναι απλά και κατανοητά. (Safavian και Landgrebe, 1990)

## 2.2. Ταξινομητής Random Forest

Ο ταξινομητής Random Forest (από εδώ και στο εξής θα αναφέρεται ως RF), είναι ένας συνδυασμός machine learning αλγορίθμου και DT (δέντρων απόφασης) και θεωρείται ταξινομητής συνόλων, δηλαδή χρησιμοποιεί τα αποτελέσματα από πολλά διαφορετικά μοντέλα, για να βγάλει μια απόφαση. Στην συγκεκριμένη περίπτωση, τα διαφορετικά μοντέλα αυτά είναι μερικές εκατοντάδες DT, στην πλειοψηφία των εργασιών είναι 500 δέντρα καθώς πριν επιτευχθεί ο συγκεκριμένος αριθμός δέντρων το σφάλμα σταθεροποιείται (Lawrence et al, 2006), και το αποτέλεσμα προκύπτει με βάση τα αποτελέσματα των δέντρων (Hornung, 2010).

Όπως προαναφέρθηκε ο RF είναι ένας ταξινομητής που αποτελείται από πολλά DT, όπου σε αυτά κατανέμονται ομοιόμορφα, τυχαία διανύσματα, δεδομένα δηλαδή, και καθένα από αυτά βγάζει μία “ψηφοφορία” για την πιο δημοφιλή κλάση (Nguyen et al, 2018). Όσες περισσότερες φορές, δηλαδή σε όσα περισσότερα DT, ένα αντικείμενο αποκτήσει μία κλάση,

τότε αυτή θα είναι και η τελική του κλάση (Breiman, 2001).

Είναι ένας μη-παραμετρικός ταξινομητής, δηλαδή δεν απαιτείται η εκτίμηση της στατιστικής κατανομής των δεδομένων και δεν χρειάζεται να είναι αναγνωρίσιμο το σχήμα της υποκείμενης δομής δεδομένων (δηλαδή δεν χρειάζεται τα δεδομένα να ακολουθούν κανονική κατανομή), με αποτέλεσμα να είναι πολύ ελκυστικός για σκοπούς ταξινόμησης (Toosi et al, 2019).

Η κατασκευή των DT, που χρησιμοποιούνται από τον RF, χωρίζεται σε 2 βήματα σύμφωνα με την βιβλιογραφία. Το πρώτο βήμα είναι η τυχαία επιλογή, με αντικατάσταση κάθε φορά, δεδομένων από τα παρεχόμενα δεδομένα εκπαίδευσης για την δημιουργία του δέντρου. Για κάθε δέντρο, χρησιμοποιείται ένα διαφορετικό υποσύνολο δεδομένων εκπαίδευσης. Με αυτό τον τρόπο χρησιμοποιούνται τα 2/3 των δεδομένων εκπαίδευσης, ενώ το 1/3 αυτών χρησιμοποιείται για να δοκιμαστεί η ακρίβεια του μοντέλου. Αυτό το 1/3 των δεδομένων εκπαίδευσης, απαντάται στην βιβλιογραφία ως “out-of-bag” δείγμα.

Το δεύτερο βήμα για την κατασκευή των DT, χρησιμοποιείται τυχαία δειγματοληψία, για τον προσδιορισμό της συνθήκης χωρισμού της απόφασης για κάθε κόμβο ενός δέντρου. Δηλαδή σε κάθε κόμβο ένα υποσύνολο των μεταβλητών πρόβλεψης επιλέγεται τυχαία για την δημιουργία του δυαδικού κανόνα (με την παρακολούθηση του σφάλματος, της δύναμης και της συσχέτισης, επιλέγεται ο αριθμός των χαρακτηριστικών που χρησιμοποιούνται για τον διαχωρισμό της απόφασης στους κόμβους). Ο αριθμός του υποσυνόλου μεταβλητών πρόβλεψης, μπορεί να επιλεγεί από τον χρήστη, ή να γίνει από τον ίδιο τον αλγόριθμο. Με την τυχαία επιλογή τους από τον αλγόριθμο, επιτυγχάνεται λιγότερη συσχέτιση μεταξύ των δέντρων που έχει ως αποτέλεσμα μικρότερο ποσοστό σφάλματος. Γενικά ισχύει ότι αν χρησιμοποιηθεί σε κάθε δέντρο, όλο το

πλήθος των μεταβλητών πρόβλεψης, τότε τα δέντρα θα είναι παρόμοια, άρα θα έχουν μεγάλη συσχέτιση μεταξύ τους, με αποτέλεσμα τα ποσοστά σφάλματος να είναι μεγάλα, ενώ αν χρησιμοποιηθεί πολύ μικρότερος αριθμός μεταβλητών πρόβλεψης σε κάθε δέντρο, αυτό θα έχει ως συνέπεια μικρότερη συσχέτιση και ως επακόλουθο αυτού, ο RF να έχει λιγότερη προβλεψιμότητα από τα δέντρα με περισσότερες μεταβλητές.

Από το 2ο βήμα κατασκευής των DT, του RF, προκύπτει ότι γενικά πρέπει να επιλέγεται ο αριθμός των μεταβλητών που θα χρησιμοποιηθεί και ο οποίος παρέχει χαμηλή αυτοσυσχέτιση μεταξύ των δέντρων με επαρκή προβλεπτική ισχύ όμως (Breiman, 2001).

Το βέλτιστο εύρος υποσυνόλων των μεταβλητών πρόβλεψης, είναι αρκετά ευρύ, υπάρχουν όμως εύκολοι έλεγχοι που μπορούν να γίνουν για την επιλογή του βέλτιστου μεγέθους, σύμφωνα με τον Pal (2005). Σύμφωνα με τους Gislason et al (2006), συχνά επιλέγεται ως η τετραγωνική ρίζα των δεδομένων εισόδου, δηλαδή των μεταβλητών πρόβλεψης.

Σύμφωνα με τους Liaw και Wiener (2002), υπάρχουν 4 παράμετροι που πρέπει να καθοριστούν. Η πρώτη παράμετρος είναι ότι πρέπει να έχουμε ως μεταβλητές πρόβλεψης, ως input δηλαδή, το ΨΜΕ (Ψηφιακό Μοντέλο Εδάφους, DEM) καθώς και τα κανάλια του δορυφόρου που χρησιμοποιούμε. Η δεύτερη παράμετρος που πρέπει να καθοριστεί είναι ο αριθμός των δέντρων που πρέπει να φτιαχτούν, που μπορεί να είναι τόσο μεγάλος, όσο επιτρέπει η μνήμη του υπολογιστή (Guan et al, 2013). Η τρίτη παράμετρος είναι ο αριθμός των μεταβλητών πρόβλεψης που θα χρησιμοποιηθούν για την κατασκευή του δυαδικού κανόνα για κάθε χώρισμα (split). Η τέταρτη παράμετρος που πρέπει να καθοριστεί είναι οι παράμετροι για τον υπολογισμό των βοηθητικών πληροφοριών σχετικά με τα σφάλματα ταξινόμησης και την σημαντικότητα των μεταβλητών. Τέλος πρέπει να σημειωθεί πως οι βοηθητικές πληροφορίες, μαζί με τις ακραίες

τιμές, μπορούν να χρησιμοποιηθούν για την αξιολόγηση της απόδοσης του μοντέλου και για αλλαγές στα δεδομένα εκπαίδευσης, αν είναι απαραίτητο.

Μέσω του διαγράμματος του ποσοστού σφαλμάτων με το πλήθος δέντρων του δάσους, μας δίνεται μια ένδειξη για το αν ο αριθμός των δέντρων είναι επαρκής. Για να είναι επαρκής ο αριθμός των δέντρων, πρέπει το ποσοστό σφαλμάτων να είναι σταθερό (Hornig, 2010).

Τα πλεονεκτήματα του RF, είναι ότι ξεπερνά τα προβλήματα των DT, καθώς χρησιμοποιεί εκατοντάδες DT, για να πάρει μια απόφαση, ενώ ταυτόχρονα διατηρεί τα οφέλη τους. Επίσης οι παράμετροι που πρέπει να χρησιμοποιηθούν είναι εύκολο να προσδιοριστούν, ενώ δεν είναι πολύ ευαίσθητος στις παραμέτρους που χρησιμοποιούνται για την εκτέλεσή του, σε σχέση με άλλους ταξινομητές αλγορίθμων εκμάθησης.

Ακόμη, είναι εύκολα παραμετροποιήσιμος, ενώ μπορεί να χρησιμοποιεί είτε συνεχόμενα είτε κατηγορικά σύνολα δεδομένων, είτε και τα δύο σύνολα δεδομένων ταυτόχρονα. Επίσης μπορεί να διαχειριστεί χιλιάδες μεταβλητές εισόδου. Επιπλέον, ο RF, είναι ικανός να αντιμετωπίσει ακραίες τιμές, ενώ παράλληλα δεν είναι ευαίσθητος στον θόρυβο και στις ακραίες τιμές των δεδομένων εκπαίδευσης. Επιπροσθέτως, υπολογίζει βοηθητικές πληροφορίες, όπως σφάλματα ταξινόμησης αλλά και την σημαντικότητα των μεταβλητών. Τέλος ο RF, δημιουργεί μια εσωτερική αμερόληπτη εκτίμηση του σφάλματος γενίκευσης, σφάλμα OOB (Hornig, 2010; Rodriguez-Galiano et al,2012; Mellor et al,2012; Ghosh et al, 2014).

Εν κατακλείδι αξίζει να σημειωθεί πως δεν βρέθηκαν μειονεκτήματα για τον συγκεκριμένο αλγόριθμο, RF, στην βιβλιογραφία. Μόνο οι Nitze et al (2012), αναφέρουν ως μειονέκτημα, γενικά για τους αλγορίθμους εκπαίδευσης, ότι η υπολογιστική τους πολυπλοκότητα είναι υψηλότερη σε σχέση με τους απλούς επιβλεπόμενους ταξινομητές, όπως αυτός της μέγιστης πιθανοφάνειας ή του κοντινότερου γείτονα.

Πολλοί ερευνητές έχουν χρησιμοποιήσει τον RF για να παρακολουθήσουν τις χρήσεις/καλύψεις γης. Οι Stefanski et al (2013), χρησιμοποίησαν τρεις δορυφορικές εικόνες RapidEye με χωρική διακριτική ικανότητα 5 μέτρων για μια περιοχή κοντά στο Λουξεμβούργο και 2 δορυφορικές εικόνες SPOT5 με διπλάσια χωρική διακριτική ικανότητα από αυτή του RapidEye για μια περιοχή κοντά στη Βόννη της Γερμανίας με σκοπό να βρουν έναν ημιαυτόματο τρόπο εύρεσης των παραμέτρων για τον αλγόριθμο κατάτμησης εικόνων, με χρήση του RF και του αλγορίθμου Superpixel Contour (SPC). Για να πετύχουν τον σκοπό τους δημιούργησαν θεματικά επίπεδα κάλυψης γης και για τις δύο υπομελέτη περιοχές και με τους δύο προαναφερόμενους αλγορίθμους, πετυχαίνοντας πάνω από 86% ακρίβεια ταξινόμησης με τον αλγόριθμο RF.

Για το έτος 2010, παράχθηκε χάρτης εδαφοκάλυψης για όλη την Ελλάδα, από τους Gounaridis et al (2015), με χρήση του RF και 27 εικόνων Landsat 5 & 7 για τα έτη 2008 έως 2011. Οι εικόνες διορθώθηκαν ατμοσφαιρικά, ραδιομετρικά και γεωμετρικά και στην συνέχεια με ημιαυτόματο τρόπο δημιούργησαν δεδομένα εκπαίδευσης. Ως μεταβλητές πρόβλεψης χρησιμοποίησαν τις 7 μπάντες των δορυφόρων, 2 PCA (μια για το ορατό φάσμα και 1 για τις μπάντες του κοντινού υπέρυθρου), μετασχηματισμό Tasseled Cap, τον δείκτη βλάστησης NDVI, μερικές αναλογίες μπαντών και την κλίση και το υψόμετρο. Συνολικά οι μεταβλητές πρόβλεψης ήταν 20. Με την χρήση του RF, των 20 θεματικών επιπέδων και των δεδομένων εκπαίδευσης, παράχθηκε χάρτης εδαφοκάλυψης με 12 κλάσεις με ακρίβεια ταξινόμησης 83%.

Ο Colditz (2015), σύγκρινε διάφορες στρατηγικές κατανομές δεδομένων εκπαίδευσης για συνεχής ή διακριτές εδαφοκαλύψεις χρησιμοποιώντας αλγορίθμους δέντρων απόφασης ταξινόμησης και παλινδρόμησης. Συγκεκριμένα χρησιμοποιήθηκαν δορυφορικά δεδομένα MODIS, και των

δύο αισθητήρων Aqua και Terra, για την χρονική περίοδο Οκτώβριου 2005 με Μάρτιο του 2007, καθώς και δεδομένα Landsat, καθώς και δεδομένα από το NLCD (National Land Cover Data set). Από τα παραπάνω δεδομένα δημιουργήθηκαν χάρτες εδαφοκάλυψης. Οι κλάσεις ταξινόμησης, με τον RF, προέκυψαν από 1000 δέντρα και ο αριθμός των μεταβλητών για τον χωρισμό της απόφασης σε κάθε κόμβο τέθηκε στις 20 μεταβλητές, όπου ήταν η τετραγωνική ρίζα του συνόλου των μεταβλητών. Τα αποτελέσματα ταξινόμησης ήταν καλύτερα από τον RF σε σχέση με την C5.0 (boosted μέθοδος του RF).

Για τον εντοπισμό των αλλαγών της εδαφοκάλυψης της Λατινικής Αμερικής και της Καραϊβικής για τα έτη 2001-2010, χρησιμοποιήθηκαν 8970 σκηνές του MODIS (16-day composite), από τις ημερήσιες εικόνες και χρησιμοποιήθηκαν μόνο τα εικονοστοιχεία με την μεγαλύτερη ποιότητα. Επίσης χρησιμοποιήθηκαν 2 δείκτες βλάστησης (NDVI, EVI), στατιστικά από την χρονοσειρά του MOD 13, ανά τετράμηνο, εξάμηνο και ετήσια. Συνολικά είχαν 120 μεταβλητές πρόβλεψης και περίπου 45000 σημεία εκπαίδευσης, ενώ χρησιμοποίησαν 1999 δέντρα και 11 μεταβλητές για τον χωρισμό της απόφασης στον κάθε κόμβο. Και για τις 26 περιοχές η ακρίβεια της ταξινόμησης με τον RF, έφτανε για 9 κλάσεις χρήσεων/καλύψεων γης περίπου στο 80,2%+-8,1%, και για τις 5 γενικευμένες κλάσεις κατά μέσο όρο στο 84,6%+-6,5% (Clark et al, 2012).

Οι Clark et al (2010), χρησιμοποίησαν 184 δεκαεξαήμερες εικόνες του MODIS, για τα έτη 2000 έως 2008 για την οικοπεριοχή Dry Chaco, η οποία καταλαμβάνει εδάφη σε Αργεντινή, Βολιβία και Παραγουάη, για να παρουσιάσουν μία μέθοδο για την ετήσια χαρτογράφηση των χρήσεων/καλύψεως γης. Η ταξινόμηση έγινε με χρήση του RF, με 1000 δέντρα για 9 κλάσεις χρήσης/κάλυψης γης. Στον τελικό χάρτη καλύψεων/χρήσεων γης η συνολική ακρίβεια ήταν 74,2%.

Ακόμη έγινε εντοπισμός αλλαγών εδαφοκάλυψης με PCC (Post-Classification Comparison) και ταξινόμηση 5 κλάσεων, για την περιοχή Huntington Wildfire Forest, με χρήση του αλγορίθμου RF σε 6 εικόνες Landsat 2 για το 1991, μια χειμερινή και μία εαρινή, με βάση τον κύκλο ανάπτυξης της φυτοκάλυψης που υπήρχε στην περιοχή μελέτης, 2 για το 2001 και 2 για το 2010. Τα αποτελέσματα της ταξινόμησης έδειξαν σαφώς καλύτερα αποτελέσματα για τον RF σε σχέση με τα DT και τα SVM, για τις χρονιές 1991 (84,5%) και 2001 (86,4%), ενώ για το 2011 τα SVM έβγαλαν καλύτερα αποτελέσματα (88,2%) σε σχέση με τον RF (87,3%). Στην συνέχεια εντόπισαν τις αλλαγές στην εδαφοκάλυψη με βάση τις ταξινομημένες εικόνες του 1991, 2001 και 2011. Ωστόσο οι συγγραφείς αναφέρουν ότι ο εντοπισμός αλλαγών έχει διάφορους περιορισμούς, καθώς σε τόσο μικρά χρονικά διαστήματα, υπάρχει πολύ μικρή πιθανότητα να έχει αλλάξει ο τύπος του δάσους μία υποπεριοχή καθώς και ότι ο μικρός αριθμός περιοχών δειγματοληψίας δεν θα μπορούσε να παράσχει αξιόπιστα στατιστικά αποτελέσματα (Li et al, 2013).

Οι Sesnie et al (2008) χρησιμοποίησαν εικόνες Landsat για τα έτη 1986 και 2001 και Ψηφιακό Μοντέλο Εδάφους μίας δασικής έκτασης ανάμεσα σε Νικαράγουα και Κόστα Ρίκα. Στην συνέχεια δημιούργησαν ένα μωσαϊκό για κάθε χρονιά. Το κάθε μωσαϊκό, μαζί με δείκτες βλάστησης και άλλα δεδομένα, συνολικά 35 μεταβλητές πρόβλεψης, ταξινομήθηκαν με χρήση του RF και έπειτα με PCC εντόπισαν τις αλλαγές στους τύπους βλάστησης του δάσους.

Τέλος για την πρόβλεψη της αστικής εξάπλωσης το έτος 2045, οι Gounaridis et al (2018) χρησιμοποίησαν εικόνες των δορυφόρων Landsat 4, 5, 7 και 8. Για να δημιουργήσουν το μοντέλο πρόβλεψης, χρειάστηκε να γίνει εντοπισμός των αλλαγών στην χρήση/κάλυψη γης από το 1991 μέχρι το 2015, των Μεσογείων Αττικής. Οι χρήσεις/καλύψεις γης ήταν αστικός



ιστός, τεχνητός μη αστικός ιστός και άλλες χρήσεις/καλύψεις. Αρχικά διόρθωσαν ραδιομετρικά και ατμοσφαιρικά τις δορυφορικές εικόνες. Στην συνέχεια χρησιμοποιήθηκαν ως δεδομένα εισόδου στον ταξινομητή Random Forest, με 500 δέντρα και 4 μεταβλητές πρόβλεψης για κάθε κόμβο, όλα τα κανάλια, και τα θερμικά, η πρώτη συνιστώσα από την Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis, PCA) για τα οπτικά κανάλια και η πρώτη για τα θερμικά κανάλια, ο ενισχυμένος δείκτης βλάστησης (EVI) καθώς και ο κανονικοποιημένος δείκτης βλάστησης (NDVI), ο κανονικοποιημένος δείκτης υγρασίας (NDMI), ο κανονικοποιημένος δείκτης γυμνού εδάφους (NDBal), ο κανονικοποιημένος δείκτης δομημένης επιφάνειας (NDBI), ο ενισχυμένος δείκτης δομημένης επιφάνειας και γυμνού εδάφους (EBBI), οι μετασχηματισμοί Tasseled Cap και το Ψηφιακό Μοντέλο Εδάφους (ΨΜΕ) και η κλίση του. Έπειτα αντικαταστάθηκαν στις ταξινομημένες εικόνες οι κλάσεις των εικονοστοιχείων εκείνων που δεν είχαν γειτονικά εικονοστοιχεία με ίδια κλάση, με την χρήση παράθυρου μεγέθους 3X3 εικονοστοιχεία, ενώ η νέα κλάση ήταν η επικρατούσα τιμή του παραθύρου. Η ακρίβεια των ταξινομήσεων εκτεινόταν από το 90,5% έως το 92,3%. Στην συνέχεια με την χρήση διασταύρωσης πινάκων εντόπισαν τις αλλαγές χρήσεων/καλύψεων γης που επήλθαν τα χρόνια μελέτης. Τέλος έγινε πρόβλεψη της αστικής εξάπλωσης για το 2045, ενώ για τον εντοπισμό του καλύτερου μοντέλου πρόβλεψης χρησιμοποιήθηκε η ταξινόμηση του 2015.

### **3. Μέθοδοι εντοπισμού αλλαγών χρήσεων/καλύψεων γης**

Για τον εντοπισμό, και κατά συνέπεια την χαρτογράφηση, των αλλαγών κάλυψης και χρήσης γης έχουν χρησιμοποιηθεί δορυφορικά δεδομένα διαφόρων χωρικών αναλύσεων. Οι Lu et al (2004), έχουν καταγράψει και

ομαδοποιήσει τις τεχνικές εντοπισμού αλλαγών που έχουν χρησιμοποιηθεί. Οι τεχνικές εντοπισμού αλλαγών έχουν ομαδοποιηθεί σε 7 κατηγορίες οι οποίες είναι: (1) οι τεχνικές με χρήση άλγεβρας, (2) με χρήση μετατροπών, (3) με χρήση μεθόδων ταξινόμησης, (4) με χρήση προηγμένων μοντέλων, (5) με χρήση Συστημάτων Γεωγραφικών Πληροφοριών, (6) με οπτική ανάλυση και (7) με άλλες προσεγγίσεις. Οι πρώτες 6 κατηγορίες έχουν χρησιμοποιηθεί ευρέως τα τελευταία χρόνια, ενώ η έβδομη κατηγορία περιλαμβάνει όλες αυτές τις μεθόδους που δεν μπορούν να συμπεριληφθούν σε καμία από τις προηγούμενες 6, ενώ ταυτόχρονα είναι και οι τεχνικές που δεν χρησιμοποιούνται συχνά για την ανίχνευση αλλαγών. Κάθε μία από τις μεθόδους που περιλαμβάνονται στις παραπάνω κατηγορίες έχει πλεονεκτήματα και μειονεκτήματα, τα οποία συγκεντρώνονται στην έρευνα των Lu et al (2004), ενώ στην έρευνα τους υπάρχει και βαθμολογία για κάθε μέθοδο ανάλογα την ευκολία της χρήσης, με 1 την πιο εύκολη και 5 την πιο δύσκολη. Ο παρακάτω πίνακας παρουσιάζει τις 7 προαναφερόμενες κατηγορίες μεθόδων εντοπισμού αλλαγής και ενδεικτικές μεθόδους που ανήκουν σε αυτές τις κατηγορίες.

*Πίνακας 1: Κατηγοριοποιημένες τεχνικές εντοπισμού αλλαγών*

| <b>Κατηγορία</b>                  | <b>Τεχνικές εντοπισμού αλλαγών</b>   |
|-----------------------------------|--|
| Άλγεβρα                           | Διαίρεση εικόνων, Διαφορά δεικτών βλάστησης, Change Vector Analysis.                           |
| Μετασηματισμοί                    | Ανάλυση Κύριων Συνιστωσών, Tasseled Cap,.  |
| Ταξινόμηση                        | Σύγκριση ταξινομημένων εικόνων, Συνδυασμός πολυεποχιακών εικόνων.                              |
| Προηγμένα μοντέλα                 | Μοντέλο ανακλαστικότητας Li-Strahler, Μοντέλο φασματικής μίξης, Μοντέλο βιοφυσικών παραμέτρων. |
| Συστήματα Γεωγραφικών πληροφοριών | Ολοκληρωμένη μέθοδος GIS και τηλεπισκόπησης, Προσέγγιση με GIS                                 |

|                    |  |
|--------------------|--|
| Οπτική Ανάλυση     | Οπτική ερμηνεία  |
| Άλλες Προσεγγίσεις | Μετρήσεις χωρικών εξαρτήσεων, Σύστημα βασιζόμενο στην γνώση του ερευνητή, Μέθοδος Παραγωγής Περιοχής, Συνδυασμός δεικτών βλάστησης, θερμοκρασίας εδάφους και χωρικής δομής, Καμπύλες αλλαγής, Γενικευμένα γραμμικά μοντέλα, Μέθοδος βασιζόμενη σε χωρική στατιστική. |

*(Δημιουργήθηκε από τον συγγραφέα, σύμφωνα με τον πίνακα των Lu et al. (2004))*

Ο Mas (1999) παρακολούθησε τις αλλαγές στην εδαφοκάλυψη με χρήση διαφόρων μεθόδων εντοπισμού αλλαγής, για τμήμα της περιοχής της λιμνοθάλασσας TeÂrminos, που βρίσκεται στην πολιτεία Campeche, του Μεξικού. Χρησιμοποίησε 3 δορυφορικές εικόνες Landsat, για τα έτη 1970, 1974 και 1992. Οι μέθοδοι που χρησιμοποίησε για τον εντοπισμό αλλαγών ήταν οι διαφορές εικόνων, η διαφορά των δεικτών βλάστησης, η επιλογή κύριας συνιστώσας, ταυτόχρονη ταξινόμηση των εικόνων χωρίς την επιλογή περιοχών που εντοπίζεται σίγουρα η αλλαγή, η μέθοδος της σύγκρισης ταξινομημένων εικόνων και ένας συνδυασμός βελτιστοποίηση εικόνας με σύγκριση ταξινομημένων εικόνων. Η μέθοδος ταξινόμησης που χρησιμοποιήθηκε ήταν η επιβλεπόμενη ταξινόμηση μέγιστης πιθανοφάνειας, ενώ τα καλύτερα αποτελέσματα εντοπισμού αλλαγών (change- no change) προέκυψαν από την μέθοδο σύγκρισης ταξινομημένων εικόνων, με ποσοστό ακρίβειας 86,87%, και 82,42% για τον εντοπισμό αλλαγών στους τύπους εδαφοκάλυψης (from-to).

Για την ανάλυση ακρίβειας του εντοπισμού αλλαγών με χρήση κανόνων βασισμένων στην λογική αλλά και επαλήθευση αυτής με την μέθοδο σύγκριση ταξινομημένων εικόνων, οι Liu και Zhou (2004), χρησιμοποίησαν 5 δορυφορικές εικόνες Landsat, τις οποίες τις ταξινόμησαν με 2 ταξινομητές, μέγιστης πιθανοφάνειας και νευρικών δικτύων για τον εντοπισμό των κατευθύνσεων των αλλαγών σε μια αστική περιφέρεια του Πεκίνο. Τελικά χρησιμοποίησαν τα θεματικά επίπεδα που προέκυψαν από τα νευρωνικά δίκτυα, καθώς παρείχαν μεγαλύτερα ποσοστά ακρίβειας. Στην

συνέχεια εντόπισαν τις αλλαγές των εδαφοκαλύψεων ως αλλαγή εδαφοκάλυψης ή μή. Τέλος με την χρήση λογικών κανόνων εντόπισαν τις αλλαγές στους τύπους των εδαφοκαλύψεων για κάθε έτος που εξετάζαν.

Η συγκεκριμένη μέθοδος χρησιμοποιήθηκε για τον εντοπισμό αλλαγών στο Δυτικό Δέλτα του ποταμού Νείλου της Αιγύπτου (Abd El-Kawy et al, 2011). Συγκεκριμένα χρησιμοποιήθηκαν 4 δορυφορικές εικόνες Landsat, μια για το 1984, μια για το 1999, μια για το 2005 και μια για το 2009, κατά την χρονική περίοδο της ανάπτυξης των φυτών, ενώ χρησιμοποιήθηκαν μόνο οι ορατές μπάντες και οι υπέρυθρες. Επίσης έγινε χρήση και συμπληρωματικών δεδομένων για την επιβλεπόμενη ταξινόμηση των δορυφορικών εικόνων με χρήση της μεθόδου μέγιστης πιθανοφάνειας. Στην συνέχεια εντοπίστηκαν οι αλλαγές στην εδαφοκάλυψη με την προαναφερόμενη μέθοδο και προέκυψαν 3 θεματικά επίπεδα αλλαγών (1984-1999, 1999-2005, 2005-2009).

Οι Mohamed και El-Raey (2019), χρησιμοποίησαν 2 δορυφορικές εικόνες Landsat για τα έτη 2000 και 2017 αντίστοιχα, από την περιοχή El-Fayoum, μία όαση που βρίσκεται στο βόρειο μέρος την Αιγυπτιακής δυτικής ερήμου, με σκοπό τον εντοπισμό της αλλαγής κάλυψης και χρήσης γης. Αρχικά με μη επιβλεπόμενη ταξινόμηση και με την μέθοδο ISODATA δημιούργησαν 2 θεματικά επίπεδα χρήσεων/καλύψεων γης και έπειτα δημιούργησαν δεδομένα εκπαίδευσης, από δεδομένα πεδίου και συμπληρωματικά δεδομένα. Στην συνέχεια χρησιμοποίησαν την μη επιβλεπόμενη ταξινόμηση καθώς και τα δεδομένα εκπαίδευσης και με την μέθοδο μέγιστης πιθανοφάνειας διενέργησαν μια επιβλεπόμενη ταξινόμηση δημιουργώντας ξανά 2 νέα θεματικά επίπεδα χρήσεων/καλύψεων γης που προερχόντουσαν πλέον από μία υβριδική μέθοδο ταξινόμησης. Έπειτα με την μέθοδο PCC, εντόπισαν τις αλλαγές στις χρήσεις/καλύψεις και με την χρήση Μαρκοβιανών μοντέλων προέβλεψαν την μελλοντική εδαφοκάλυψη

για το 2050.

Οι Rodriguez-Galiano και Chica-Olmo (2012) χρησιμοποιώντας 4 δορυφορικές εικόνες Landsat 5, 2 για το έτος 1998 και 2 για το έτος 2004, για την περιοχή Γρανάδα της Ισπανίας. Για κάθε έτος οι εικόνες πάρθηκαν με βάση την φαινολογία των φυτών, συγκεκριμένα για τους μήνες Απρίλιο και Αύγουστο. Αρχικά έγινε γεωαναφορά των δορυφορικών εικόνων και μετατροπή των ψηφιακών τιμών σε ανακλαστικότητα. Στην συνέχεια χρησιμοποίησαν τα δορυφορικά δεδομένα, οπτικά και τα θερμικά κανάλια, το ψηφιακό μοντέλο εδάφους (ΨΜΕ) και την κλίση του εδάφους, καθώς και τους μετασχηματισμούς Tasseled Cap. Όλα τα παραπάνω χρησιμοποιήθηκαν ως δεδομένα εισόδου στον ταξινομητή Random Forest με 5000 δέντρα και οι ταξινομημένοι χάρτες εδαφοκάλυψης που προέκυψαν είχαν τιμές  $k$  (kappa) 0,92 και 0,91 για το 1998 και 0,93 και 0,92 για το 2004. Στην συνέχεια με την διασταύρωση πινάκων εντοπίστηκαν οι μεταβολές της εδαφοκάλυψης από το 1998 στο 2004. Ο χάρτης εντοπισμού αλλαγών είχε ακρίβεια 86%, το οποίο προέκυψε από τον πολλαπλασιασμό των ακριβειών των ταξινομημένων χαρτών του 1998 και 2004 αντίστοιχα.

Για την παρούσα εργασία χρησιμοποιήθηκε η μέθοδος εντοπισμού αλλαγών με σύγκριση ταξινομημένων εικόνων (Post-Classification Comparison, PCC). Για την χρήση αυτής της μεθόδου χρειάζεται να γίνει ταξινόμηση των δορυφορικών εικόνων (τουλάχιστον 2 δορυφορικές εικόνες), με την χρήση κάποιου ταξινομητή, ώστε να μετατραπούν σε θεματικά επίπεδα πληροφορίας. Στην συνέχεια γίνεται σύγκριση των θεματικών επιπέδων πληροφορίας σε επίπεδο εικονοστοιχείου (pixel by pixel) με τελικό αποτέλεσμα τον εντοπισμό των αλλαγών. Με τον τρόπο αυτό μειώνονται οι επιδράσεις των ατμοσφαιρικών συνθηκών, που επικρατούσαν την χρονική στιγμή της λήψης των δορυφορικών σκηνών, καθώς και οι περιβαλλοντικές συνθήκες μεταξύ των εικόνων. Επίσης

μειώνονται και οι επιδράσεις των δεκτών, όταν προέρχονται από διαφορετικούς δέκτες.

Η μέθοδος PCC είναι χρονοβόρα, καθώς χρειάζεται πολύ χρόνο για την ταξινόμηση των δορυφορικών εικόνων, αφού η ακρίβεια της συγκεκριμένης μεθόδου, εμπίπτει στην ακρίβεια της ταξινόμησης της κάθε εικόνας, καθώς και στην ορθή γεωμετρική αναφορά της κάθε εικόνας αλλά και στην ορθή γεωμετρική αναφορά των εικόνων μεταξύ τους. Επίσης με την προαναφερόμενη μέθοδο και αναλύοντας τα αποτελέσματα των θεματικών επιπέδων προκύπτει ένα πίνακας αλλαγών, που μας δείχνει την κατάσταση του κάθε εικονοστοιχείου την χρονική στιγμή  $t_1$  και σε ποια κατάσταση βρέθηκε την χρονική στιγμή  $t_2$ . Παρόλα αυτά η συγκεκριμένη μέθοδος μπορεί να αποδώσει αρκετά εσφαλμένα αποτελέσματα αλλαγής (Singh, 1989).

## ΚΕΦΑΛΑΙΟ 2: Περιοχή Μελέτης

Η Λέσβος βρίσκεται στην βορειοδυτικό τμήμα του Αιγαίου Πελάγους της Ελλάδας. Έχει έκταση περίπου 1640 τετραγωνικά χιλιόμετρα, ενώ είναι το 3ο μεγαλύτερο νησί της Ελλάδας και η ακτογραμμή της έχει μήκος περίπου 370 χιλιόμετρα. Το μέγιστο υψόμετρο που παρατηρείται είναι τα 969 μέτρα, στο βουνό Λεπέτυμνος στα βορειοανατολικά, και τα 968 μέτρα στο βουνό Όλυμπος στα νότια, ενώ σύμφωνα με τους Marathianou et al (2000) στο 63% του εδάφους της Λέσβου συναντώνται κλίσεις εδάφους μεγαλύτερες των 18% (10,2°). Σύμφωνα με τους Symeonakis et al (2016) μεγάλες εκτάσεις της Λέσβου έχουν απότομες πλαγιές με αποτέλεσμα να ευνοείται η ανάπτυξη της απορροής των υδάτων, αλλά και η διάβρωση.

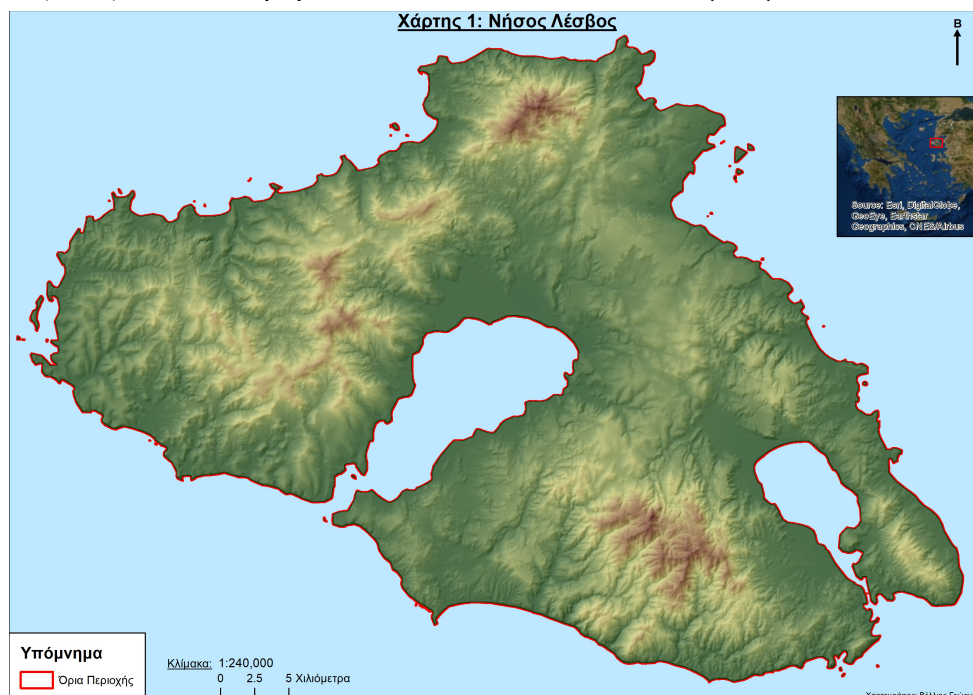
Το κλίμα του νησιού, όπου και επηρεάζεται από την περιφερειακή επίδραση της ορογραφίας και των μορφών της ατμοσφαιρικής κυκλοφορίας, χαρακτηρίζεται από έντονες εποχιακές και χωρικές διακυμάνσεις των βροχοπτώσεων καθώς και από διακυμάνσεις της μέγιστης και ελάχιστης ημερήσιας θερμοκρασίας. Η ετήσια βροχόπτωση έχει εύρος από 701mm (στην Αγιάσσο, στα ανατολικά του νησιού) μέχρι 479mm (στο Σίγρι, στα δυτικά, όπου και θεωρείται η ημίξηρη πλευρά του νησιού) (Symeonakis et al., 2016), ενώ η μέση ετήσια θερμοκρασία αέρα είναι περίπου 17,7°C σχεδόν σε ολόκληρο το νησί (Marathianou et al,2000).

Η φυτοκάλυψη της Λέσβου είναι κυρίως η μεσογειακή μακία (Mediterranean maquis), τα φρύγανα (phrygana), τα δάση Πεύκης (pine) και φυλλοβόλων δρυών (deciduous oak), ελαιώνες (olive groves) και οι αρδευόμενες και μη αρδευόμενες καλλιέργειες (Symeonakis et al, 2012; Symeonakis et al, 2016). Επίσης στο νησί υπάρχουν δυο αλυκές, της Καλλονής, όπου είναι η 3η μεγαλύτερη της χώρας σε έκταση (2630 στρέμματα), και του Πολιχνίτου που είναι η μικρότερη της χώρας (600

στρέμματα), ενώ εντοπίζονται και μεγάλης έκτασης ανθρωπογενείς κατασκευές, όπως είναι οι οικισμοί, οι δρόμοι και οι κατασκευές μεταφορών (αεροδρόμιο, λιμενοβραχίονες κτλ.), ενώ εντοπίζονται και ανθρωπογενείς παρεμβάσεις στο τοπίο, όπως είναι τα ΧΥΤΑ και τα λατομεία.

Το γεωλογικό υπόβαθρο του νησιού αποτελείται κυρίως από μητρικά πετρώματα, όπως όξινα πυριγενή πετρώματα (λάβες, πυροκλαστικά κτλ.), από υπερβασικά πετρώματα (περιδοντίτες, ιγνιμβρίτες), μεταμορφωμένα πετρώματα (μάρμαρα και σχιστόλιθοι), ιζηματογενείς αποθέσεις (ασβεστόλιθοι κτλ.), μη ενοποιημένες αποθέσεις (συσσωματώματα, άργιλοι κτλ.) και πρόσφατα προσχωσιγενή κοιτάσματα (Kosmas et al, 2000).

Μόνιμος πληθυσμός της Λέσβου, σύμφωνα με την απογραφή της ΕΛΣΤΑΤ, που έγινε το 2011, είναι 86436 , ενώ σύμφωνα με τους Kontos et al (2003), το 2020 προβλέπεται ότι θα είναι 126000 οι μόνιμοι κάτοικοι.



*Χάρτης 1: Περιοχή μελέτης και η θέση της στην ευρύτερη περιοχή*



### **ΚΕΦΑΛΑΙΟ 3: Δεδομένα**

Το παρόν κεφάλαιο χωρίζεται σε δύο επιμέρους υποκεφάλαια, στα οποία θα παρουσιαστούν και θα αναλυθούν τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα μελέτη, καθώς και από που αντλήθηκαν αυτά, και η μεθοδολογία που ακολουθήθηκε για να εξαχθούν ορθά και επιστημονικά αποτελέσματα, ώστε να απαντηθούν τα ερευνητικά ερωτήματα.

#### **1. Συλλογή δεδομένων**

Στο παρόν υποκεφάλαιο θα παρουσιαστούν τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα μεταπτυχιακή διπλωματική διατριβή, κάποια από τα μεταδεδομένα τους καθώς και οι πηγές άντλησης αυτών. Τα δεδομένα κατηγοριοποιούνται σε τρεις κατηγορίες, τα διανυσματικά, τα δορυφορικά και τα μετεωρολογικά δεδομένα, τα οποία αναλύονται στα παρακάτω ομώνυμα υποκεφάλαια.

##### **1.1 Διανυσματικά Δεδομένα**

Τα διανυσματικά δεδομένα που χρησιμοποιήθηκαν στην παρούσα μεταπτυχιακή διπλωματική διατριβή, είναι η ακτογραμμή της Νήσου Λέσβου καθώς και των νησίδων που βρίσκονται κοντά στην ακτογραμμή και το οδικό δίκτυο.

Η ακτογραμμή της Ν. Λέσβου, σε μορφή πολύγωνου, ανακτήθηκε από την διαδικτυακή πλατφόρμα [geodata.gov.gr](http://geodata.gov.gr), η οποία είναι μια Υποδομή Γεωχωρικών Πληροφοριών, όπου παρέχει δεδομένα σύμφωνα με την ευρωπαϊκή οδηγία INSPIRE. Συγκεκριμένα χρησιμοποιήθηκε το διανυσματικό αρχείο των Καλλικρατικών Δήμων, το οποίο παρεχόταν στο

Ελληνικό Γεωδαιτικό Σύστημα Αναφοράς (ΕΓΣΑ87), και στην συνέχεια με χωρικές ερωτήσεις απομονώθηκε ο Δήμος Λέσβου.

Το οδικό δίκτυο της Ν. Λέσβου, αποκτήθηκε μέσω του OpenStreetMap, όπου παρέχεται η δυνατότητα απόκτησης διάφορων θεματικών επιπέδων από οποιοδήποτε μέρος του κόσμου. Συγκεκριμένα εξήχθησαν 30 αρχεία από το OSM, τα οποία περιλάμβαναν το οδικό δίκτυο του νησιού, σημεία αναφοράς, δίκτυα μεταφορών που συνδέουν την Λέσβο με άλλα νησιά και άλλα. Τα 30 αυτά osm αρχεία, μετατράπηκαν σε γεωχωρικά δεδομένα (shapefile), ενώθηκαν σε ένα τελικό διανυσματικά αρχείο και στην συνέχεια αφαιρέθηκε όλη η περιττή πληροφορία με σκοπό να μείνει μόνο το οδικό δίκτυο του νησιού. Τέλος το αρχείο αυτό επαναπροβλήθηκε από το παγκόσμιο γεωδαιτικό σύστημα (WGS84) στο ελληνικό (ΕΓΣΑ 87).

Τα όρια της περιοχής χρησιμοποιήθηκαν για να απομονώσουμε μόνο την περιοχή μελέτης, ενώ το οδικό δίκτυο χρησιμοποιήθηκε ως ένα ακόμη επίπεδο εκπαίδευσης του αλγορίθμου ταξινόμησης.

*Πίνακας 1: Πίνακας χαρακτηριστικών διανυσματικών δεδομένων*

| <b>Όνομα Αρχείου</b> | <b>Γεωμετρία</b> | <b>Πηγή</b>    | <b>Τύπος Αρχείου</b> | <b>Προβολικό</b> |
|----------------------|------------------|----------------|----------------------|------------------|
| Οδικό Δίκτυο         | Γραμμή           | OpenStreetMap  | osm                  | WGS84            |
| Ν. Λέσβου            | Πολύγωνο         | Geodata.gov.gr | shapefile            | ΕΓΣΑ87           |

## **1.2. Δορυφορικά Δεδομένα**

Τα δορυφορικά δεδομένα που χρησιμοποιήθηκαν, αντλήθηκαν από την διαδικτυακή πλατφόρμα του προγράμματος Copernicus, της Ευρωπαϊκής Υπηρεσίας Διαστήματος (ESA), την εφαρμογή EOli-SA (από τον Ιανουάριο του 2019 καταργήθηκε η συγκεκριμένη εφαρμογή και πλέον τα δεδομένα παρέχονται στην ιστοσελίδα της ESA, όπου μας παρέχεται η

δυνατότητα αναζήτησης σε δεδομένα παλαιότερων αποστολών πριν του Sentinel) και από την ιστοσελίδα της Γεωλογικής Υπηρεσίας των Ηνωμένων Πολιτειών της Αμερικής (USGS).

Συγκεκριμένα από την εφαρμογή EOLi-SA, ανακτήθηκαν τέσσερις δορυφορικές εικόνες του ενεργητικού δορυφόρου (Radar) ERS, από την ιστοσελίδα της ESA, μια δορυφορική εικόνα του Radar δορυφόρου Sentinel 1 καθώς και ψηφιακό μοντέλο εδάφους και από την USGS δύο δορυφορικές εικόνες του Landsat, μία από τον Landsat 5 και μια από τον Landsat 8.

Τα δεδομένα Landsat αποκτήθηκαν σε επίπεδο επεξεργασίας 2 (LVL 2 Products), όπου είναι ραδιομετρικά, γεωμετρικά και ατμοσφαιρικά διορθωμένα, ενώ η αλλαγή μεγέθους του εικονοστοιχείου γίνεται με την μέθοδο Cubic Convolution (Zanter,2018<sup>a</sup> ; Zanter,2019).

Η ατμοσφαιρική διόρθωση των εικόνων του Landsat 5, όπως και για τον 4 και τον 7, γίνεται με βάση το πρόγραμμα LEDAPS (Landsat Ecosystem Disturbance Adaptive Processing System) το οποίο δημιουργήθηκε από την NASA, την Goddard Space Flight Center (GSFC) και το Πανεπιστήμιο του Μέριλαντ το 2006. Το 2011 η NASA έβγαλε μία καινούρια αναβάθμιση του συγκεκριμένου προγράμματος, το οποίο και χρησιμοποιεί η USGS, που ονομάζεται LEDAPS 1.0.0 .

Το LEDAPS 1.0.0 παράγει την ανακλαστικότητα των στόχων, χωρίς τις επιδράσεις της ατμόσφαιρας, κάνει ατμοσφαιρικές διορθώσεις που βασίζονται στο 6S αλγόριθμο (Second Simulation of a Satellite Signal in the Solar Spectrum) και δημιουργεί μάσκα στα σύννεφα (USGS,2012).

Με τον αλγόριθμο 6S πλέον οι ραδιομετρικές ιδιότητες των βασικών συστατικών (αιθάλης, ωκεανών, σκόνη, υδατοδιαλύτης) εντοπίζονται καλύτερα στην ακτίνα των σωματιδίων, έχουν πραγματοποιηθεί βελτιώσεις όσο αφορά την ανάλυση και την ακρίβεια των φασματοσκοπικών δεδομένων σε σχέση με αυτές του αλγορίθμου 5S. Τέλος ο αλγόριθμος 6S

χρησιμοποιεί δεδομένα αερολυμάτων, μορίων της ατμόσφαιρας καθώς και το υψόμετρο για την ατμοσφαιρική διόρθωση (Vermote et al, 1997).

Η ατμοσφαιρική διόρθωση του Landsat 8, επιτυγχάνεται μέσω του αλγορίθμου LaSRC (Landsat 8 Surface Reflectance Code). Ως δεδομένα εισόδου ο συγκεκριμένος αλγόριθμος δέχεται τα επιπέδου 1 (LVL 1 products) του Landsat 8 και συγκεκριμένα τις μπάντες του δορυφόρου καθώς και την μπάντα αξιολόγησης ποιότητας των εικονοστοιχείων, καθώς και τα μεταδεδομένα των εικόνων, ενώ πληροφορίες για τα αερολύματα παίρνει από τους αισθητήρες του MODIS, Terra και Aqua, και είναι σχετικοί με το όζον, τους υδρατμούς καθώς και το υψόμετρο (Vermote et al, 2016; Zanter, 2018<sup>b</sup>). Περισσότερες διαφορές μεταξύ των δύο αλγορίθμων φαίνονται στον Πίνακα 7 (βλ. Παράρτημα Πινάκων).

*Πίνακας 2: Πίνακας χαρακτηριστικών δορυφορικών δεδομένων*

| <b>Αισθητήρας</b> | <b>Ημερομηνία</b> | <b>Επίπεδο Επεξεργασίας</b> | <b>Τρόπος Λειτουργίας</b> | <b>Πόλωση</b> | <b>Μέγεθος Εικονοστοιχείου</b> |
|-------------------|-------------------|-----------------------------|---------------------------|---------------|--------------------------------|
| ERS-1             | 10/06/93          | -                           | PI                        | VV            | 25*25                          |
| ERS-1             | 10/06/93          | -                           | PI                        | VV            | 25*25                          |
| ERS-1             | 15/07/09          | -                           | PI                        | VV            | 25*25                          |
| ERS-1             | 31/07/09          | -                           | PI                        | VV            | 25*25                          |
| Sentinel 1        | 30/06/18          | -                           | IW                        | VV+ VH        | 10*10                          |
| Landsat 5         | 12/06/92          | Level 2                     | -                         | -             | 30*30                          |
| Landsat 5         | 26/07/09          | Level 2                     | -                         | -             | 30*30                          |
| Landsat 8         | 03/07/18          | Level 2                     | -                         | -             | 30*30                          |

### 1.3. Μετεωρολογικά δεδομένα

Για την επιλογή των δορυφορικών εικόνων χρειαζόταν να γνωρίζουμε τις καιρικές συνθήκες που επικρατούσαν κατά την περίοδο που βρίσκονταν οι αισθητήρες πάνω από την περιοχή μελέτης.

Τα μετεωρολογικά δεδομένα χρησιμοποιήθηκαν με σκοπό την επιλογή των δορυφορικών δεδομένων από τους οπτικούς δέκτες, οι οποίοι καταγράφουν ακτινοβολία με μικρό μήκος κύματος όπου δεν διαπερνά τα σύννεφα, με αποτέλεσμα να μην υπάρχει διαθέσιμη πληροφορία για την περιοχή που βρίσκεται κάτω από αυτά, και για τα εικονοληπτικά Radar.

Σύμφωνα με την βιβλιογραφία (Καρτάλης και Φείδας (2015), σελ.1-5; Παρχαρίδης (2015), σελ. 21-22; Richard (2009), σελ. 5), τα εικονοληπτικά Radar, δεν επηρεάζονται από την νεφοκάλυψη, σε αντίθεση με τους οπτικούς δέκτες, καθώς το μήκος κύματος των παλμών, είναι ικανό να διαπεράσει τα νέφη. Τα Radar όμως όπως αναφέρθηκε και σε άλλο κεφάλαιο, καταγράφουν και την περιεκτικότητα νερού σε βλάστηση και έδαφος. Με αυτό τον τρόπο επηρεάζονται οι τιμές οπισθοσκέδασης. Γι' αυτό τον λόγο, πριν την απόκτηση των δορυφορικών εικόνων, επιλέχθηκαν μέρες με μηδενικές ποσότητες κατακρημνισμάτων για τουλάχιστον τις έξι προηγούμενες μέρες από την ημερομηνία λήψης των δορυφορικών δεδομένων.

Τα μετεωρολογικά δεδομένα που χρησιμοποιήθηκαν αντλήθηκαν από την Εθνική Μετεωρολογική Υπηρεσία (EMY), τα οποία διανέμονται δωρεάν μέσω της ιστοσελίδας της EMY. Τα δεδομένα αυτά δίνονται ανά πενταετία και ανά μετεωρολογικό σταθμό. Υπάρχουν δεδομένα από το 1975 για την Μυτιλήνη, ενώ παρέχονται δεδομένα καταγραφών μέχρι το 2004 σε έτοιμους τυποποιημένους πίνακες ανά 3 ώρες, για την ατμοσφαιρική πίεση, την μέγιστη και ελάχιστη θερμοκρασία, την υγρασία, την ταχύτητα του

ανέμου, την κατεύθυνση του αέρα, το ποσοστό της νεφοκάλυψης και της ορατότητας. Από το 2004 και έπειτα οι τυποποιημένοι πίνακες δεν παρέχονται δωρεάν, δίνεται όμως η δυνατότητα παραγγελίας τους.

Για τον λόγο αυτό χρησιμοποιήθηκε η πλατφόρμα *weatherundeground*, όπου μας παρέχει δωρεάν μετεωρολογικά δεδομένα από όλο τον κόσμο από σταθμούς προσωπικούς, αεροδρομίων και επιστημονικών φορέων. Τα δεδομένα που παρέχονται είναι ανά μισή ώρα, ενώ παρέχουν πληροφορίες σχετικά με την θερμοκρασία, την υγρασία, την κατεύθυνση και την ταχύτητα του αέρα, την ατμοσφαιρική πίεση, την βροχόπτωση και την νεφοκάλυψη.

Για την διασταύρωση των δεδομένων που αντλήθηκαν από την πλατφόρμα *weatherundeground*, χρησιμοποιήθηκαν μετεωρολογικά δεδομένα που παραχωρήθηκαν από τον μετεωρολογικό σταθμό, Μυτιλήνης, του Εργαστηρίου Φυσικών Καταστροφών του Τμήματος Γεωγραφίας του Πανεπιστημίου Αιγαίου και συγκεκριμένα από τον επιστημονικό υπεύθυνο και καθηγητή του τμήματος κύριο Καλαμποκίδη Κωσταντίνο ([meteo.aegean.gr](http://meteo.aegean.gr)). Τα δεδομένα αυτά ήταν σε μορφή *dat*, ενώ υπήρχαν καταγραφές από την 01/01/2015 μέχρι την 31/12/2018.

Επίσης, οι μετεωρολογικοί σταθμοί κατέγραφαν δεδομένα σχετικά με την σχετική υγρασία, την βροχόπτωση, την θερμοκρασία, την ταχύτητα και την διεύθυνση του ανέμου, την ατμοσφαιρική πίεση, την θερμοκρασία του εδάφους και την διάρκεια της βροχόπτωσης ανά δέκα λεπτά και ανά μία ώρα.

Τέλος πρέπει να τονιστεί πως τα μετεωρολογικά δεδομένα δεν χρησιμοποιήθηκαν για την εκπαίδευση του ταξινομητή που χρησιμοποιήθηκε στην παρούσα εργασία.

## ΚΕΦΑΛΑΙΟ 4: Μεθοδολογία

Στο παρών κεφάλαιο θα παρουσιαστεί η μεθοδολογία που ακολουθήθηκε στην παρούσα μεταπτυχιακή διπλωματική διατριβή, ώστε να προκύψουν ορθά και επιστημονικά τεκμηριωμένα αποτελέσματα. Η μεθοδολογία χωρίζεται σε έξι (6) στάδια όσα δηλαδή και τα υποκεφάλαια που ακολουθούν. Αρχικά αναλύεται η προεπεξεργασία των δεδομένων, ώστε να μπορούν να χρησιμοποιηθούν για την περαιτέρω ανάλυσή τους, έπειτα αναφέρεται ο τρόπος με τον οποίο δημιουργήθηκαν τα δειγματοληπτικά σημεία για την εκπαίδευση του αλγορίθμου ταξινόμησης, ακόμη έγινε προσπάθεια σύζευξης των δορυφορικών δεδομένων των εικονοληπτικών Radar και των οπτικών δεκτών, έπειτα η τεχνική εκπαίδευσης του αλγορίθμου ταξινόμησης Random Forest με χρήση των δειγματοληπτικών σημείων, αξιολόγηση της ακρίβειας ταξινόμησης και στο τέλος ο εντοπισμός αλλαγών εδαφοκάλυψης.

### 1. Προεπεξεργασία δεδομένων

Αφού αποκτήθηκαν τα διανυσματικά και δορυφορικά δεδομένα, ξεκίνησε η επεξεργασία τους. Αρχικά από το διανυσματικό αρχείο των Καλλικρατικών Δήμων, με χωρικά ερωτήματα, απομονώθηκε η περιοχή μελέτης, Δήμος Λέσβου, όπου περιλαμβάνει μόνο την Νήσο Λέσβο καθώς και τις νησίδες που βρίσκονται σε κοντινή απόσταση από την ακτογραμμή και είχαν αποτυπωθεί στο διανυσματικό αρχείο.

Επίσης αφού ανακτήθηκε το πολυθεματικό επίπεδο osm, σε τριάντα (30) επιμέρους αρχεία, απομονώθηκε το οδικό δίκτυο της περιοχής μελέτης, ενώθηκε σε ένα διανυσματικό αρχείο, shape file, και ξεκίνησε η αφαίρεση δεδομένων που δεν χρειάζονταν στο οδικό δίκτυο, όπως δασικοί δρόμοι,

υδάτινοι οδοί (διασύνδεση με άλλες περιοχές μέσω θαλάσσιων μεταφορών, κ.ά.). Έπειτα το αρχείο αυτό επαναπροβλήθηκε στο προβολικό που βρισκόταν και το αρχείο των ορίων της περιοχής μελέτης και στην συνέχεια υπολογίστηκε η ευκλείδεια απόσταση από το οδικό δίκτυο του νησιού και περιορίστηκε το αποτέλεσμα στα όρια της περιοχής μελέτης.

Στην συνέχεια ξεκίνησε η επεξεργασία των δορυφορικών δεδομένων, η οποία πραγματοποιήθηκε εξολοκλήρου στο πρόγραμμα SNAP, το οποίο είναι λογισμικό ανοικτού κώδικα. Αρχικά έγινε επαναπροβολή των δορυφορικών εικόνων Landsat 5 & 8 στο ίδιο προβολικό με αυτό των διανυσματικών δεδομένων. Τέλος υπολογίστηκε ο κανονικοποιημένος δείκτης βλάστησης NDVI, οι μετασχηματισμοί Tasseled Cap Greenness και Brightness, λόγοι των καναλιών κοντινό υπέρυθρο/κόκκινο, κόκκινο/πράσινο, μέσο υπέρυθρο/πράσινο. Στο τέλος απομονώθηκε η περιοχή μελέτης.

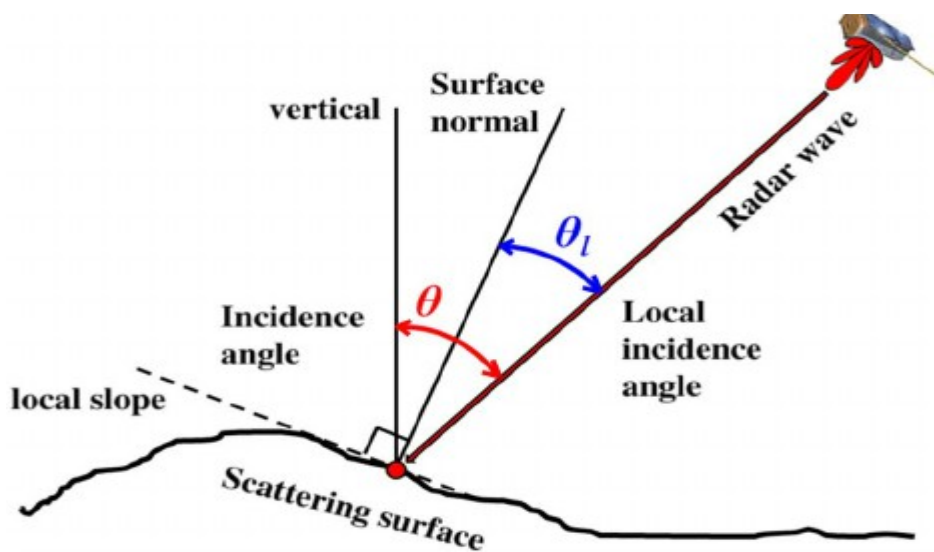
Ύστερα πραγματοποιήθηκε απομάκρυνση των θερμικών θορύβων (Thermal Noise Removal) της δορυφορικής εικόνας του Sentinel 1. Αυτό έγινε καθώς ο παλμός επιδρά με τον στόχο και επιστρέφει ίδιας ενέργειας παλμός αν και μόνο αν ο στόχος βρίσκεται σε θερμική ισορροπία. Αν ο στόχος δεν βρίσκεται σε θερμική ισορροπία, τότε είτε απορροφά ενέργεια και άρα ο παλμός επιστρέφει με λιγότερη ενέργεια, είτε αντανακλά περισσότερη ενέργεια με την μορφή θερμοκρασίας. Έτσι με την τεχνική της απομάκρυνσης θερμικού θορύβου από την εικόνα, επιτεύχθει η διόρθωση των προαναφερθέντων ραδιομετρικών παραμορφώσεων (Richards (2009), σελ. 312-314).

Στην συνέχεια χρησιμοποιήθηκαν στοιχεία των μεταδεδομένων, συγκεκριμένα τροχιακά δεδομένα, για την υποβοήθηση των επόμενων “βημάτων” της μεθοδολογίας σχετικά με την γεωκωδικοποίηση.

Έπειτα πραγματοποιήθηκε βαθμονόμηση με βάση το Sigma, όπου είναι



ένα προηγμένο μέτρο βαθμονόμησης που λαμβάνει υπόψη τη σταθερά βαθμονόμησης και τη γωνία πρόσπτωσης (η γωνία μεταξύ της κατακόρυφης επιφάνειας της γης και του εισερχόμενου σήματος, η κόκκινη γωνία στο Σχήμα 1) οι οποίες καταγράφονται στα μεταδεδομένα της εικόνας. Με αυτή την βαθμονόμηση λαμβάνουμε τις τιμές της οπισθοσκέδασης ανά μονάδα επιφανείας στο έδαφος.



Σχήμα 1: Γωνία πρόσπτωσης  $\theta_0$  (κόκκινο) και τοπική γωνία πρόσπτωσης  $\theta_l$  (μπλε)

πηγή: Rizzoli P., Brautigam B., (2014)

Στην συνέχεια πραγματοποιήθηκε ραδιομετρική διόρθωση gamma0 με βάση το ανάγλυφο (Radiometric Terrain Flattening), το οποίο μας παρέχει δωρεάν η ESA σε ανάλυση 1 sec. Η διόρθωση αυτή έγινε με βάση τον αλγόριθμο του Small David (2011). Οι διακυμάνσεις του εδάφους επηρεάζουν όχι μόνο την θέση ενός γνωστού σημείου στη γη, όπως αναφέρθηκε και σε προηγούμενο κεφάλαιο, αλλά και την φωτεινότητα των εικονοστοιχείων (υφή) που σχετίζεται με την γεωμετρία της περιοχής. Χωρίς αυτή την διόρθωση του συστημικού λάθους, απειλούνται τα σημεία με μικρή κλίση εδάφους από τα σημεία με μεγάλες κλίσεις, όπου θα φαίνονται εντονότερα. Με αυτή την ραδιομετρική διόρθωση λαμβάνουμε

τις τιμές της οπισθοσκεδασης ανά μονάδα επιφάνειας κάθετης στην δέσμη ακτινών της κεραίας του πομπού-δέκτη.

Έπειτα πραγματοποιήθηκε γεωμετρική διόρθωση (Range Doppler Terrain Correction) για απαλοιφή του φαινομένου των παραμορφώσεων της κεκλιμένης απόστασης (slant range distortion).

Τέλος πραγματοποιήθηκε επαναπροβολή της εικόνας στο Ελληνικό Γεωδαιτικό Σύστημα Αναφοράς (ΕΓΣΑ 87), ενώ έγινε ταυτόχρονα και αλλαγή του μεγέθους του εικονοστοιχείου σε 30\*30 μέτρα από τα 10\*10 μέτρα, με την μέθοδο Cubic Convolution, την ίδια που χρησιμοποιούν και στα δεδομένα Landsat η USGS, και στην συνέχεια απομονώθηκε η περιοχή μελέτης με βάση τα όριά της.

Έπειτα πραγματοποιήθηκε ραδιομετρική βαθμονόμηση, όλων των εικόνων του ERS, με βάση το Sigma, όπως και στον Sentinel 1. Με αυτή την βαθμονόμηση λαμβάνουμε τις τιμές της οπισθοσκεδασης ανά μονάδα επιφάνειας στο έδαφος. Ακόμη έγινε ραδιομετρική διόρθωση  $\gamma_0$  με βάση το ανάγλυφο (Radiometric Terrain Flattening), το οποίο μας παρείχε δωρεάν η ESA σε ανάλυση 1 sec. Η διόρθωση αυτή έγινε με βάση τον αλγόριθμο του Small David (2011). Στην συνέχεια με την μέθοδο InSAR έγινε ένα μωσαϊκό των χρονιών 1993 και 2008 που αποτελούνταν από 2 σκηνές η κάθε χρονιά. Τέλος έγινε επαναπροβολή των εικόνων σε ΕΓΣΑ87, αλλαγή του μεγέθους του εικονοστοιχείου σε 30\*30 μέτρα και απομόνωση της περιοχής μελέτης.

## **2. Δημιουργία δειγματοληπτικών σημείων**

Για την δημιουργία των δειγματοληπτικών σημείων χρησιμοποιήθηκαν τα όρια της περιοχής μελέτης και δημιουργήθηκαν 2500 σημεία με τυχαίο τρόπο. Στην συνέχεια χρησιμοποιήθηκαν μέσω Google Earth για το 2018,

μέσω των αεροφωτογραφιών του Κτηματολογίου και συμπληρωματικά του Google Earth για το 2009 και τέλος για το 1993 λόγω έλλειψης υποβάθρου υψηλότερης χωρικής διακριτικής ικανότητας από αυτές των δορυφορικών εικόνων μας ελέγχθηκαν οπτικά με συνδυασμό των εικόνων του 2008 και 1993. Τέλος καταγράφηκε στα σημεία η εδαφοκάλυψη, όπως φαίνεται και στον παρακάτω πίνακα (Πίνακας 3), και όπου κρίθηκε σκόπιμο τα σημεία μετακινήθηκαν ώστε να βρίσκονται σε ομοιογενείς περιοχές. Τέλος πρέπει να προστεθεί πως η περιγραφή των κλάσεων προέρχεται από τους Brossard et al (2000).

*Πίνακας 3: Τύποι εδαφοκάλυψης περιοχής μελέτης*

| <b>A/A</b> | <b>Εδαφοκάλυψη</b>     |
|------------|------------------------|
| 1          | Δομημένο Περιβάλλον    |
| 2          | Καλλιέργειες           |
| 3          | Ελαιώνες               |
| 4          | Γυμνό Έδαφος           |
| 5          | Κωνοφόρα Βλάστηση      |
| 6          | Πλατύφυλλη Βλάστηση    |
| 7          | Φρύγανα                |
| 8          | Σκληροφυλλική Βλάστηση |
| 9          | Αλυκές                 |
| 10         | Υδάτινες Περιοχές      |

Με τον όρο “Δομημένο περιβάλλον”, νοείται περιοχή που καλύπτεται κυρίως από οικίες και κτίρια, δρόμους, θέσεις πάρκινγκ, αλλά και περιοχές που αναμειγνύεται αστικός ιστός με αγροτεμάχια, δηλαδή περιοχές με 30% και πάνω αστικές κατασκευές, περιοχές που ασχολούνται κυρίως με βιομηχανικές δραστηριότητες μεταποίησης αλλά και παραγωγής, εμπορικές συναλλαγές, χρηματοπιστωτικές δραστηριότητες και υπηρεσίες, υποδομές

μεταφορών όπως λιμάνια, αεροδρόμια.

Ως “Καλλιέργειες”, νοείται οποιαδήποτε περιοχή χρησιμοποιείται για καλλιέργεια φρούτων, σιτηρών και οσπρίων, καθώς και εκτάσεις οπωροφόρων δέντρων.

Με τον όρο “Ελαιώνες”, εννοούνται περιοχές στις οποίες κατά κύριο λόγο γίνεται καλλιέργεια ελιών. Στην Λέσβο είναι η βασική καλλιέργεια, ενώ καταλαμβάνει μεγάλες εκτάσεις, με κύριες γύρω από την Καλλονή, την Αγία Παρασκευή, την Συκαμνιά, τα Παράκοιλα, την Ερεσό, την Άγρα, τα Χίδηρα και την περιοχή ανάμεσα στην Μυτιλήνη και τα Μυστεργνά.

Με τον όρο “Γυμνό Έδαφος”, ομαδοποιούμε τις περιοχές αυτές οι οποίες δεν περιέχουν ούτε βλάστηση, αλλά ούτε κάποια ανθρωπογενής κατασκευή (παραδείγματος χάριν οδικό δίκτυο, γέφυρες και άλλα). Δηλαδή περιοχές όπως οι κορυφές των βουνών, τα λατομεία και η ακτογραμμή.

Ως “Κωνοφόρα Βλάστηση”, νοούνται δασικές περιοχές (είτε πυκνής βλάστησης, είτε αραιής) που περιέχουν πεύκα, τραχεία και μαύρη Πεύκη.

Ως “Πλατύφυλλη Βλάστηση” έχουν κατηγοριοποιηθεί οι Καστανιόνες της Λέσβου οι οποίοι μπερδεύονται φασματικά με τις κλάσεις των Ελαιώνων. Επίσης σε αυτήν την κλάση περιλήφθηκαν και τα πλατάνια, λεύκες, φτελιές και άλλα είδη δέντρων που εντοπίζονται στις όχθες ποταμών, λιμνών, χειμάρρων και ρεμάτων. Τα παραπάνω φυτά έχουν παρόμοιες φασματικές υπογραφές και για τον λόγο αυτό τοποθετήθηκαν στην ίδια κλάση.

“Φρύγανα” ονομάζονται οι θάμνοι με ύψος από 20 έως 70 εκατοστά, πολύ μικρότερη βιομάζα από την σκληροφυλλική βλάστηση, αλλά με παρόμοια κάλυψη κομοστέγης, το καλοκαίρι φύλλα και βλαστοί πεθαίνουν (εν μέρει ή συνολικά) λόγω ακραίων θερμοκρασιών ή της κατανομής της βροχόπτωσης. Αυτός ο τύπος βλάστησης θεωρείται ότι είναι συνήθως αποτέλεσμα της υποβάθμισης προϋπάρχουσας βλάστησης (Blondel et al,

2010).

Ως “Σκληροφυλλική βλάστηση”, κατηγοριοποιούμε την μακία βλάστηση, που αποτελείται από πυκνούς θάμνους μέχρι 2 μέτρα σε ύψος, την αγριελιά, τα πουρνάρια, ο σκίνος, η χαρουπιά, η κουμαριά, η πικροδάφνη, η μυρτιά. Ο συγκεκριμένος τύπος βλάστησης εμφανίζεται στις περιοχές βόρεια και δυτικά του όρους Λεπέτυμνος, νότια του όρους Όλυμπος.

Ως “Αλυκές”, ορίζονται περιοχές που είναι επιρρεπείς σε πλημμύρες με θαλασσινό νερό, σε κάποιο στάδιο του ετήσιου κύκλου, λόγω παλίρροιας, όπου και αποικούνται τελικά από αλοφιλικά φυτά, περιοχές για την παραγωγή αλατιού με εξάτμιση, ενώ περιέχονται και οι περιοχές που γίνεται οργανωμένη εκτροφή οστρακοειδών και ψαριών στην ακτή και όχι στην ανοικτή θάλασσα (ιχθυοκαλλιέργειες).

Ως “Υδάτινες επιφάνειες” ορίζονται όλες οι περιοχές εκείνες που έχουν νερό καθ' όλη την διάρκεια του έτους, όπως οι λίμνες, φυσικές και τεχνητές, τα ποτάμια, και τα δέλτα αυτών, τα ρυάκια και τα ρέματα, τα θαλάσσια ύδατα, οι δεξαμενές και τα αρδευτικά κανάλια, καθώς και όλες τις περιοχές εκείνες που καλύπτονται από ρηχές υδάτινες επιφάνειες, όπως είναι οι περιοχές στον Κόλπο Καλλονής.

### **3. Σύζευξη εικονοληπτικών Radar και οπτικών δεκτών**

Για την σύζευξη των εικονοληπτικών Radar και των οπτικών δεκτών που χρησιμοποιήθηκαν στην παρούσα διατριβή, χρησιμοποιήθηκε η μέθοδος αντίστροφης ανάλυσης κύριων συνιστωσών (Inverse Principal Component Analysis, IPCA) με χρήση της γλώσσας R.

Αρχικά γίνεται ανάλυση των κύριων συνιστωσών (PCA) των καναλιών του οπτικού δέκτη. Στην συνέχεια εφαρμόζεται η PCA σε κάθε

εικονοστοιχείο (loadings) του κάθε καναλιού.

Έπειτα ερευνάται η συσχέτιση των τιμών των loadings, με τις αντίστοιχες τιμές του εικονοληπτικού Radar. Αφού βρεθεί η συνιστώσα με την μεγαλύτερη συσχέτιση με το κανάλι του εικονοληπτικού Radar αντικαθίσταται από αυτό.

Τέλος εφαρμόζεται η IPCA, έχοντας πλέον τις  $n-1$  κύριες συνιστώσες μαζί με το εικονοληπτικό Radar, αφαιρείται από τις κύριες συνιστώσες του οπτικού δέκτη ο μέσος όρος της τιμής των εικονοστοιχείων του καναλιού από τα loadings, διαιρούνται με την τυπική απόκλιση των αρχικών τιμών του κάθε καναλιού του οπτικού δέκτη και του εικονοληπτικού Radar.

#### **4. Ταξινόμηση με τον αλγόριθμο Random Forest**

Με χρήση των σημείων δειγματοληψίας, καθώς και των δεδομένων που αναλύθηκαν σε προηγούμενα υποκεφάλαια, έγινε εκπαίδευση του αλγορίθμου Random Forest, ώστε να προκύψει ένας θεματικός χάρτης εδαφοκάλυψης για κάθε έτος που εξετάστηκε. Ο ταξινομητής Random Forest εκπαιδεύτηκε μέσω της συνάρτησης RandomForest στο περιβάλλον της R. Εδώ πρέπει να τονιστεί πως οι Gislason et al (2006), υποστηρίζουν ότι η μεταβλητή του υψομέτρου είναι η μόνη τοπογραφική μεταβλητή που είναι στατιστικά σημαντική για τον εντοπισμό κλάσεων εδαφοκάλυψης. Παρ' όλα αυτά στην παρούσα εργασία χρησιμοποιήθηκε και η κλίση και ο προσανατολισμός.

Για τον εντοπισμό του κατάλληλου αριθμού δέντρων για τον αλγόριθμο έγινε η εκπαίδευση του αλγορίθμου με 500 έως 1600 δέντρα και επιλέχθηκε αυτός με το μικρότερο OOB σφάλμα.

Για να βρεθεί η συνεισφορά του εικονοληπτικού Radar στην ακρίβεια της ταξινόμησης εκπαιδεύτηκαν δύο αλγόριθμοι, με τα ίδια

δειγματοληπτικά σημεία. Ο πρώτος αλγόριθμος είχε ως δεδομένα εισόδου τα κανάλια του οπτικού δέκτη, διάφορους λόγους καναλιών από τον οπτικό δέκτη, τους μετασχηματισμούς TC Greenness και Brightness, τον δείκτη NDVI, το κανάλι του εικονοληπτικού Radar, την απόσταση από το οδικό δίκτυο, το ΨΜΕ, την κλίση και τον προσανατολισμό. Ο δεύτερος αλγόριθμος εκπαίδευσης περιείχε τα ίδια δεδομένα εισόδου χωρίς το εικονοληπτικό Radar. Στην συνέχεια για να διερευνήσουμε και την συνεισφορά των τοπογραφικών μεταβλητών, συγκεκριμένα του προσανατολισμού, εκπαιδεύτηκαν άλλοι 3 αλγόριθμοι. Ο τρίτος αλγόριθμος εκπαίδευσης περιείχε τα ίδια δεδομένα εισόδου με τον πρώτο αλγόριθμο εκπαίδευσης χωρίς την κλίση και τον προσανατολισμό. Ο τέταρτος αλγόριθμος είχε ίδια δεδομένα εισόδου με τον δεύτερο αλγόριθμο εκπαίδευσης, αλλά χωρίς την κλίση και τον προσανατολισμό και ο πέμπτος αλγόριθμος εκπαίδευσης είχε τα ίδια δεδομένα εισόδου με τον πρώτο αλγόριθμο, χωρίς την μεταβλητή του προσανατολισμού. Για να δούμε την συνεισφορά των μεταβλητών του εικονοληπτικού Radar, καθώς και του προσανατολισμού, ελέγχθηκε μόνο η ολική ακρίβεια του κάθε μοντέλου και ο συντελεστής kappa.

Για την βαθμολογία των ανεξάρτητων μεταβλητών οι οποίες συνέβαλαν στην τελική “ψηφοφορία” σε κάθε δέντρο υπολογίζεται ο συντελεστής Gini, όπου με αυτό τον τρόπο αυξάνεται η αποτελεσματικότητα της ταξινόμησης (Menze et al, 2009).

(Πίνακας 4: Δεδομένα Εισόδου των 5 μοντέλων ταξινόμησης που χρησιμοποιήθηκαν)

|   | <b>Δεδομένα Εισόδου</b> | <b>RF- 1</b> | <b>RF- 2</b> | <b>RF- 3</b> | <b>RF- 4</b> | <b>RF- 5</b> |
|---|-------------------------|--------------|--------------|--------------|--------------|--------------|
| 1 | Blue band               | ✓            | ✓            | ✓            | ✓            | ✓            |
| 2 | Green band              | ✓            | ✓            | ✓            | ✓            | ✓            |
| 3 | Red band                | ✓            | ✓            | ✓            | ✓            | ✓            |
| 4 | NIR band                | ✓            | ✓            | ✓            | ✓            | ✓            |

|    |                            |   |   |   |   |   |
|----|----------------------------|---|---|---|---|---|
| 5  | SWIR1 band                 | ✓ | ✓ | ✓ | ✓ | ✓ |
| 6  | SWIR2 band                 | ✓ | ✓ | ✓ | ✓ | ✓ |
| 7  | Red/Green band             | ✓ | ✓ | ✓ | ✓ | ✓ |
| 8  | NIR/Red band               | ✓ | ✓ | ✓ | ✓ | ✓ |
| 9  | SWIR2/Green band           | ✓ | ✓ | ✓ | ✓ | ✓ |
| 10 | NDVI                       | ✓ | ✓ | ✓ | ✓ | ✓ |
| 11 | TCBVI                      | ✓ | ✓ | ✓ | ✓ | ✓ |
| 12 | TCGVI                      | ✓ | ✓ | ✓ | ✓ | ✓ |
| 13 | Radar band                 | ✓ | ✗ | ✓ | ✗ | ✓ |
| 14 | Distance from road network | ✓ | ✓ | ✓ | ✓ | ✓ |
| 15 | DEM                        | ✓ | ✓ | ✓ | ✓ | ✓ |
| 16 | Slope                      | ✓ | ✓ | ✗ | ✗ | ✓ |
| 17 | Aspect                     | ✓ | ✓ | ✗ | ✗ | ✗ |

## 5. Αξιολόγηση ακρίβειας ταξινόμησης

Η ταξινόμηση που προέκυψε, αξιολογήθηκε με βάση την συνολική ακρίβεια ταξινόμησης, αλλά και με το στατιστικό kappa. Και οι 2 τιμές μας παρέχονταν αυτόματα ύστερα από την εκπαίδευση του αλγορίθμου από την R.

Η ολική ακρίβεια ταξινόμησης υπολογίζεται από τον πίνακα σφαλμάτων, αθροίζοντας όλο το πλήθος των εικονοστοιχείων που έχουν ταξινομηθεί σωστά (η διαγώνιος του πίνακα δηλαδή), και διαιρώντας το άθροισμα αυτό με το συνολικό αριθμό των εικονοστοιχείων που έχουν χρησιμοποιηθεί για την αξιολόγηση της ακρίβειας.

Το στατιστικό  $k$ , είναι ένα στατιστικό που μας δείχνει πόσο καλύτερη είναι η ταξινόμηση που έγινε, από αυτό που θα περιμέναμε από μια τυχαία ταξινόμηση των εικονοστοιχείων. Για να υπολογιστεί το στατιστικό  $k$ , πρέπει να υπολογιστεί το άθροισμα του γινομένου των εικονοστοιχείων των



γραμμών και των στηλών, από τον πίνακα σφαλμάτων, προς το αθροιστικό άθροισμα του πίνακα σφαλμάτων. Αυτή η τιμή που προκύπτει είναι η εκτιμώμενη ακρίβεια. Στην συνέχεια η εκτιμώμενη ακρίβεια αφαιρείται από την ολική ακρίβεια, και το αποτέλεσμα, διαιρείται με 1-την εκτιμώμενη ακρίβεια, δίνοντάς μας έτσι το στατιστικό k. (Campbell και Wynne (2011), σελ.419-422)

## **6. Εντοπισμός αλλαγών εδαφοκάλυψης**

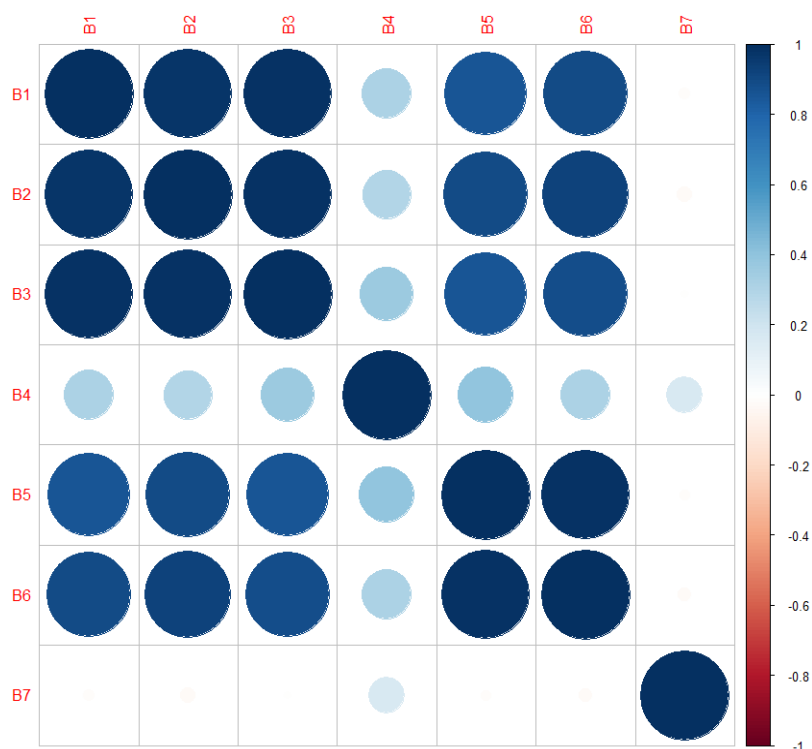
Από τους τρεις θεματικούς χάρτες που προέκυψαν και με την μέθοδο σύγκρισης ταξινομημένων εικόνων, PCC, και συγκεκριμένα με την διασταύρωση πινάκων, εντοπίστηκαν οι αλλαγές στην εδαφοκάλυψη και υπολογίστηκε η έκταση που “χάθηκε” από κάθε κλάση, καθώς και η κατεύθυνση της αλλαγής.

Τέλος επαναταξινομήθηκαν τα εικονοστοιχεία εκείνα τα οποία είχαν αποκλίνουσα τιμή σε σχέση με τα γειτονικά τους, και τα οποία είχαν έκταση 0,1 εκτάρια (900 τετραγωνικά μέτρα). Η επαναταξινόμηση έγινε με βάση την επικρατούσα τιμή των 8 γειτονικών του εικονοστοιχείων. Αυτό έγινε καθώς συνήθως είναι λάθη ταξινομήσεων από την ραδιομετρική μίξη των κλάσεων.

## ΚΕΦΑΛΑΙΟ 5: Αποτελέσματα

Σε αυτό το κεφάλαιο θα παρουσιαστούν τα αποτελέσματα που προέκυψαν ύστερα από την επεξεργασία των δεδομένων, ακολουθώντας την μεθοδολογία που παρουσιάστηκε.

Αρχικά, πρέπει να τονίσουμε πως δεν πραγματοποιήθηκε η σύζευξη των εικονοληπτικών Radar με τους οπτικούς δέκτες μέσω της IPCA, καθώς οι συσχετίσεις μεταξύ των loadings της PCA των οπτικών δεκτών με τα εικονοληπτικά Radar, είχαν τιμές από  $-0,0284$  έως  $0,1694$ , όπως φαίνεται και στο Γράφημα 1, τιμές δηλαδή που μας δείχνουν ότι δεν υπάρχει καθόλου γραμμική συσχέτιση. Έτσι τα δεδομένα των οπτικών δεκτών και



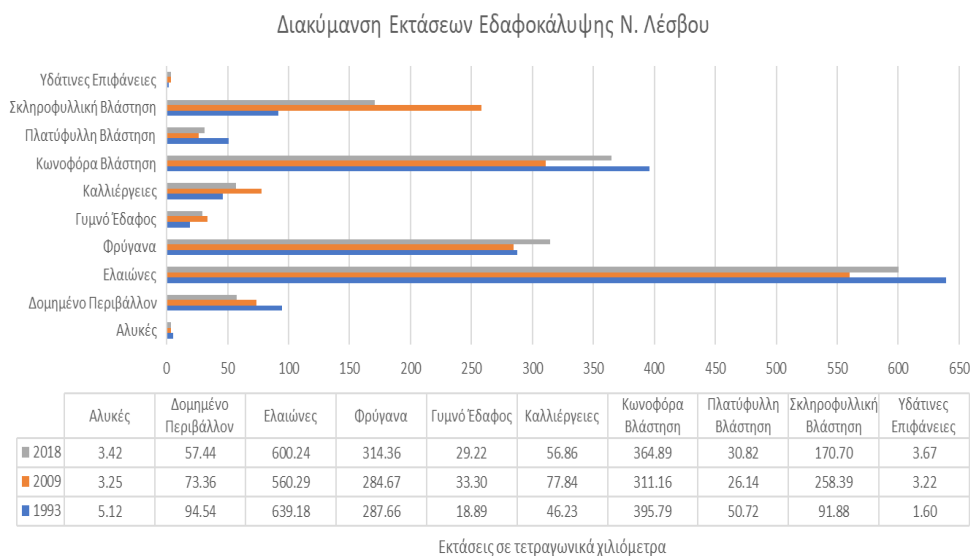
Γράφημα 1: Συσχέτιση των loadings των καναλιών των οπτικών δεκτών με το εικονοληπτικό Radar

των εικονοληπτικών Radar χρησιμοποιήθηκαν σαν δεδομένα εισόδου των μοντέλων ταξινόμησης που έχουν προαναφερθεί, χωρίς κάποια περαιτέρω επεξεργασία. Με τον τρόπο αυτό η σύζευξη πραγματοποιήθηκε μέσω της χρήσης του ταξινομητή Random Forest.

Στους παρακάτω χάρτες (βλ. Χάρτης 2; Χάρτης 3; Χάρτης 4) παρουσιάζεται η εδαφοκάλυψη της περιοχής μελέτης για τα έτη 1993 (ακρίβεια 81,89%), 2009 (ακρίβεια 88,63%) και 2018 (ακρίβεια 86,72%) αντίστοιχα, όπως προέκυψε ύστερα από την εκπαίδευση του αλγορίθμου Random Forest.

Επίσης με αποχρώσεις του κόκκινου εμφανίζεται το Δομημένο Περιβάλλον, με αποχρώσεις του μπλε εμφανίζονται οι Υδάτινες Επιφάνειες, με αποχρώσεις του λευκού εμφανίζονται οι Αλυκές, με αποχρώσεις του γκρι το Γυμνό Έδαφος με αποχρώσεις του κίτρινου εμφανίζονται οι Καλλιεργήσιμες Εκτάσεις και τα Φρύγανα, ενώ με αποχρώσεις του πράσινου εμφανίζεται η βλάστηση (Ελαιώνες, Πλατύφυλλη Βλάστηση, Σκληροφυλλική Βλάστηση και Κωνοφόρα Βλάστηση).

Στο γράφημα (Γράφημα 2) που ακολουθεί παρουσιάζονται οι εκτάσεις των κλάσεων χρήσεων/καλύψεων γης, που εξετάσαμε στην παρούσα διατριβή, σε τετραγωνικά χιλιόμετρα, για τις 3 χρονικές στιγμές κατά την διάρκεια μελέτης.



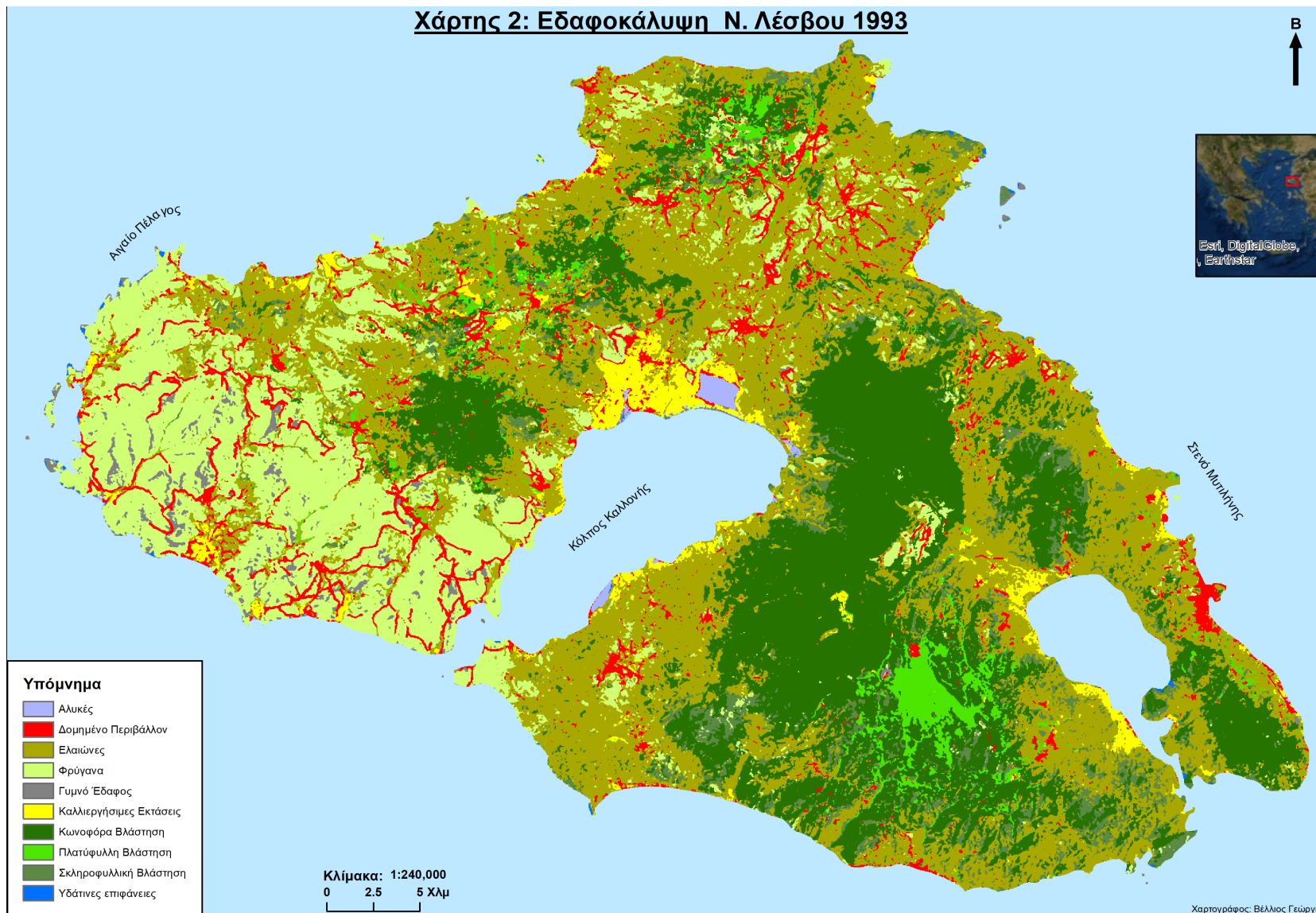
Γράφημα 2: Γραφική παρουσίαση των εκτάσεων εδαφοκάλυψης της Ν. Λέσβου σε τετραγωνικά χιλιόμετρα για τα έτη 1993, 2009 και 2018.

Στους Χάρτες 5 και Χάρτης 6 εμφανίζονται οι αλλαγές εδαφοκάλυψης των καλλιεργήσιμων εκτάσεων και των ελαιώνων από το 1993 μέχρι το 2009 και από το 2009 μέχρι το 2018 αντίστοιχα. Συγκεκριμένα εμφανίζονται οι εκτάσεις που προστέθηκαν και χάθηκαν από τις δύο αυτές εδαφοκαλύψεις που εξετάζουμε. Οι αποχρώσεις των χρωμάτων για τις εκτάσεις που προστέθηκαν στις συγκεκριμένες χρήσεις/καλύψεις γης, είναι πράσινες, ενώ για αυτές που χάθηκαν εμφανίζονται με κόκκινες αποχρώσεις.

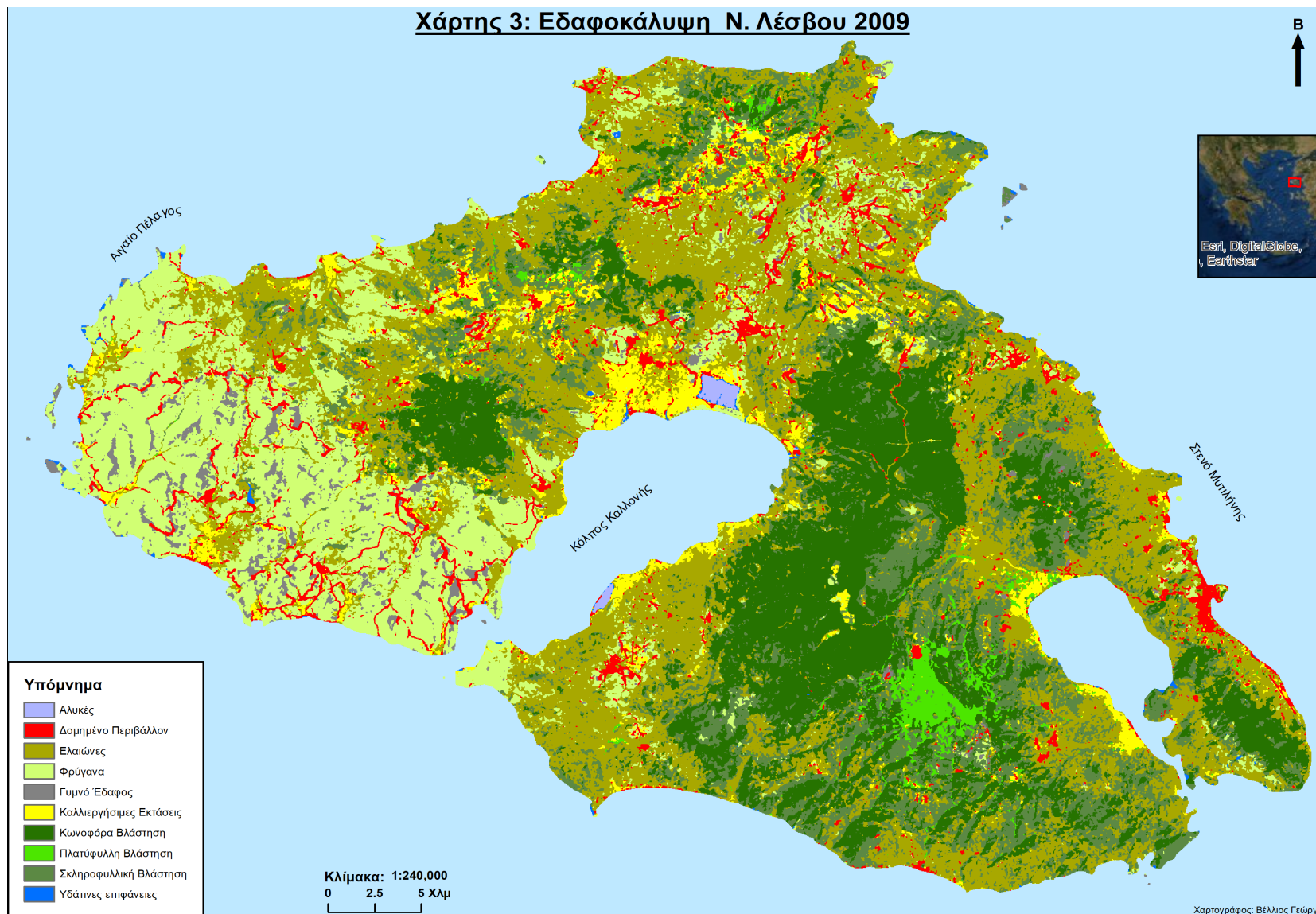
Στους Χάρτες 7 και Χάρτης 8, εμφανίζονται οι αλλαγές εδαφοκάλυψης του δομημένου περιβάλλοντος, συγκεκριμένα εμφανίζεται η επέκταση του δομημένου περιβάλλοντος για τις 2 περιόδους μελέτης, δηλαδή για 1993-2009 και 2009-2018, ενώ στους χάρτες δίνεται έμφαση στις πιέσεις που ασκούν τα δύο αστικά κέντρα, που εμφανίζουν την μεγαλύτερη συγκέντρωση πληθυσμού, η Μυτιλήνη και το Πλωμάρι, στις υπόλοιπες εδαφοκαλύψεις. Για την οπτικοποίηση της προέλευσης της επέκτασης του δομημένου περιβάλλοντος, χρησιμοποιήθηκαν οι αποχρώσεις των κλάσεων που έγιναν δομημένο περιβάλλον.

Στους Χάρτες 9 και Χάρτης 10, εμφανίζονται οι αλλαγές εδαφοκάλυψης των αλυκών και των υδάτινων επιφανειών. Συγκεκριμένα εμφανίζονται οι τεχνητές λίμνες που δημιουργήθηκαν κατά το διάστημα 1993-2009, ενώ παρατηρούμε και την μίξη των κλάσεων αυτών στις δύο αλυκές του νησιού, στην Καλλονή και στην Σκάλα Πολιχνίτου .

Οι χάρτες αλλαγών εδαφοκάλυψης, για την περίοδο 1993-2009 έχουν ακρίβεια εντοπισμού αλλαγών 72,58% και για την περίοδο 2009-2018 έχουν ακρίβεια 76,86%.

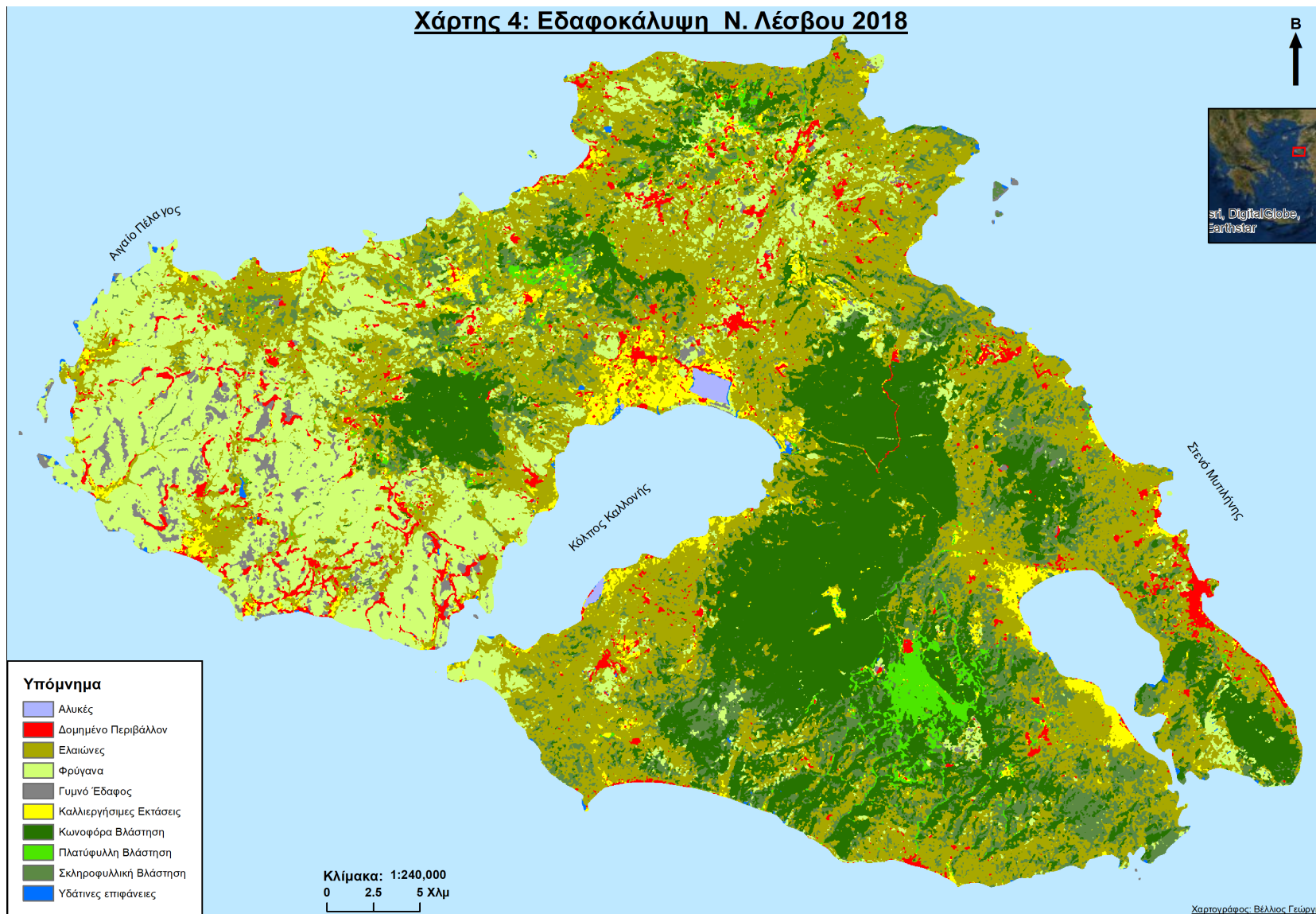


Χάρτης 2: Εδαφοκάλυψη Νήσου Λέσβου, για το έτος 1993 όπως προέκυψε από την ταξινόμηση



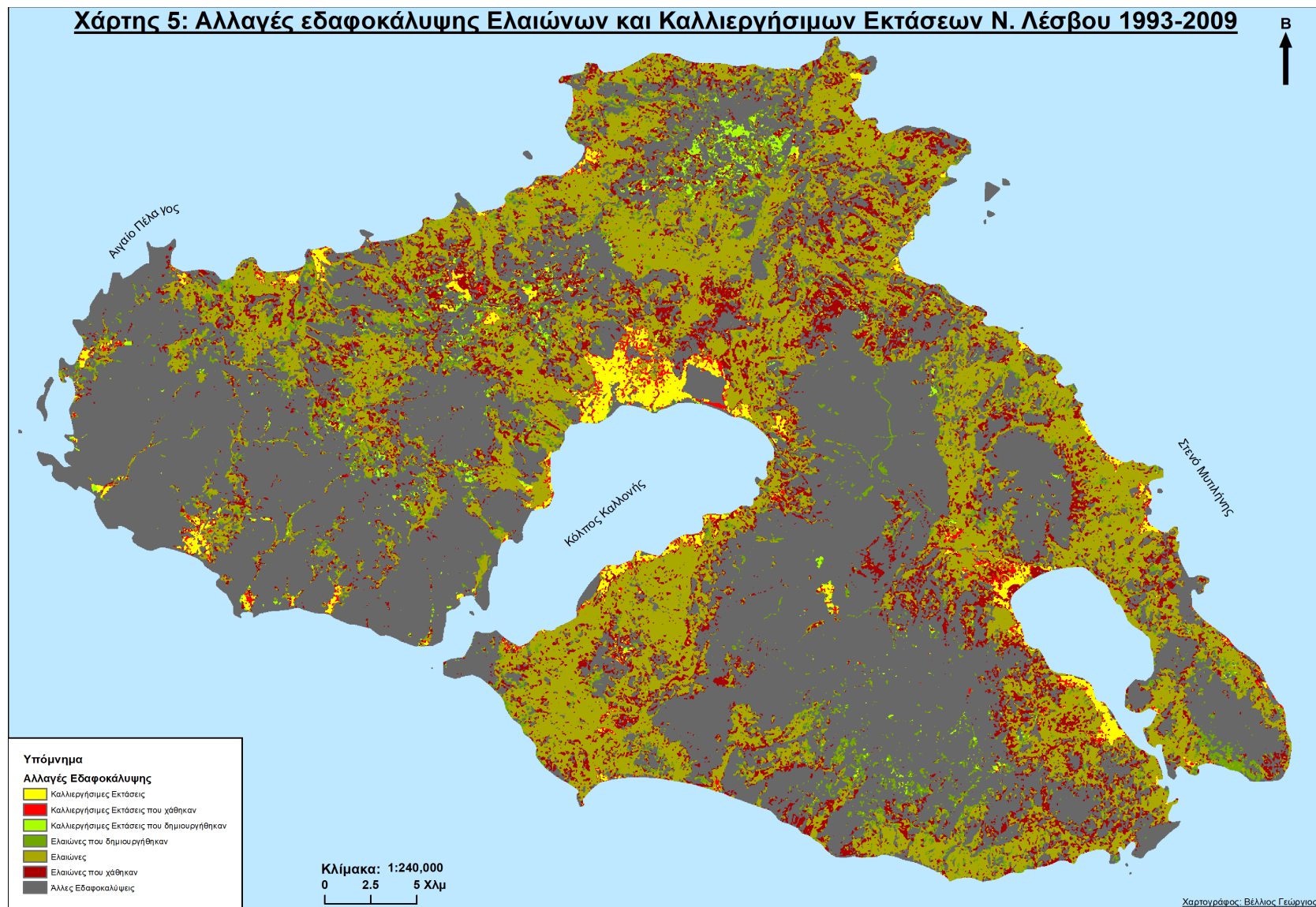
Χάρτης 3: Εδαφοκάλυψη Νήσου Λέσβου, για το έτος 2009 όπως προέκυψε από την ταξινόμηση



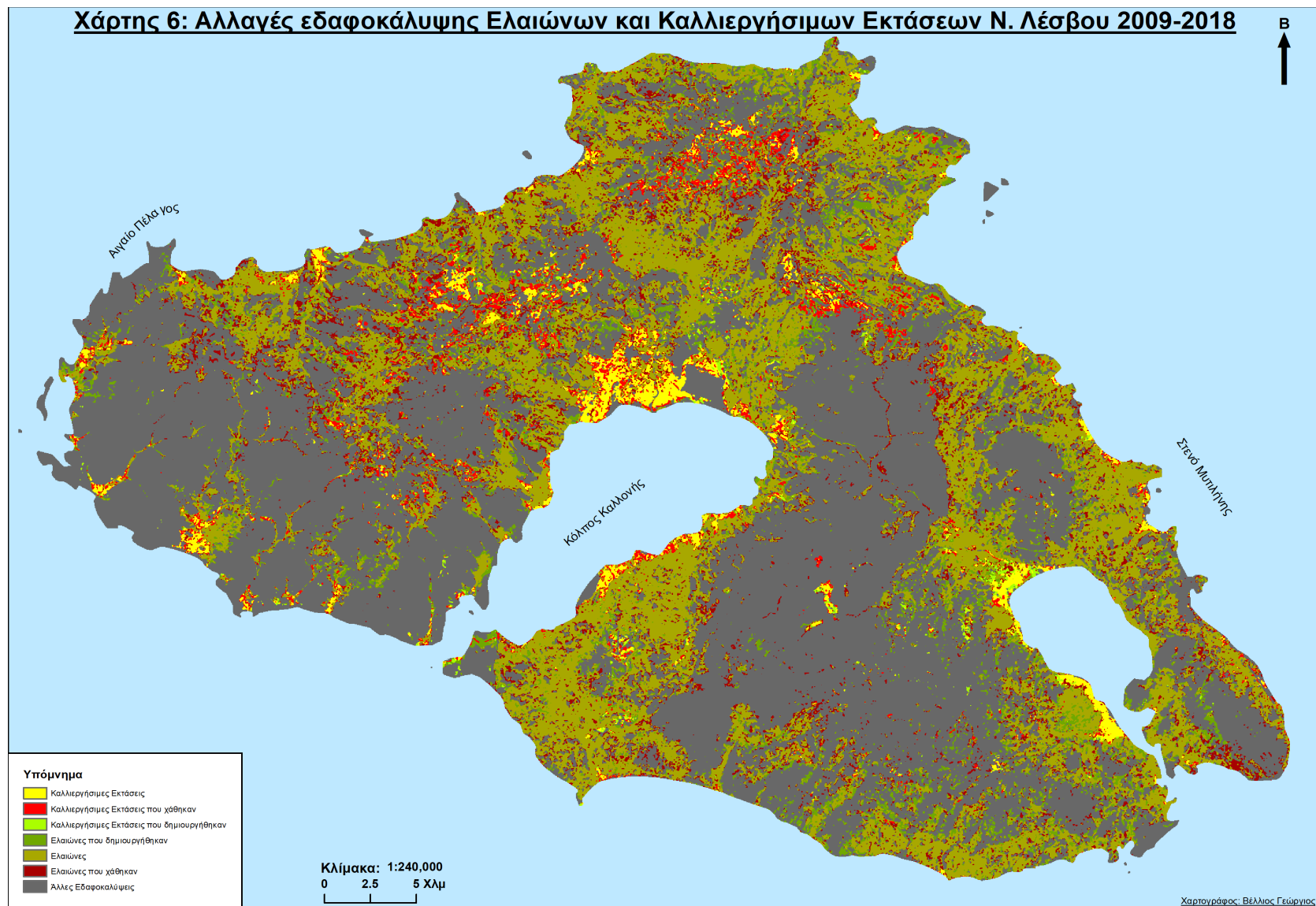


Χάρτης 4: Εδαφοκάλυψη Νήσου Λέσβου, για το έτος 2018 όπως προέκυψε από την ταξινόμηση

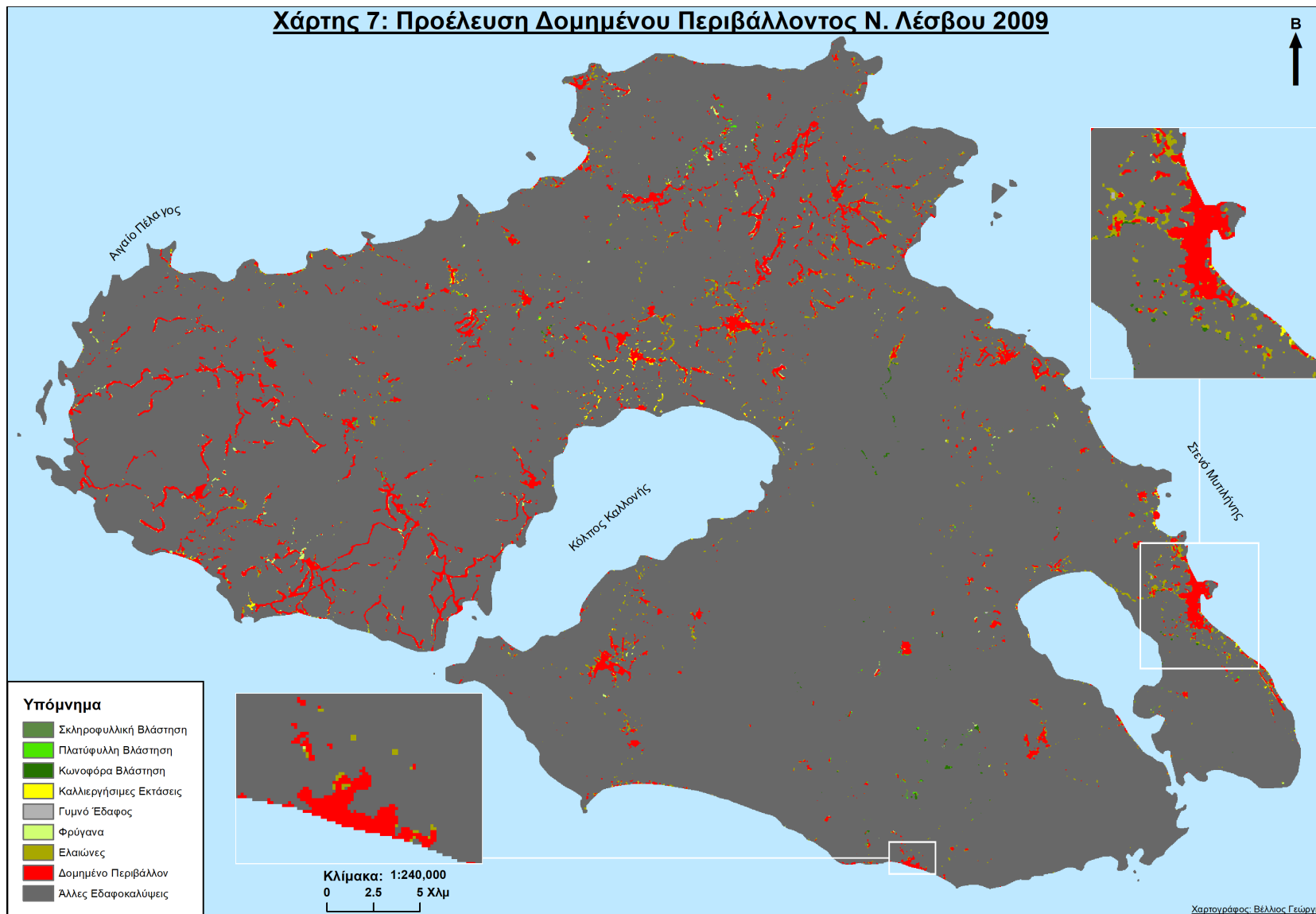




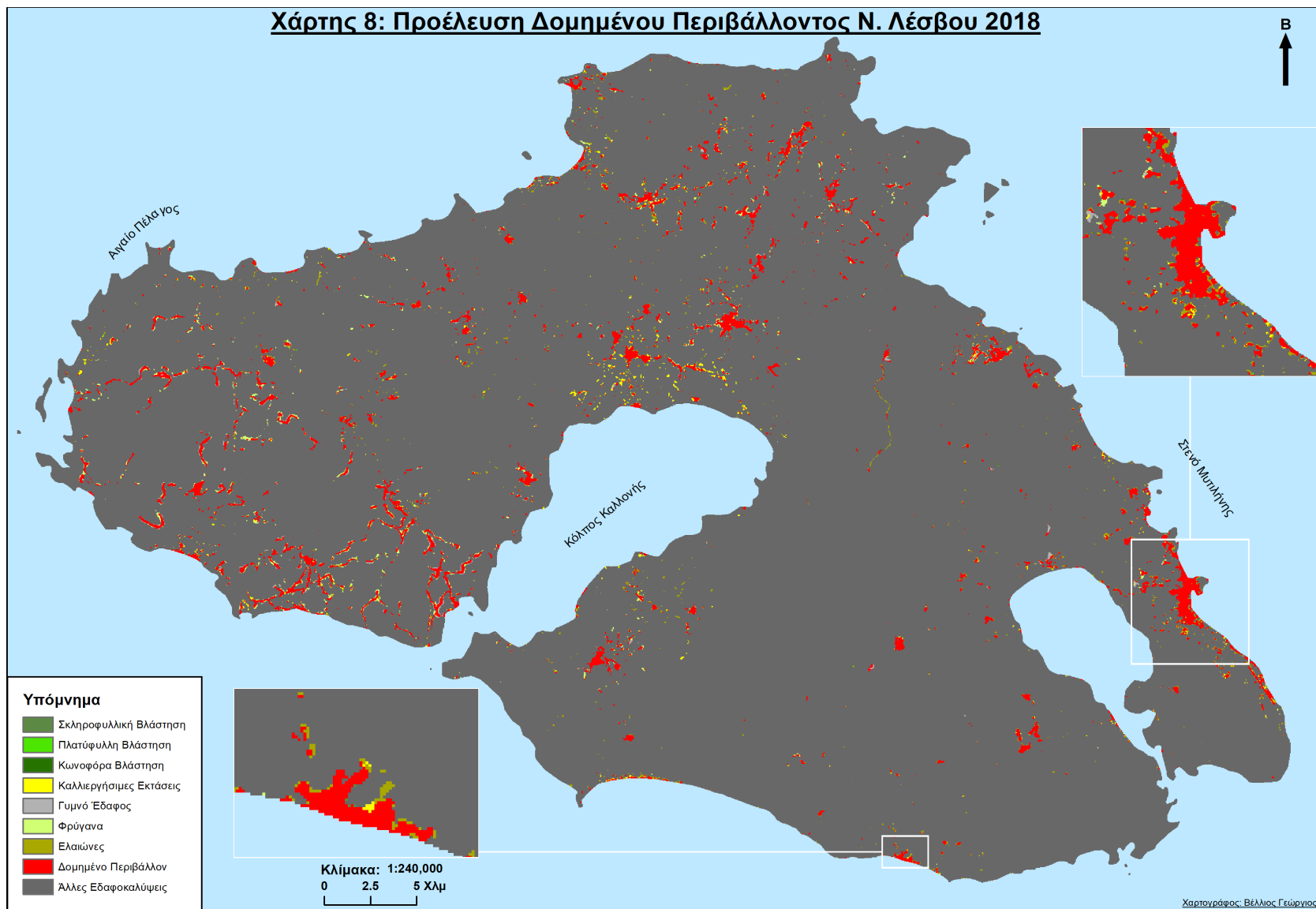
Χάρτης 5: Επέκταση και συρρίκνωση Ελαιώνων και Καλλιεργήσιμων Εκτάσεων Ν. Λέσβου, για το διάστημα 1993-2009



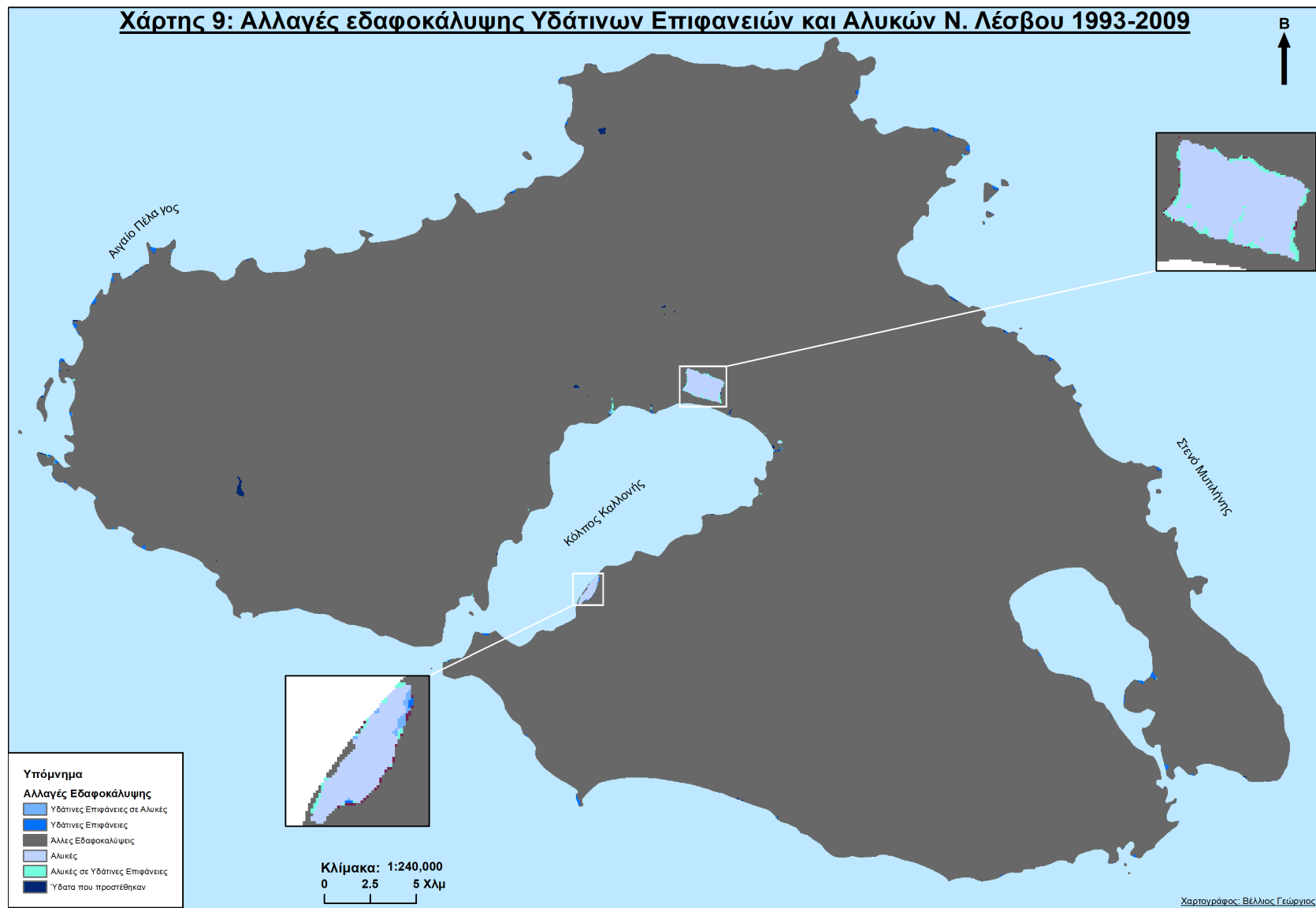
Χάρτης 6: Επέκταση και συρρίκνωση Ελαιώνων και Καλλιεργήσιμων Εκτάσεων Ν. Λέσβου, για το διάστημα 2009-2018



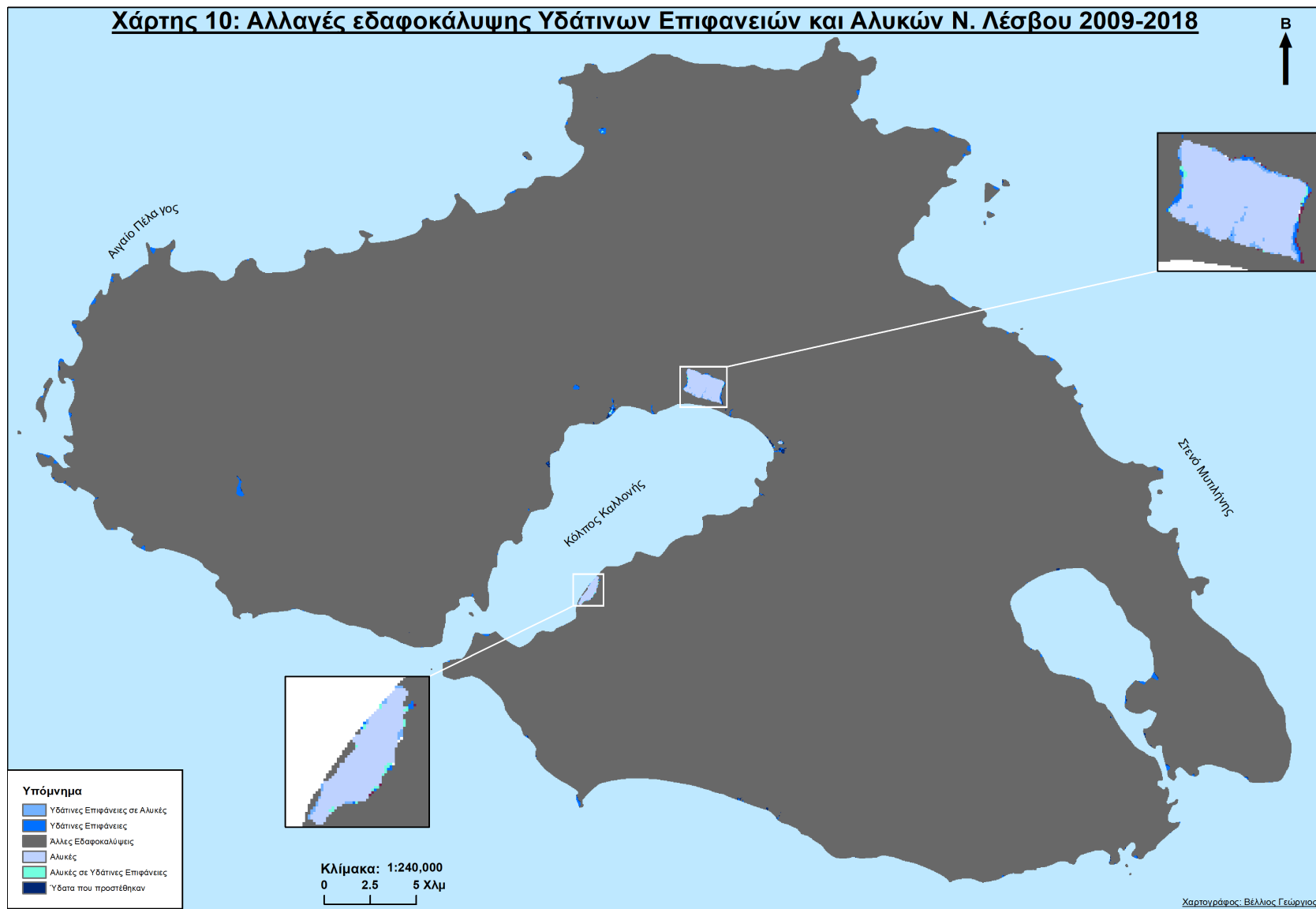
Χάρτης 7: Προέλευση Δομημένου Περιβάλλοντος 2009, Ν. Λέσβου



Χάρτης 8: Προέλευση Δομημένου Περιβάλλοντος 2018, Ν. Λέσβου



Χάρτης 9: Αλλαγές εδαφοκάλυψης σε Υδάτινες Επιφάνειες και Αλυκές Ν. Λέσβου, το διάστημα 1993-2009



Χάρτης 10: Αλλαγές εδαφοκάλυψης σε Υδάτινες Επιφάνειες και Αλυκές Ν. Λέσβου, το διάστημα 2009-2018

Η ολική ακρίβεια του μοντέλου της ταξινόμησης της χρήσης/κάλυψης γης, που επιλέχθηκε τελικά, του 1993 ήταν 81,89%, το στατιστικό  $k$  ήταν 79,29%, ενώ το σφάλμα OOB ήταν 15,49. Η ολική ακρίβεια για το 2009 ήταν 88,63%, το  $k$  ήταν 87,11% και το OOB ήταν 10,11, ενώ για την ταξινόμηση του 2018 η ολική ακρίβεια ήταν 86,72%, το  $k$  ήταν 84,83% με OOB σφάλμα 11,96. Αυτό το μοντέλο περιείχε ως δεδομένα εισόδου όλες τις μεταβλητές που είχαν συλλεχθεί και μας έδωσε τα μεγαλύτερα ποσοστά ακρίβειας στην ταξινόμηση, με χρήση 1600 δέντρων και 8 μεταβλητών για κάθε κόμβο.

Στο μοντέλο που δεν χρησιμοποιήθηκε το εικονοληπτικό Radar ως δεδομένο εισόδου, με 1600 δέντρα και 8 μεταβλητές για κάθε κόμβο, είχαμε 86,56%, 87,99% και 81,57% ολική ακρίβεια ταξινόμησης για τα έτη 2018, 2009 και 1993 αντίστοιχα, ενώ οι αντίστοιχες τιμές του  $k$  ήταν 84,65%, 86,39 και 81,57%.

Στο μοντέλο που δεν χρησιμοποιήθηκε η μεταβλητή του προσανατολισμού, θέσαμε 1200 δέντρα και 4 μεταβλητές για τον κάθε κόμβο, ενώ η ολική ακρίβεια του μοντέλου μας ήταν 85,62%, 87,88%, 81,47%, με  $k$  83,57%, 86,37% και 78,91% για τα έτη 2018, 2009 και 1993 αντίστοιχα.

Στο μοντέλο που από τις τοπογραφικές μεταβλητές, χρησιμοποιήθηκε μόνο η μεταβλητή του υψομέτρου, θέσαμε 1200 δέντρα και 6 μεταβλητές για τον κάθε κόμβο, η ολική ακρίβεια του μοντέλου ταξινόμησης, ήταν για το 2018 84,37%, για το 2009 85,67% και για το 1993 81,32%. Οι τιμές του στατιστικού  $k$ , ήταν 82,35%, 83,78% και 78,85% αντίστοιχα.

Στο τελευταίο μοντέλο, που χρησιμοποιήθηκε μόνο το υψόμετρο από τις τοπογραφικές μεταβλητές και δεν χρησιμοποιήθηκαν τα δεδομένα του εικονοληπτικού Radar, θέσαμε 1300 δέντρα και 6 μεταβλητές για κάθε κόμβο, η ολική ακρίβεια της ταξινόμησης ήταν για το 2018 84,08% , για το

2009 85,24% και για το 1993 80,89%, ενώ το κ ήταν 82,03%, 83,29% και 78,36% αντίστοιχα.



## ΚΕΦΑΛΑΙΟ 6: Συζήτηση

Στο κεφάλαιο αυτό θα εξαχθούν κάποια συμπεράσματα από τα αποτελέσματα που προέκυψαν. Η καλύτερη ταξινόμηση, αυτή που δηλαδή παρείχε καλύτερη ακρίβεια ταξινόμησης και καλύτερες τιμές στο  $k$  στατιστικό και στις τρεις χρονιές μελέτης, ήταν αυτή που χρησιμοποιήθηκαν όλα τα δεδομένα που είχαν συλλεχθεί.

Για τον χάρτη εδαφοκάλυψης του 1993 η ολική ακρίβεια ταξινόμησης ήταν 81,89% με  $k$  79,29%, ενώ όπως φαίνεται και στον Πίνακα 5 οι κλάσεις Σκληροφυλλική Βλάστηση (περίπου 0,34) και Καλλιέργειες (περίπου 0,30), εμφανίζουν τα μεγαλύτερα ποσοστά λαθών. Η κλάση

(Πίνακας 5: Πίνακας ακρίβειας ταξινόμησης και σφαλμάτων ανά κλάση, για το έτος 1993)

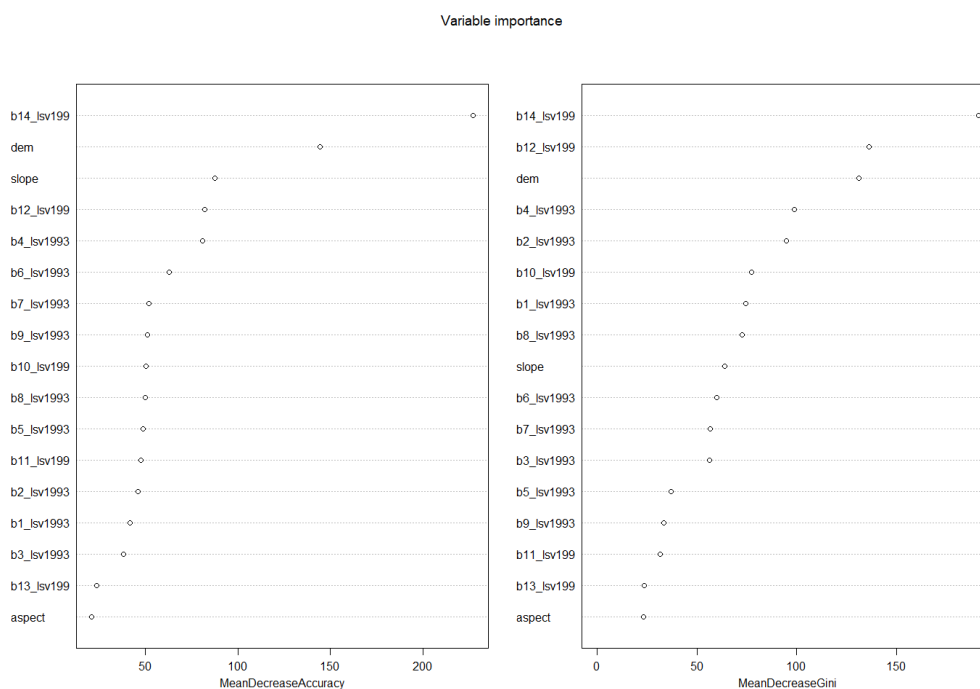
| Prediction    | 1993   |          |         |         |       |              |          |            |               |          | Σφάλμα Κλάσης |
|---------------|--------|----------|---------|---------|-------|--------------|----------|------------|---------------|----------|---------------|
|               | Αλικές | Δομημένο | Ελαιώες | Φρύγανα | Γυμνό | Καλλιέργειες | Κωνοφόρα | Πλατύφυλλη | Σκληροφυλλική | Υδάτινες |               |
| Αλικές        | 80     | 0        | 0       | 0       | 0     | 0            | 0        | 0          | 0             | 0        | 0             |
| Δομημένο      | 1      | 150      | 6       | 0       | 2     | 1            | 0        | 0          | 0             | 0        | 0,0625        |
| Ελαιώες       | 0      | 5        | 169     | 3       | 0     | 13           | 4        | 0          | 7             | 0        | 0,1592        |
| Φρύγανα       | 0      | 2        | 8       | 163     | 10    | 0            | 1        | 0          | 2             | 0        | 0,1237        |
| Γυμνό         | 1      | 2        | 2       | 17      | 88    | 0            | 1        | 0          | 9             | 0        | 0,2667        |
| Καλλιέργειες  | 0      | 3        | 18      | 3       | 0     | 65           | 2        | 1          | 1             | 0        | 0,3011        |
| Κωνοφόρα      | 0      | 0        | 4       | 0       | 0     | 0            | 271      | 1          | 13            | 0        | 0,0623        |
| Πλατύφυλλη    | 0      | 3        | 5       | 0       | 0     | 6            | 2        | 104        | 1             | 0        | 0,1405        |
| Σκληροφυλλική | 0      | 1        | 10      | 4       | 8     | 0            | 28       | 3          | 103           | 0        | 0,3439        |
| Υδάτινες      | 2      | 1        | 1       | 2       | 2     | 0            | 1        | 0          | 0             | 24       | 0,2727        |

|                 |        |
|-----------------|--------|
| Ολική Ακρίβεια: | 0,8189 |
| $k$ :           | 0,7929 |

Σκληροφυλλική Βλάστηση εμφανίζει να μπερδεύεται με τους Ελαιώες, κάτι το οποίο είναι λογικό, καθώς οι άγριες ελιές έχουν ταξινομηθεί ως Σκληροφυλλική Βλάστηση, και μπερδεύεται και με την Κωνοφόρα Βλάστηση, κάτι το οποίο χρήζει περαιτέρω διερεύνησης, καθώς οι δύο αυτές κλάσεις έχουν διαφορετικές φασματικές υπογραφές. Η κλάση Καλλιέργειες εμφανίζεται να μπερδεύεται με τους Ελαιώες, που αυτό μπορεί να οφείλεται στο ότι ως Καλλιέργειες έχουν ταξινομηθεί και

οπωροφόρα δέντρα, τα οποία είναι φυτεμένα με τον ίδιο τρόπο που φυτεύονται οι ελιές (σε κάνναβο). Το μικρότερο ποσοστό λάθους εμφανίζεται στην κλάση Αλυκές, όπου και είναι μηδενικό, και ακολουθούν η Κωνοφόρα Βλάστηση (0,0623) και το Δομημένο Περιβάλλον (0,0625).

Τέλος παρατηρούμε από το Γράφημα 3, ότι η μεταβλητή της απόστασης από το οδικό δίκτυο είναι η σημαντικότερη για την ταξινόμηση, καθώς το Mean Decrease in Accuracy, είναι μια βαθμολογία η οποία δείχνει το κατά πόσο θα μειωθεί η ακρίβεια του μοντέλου ταξινόμησης αν αφαιρεθεί μία μεταβλητή και έχει σχέση με το ποσοστό σφάλματος OOB, και η αμέσως επόμενη σημαντικότερη μεταβλητή είναι το ΨΜΕ, ενώ η λιγότερο σημαντική μεταβλητή είναι ο προσανατολισμός. Επίσης παρατηρούμε πως



Γράφημα 3: Σημαντικότητα μεταβλητών για την ταξινόμηση του 1993.

η ίδια μεταβλητή είναι αυτή με την μεγαλύτερη βαθμολογία στο Mean Decrease Gini, όπου μας δείχνει ότι είναι η μεταβλητή που μπορεί να διαχωρίσει τις κλάσεις ανά κόμβο και άρα συμβάλλει περισσότερο στην

ομοιογένεια του μοντέλου, ενώ η αμέσως επόμενη μεταβλητή είναι το NDVI και η μεταβλητή που συμβάλλει λιγότερο στον διαχωρισμό των κλάσεων είναι ο προσανατολισμός (Nguyen et al, 2018).

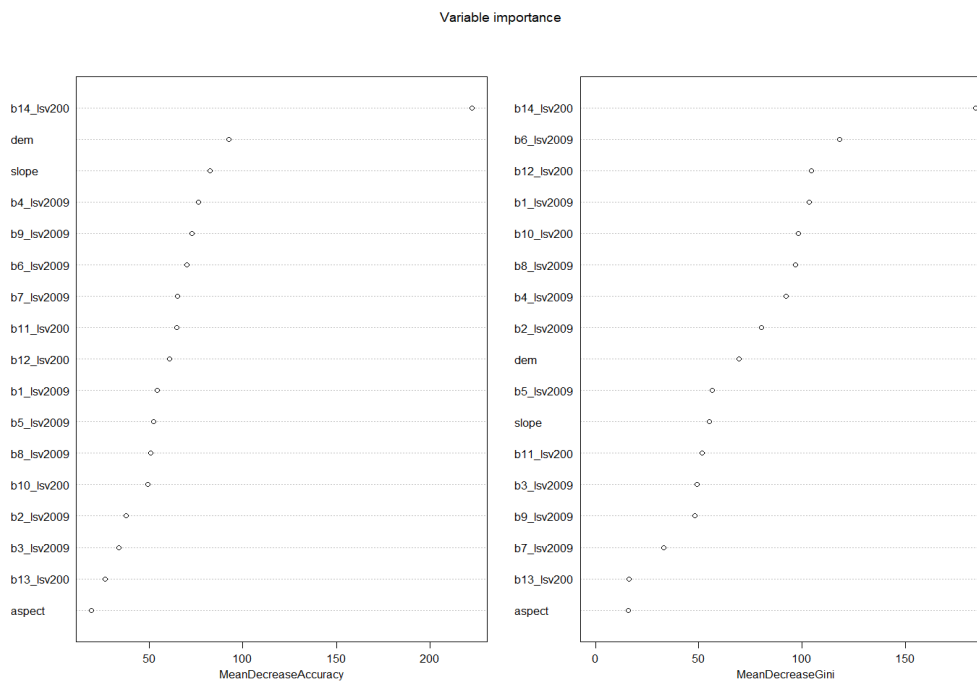
Για το 2009 η ολική ακρίβεια της ταξινόμησης είναι 88,63% και με Kappa 87,11% ενώ όπως φαίνεται από τον Πίνακα 6 πάλι οι κλάσεις Καλλιέργειες (0,2809) και Σκληροφυλλική Βλάστηση (0,2125) έχουν τα μεγαλύτερα ποσοστά λάθους, ενώ ισχύουν τα ίδια που προαναφέρθηκαν για την ταξινόμηση του 1993. Τα μικρότερα ποσοστά λαθών εμφανίζουν πάλι οι Αλυκές (0,0149) και το Δομημένο Περιβάλλον (0,0159).

(Πίνακας 6: Πίνακας ακρίβειας ταξινόμησης και σφαλμάτων ανά κλάση, για το έτος 2009)

| Prediction    | Reference |          |         |         |       |              |          |            |               |          | Σφάλμα Κλάσης |
|---------------|-----------|----------|---------|---------|-------|--------------|----------|------------|---------------|----------|---------------|
|               | Αλυκές    | Δομημένο | Ελαιώες | Φρύγανα | Γυμνό | Καλλιέργειες | Κωνοφόρα | Πλατύφυλλη | Σκληροφυλλική | Υδάτινες |               |
| Αλυκές        | 66        | 0        | 0       | 0       | 0     | 0            | 0        | 0          | 0             | 1        | 0,0149        |
| Δομημένο      | 0         | 186      | 3       | 0       | 0     | 0            | 0        | 0          | 0             | 0        | 0,0159        |
| Ελαιώες       | 0         | 2        | 187     | 5       | 0     | 5            | 1        | 0          | 7             | 0        | 0,0967        |
| Φρύγανα       | 0         | 1        | 4       | 165     | 9     | 0            | 0        | 0          | 2             | 0        | 0,0884        |
| Γυμνό         | 0         | 3        | 0       | 10      | 103   | 0            | 0        | 0          | 1             | 0        | 0,1197        |
| Καλλιέργειες  | 0         | 3        | 15      | 4       | 0     | 64           | 0        | 2          | 1             | 0        | 0,2809        |
| Κωνοφόρα      | 0         | 0        | 1       | 0       | 0     | 0            | 273      | 3          | 11            | 0        | 0,052         |
| Πλατύφυλλη    | 0         | 0        | 2       | 0       | 0     | 4            | 4        | 104        | 3             | 0        | 0,112         |
| Σκληροφυλλική | 0         | 0        | 9       | 2       | 2     | 0            | 17       | 4          | 126           | 0        | 0,2125        |
| Υδάτινες      | 3         | 0        | 1       | 1       | 0     | 0            | 1        | 0          | 0             | 34       | 0,15          |

|                 |        |
|-----------------|--------|
| Ολική Ακρίβεια: | 0,8863 |
| k:              | 0,8711 |

Επίσης από το Γράφημα 4 παρατηρούμε ότι η μεταβλητή απόσταση από το οδικό δίκτυο παραμένει να είναι η πιο σημαντική για την ταξινόμηση, αμέσως επόμενη είναι το ΨΜΕ, καθώς και να συμβάλλει περισσότερο στην ομοιογένεια του μοντέλου, με αμέσως επόμενη μεταβλητή το NDVI.



Γράφημα 4: Σημαντικότητα μεταβλητών για την ταξινόμηση του 2009.

Για το 2018 η ολική ακρίβεια της ταξινόμησης είναι 86,72% και Καρρα 84,83%, ενώ στον Πίνακα 7 παρατηρούμε ότι οι Καλλιέργειες (0,3) έχουν τα μεγαλύτερα ποσοστά σφαλμάτων, με αμέσως επόμενη να εμφανίζεται η κλάση Υδάτινες Επιφάνειες (0,1875). Οι Υδάτινες Επιφάνειες παρατηρούμε να μπερδεύονται αρκετά με τις Αλυκές. Οι Αλυκές εμφανίζουν το μικρότερο ποσοστό σφαλμάτων (0,0133) και ακολουθεί το Δομημένο Περιβάλλον (0,0348).

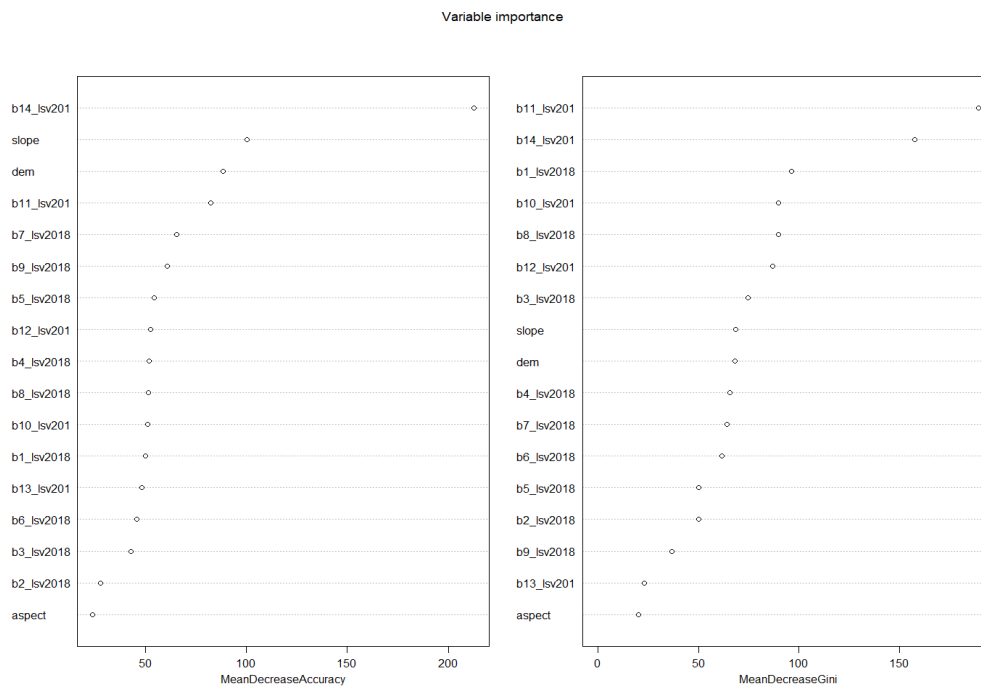
(Πίνακας 7: Πίνακας ακρίβειας ταξινόμησης και σφαλμάτων ανά κλάση, για το έτος 2018)

| Prediction    | Reference |          |         |         |       |              |          |            |               |          | Σφάλμα Κλάσης |        |
|---------------|-----------|----------|---------|---------|-------|--------------|----------|------------|---------------|----------|---------------|--------|
|               | Αλυκές    | Δομημένο | Ελαιώες | Φρύγανα | Γυμνό | Καλλιέργειες | Κωνοφόρα | Πλατύφυλλη | Σκληροφυλλική | Υδάτινες |               |        |
| Αλυκές        | 74        | 0        | 0       | 0       | 0     | 0            | 0        | 0          | 0             | 0        | 1             | 0,0133 |
| Δομημένο      | 0         | 194      | 3       | 1       | 0     | 3            | 0        | 0          | 0             | 0        | 0             | 0,0348 |
| Ελαιώες       | 0         | 3        | 171     | 4       | 0     | 5            | 1        | 0          | 10            | 0        | 0             | 0,1185 |
| Φρύγανα       | 0         | 0        | 8       | 161     | 15    | 0            | 0        | 0          | 2             | 0        | 0             | 0,1344 |
| Γυμνό         | 0         | 1        | 0       | 13      | 99    | 1            | 0        | 0          | 2             | 1        | 0             | 0,1538 |
| Καλλιέργειες  | 0         | 7        | 15      | 0       | 1     | 70           | 0        | 4          | 2             | 1        | 0             | 0,3    |
| Κωνοφόρα      | 0         | 0        | 3       | 1       | 0     | 0            | 264      | 4          | 14            | 0        | 0             | 0,0769 |
| Πλατύφυλλη    | 0         | 0        | 8       | 0       | 0     | 6            | 3        | 97         | 3             | 0        | 0             | 0,1491 |
| Σκληροφυλλική | 0         | 0        | 8       | 0       | 1     | 1            | 12       | 2          | 127           | 0        | 0             | 0,1589 |
| Υδάτινες      | 4         | 0        | 0       | 2       | 1     | 1            | 1        | 0          | 0             | 39       | 0             | 0,1875 |

|                 |        |
|-----------------|--------|
| Ολική Ακρίβεια: | 0,8672 |
| k:              | 0,8483 |

Το Γράφημα 5 μας παρουσιάζει την σημαντικότητα των μεταβλητών καθώς και το πόσο συμβάλει η κάθε μεταβλητή στην ομοιογένεια του μοντέλου. Η απόσταση του οδικού δικτύου εμφανίζεται να είναι η σημαντικότερη μεταβλητή για την ταξινόμηση και να ακολουθεί η κλίση του εδάφους, ενώ για την συμβολή στην ομοιογένεια του μοντέλου παρατηρούμε πως η μεταβλητή που συμβάλει περισσότερο είναι ο μετασχηματισμός Tasseled Cap, συγκεκριμένα αυτός του Brightness και ακολουθεί η απόσταση του οδικού δικτύου.

Στο σύνολο των μοντέλων ταξινόμησης, παρατηρούμε ότι με βάση το MDA, ο προσανατολισμός του εδάφους είναι η μεταβλητή που έχει συμβάλει λιγότερο στην ακρίβεια της ταξινόμησης για τα έτη 1993, 2009 και 2018, ενώ αμέσως επόμενη είναι τα δεδομένα του εικονοληπτικού Radar για τα έτη 1993 και 2009, και για το 2018 είναι τα δεδομένα του καναλιού που καταγράφει στο κόκκινο φάσμα της ηλεκτρομαγνητικής ακτινοβολίας.



Γράφημα 5: Σημαντικότητα μεταβλητών για την ταξινόμηση του 2018.

Επίσης στο σύνολο των μοντέλων ταξινόμησης, παρατηρούμε ότι με βάση το MDG, ο προσανατολισμός και τα εικονοληπτικά Radar, συμβάλλουν λιγότερο από τις υπόλοιπες μεταβλητές στην ομοιογένεια του μοντέλου ταξινόμησης για όλες τις χρονιές.

## ΚΕΦΑΛΑΙΟ 7: Συμπεράσματα

Από τα αποτελέσματα που παρουσιάστηκαν και αναλύθηκαν στις προηγούμενες ενότητες συμπεραίνουμε ότι για την ταξινόμηση του έτους 1993, και συγκεκριμένα από την ακρίβεια της ταξινόμησης, από όλα τα μοντέλα που χρησιμοποιήθηκαν, προκύπτει η ανάγκη χρήσης υποβάθρου υψηλής χωρικής διακριτικής ικανότητας, καθώς με τον οπτικό έλεγχο που διενεργήθηκε σε αυτές, με βάση τις δορυφορικές εικόνες του 2009, δεν εξήχθησαν τα επιθυμητά αποτελέσματα.

Επίσης συμπεραίνουμε ότι είναι αναγκαία η χρήση επιπλέον δεικτών που να βοηθάνε περισσότερο στον διαχωρισμό του γυμνού εδάφους με το δομημένο περιβάλλον, καθώς παρατηρήθηκε ότι υπάρχει μίξη των δύο κλάσεων σε όλες τις χρονιές, για να βοηθηθεί ο αλγόριθμος στο να ξεχωρίσει τις δύο αυτές κλάσεις. Ακόμη είναι αναγκαία και η χρήση δεικτών που να εντοπίζουν το νερό, ώστε να αποφευχθεί η λάθος ταξινόμηση περιοχών με ρηγά ύδατα ως γυμνό έδαφος.

Προτείνεται η χρήση διπλής, τετραπλής πόλωσης δεδομένων εικονοληπτικών Radar, καθώς και της αποσύνθεσης της πόλωσης (Polarization Decomposition), ώστε να αυξηθούν οι διαστάσεις της πληροφορίας των Radar για να ερευνηθεί περισσότερο η συμβολή της πληροφορίας τους στο τελικό μοντέλο ταξινόμησης.

Τέλος προτείνεται η επανάληψη της παρούσας μελέτης με χρήση πολυεποχικών δεδομένων και συγκεκριμένα με απόκτηση δεδομένων για τους μήνες Ιούλιο και Φεβρουάριο, όπου αλλάζει η φαινολογία της φυτοκάλυψης της περιοχής μελέτης με αποτέλεσμα να μπορούμε να ξεχωρίσουμε καλύτερα κλάσεις οι οποίες μπερδεύονται μεταξύ τους.

## ΚΕΦΑΛΑΙΟ 8: Βιβλιογραφία

### 1. Βιβλία

Καρτάλης Κ., Φείδας Χ., (2015), Αρχές και Εφαρμογές Δορυφορικής Τηλεπισκόπησης, ΤΖΙΟΛΑ, σελ.1-5, 45-46, 50-53, 60-64, 138-147.

Παρχαρίδης Ι., (2015), Αρχές δορυφορικής Τηλεπισκόπησης Θεωρία και εφαρμογές, Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, σελ. 21-22.

Blondel J., Aronson J., Bodiou J.Y., Boeuf G., (2010), The Mediterranean Region: Biological Diversity in Space and Time, 2<sup>nd</sup> Edition, Oxford University Press.

Campbell J.B., Wynne R.H., (2011), Introduction to Remote Sensing, 5<sup>th</sup> edition, The Guilford Press, New York, London, σελ. 419-422, 585-586.

Richards J.A., Jia X., (2006), Remote Sensing Digital Image Analysis An Introduction, 4<sup>th</sup> Edition, Springer, σελ.9.

Richards J.A., (2009), Remote Sensing with Imaging Radar, Springer, σελ 28-29, 62, 111-112, 312-314.

### 2. Επιστημονική βιβλιογραφία

Abd El-Kawy O.R., Rod J.K., Ismail H.A., Suliman A.S., (2011), Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data, Applied Geography, 483-494.

Amarsaikhan D., Saandar M., Ganzorig M., Blotevogel H. H., Egshiglen E., Gantuyal R., Nergui B., Enkhjargal D., (2012), Comparison of multisource image fusion methods and land cover classification, International Journal of Remote Sensing, 33, 2532-2550.

Bossard M., Feranec J., Otahel J., (2000), CORINE land cover technical guide – Addendum 2000 (Report No. 40), Copenhagen European Environment Agency.



- Breiman L., (2001), Random Forests, *Machine Learning*, 45, 5-32 .
- Clark M. L., Aide T. M., Riner G., (2012), Land change for all municipalities in Latin America and the Caribbean assessed from 250-m MODIS imagery (2001-2010), *Remote Sensing of Environment*, 126, 84-103.
- Clark M. L., Aide T. M., Riner G., (2010), A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America, *Remote Sensing of Environment*, 114, 2816-2832.
- Colditz R. R., (2015), An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms, *Remote Sensing* , 7, 9655-9681.
- Gislason P. A., Benediktsson J. A., Sveinsson J. R., (2006), Random forests for land cover classification, *Pattern Recognition Letters*, 27, 294-300.
- Gounaridis D., Apostolou A., Koukoulas S., (2015), Land cover of Greece, 2010: a semi-automated classification using random forests, *Journal of Maps*.
- Guan H., Chapman M., Deng F., Ji Z., Yang X., (2013), Integration of orthoimagery and lidar data for object-based urban thematic mapping using random forests, *International Journal of Remote Sensing*, 34, 5166-5186.
- Horning H., (2010) Random Forests: An algorithm for image classification and generation of continuous fields data sets, *International Conference in Geoinformatics for Spatial Infrastructure Development in Earth and Applied Science*.
- Kontos T.D., Komilis D.P., Halvadakis C.P., (2003), Siting MSW landfills on Lesbos island with a GIS-based methodology, *Waste Management and Research*, 21, 262-277.
- Kosmas C., Danalatos N.G., Gerontidis St., (2000), The effect of land parameters on vegetation performance and degree of erosion under Mediterranean conditions, *Catena*, 40, 3-17.
- Lawrence R. L., Wood S. D., Sheley R. L., (2006), Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications

(randomForest), *Remote Sensing of Environment*, 100, 356-362.

Li M., Im J., Beier C., (2013), Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington Wildfire Forest, *GIScience & Remote Sensing*, 50, 361-384.

Liaw A., Wiener M., (2002), Classification and regression by randomForest, *R News*, 2, 18-22.

Liu H., Zhou Q., (2004), Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison, *International Journal of Remote Sensing*, 25:5, 1037-1050.

Lu D., Mausel P., Brondizio E., Moran E., (2004), Change detection techniques, *International Journal of Remote Sensing*, 25:12, 2365-2401.

Marathianou M., Kosmas C., Georntidis St., Detsis V., (2000), Land-use evolution and degradation in Lesvos (Greece): A historical approach, *Land Degradation and Development*, 11, 63-73.

Mas J.F., (1999), Monitoring land-cover changes: A comparison of change detection techniques, *International Journal of Remote Sensing*, 20:1, 139-152.

Menze B. H., Kelm M. B., Masuch R., Himmelreich U., Bachert P., Petrich W., Hamprecht A. F., (2009), A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data, *BMC Bioinformatics*. 10:213.

Mohamed S.A., El-Raey M.E., (2019), Land cover classification and change detection analysis of Qaroun and Wadi El-Rayyan lakes using multi-temporal remotely sensed imagery, *Environmental Monitoring and Assessment*, 191:229.

Nguyen H.T.T, Doan T.M., Radeloff V., (2018), Applying Random Forest classification to map land use/cover using Landsat 8 OLI, *International Society for Photogrammetry and Remote Sensing*, XLII-3/W4, 363-367.

Nitze I., Schulthess U., Asche H., (2012), Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification,

Proceedings of the 4<sup>th</sup> GEOBIA, 7-9.

Pal M., (2005) Random Forest classifier for remote sensing classification, International Journal of Remote Sensing, 26, 217-222.

Rizzoli P., Brautigam B.,(2014), Radar Backscatter Modeling Based on Global TanDEM-X Mission Data, IEEE Transactions on Geoscience and Remote Sensing. (Διαθέσιμο σε [https://www.researchgate.net/publication/259904538\\_Radar\\_Backscatter\\_Modeling\\_Based\\_on\\_Global\\_TanDEM-X\\_Mission\\_Data](https://www.researchgate.net/publication/259904538_Radar_Backscatter_Modeling_Based_on_Global_TanDEM-X_Mission_Data) )

Sesnie S. E., Gessler P. E., Finegan B., Thessler S., (2008), Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments, Remote Sensing of Environment, 112, 2145-2159.

Singh A., (1989), Review Article Digital change detection techniques using remotely-sensed data, International Journal of Remote Sensing, 10:6, 989-1003.

Small D., (2011), Flattening Gamma: Radiometric Terrain Correction for Sar Imagery, IEEE Transactions on Geoscience and Remote Sensing, 49, 3081-3093.

Stefanski J., Mack B., Waske B., (2013), Optimization of object-based image analysis with Random Forests for land cover mapping, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6, 2492-2504.

Symeonakis E., Caccetta P., Koukoulas S., Furby S., Karathanasis N., (2012), Multi-temporal land-cover classification and change analysis with conditional probability networks: the case of Lesbos Island (Greece), International Journal of Remote Sensing, 33, 4075-4093.

Symeonakis E., Karathanasis N., Koukoulas S., Panagopoulos G., (2016), Monitoring sensitivity to land degradation and desertification with the environmentally sensitive area index: The case of Lesbos Island, Land Degradation and Development, 27, 1562-1573.

Toosi N.B., Fakheran S., Pourmanafi S., Ginzier C., Waser L. T., (2019), Comparing different classification algorithms for monitoring mangrove cover changes in southern Iran, Global Ecology and Conservation.

USGS, (2012), Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS): Algorithm Description Document, Version 1.3, Department of the Interior U.S. Geological Survey.

Vermote E.F., Tanre D., Deuze J.L., Herman M., Morcrette J.J., (1997), Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An Overview, IEEE Transactions on Geoscience and Remote Sensing, 35, 675-685.

Vermote E.F., Justice C., Claviere M., Franch B., (2016), Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product, Remote Sensing of Environment, 185, 46-56.

Zanter K., (2018<sup>a</sup>), Landsat 7(L7) Data Users Handbook, Version 1.0, Department of the Interior U.S. Geological Survey.

Zanter K., (2018<sup>b</sup>), Landsat 8 Surface Reflectance Code (LaSRC) Product Guide), Version 1.0, Department of the Interior U.S. Geological Survey.

Zanter K., (2019), Landsat 8 (L8) Data Users Handbook, Version 4.0, Department of the Interior U.S. Geological Survey.

### 3. Διαδικτυακές πηγές

[https://www.usgs.gov/land-resources/nli/landsat/landsat-5?qt-science\\_support\\_page\\_related\\_con=0#qt-science\\_support\\_page\\_related\\_con](https://www.usgs.gov/land-resources/nli/landsat/landsat-5?qt-science_support_page_related_con=0#qt-science_support_page_related_con), Τελευταία πρόσβαση 09/06/2019

[https://www.usgs.gov/land-resources/nli/landsat/landsat-8?qt-science\\_support\\_page\\_related\\_con=0#qt-science\\_support\\_page\\_related\\_con](https://www.usgs.gov/land-resources/nli/landsat/landsat-8?qt-science_support_page_related_con=0#qt-science_support_page_related_con), Τελευταία πρόσβαση 09/06/2019

<https://sentinel.esa.int/web/sentinel/missions/sentinel-2>, Τελευταία πρόσβαση 09/06/2019

<https://sentinel.esa.int/web/sentinel/document-library/content/-/article/sentinel-2-level-1-to-level-1c-product-specifications>, Τελευταία πρόσβαση 09/06/2019

<https://sentinel.esa.int/web/sentinel/missions/sentinel-1/observation->

[scenario](#), Τελευταία πρόσβαση 10/06/2019

<https://earth.esa.int/web/guest/-/sar-single-look-complex-image-product-1368>, Τελευταία πρόσβαση 10/06/2019

<https://earth.esa.int/web/guest/-/sar-precision-image-product-1477>,  
Τελευταία πρόσβαση 10/06/2019

<http://www.hnms.gr/emv/el/> , Τελευταία πρόσβαση 10/06/2019

<https://www.wunderground.com>, Τελευταία πρόσβαση 10/06/2019

<http://meteo.aegean.gr/>, Τελευταία πρόσβαση 10/06/2019