



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ –  
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ  
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ – ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

## Διπλωματική Εργασία

Διερεύνηση μεθόδων εκτίμησης επιπέδων οργανικού  
άνθρακα στο έδαφος από τηλεπισκοπικά  
πολυφασματικά δεδομένα

Καβαθά Νεφέλη

Αθήνα, Οκτώβριος 2023





NATIONAL TECHNICAL UNIVERSITY OF ATHENS

SCHOOL OF RURAL, SURVEYING AND GEOINFORMATICS ENGINEERING

DEPARTMENT OF TOPOGRAPHY – REMOTE SENSING LABORATORY

# Exploration of methods for estimating soil organic carbon levels using multi-spectral remote sensing data

Diploma Thesis

Kavvatha Nefeli

Athens, October 2023







ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ –  
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ  
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ – ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

## Διπλωματική Εργασία

# Διερεύνηση μεθόδων εκτίμησης επιπέδων οργανικού άνθρακα στο έδαφος από τηλεπισκοπικά πολυφασματικά δεδομένα

Τριμελής εξεταστική επιτροπή

Κ. Καράντζαλος

Α. Δουλάμης

Ι. Παπουτσής

.....

.....

.....

Καθηγητής ΕΜΠ

Αν. Καθηγητής ΕΜΠ

Επικ. Καθηγητής ΕΜΠ

*Επιβλέπων*

Αθήνα, Οκτώβριος 2023



# Ευχαριστίες

Θα ήθελα να ευχαριστήσω, τον κ. Κων/νο Καράντζαλο, Καθηγητή Ε.Μ.Π., για την καθοδήγησή του σε όλα τα στάδια εκπόνησης της παρούσας εργασίας.

Ακόμα, θα ήθελα να ευχαριστήσω τον Αλέξανδρο Φάλαγγα, την Ελένη Σοφικίτη και τη Χριστίνα Καρακίζη από το εργαστήριο τηλεπισκόπησης της ΣΑΤΜ ΕΜΠ, για την πολύτιμη συμβολή τους στη διεκπεραίωση της εργασίας, καθώς και για την προθυμία που έδειχναν, κάθε φορά που χρειάστηκαν τη βοήθεια τους.

Τέλος, ευχαριστώ την οικογένεια και τους φίλους μου, για την υποστήριξή τους, καθ' όλη τη διάρκεια εκπόνησης της παρούσας διπλωματικής εργασίας.

# Περίληψη

Η μελέτη και παρακολούθηση του οργανικού άνθρακα στο έδαφος είναι καθοριστική για την προστασία του περιβάλλοντος, τη διατήρηση της αειφορίας, την ενίσχυση της γεωργικής παραγωγής και την αντιμετώπιση της κλιματικής αλλαγής. Η γεωργία του άνθρακα έχει ως στόχο να συμβάλλει στη βελτιστοποίηση της δέσμευσης του άνθρακα από τις καλλιέργειες. Η παρούσα εργασία εξετάζει τις δυνατότητες των μεθόδων εκτίμησης των επιπέδων του οργανικού άνθρακα, με τη χρήση πολυφασματικών εικόνων. Η μέθοδος που εφαρμόστηκε, βασίζεται στο Σύστημα Γεωχωρικής Ανίχνευσης Εδάφους (GEOS3) για τον εντοπισμό των περιοχών γυμνού εδάφους σε χρονοσειρές εικόνων του δορυφόρου Landsat 5. Δημιουργήθηκαν συνθετικές εικόνες εδάφους και ερευνήθηκε ο βαθμός συσχέτισης της ανακλώμενης ακτινοβολίας, με τις τιμές του οργανικού άνθρακα, με τη χρήση τηλεπισκοπικών δεικτών, όπως δείκτες βλάστησης και εδάφους. Η μελέτη εφαρμόστηκε σε γεωργική έκταση στην περιοχή του Ξηρόμερου Αιτωλοακαρνανίας. Χρησιμοποιήθηκαν δείγματα εδάφους από την εδαφολογική μελέτη της περιοχής Ξηρομέρου Ν. Αιτωλοακαρνανίας, του έτους 1995, του Ινστιτούτου Χαρτογράφησης και Ταξινόμησης Εδαφών Λάρισας, του Εθνικού Ιδρύματος Αγροτικής Έρευνας. Για την αναζήτηση της συσχέτισης μεταξύ των μεταβλητών χρησιμοποιήθηκαν η γραμμική παλινδρόμηση και ο αλγόριθμος των τυχαίων δασών (Random Forest).



# Abstract

Studying and monitoring soil organic carbon is crucial to protect the environment, maintain sustainability, enhance agricultural production and addressing climate change. Carbon farming aims to contribute to the optimization of carbon dioxide sequestration from crops. This paper examines the potential of remote sensing methods for estimating organic carbon levels using multispectral imagery. The method applied is based on the Geospatial Earth Observation System (GEOS3) to identify bare soil areas in image time-series of Landsat 5 satellite. Bare soil composites were generated and the level of correlation of reflected radiation, with organic carbon values, was investigated using remote sensing indices such as vegetation and soil indices. The study was applied to an agricultural area in the region of Xiromero, Aitolokarnania, Greece. Soil samples from the soil study of the area of Xiromero, Aitolokarnania, of the year 1995, by the Institute of Soil Mapping and Classification of Larissa, National Agricultural Research Foundation, were used. Linear regression and Random Forest algorithm were used to evaluation of the correlation between variables.

# Περιεχόμενα

Ευχαριστίες.....	i
Περίληψη.....	ii
Abstract .....	iii
Κεφάλαιο 1 .....	3
1.1. Εισαγωγή .....	3
1.2. Αντικείμενο και Στόχοι .....	4
Κεφάλαιο 2 - Θεωρητικό Υπόβαθρο .....	4
2.1. Ροές άνθρακα σε χερσαίο οικοσύστημα.....	4
2.1.1. Οργανικός άνθρακας.....	5
2.1.2. Υπολογισμός οργανικού άνθρακα στο έδαφος .....	6
2.2. Προγενέστερες εργασίες.....	7
2.2.1. Δεδομένα.....	8
2.2.1.1. Υπερφασματικά δεδομένα.....	8
2.2.1.2. Πολυφασματικές εικόνες.....	9
Sentinel 2.....	9
Landsat 8 .....	9
Landsat 5 .....	9
2.2.1.3. Δείγματα εδάφους .....	10
2.2.1.4. Το πρόγραμμα LUCAS .....	11
2.2.2. Μέθοδοι εκτίμησης του οργανικού άνθρακα.....	11
2.2.2.1. Το μοντέλο SAFY.....	11
2.2.2.2. Μέθοδοι παλινδρόμησης.....	14
Γραμμική παλινδρόμηση.....	14
PLSR .....	15
ELM.....	16

SVMR .....	16
Τυχαία Δάση (Random Forest) .....	17
Gradient Boosting Regressor .....	18
Δέντρο απόφασης .....	18
Κεφάλαιο 3 – Μεθοδολογία .....	21
3.1. Η περιοχή μελέτης.....	21
3.2. Δεδομένα.....	22
3.3. Μεθοδολογία .....	23
3.3.1. Google Earth Engine .....	23
3.3.2. Δείκτες .....	28
3.3.3. Μέθοδοι εκτίμησης SOC .....	29
Κεφάλαιο 4 – Αποτελέσματα και αξιολόγηση .....	30
4.1. Αποτελέσματα ανίχνευσης περιοχών γυμνου εδάφους.....	30
4.2. Μοντέλα εκτίμησης SOC .....	39
4.2.1 Γραμμική παλινδρόμηση.....	39
4.2.2 Αποτελέσματα αλγορίθμου τυχαίων δασών (Random Forest) .....	49
Κεφάλαιο 5 – Συμπεράσματα και προτάσεις.....	53
5.1. Συμπεράσματα .....	53
5.2. Προτάσεις για μελλοντική έρευνα .....	54
Βιβλιογραφία .....	55
Διαγράμματα.....	58
Παράρτημα εικόνων.....	59
Παράρτημα πινάκων .....	60

# Κεφάλαιο 1

## 1.1. Εισαγωγή

Σε παγκόσμιο επίπεδο, η ποσότητα του άνθρακα που αποθηκεύεται στο έδαφος είναι η δεύτερη μεγαλύτερη μετά από εκείνη του ωκεανού (38.400 Gt). Η χερσαία βιοτική δεξαμενή άνθρακα είναι περίπου 560 Gt οργανικού άνθρακα, ενώ η δεξαμενή άνθρακα του εδάφους υπολογίζεται ως υπερτετραπλάσια του αριθμού αυτού. Η χωρητικότητα της δεξαμενής οργανικού άνθρακα των παγκόσμιων εδαφών έχει εκτιμηθεί ποικιλοτρόπως για τα κύρια οικοσυστήματα (Stockmann et al., 2013). Για παράδειγμα, περίπου 2344 Gt του οργανικού άνθρακα αποθηκεύονται στα τρία πρώτα μέτρα του εδάφους, με το 54% ή 1500 Gt του οργανικού άνθρακα (C) να αποθηκεύεται στο πρώτο μέτρο του εδάφους και περίπου 615 Gt αποθηκεύονται στα πρώτα 20 cm (Jobbágy and Jackson, 2000, Guo and Gifford, 2002). Αυτό έρχεται σε αντίθεση με την προσθήκη ~9 Gt ανθρωπογενώς απελευθερωμένου "νέου" C που προστίθεται στην ατμόσφαιρα ετησίως από ορυκτές πηγές C (άνθρακας, πετρέλαιο και φυσικό αέριο) και μέσω της υποβάθμισης των οικοσυστημάτων (Stockmann et al., 2013).

Γενικά, η φυτική παραγωγή και τα πρότυπα κατανομής της βιομάζας επηρεάζουν έντονα τη σχετική κατανομή του C σε σχέση με το βάθος του εδάφους (Jobbágy and Jackson, 2000), που σημαίνει ότι όσο μεγαλύτερα βάθη είναι παλαιότερα αποθηκευμένος ο άνθρακας.

Η γεωργία άνθρακα είναι μια ολόπλευρη προσέγγιση γεωργικής εκμετάλλευσης για τη βελτιστοποίηση της δέσμευσης του άνθρακα από τις καλλιέργειες με την εφαρμογή πρακτικών που είναι γνωστό ότι βελτιώνουν το ρυθμό με τον οποίο το CO<sub>2</sub> απομακρύνεται από την ατμόσφαιρα και αποθηκεύεται στο ως οργανική ύλη στο έδαφος. Με δεδομένο ότι η γεωργία είναι ένας από τις κύριες πηγές εκπομπής αερίων του θερμοκηπίου, η εφαρμογή βιώσιμων πρακτικών και τεχνολογιών μπορεί να έχει θετική συμβολή στην κλιματική αλλαγή.

Η Ευρωπαϊκή Ένωση δίνει ιδιαίτερη βαρύτητα στον τομέα της γης, καθώς τον θεωρεί το κλειδί για την επίτευξη μιας κλιματικά ουδέτερης οικονομίας. Για το λόγο αυτό, η ευρωπαϊκή πράσινη συμφωνία, η οποία δίνει κατεύθυνση στην δημιουργία κινήτρων για υιοθέτηση πρακτικών φιλικών προς το κλίμα, προωθεί στα κράτη μέλη τη γεωργία άνθρακα. Το 2021 εκδόθηκε η ανακοίνωση για τους βιώσιμους κύκλους του άνθρακα, η οποία καθορίζει δράσεις που αποσκοπούν στην αντιμετώπιση των σημερινών προκλήσεων για τη γεωργία άνθρακα, προκειμένου να αναβαθμιστεί αυτό το πράσινο επιχειρηματικό μοντέλο που ανταμείβει τους διαχειριστές γης για την υιοθέτηση πρακτικών που οδηγούν στη δέσμευση άνθρακα, σε συνδυασμό με ισχυρά οφέλη για τη βιοποικιλότητα.

Τέτοιες πρακτικές είναι η αναδάσωση που σέβεται τη βιοποικιλότητα και στοχεύει στην ενίσχυση της αειφορικής διαχείρισης των δασών και την προσαρμογή τους στην κλιματική αλλαγή, μορφές μικτής γεωργίας όπως η αγροδασοπονία, χρήση ενδιάμεσων καλλιεργειών και εναλλαγή καλλιεργειών κάλυψης, κατεργασίας εδάφους και βελτίωσης των χαρακτηριστικών του εδάφους, η στοχευμένη μετατροπή εκτάσεων από καλλιέργεια σε αγρανάπαυση, ή μη καλλιεργούμενων σε βοσκοτόπους και η προστασία και αποκατάσταση τυρφώνων και υγροτόπων με σκοπό τη μείωση της οξείδωσης του υπάρχοντος αποθέματος άνθρακα και αυξάνει τη δυνατότητα δέσμευσης άνθρακα.

## 1.2. Αντικείμενο και Στόχοι

Αντικείμενο της παρούσας εργασίας είναι αξιοποίηση μεθόδων της τηλεπισκόπησης για την ανίχνευση των επιπέδων του οργανικού άνθρακα του εδάφους με τη χρήση πολυφασματικών εικόνων, με σκοπό την εφαρμογή της συγκεκριμένης τεχνικής σε μεγάλες γεωργικές εκτάσεις. Η εκτίμηση του οργανικού άνθρακα του εδάφους με τηλεπισκοπικές μεθόδους είναι καθοριστική για τη διαχείριση των καλλιεργειών στην κατεύθυνση της γεωργίας του άνθρακα και κατ' επέκταση για την επίτευξη μιας κλιματικά ουδέτερης οικονομίας.

# Κεφάλαιο 2 - Θεωρητικό Υπόβαθρο

## 2.1. Ροές άνθρακα σε χερσαίο οικοσύστημα

Η εισροή του άνθρακα στα οικοσυστήματα πραγματοποιείται με δέσμευση του διοξειδίου του άνθρακα από τα φυτά, μέσω της διαδικασίας της φωτοσύνθεσης. Η ποσότητα αυτή, στην κλίμακα των οικοσυστημάτων ονομάζεται ακαθάριστη πρωτογενής παραγωγή (Gross Primary Production, GPP). Η διαπνοή των φυτών, επιστρέφει στην ατμόσφαιρα μέρος της GPP. Η ποσότητα του άνθρακα που προκύπτει από τη διαφορά της αναπνοής των φυτών από την GPP ονομάζεται καθαρή πρωτογενής παραγωγή (NPP).

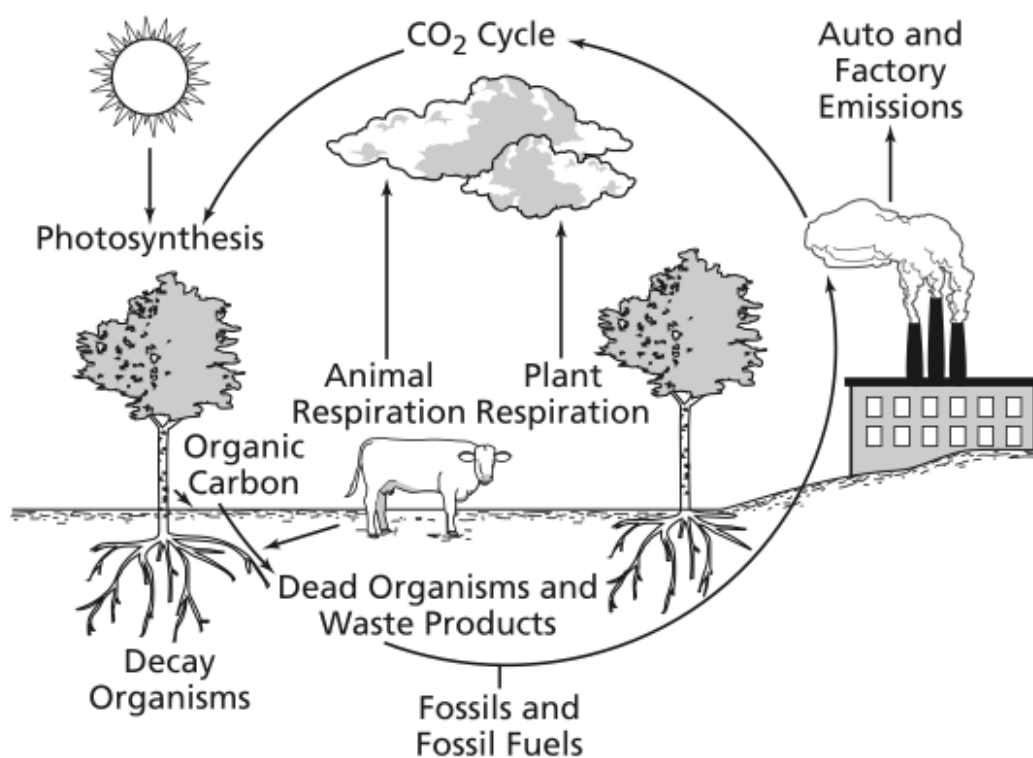
Η διαπνοή των φυτών παρέχει την ενέργεια για την απορρόφηση των θρεπτικών συστατικών και την παραγωγή και διατήρηση της βιομάζας. Τα περισσότερα οικοσυστήματα παρουσιάζουν παρόμοια αποτελεσματικότητα στην μετατροπή της GPP σε NPP, με την ποσότητα της NPP να υπολογίζεται στο μισό της GPP. (Charin et al. 2012)

Η μεταφορά άνθρακα από τα φυτά στο έδαφος, μέσω των υπολειμμάτων τους, των εκκρίσεων ενζύμων των ριζών, των φύλλων και των βλαστών, καθώς και υπολειμμάτων σιτηρών από τις καλλιέργειες αποτελεί τον κυριότερο παράγοντα για τη δημιουργία της οργανικής ουσίας του εδάφους (SOM). Άλλες σημαντικές πηγές είναι τα λιπάσματα, η κοπριά από τα ζώα, οργανικά υλικά που μετατρέπονται σε χούμο. Η περιεκτικότητα του εδάφους σε οργανική ουσία εξαρτάται από παράγοντες όπως: η φυτοκάλυψη του, η εκμετάλλευση του εδάφους, η περιοχή, το κλίμα και η υγρασία. Για παράδειγμα οι δασικές περιοχές είναι πλουσιότερες σε οργανική ουσία από τις καλλιεργούμενες εκτάσεις. Περισσότερη οργανική ουσία παρατηρείται, επίσης, στα ψυχρότερα και πιο υγρά κλίματα, όπως στα εδάφη της Βόρειας και Δυτικής Ευρώπης, σε σύγκριση με τις περιοχές της Μεσογείου. Οι εκτάσεις με φυσική βλάστηση και τα εδάφη που συγκρατούν περισσότερη υγρασία παρουσιάζουν υψηλότερη συγκέντρωση οργανικής ουσίας. Σημαντική παράμετρος είναι και η σύσταση του εδάφους. Συγκεκριμένα τα πιο βαριά εδάφη με υψηλή περιεκτικότητα σε άργιλο, έχουν υψηλότερη περιεκτικότητα από ό,τι τα αμμώδη, τα οποία είναι ελαφρύτερα.

Η περιεκτικότητα σε οργανική ουσία εξαρτάται και από εξωτερικούς παράγοντες, οι οποίοι έχουν τη δυνατότητα να τη μεταβάλλουν. Ο τρόπος εκμετάλλευσης του εδάφους, όπως το σύστημα καλλιέργειας που εφαρμόζεται, μπορεί να εμπλουτίσει ή να μειώσει την οργανική ουσία στο έδαφος.

Βάσει της περιεκτικότητας τους σε οργανική ουσία, τα εδάφη κατατάσσονται σε 5 κατηγορίες: χαμηλό για περιεκτικότητα μικρότερη από 1%, μέτρια χαμηλό για 1% έως 2%, μέτριο για 2% έως 4%, υψηλό για 4% έως 8% και πολύ υψηλό για 8% έως 16%. Οι καλλιεργούμενες εκτάσεις στην Ελλάδα έχουν κατά κύριο λόγο χαμηλή συγκέντρωση οργανικής ουσίας, καθώς το ποσοστό της είναι κατά μέσο όρο 1 έως 2,5%. (Σινάνης, 2022)

Η σημασία της έγκειται στην πολλαπλή επίδραση που έχει στο έδαφος, συμβάλλοντας στην ανάπτυξη των φυτών, την αποσάθρωση των ορυκτών, δέσμευση βαρέων μετάλλων, την ανάπτυξη των μικροοργανισμών, καθώς και την ικανότητα του εδάφους να συγκρατεί το νερό και να διατηρεί σταθερή τη θερμοκρασία του. Κατά την αποσύνθεση της, η νεκρή οργανική ουσία μετατρέπεται σε CO<sub>2</sub>, ανόργανα στοιχεία και χούμο. Η μετατροπή σε CO<sub>2</sub> αποτελεί τον κυριότερο τρόπο διαρροής άνθρακα από τα οικοσυστήματα. Επειδή η άμεση μέτρηση της οργανικής ύλης δεν είναι εύκολη διαδικασία, ο προσδιορισμός της γίνεται μέσω της μέτρησης του οργανικού άνθρακα στο έδαφος, ο οποίος αποτελεί το κύριο συστατικό της οργανικής ουσίας του εδάφους, εξαιρετικά σημαντικό για όλες τις εδαφικές διεργασίες.



Εικόνα 1. Ροές άνθρακα σε χερσαίο οικοσύστημα

### 2.1.1. Οργανικός άνθρακας

Ο οργανικός άνθρακας του εδάφους (SOC) είναι το μεγαλύτερο απόθεμα άνθρακα (C) στα περισσότερα χερσαία οικοσυστήματα (Lal, 2008a), το οποίο περιέχει περίπου 2344 Gt οργανικού C παγκοσμίως (Stockmann et al., 2013). Επιπλέον, το έδαφος αναγνωρίζεται ως η δεύτερη μεγαλύτερη δεξαμενή C μετά τους ωκεανούς και ένα από τα σημαντικότερα συστατικά της βιόσφαιρας, παρέχοντας σημαντικές υπηρεσίες και λειτουργίες του οικοσυστήματος (Ogle & Paustian, 2005). Τα πλαίσια πολιτικής που επηρεάζουν τις χρήσεις γης και τις αλλαγές στις χρήσεις γης θα μπορούσαν να προκαλέσουν δραματικές αλλαγές στα

επίπεδα SOC. Σε παγκόσμια κλίμακα, οι κατευθυντήριες γραμμές για την ποσοτικοποίηση των αλλαγών στις εκπομπές αερίων του θερμοκηπίου από τα γεωργικά εδάφη, τις αλλαγές στις χρήσεις γης και τις δασικές κατηγορίες έχουν υιοθετηθεί στις δραστηριότητες "Χρήση γης, αλλαγή χρήσης γης και δασοκομία" (LULUCF), οι οποίες απαριθμούνται στα άρθρα 3.3 και 3.4 του Πρωτοκόλλου του Κιότο (UN, 1997). Μετά την απόφαση LULUCF 529/2013, η λογιστική καταγραφή της διαχείρισης των καλλιεργούμενων και βοσκήσιμων εκτάσεων έχει καταστεί υποχρεωτική για τα κράτη μέλη της ΕΕ.

Ο οργανικός άνθρακας του εδάφους (SOC) είναι το βασικό στοιχείο που καθορίζει την ποιότητα του εδάφους, τη γονιμότητα, τη γεωργική κερδοφορία και τη δέσμευση του διοξειδίου του άνθρακα (CO<sub>2</sub>) στην ατμόσφαιρα. Το SOC επηρεάζει τις φυσικοχημικές και βιολογικές ιδιότητες του εδάφους, γεγονός που βελτιώνει ταυτόχρονα τη δομή του εδάφους, την ικανότητα συγκράτησης νερού και θρεπτικών συστατικών.

### 2.1.2. Υπολογισμός οργανικού άνθρακα στο έδαφος

Η ποσότητα του οργανικού άνθρακα μπορεί να υπολογιστεί άμεσα με εργαστηριακές μεθόδους. Η ακριβέστερη μέθοδος για τον υπολογισμό του οργανικού άνθρακα στο έδαφος είναι η συλλογή δειγμάτων εδάφους με επιτόπιο έλεγχο και προσδιορισμό της περιεκτικότητας τους σε C στο εργαστήριο. Η συγκεκριμένη διαδικασία επιτυγχάνεται με δύο μεθόδους: της ξηρής καύσης και της υγρής οξειδωσης. Κατά την ξηρή καύση, το δείγμα αρχικά ξηραίνεται και στη συνέχεια τοποθετείται σε φούρνο αντίστασης, σκοπό την καύση των οργανικών ουσιών που περιέχει. Με αυτή τη διαδικασία, ο άνθρακας μετατρέπεται σε διοξείδιο του άνθρακα (CO<sub>2</sub>), το οποίο παγιδεύεται και προσδιορίζεται ογκομετρικά και βαρυμετρικά. Τα κύρια πλεονεκτήματα της μεθόδου είναι η υψηλή ακρίβεια, ταχύτητα εκτέλεσης και αξιοπιστία, καθώς και το γεγονός ότι δεν παράγει επικίνδυνα απόβλητα. Στα μειονεκτήματα της συγκαταλέγονται: το υψηλό κόστος, καθώς απαιτεί εργαστήριο και υψηλό επίπεδο εκπαίδευσης του μελετητή και η αδυναμία χρήσης της μεθόδου για μεγάλες εκτάσεις. Η μέθοδος που χρησιμοποιείται συχνότερα είναι η μέθοδος της υγρής οξειδωσης - Walkley Black, η οποία βασίζεται στην οξείδωση του C με τη χρήση διχρωμικού καλίου (K<sub>2</sub>Cr<sub>2</sub>O<sub>7</sub>) σε θειικό οξύ (H<sub>2</sub>SO<sub>4</sub>) και θειικό σίδηρο (FeSO<sub>4</sub>). Ο οργανικός άνθρακας προσδιορίζεται έμμεσα με δύο διαδοχικές αντιδράσεις οξειδοαναγωγής. Η υγρή οξείδωση έχει χαμηλό κόστος και ευκολία στην εκτέλεση, διότι απαιτεί ελάχιστο και απλό εξοπλισμό, όμως επιτυγχάνει χαμηλότερη ακρίβεια, καθώς παρουσιάζει μικρότερη ευαισθησία στον εντοπισμό μικρών αλλαγών και κατά την εκτέλεση της, χρησιμοποιεί και παράγει επικίνδυνες χημικές ουσίες.

Η εκτίμηση του SOC με συμβατικές τεχνικές πεδίου είναι μια αρκετά χρονοβόρα και κοστοβόρα διαδικασία, η οποία παρέχει δεδομένα που καλύπτουν πολύ περιορισμένες εκτάσεις. Λόγω των δυσκολιών που παρουσιάζουν οι επίγειες μέθοδοι, δεν συνηθίζεται η κάλυψη μεγάλων εκτάσεων από τέτοιου είδους εδαφολογικές μελέτες, παρά μόνο η λήψη μικρού αριθμού δειγμάτων για τις ανάγκες της εκάστοτε έρευνας. Η χωρική μεταβολή του SOC καθιστά αναξιόπιστη τη γενίκευση των αποτελεσμάτων των εργαστηριακών μελετών χωρίς επιβεβαίωση τους από επιπλέον δεδομένα.

Οι πλατφόρμες τηλεπισκόπησης (RS), όπως οι εναέρια και οι δορυφορικές, χρησιμοποιούνται ευρέως για την εκτίμηση του SOC σε διαφορετικές φασματικές αναλύσεις, αλλά

δυσκολεύονται στη διάκριση του SOC από την επιφάνεια του εδάφους και της βλάστησης. Επίσης, παρατηρούνται εμπόδια στην εκτίμηση του SOC όταν η συγκέντρωση του SOC είναι μικρή στο έδαφος.

## 2.2. Προγενέστερες εργασίες

Οι Ben-Dor et al., 1997 και Nocita et al., 2015, επεσήμαναν ότι τα σημαντικότερα χαρακτηριστικά που συνδέονται με το SOC, εντοπίζονται στις φασματικές περιοχές 450, 590 και 664 nm, οι οποίες για τον Landsat 5 βρίσκονται εντός των καναλιών του ορατού φάσματος που έχουν κέντρο 485, 560 και 660 nm για τα bands 1, 2, 3, αντίστοιχα. Επίσης χαρακτηριστικά του εδάφους ανιχνεύονται στις περιοχές SWIR, από 1600 έως 1900 nm που για τον Landsat 5 καλύπτεται μερικώς από το εύρος του band 5 (SWIR1) και στην περιοχή 2100 έως 2300 nm που καλύπτεται εξ ολοκλήρου από το Band 7 (SWIR 2).

Οι Gholizadeh et al., (2018) πραγματοποίησαν μια μελέτη σε 4 περιοχές με καλλιέργειες στην Τσεχία, από τις οποίες συνέλεξαν και επεξεργάστηκαν 200 δείγματα εδάφους (0-10cm) και σύγκριναν την περιεκτικότητά τους σε άνθρακα με τα δεδομένα από φασματοραδιόμετρο ASD FieldSpec, αερομεταφερόμενο υπερφασματικούς δέκτες (CASI/SASI sensors) και πολυφασματικές εικόνες Sentinel-2. Χρησιμοποίησαν 18 φασματικούς δείκτες για την ανίχνευση του SOC και της υφής του εδάφους (περιεκτικότητα σε άργιλο, ιλύ και άμμο) και εφάρμοσαν τη μέθοδο SVMR για τη σύγκριση των τιμών των δεδομένων τους. Συγκεκριμένα οι δείκτες που χρησιμοποιήθηκαν ήταν ο δείκτης κανονικοποιημένης διαφοράς βλάστησης (NDVI), ο μετασχηματισμένος δείκτης βλάστησης (TVI), ο ενισχυμένος δείκτης βλάστησης (EVI), ο προσαρμοσμένος στο έδαφος δείκτης ολικής βλάστησης (SATVI), ο προσαρμοσμένος στο έδαφος δείκτης βλάστησης (SAVI), ο δείκτης καταπόνησης υγρασίας (MSI), ο δείκτης πράσινης κανονικοποιημένης διαφοράς βλάστησης (GNDVI), ο δείκτης πράσινης-κόκκινης βλάστησης (GRVI), Δείκτης επιφανειακών υδάτων (LSWI), μετασχηματισμένος δείκτης βλάστησης προσαρμοσμένος στο έδαφος (TSAVI), τροποποιημένος δείκτης βλάστησης προσαρμοσμένος στο έδαφος (MSAVI), ο δεύτερος τροποποιημένος δείκτης βλάστησης προσαρμοσμένος στο έδαφος (MSAVI2), σταθμισμένος δείκτης βλάστησης (WDVI), δείκτης φωτεινότητας (BI), ο δεύτερος δείκτης φωτεινότητας (BI2), δείκτης ερυθρότητας (RI), δείκτης χρώματος (CI) και βλάστησης (V). Οι καλύτερες συσχετίσεις SOC και φασματικών ζωνών του Sentinel-2 προέκυψαν από τις περιοχές B4 (κόκκινο) και B5 (NIR), ακολουθούμενες από τις περιοχές B11 και B12 (SWIR). Μεταξύ όλων των φασματικών δεικτών, οι δείκτες BI, BI2, GNDVI και SATVI παρείχαν τις ισχυρότερες συσχετίσεις με το SOC. Η εν λόγω έρευνα έδειξε ότι ο αισθητήρας Sentinel-2 μπορεί να χρησιμοποιηθεί για την εκτίμηση και χαρτογράφηση του SOC σε μεγάλες γεωγραφικές περιοχές. Οι χάρτες SOC που δημιουργήθηκαν με τη χρήση του Sentinel-2 ήταν πιο ακριβείς σε σχέση με τους αερομεταφερόμενους αισθητήρες. Ωστόσο, υπήρξε μείωση της ακρίβειας του μοντέλου και του χάρτη για περιοχές με χαμηλές περιεκτικότητες SOC. Τα αποτελέσματα της μελέτης υπογράμμισαν τη σημασία των υπερφασματικών δεδομένων Sentinel-2 για την αξιολόγηση των χαρακτηριστικών του εδάφους, που μπορεί να γίνει με συχνότερη επανεξέταση και σε μεγαλύτερη κλίμακα από τα εργαστηριακά και αερομεταφερόμενα όργανα που χρησιμοποιούνται σήμερα.

Οι Long Guo et al., (2021) υπολόγισαν την περιεκτικότητα του εδάφους σε οργανικό άνθρακα, σε αγροτικές εκτάσεις 385,45 ha, νοτιοδυτικά της Αϊόβα (Iowa), ΗΠΑ. Για τη μελέτη τους



χρησιμοποίησαν 195 δείγματα εδάφους (0-15 cm), τα οποία σύγκριναν με υπερφασματικές εικόνες (Headwall-Hyperspec, 380–1700 nm) και χρονοσειρές από πολυφασματικές εικόνες Sentinel-2 και Landsat-8. Υπολόγισαν αρχικά το δείκτη NDVI, τον οποίο στη συνέχεια αντικατέστησαν με τους GNDVI και WNDVI. Πρόκειται για δείκτες παρόμοιους με τον NDVI, οι οποίοι προσπαθούν να διορθώσουν τον ασυμπτωτικό κορεσμό που δημιουργείται από το συνδυασμό του κόκκινου με το εγγύς υπέρυθρο, ο GNDVI υπολογίζεται με την αντικατάσταση του κόκκινου καναλιού από το πράσινο:  $GNDVI = \frac{NIR-Green}{NIR+Green}$  και ο WNDVI με την προσθήκη ενός συντελεστή στάθμισης  $\alpha = 0,2$  στην εξίσωση του NDVI,  $WNDVI = \frac{\alpha * NIR-Red}{\alpha * NIR+Red}$ . Οι συγκρίσεις πραγματοποιήθηκαν με παλινδρόμηση μερικών ελαχίστων τετραγώνων (PLSR) και ακραίας μηχανικής μάθησης (ELM). Τα αποτελέσματα έδειξαν ότι ο οργανικός άνθρακας στο έδαφος μπορεί να ανιχνευτεί μέσω των υπερφασματικών εικόνων με RMSE=0,18, R<sup>2</sup>=0,61, RPIQ=1,83, για το PLSR και RMSE=0,16, R<sup>2</sup>=0,72, RPIQ=2,03 για το ELM. Τα μοντέλα που δημιουργήθηκαν από τις εικόνες Sentinel-2 και Landsat-8 έχουν πιο περιορισμένες δυνατότητες, με RPIQ=1,45 για τον S-2 και 1,19 για τον L-8, R<sup>2</sup>= 0,42 για τον S-2 και 0,35 για τον L-8, και RMSE=0,23 για τον S-2 και 0,28 για τον L-8, αντίστοιχα. Επισημαίνουν, ωστόσο ότι οι χάρτες που δημιουργήθηκαν από τα πολυφασματικά δεδομένα, παρουσιάζουν παρόμοια χαρακτηριστικά χωρικής κατανομής με τα υπερφασματικά, εκτός από τις περιοχές στα άκρα των γεωργικών εκτάσεων, λόγω των μεικτών εικονοστοιχείων. Τα αθροιστικά ποσοστιαία σφάλματα δείχνουν ότι η πλειονότητα των διαφορών είναι κάτω του 20% για το SOC (78,40%, 76,71%), στις εικόνες Sentinel 2 και Landsat 8. Επομένως, οι χρονοσειρές των εικόνων Sentinel-2 και Landsat 8 μπορούν να παρέχουν πολύτιμες πληροφορίες για την απεικόνιση των χαρακτηριστικών χωρικής κατανομής των εδαφικών ιδιοτήτων και την καθοδήγηση της γεωργικής παραγωγή.

Οι Francis et al., (2020) εκτιμούν το SOC στην περιοχή Wayanad της Kerala (Ινδία) χρησιμοποιώντας δορυφορικά δεδομένα RS που λαμβάνονται από τον Landsat 8. Εφαρμόζοντας τα δεδομένα RS που αποκτήθηκαν, υπολογίστηκαν διάφοροι δείκτες όπως ο NDVI και ο δείκτης γυμνού εδάφους (BSI) για τον προσδιορισμό του SOC στην περιοχή μελέτης. Ομοίως, πολλοί πολυφασματικοί και υπερφασματικοί αισθητήρες χρησιμοποιούνται για την εκτίμηση του SOC και την παραγωγή εδαφικών χαρτών υψηλής ποιότητας. Τα δεδομένα που ελήφθησαν από διάφορες διαστημικές πλατφόρμες όπως οι αισθητήρες Hyperion, Landsat, EnMAP και Sentinel 2 εφαρμόστηκαν με μοντέλα SVM, μερικής παλινδρόμησης ελαχίστων τετραγώνων (PLSR) και τυχαίου δάσους (RF) για την εκτίμηση της μεταβλητότητας του SOC σε τοπικό και περιφερειακό επίπεδο (Gomez et al., 2008- Mirzaee et al., 2016- Castaldi et al., 2019).

## 2.2.1. Δεδομένα

### 2.2.1.1. Υπερφασματικά δεδομένα

Για τη μελέτη του SOC, πολλές μελέτες χρησιμοποιούν υπερφασματικούς δείκτες, όπως ο Airborne Prism Experiment (APEX) και τα φασματόμετρα απεικόνισης Micro-Hyperspec A-Series και X-Series των αερομεταφερόμενων αισθητήρων Headwall-Hyperspec. Ο APEX αποτελεί ένα αερομεταφερόμενο υπερφασματικό φασματόμετρο τύπου push-broom, το οποίο καταγράφει δεδομένα σε 313 κανάλια μεταξύ 400nm και 2500nm και είναι μέρος του

προγράμματος Copernicus της ESA. Τα φασματόμετρα απεικόνισης Micro-Hyperspec A-Series και X-Series καλύπτουν το ορατό και το εγγύς υπέρυθρο τμήμα του οπτικού φάσματος, συγκεκριμένα τις φασματικές περιοχές 380-1000 nm (325 κανάλια) και 900-1700nm (67 κανάλια) για τα A- και X-series spectrometers και φασματική ανάλυση 1.9 και 12.9nm αντίστοιχα.

#### 2.2.1.2. Πολυφασματικές εικόνες

Στη βιβλιογραφία, έχει επιχειρηθεί η ανίχνευση του οργανικού άνθρακα στο έδαφος με τη χρήση πολυφασματικών εικόνων. Πρόκειται για εύκολα προσβάσιμα δεδομένα, με παγκόσμια κάλυψη και υψηλή επαναληπτικότητα. Οι δορυφόροι που χρησιμοποιούνται συχνότερα είναι ο Sentinel 2 και οι δορυφόροι της σειράς Landsat.

#### Sentinel 2

Η αποστολή Sentinel 2 του ευρωπαϊκού προγράμματος Copernicus αποτελείται από έναν αστερισμό δύο δορυφόρων, πανομοιότυπων μεταξύ τους, οι οποίοι βρίσκονται στην ίδια ηλιοσύγχρονη τροχιά με διαφορά φάσης 180°. Οι δέκτες τους συλλέγουν δεδομένα μεταξύ 443 και 2190nm με σύντομη περίοδο επαναφοράς, 10 ημέρες στον ισημερινό με έναν δορυφόρο και 5 ημέρες με 2 δορυφόρους υπό συνθήκες χωρίς σύννεφα, που καταλήγει σε 2-3 ημέρες στα μεσαία γεωγραφικά πλάτη. Τα κανάλια B2 (490 nm), B3 (560 nm), B4 (665 nm) και B8 (842 nm) έχουν χωρική ανάλυση 10m, τα κανάλια B5 (705 nm), B6 (740 nm), B7 (783 nm), B8a (865 nm), B11 (1610 nm) και B12 (2190 nm) έχουν 20m. Τέλος, τα κανάλια διόρθωσης της ατμόσφαιρας B1, B9 και B10 έχουν 60 m χωρική ανάλυση.

#### Landsat 8

Ο Landsat 8 είναι σε τροχιά από το 2013 και παρέχει πολυφασματικά δεδομένα υψηλής ανάλυσης στο ορατό και το υπέρυθρο φάσμα. Ο δορυφόρος βρίσκεται σε ηλιοσύγχρονη τροχιά σε απόσταση 705km, εκτελεί κύκλο 233 τροχιών και επαναληψιμότητα κάθε 16 μέρες. Είναι εξοπλισμένος με δύο δέκτες, τον Operational Land Imager (OLI) και τον Thermal Infrared Sensor (TIRS). Ο πρώτος διαθέτει εννέα φασματικά κανάλια: 1) Coastal Aerosol (430 - 450 nm), 2) Blue (450 - 510 nm), 3) Green (530 - 590 nm), 4) Red (640 - 670 nm), 5) Near-Infrared (NIR) (850 - 880 nm), 6) Shortwave Infrared (SWIR) 1 (1570 - 1650 nm), 7) Shortwave Infrared (SWIR) 2 (2110 - 2290 nm), 8) Panchromatic (PAN) (500 - 680 nm), 9) Cirrus (1360 - 1380 nm). Όλα τα κανάλια έχουν χωρική ανάλυση 30m εκτός από το πανχρωματικό που έχει 15m. Ο θερμικός δέκτης περιλαμβάνει δύο κανάλια: 10) Thermal Infrared (TIR) 1 (10.6 - 11.19 μm), 11) Thermal Infrared (TIR) 2 (11.5 - 12.51 μm), με χωρική ανάλυση 100 m.

#### Landsat 5

Ο Landsat 5, προγενέστερος του Landsat 8, εκτοξεύτηκε από τη NASA την 1η Μαρτίου 1984 και παρέμεινε ενεργός για σχεδόν 29 χρόνια, μέχρι το 2013, γεγονός που τον καθιστά τον μακροβιότερο δορυφόρο στον κόσμο. Σχεδιάστηκε όπως και ο Landsat 4 και έφερε τα όργανα Multispectral Scanner (MSS) και Thematic Mapper (TM). Λόγω βλάβης στους πομπούς του

Landsat 5, σταμάτησε η αποστολή δεδομένων MSS, αρχικά για τις ΗΠΑ το 1992 και στη συνέχεια παγκόσμια, με αποτέλεσμα να απενεργοποιηθεί. Το όργανο MSS τέθηκε ξανά σε λειτουργία, μετά την αστοχία του αισθητήρα TM τον Νοέμβριο του 2011. Από τον Ιούνιο του 2012 έως τον Ιανουάριο του 2013 συλλέχθηκαν πάνω από 15.000 σκηνές MSS.

Πίνακας 1. Τα κανάλια του Landsat 5 TM

Landsat 4-5	Wavelength (micrometers)	Resolution (meters)
Band 1 - Blue	0.45 - 0.52	30
Band 2 - Green	0.52 - 0.60	30
Band 3 - Red	0.63 - 0.69	30
Band 4 - Near Infrared (NIR)	0.76 - 0.90	30
Band 5 - Shortwave Infrared (SWIR) 1	1.55 - 1.75	30
Band 6 - Thermal	10.40 - 12.50	120 (30)
Band 7 - Shortwave Infrared (SWIR) 2	2.08 - 2.35	30

Ο Landsat 5 συνέβαλε σημαντικά στην παρακολούθηση της επιφάνειας της Γης, του κλίματος και τη διαχείριση των φυσικών πόρων, καθώς παρείχε εικόνες που χρησιμοποιήθηκαν για παρακολούθηση αλλαγών στη βλάστηση και τη χρήση του εδάφους, των υδάτινων πόρων, των δασών, των καλλιεργήσιμων εκτάσεων. Επιπλέον, αξιοποιήθηκε για τη διαχείριση καταστροφών, παρέχοντας εικόνες που χρησιμοποιήθηκαν για την αξιολόγηση των συνεπειών των φυσικών καταστροφών, όπως πλημμύρες, πυρκαγιές και σεισμοί.

Οι εικόνες του Landsat 5 παραμένουν διαθέσιμες, παρέχοντας πολύτιμες πληροφορίες για την διεξαγωγή μελετών με τη σύνθεση χρονοσειρών για τα έτη, τα οποία ήταν σε λειτουργία.

Οι εικόνες Landsat 4-5 Thematic Mapper (TM) αποτελούνται από επτά φασματικές ζώνες με χωρική ανάλυση 30 μέτρων για τις ζώνες 1 έως 5 και 7. Η χωρική ανάλυση για τη ζώνη 6 (θερμικό υπέρυθρο) είναι 120 μέτρα, αλλά γίνεται επαναδειγματοληψία σε εικονοστοιχεία 30 μέτρων. Το κατά προσέγγιση μέγεθος της σκηνής είναι 170 χιλιόμετρα βόρεια-νότια επί 183 χιλιόμετρα ανατολικά-δυτικά (106 μίλια επί 114 μίλια).

### 2.2.1.3. Δείγματα εδάφους

Η πλειοψηφία των μελετών που επιχειρούν να συσχετίσουν τον οργανικό άνθρακα με την ανακλώμενη ακτινοβολία, όπως αυτή καταγράφεται σε εικόνες από δέκτες δορυφόρων ή αερομεταφερόμενους δέκτες, χρησιμοποιεί δείγματα εδάφους ως δεδομένα εκπαίδευσης και επαλήθευσης. Τα δείγματα αυτά συλλέγονται με επιτόπιους ελέγχους και κατά κύριο λόγο περιλαμβάνουν το επιφανειακό στρώμα του εδάφους (20 – 30 cm). Η ποσότητα του οργανικού άνθρακα προσδιορίζεται με εργαστηριακές μεθόδους, όπως περιγράφονται στο κεφάλαιο 2.1.2. Ορισμένα μοντέλα, όπως το SAFY απαιτούν τη χρήση μετεωρολογικών δεδομένων. Συνήθως χρησιμοποιούνται η θερμοκρασία στα 2m  $T_{air}$  και η εισερχόμενη παγκόσμια

ακτινοβολία  $R_e$ . Για τις μεθόδους που εφαρμόζονται στις επιστημονικές μελέτες, χρησιμοποιούνται, επιπλέον, η βροχόπτωση και οι τιμές εξατμισοδιαπνοής αναφοράς ETR ( $ET_0$ ). Οι εκπομπές  $CO_2$ , ροές υδρατμών, εξατμισοδιαπνοή και λανθάνουσα θερμότητα μετρώνται με τη μέθοδο eddy covariance, με τη χρήση επίγειων σταθμών συλλογής δεδομένων (flux towers). Τα δεδομένα του εδάφους τα οποία χρησιμοποιούνται από τα μοντέλα είναι η εδαφική υγρασία στη χωρητικότητα πεδίου ( $\theta_{fc}$ ) και το μόνιμο σημείο μάρανσης ( $\theta_{wp}$ ), τα οποία απαιτούνται για κάθε εξεταζόμενο πεδίο για την εκτίμηση της ημερήσιας διαθέσιμης περιεκτικότητας του εδάφους σε νερό (AWC).

#### 2.2.1.4. Το πρόγραμμα LUCAS

Το ευρωπαϊκό κέντρο δεδομένων εδάφους (ESDAC) του JRC της ΕΕ διεξάγει την έρευνα LUCAS (Land Use/Cover Area frame statistical Survey), η οποία αφορά την κάλυψη και τη χρήση γης και τη μελέτη των ιδιοτήτων του εδάφους σε 23 κράτη μέλη της ΕΕ. Από το 2009, έχει δημιουργήσει μια χωρική βάση δεδομένων με πληροφορία για ανώτερο στρώμα του εδάφους σε ολόκληρη την ΕΕ. Περιλαμβάνει σημαντικά δεδομένα για στοιχεία του εδάφους, όπως: pH (CaCl<sub>2</sub> and H<sub>2</sub>O), περιεκτικότητα οργανικού άνθρακα, ανθρακικού ασβεστίου, αζώτου, φωσφόρου, καλίου, ηλεκτρική αγωγιμότητα, σίδηρο και αλουμίνιο.

### 2.2.2. Μέθοδοι εκτίμησης του οργανικού άνθρακα

#### 2.2.2.1. Το μοντέλο SAFY

Οι εργασίες:

1. Combining High-Resolution Remote Sensing Products with a Crop Model to Estimate Carbon and Water Budget Components: Application to Sunflower.
2. Estimation of Crop Production and CO<sub>2</sub> Fluxes Using Remote Sensing: Application to a Winter Wheat/Sunflower Rotation
3. Estimation of daily CO<sub>2</sub> fluxes and of the components of the carbon budget for winter wheat by the assimilation of Sentinel 2 - like remote sensing data into a crop model
4. Application and Evaluation of a Simple Crop Modelling Framework: A Case Study for Spring Barley, Winter Wheat and Winter Oilseed Rape over Ireland

αφορούν την εκτίμηση ροών  $CO_2$  με επεξεργασία και παραμετροποίηση του απλού αλγορίθμου για την εκτίμηση της απόδοσης (SAFY), (πίνακας 2).

Η κύρια ιδέα είναι να χρησιμοποιηθεί το μοντέλο για να αναπαραστήσει γνωστές διαδικασίες σχετικές με την ανάπτυξη των καλλιεργειών, με την προϋπόθεση ότι μπορεί να γίνει προσομοίωση των διαδικασιών αυτών με τη χρήση σταθερών δεδομένων πχ. κλιματικών δεδομένων και εικόνων (δίνουν εκτιμήσεις του δείκτη φυλλικής επιφάνειας - LAI). Το μοντέλο προσομοιώνει την αύξηση της βιομάζας στην επιφάνεια (dry above-ground phytomass), βασισμένο στη θεωρία της αποδοτικότητας χρήσης φωτός (LUE) του Monteith (1977), με αναφορά της δυναμικής των πράσινων φύλλων και της επίδρασης της θερμοκρασίας. Αντίθετα, το νερό και η μεταφορά θρεπτικών στοιχείων (πχ άζωτο) μεταξύ του εδάφους και

του φυτού δεν υπολογίζονται άμεσα, αλλά μέσω της παρατήρησης των φύλλων με τη χρήση της παραμέτρου της αποτελεσματικής απόδοσης χρήσης φωτός.

Το αρχικό μοντέλο SAFY δίνει ως αποτέλεσμα τη βιομάζα στην επιφάνεια (AGB) και την εκτίμηση της παραγωγής (yield). Για την εκτίμηση των επιπέδων του άνθρακα είναι αναγκαίες ορισμένες προσαρμογές. Οι Pique et al. (2020) εξετάζουν τις εκπομπές CO<sub>2</sub> από καλλιέργειες χειμερινού σιταριού αναπτύσσοντας το μοντέλο SAFY-CO<sub>2</sub>, το οποίο προσομοιώνει την ημερήσια ανάπτυξη των καλλιεργειών (βιομάζα, ανάπτυξη των φύλλων, κλπ), τις εκπομπές CO<sub>2</sub> στο οικοσύστημα και την ετήσια παραγωγή και ισολογισμό άνθρακα στο οικοσύστημα (net ecosystem carbon budget (NECB)). Ενώ το SAFY υπολογίζει τη χρονική εξέλιξη του GAI, DAM, Yield, με δεδομένα την παγκόσμια εισερχόμενη ακτινοβολία (απορροφούμενη ακτινοβολία από το φυτό APAR) και τη μέση θερμοκρασία, το SAFY-CO<sub>2</sub> υπολογίζει αρχικά την ακαθάριστη πρωτογενή παραγωγή (GPP) ως συνάρτηση της απορροφούμενης ακτινοβολίας (APAR) και στη συνέχεια τη βιομάζα στο έδαφος και την επιφάνεια, καθώς και τις εκπομπές CO<sub>2</sub> και περιεκτικότητα σε άνθρακα.

Οι Pique et al. (2020) <sup>2</sup> ανέπτυξαν το μοντέλο SAFYE-CO<sub>2</sub> το οποίο χρησιμοποιεί τηλεπισκοπικά δεδομένα υψηλής χωρικής ανάλυσης (HSTR) για τον υπολογισμό σε ημερήσια βάση βιομάζας, ροών νερού και εκπομπών CO<sub>2</sub>, καθώς και την ετήσια απόδοση παραγωγής και τον ισολογισμό του άνθρακα σε μεγάλες εκτάσεις. Το μοντέλο σχεδιάστηκε από τον συνδυασμό του SAFY-CO<sub>2</sub> και της ενότητας του νερού από το SAFY-WB, που υπολογίζει την εξάτμιση (Ev) από το έδαφος, τη διαπνοή (TR) των φυτών των καλλιεργούμενων εκτάσεων και της περιεκτικότητας του εδάφους σε νερό (SWC). Από αυτά τα δεδομένα μπορούν να υπολογιστούν η υδατική καταπόνηση του εδάφους και οι ανάγκες των φυτών σε νερό. Η προσθήκη αυτής της μονάδας απαιτεί δύο πρόσθετες μετεωρολογικές εισροές, τη βροχή και την ET<sub>0</sub>.

Στο SAFY-CO<sub>2</sub>, το έδαφος μελετάται ως ένα ενιαίο στρώμα του οποίου τα χαρακτηριστικά (παράμετροι) χρησιμοποιούνται για τον υπολογισμό της R<sub>H</sub>. Στην ενότητα νερού του SAFYE-CO<sub>2</sub>, το έδαφος χωρίζεται σε 3 στρώματα: Ένα επιφανειακό στρώμα σταθερού βάθους (D<sub>EVAP</sub>), στο οποίο υπάρχει Ev και TR, ένα στρώμα ριζών, που συμβαίνει μόνο TR, το οποίο αυξάνεται σε βάθος κατά τη διάρκεια της καλλιεργητικής περιόδου με ρυθμό (V<sub>R</sub>) που ακολουθεί την ανάπτυξη της καλλιέργειας και περιορίζεται από ένα μέγιστο βάθος ριζών (D<sub>ROOTS</sub>) και ένα βαθύ στρώμα σταθερού βάθους (D). Το τελευταίο χρησιμοποιείται για την αρχικοποίηση της περιεκτικότητας σε νερό του ριζικού στρώματος κατά τη διάρκεια της ανάπτυξης της ρίζας και για τον έλεγχο των επιδράσεων της διήθησης και της αποστράγγισης. Το διαθέσιμο νερό (AWC) του εδάφους ενημερώνεται ουσιαστικά καθημερινά, λαμβάνοντας υπόψη τη βροχή (εισροές), το Ev και το TR που προσομοιώνεται με βάση τις μεθόδους FAO-56. Το AWC ορίζεται από τις ροές νερού (εισόδου και εξόδου) και τη συνολική υδατική ικανότητα (total water capacity - TWC) (δηλαδή, TWC = θ<sub>FC</sub> - θ<sub>WP</sub>). Η TR εξαρτάται από έναν μέγιστο συντελεστή TR (K<sub>cbMAX</sub>), έναν συντελεστή μείωσης (K<sub>TRP</sub>), μια κρίσιμη παράμετρο σχετικής υγρασίας (D<sub>fe</sub>), τον GAI και τον μέγιστο σχετικό ρυθμό πλήρωσης μεταξύ της επιφάνειας και του στρώματος ρίζας. Η ενότητα του νερού υπολογίζει επίσης δύο συναρτήσεις υδατικής καταπόνησης, μία που οδηγεί την Ev, λαμβάνοντας υπόψη το AWC του επιφανειακού στρώματος και μία άλλη που οδηγεί την ανάπτυξη της βλάστησης, λαμβάνοντας υπόψη το επιφανειακό και το ριζικό στρώμα μέσω της εκτίμησης της GPP.

Πίνακας 2. Μελέτες που αφορούν το μοντέλο SAFY

Authors	Title	Study Area	Data	Parameters	Crop	Method
Gaétan Pique, Rémy Fieuzal, Philippe Debaeke, Ahmad Al Bitar, Tiphaine Tallec and Eric Ceschia	<b>Combining High-Resolution Remote Sensing Products with a Crop Model to Estimate Carbon and Water Budget Components: Application to Sunflower.</b>	France	in situ, HSTR satellite products (SPOT 2-4,5, formosat-2, DEIMOS-1, Landsat-8, Sentinel2-A)	air temperature ( $T_{air}$ ) at 2 m, incoming global radiation (RG), rain, and reference ETR (ETO). dry aboveground mass (DAM). NEE from GPP and Reco. data on the soil moisture at field capacity ( $\theta_{rc}$ ), daily soil available water content (AWC)	sunflower	SAFY-CO2
Gaétan Pique, Taeken Wijmert, Rémy Fieuzal and Eric Ceschia	<b>Estimation of Crop Production and CO2 Fluxes Using Remote Sensing: Application to a Winter Wheat/Sunflower Rotation</b>	France	Formosat-2, SPOT 2/4, Landsat-8, biophysical variables neural network tool		sunflower	SAFY-CO2
Gaétan Pique, Rémy Fieuzal, Ahmad Al Bitar, Amanda Veloso, Tiphaine Tallec, Aurore Brut, Morgan Ferlicoq, Bartosz Zawilski, Jean-François Dejoux, Hervé Gibrin, Eric Ceschiac	<b>Estimation of daily CO2 fluxes and of the components of the carbon budget for winter wheat by the assimilation of Sentinel 2-like remote sensing data into a crop model</b>	France	HSTR data from the Formosat-2 and SPOT satellites, a priori values from literature, in situ for validation		mainly winter wheat	SAFY-CO2 model
Deepak Upreti, Tim McCarthy, Macdara O'Neill, Kazeem Ishola and Rowan Fealy	<b>Application and Evaluation of a Simple Crop Modelling Framework: A Case Study for Spring Barley, Winter Wheat and Winter Oilseed Rape over Ireland</b>	Ireland	in situ, soilgrids.org, Irish national meteorological service		Spring Barley, Winter Wheat and Winter Oilseed Rape	SAFY model and modification the soil–water balance and carbon modules to simulate water and carbon budget

#### 2.2.2.2. Μέθοδοι παλινδρόμησης

Στη βιβλιογραφία είναι συχνή η χρήση μοντέλων παλινδρόμησης για την εκτίμηση του ισοδυναμίου του άνθρακα στις καλλιέργειες. Πρόκειται για τους αλγορίθμους μηχανικής μάθησης που χρησιμοποιούνται για την πρόβλεψη της τιμής μιας μεταβλητής, η οποία εξαρτάται από ορισμένα χαρακτηριστικά ή μεταβλητές εισόδου. Τα μοντέλα που χρησιμοποιούνται περισσότερο είναι η γραμμική παλινδρόμηση, η πολυωνυμική, η λογαριθμική, η παλινδρόμηση που στηρίζεται σε δέντρα απόφασης και η μη γραμμική παλινδρόμηση.

Οι εργασίες:

1. Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas
2. Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands
3. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review
4. An advanced soil organic carbon content prediction model via fused temporal-spatial-spectral (TSS) information based on machine learning and deep learning algorithms
5. Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra
6. Soil Organic Carbon mapping of partially vegetated agricultural fields with imaging spectroscopy
7. Soil Organic Carbon Estimation in Croplands by Hyperspectral Remote APEX Data Using the LUCAS Topsoil Database
8. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel 2 spectral imaging

στηρίζουν τη μεθοδολογία που χρησιμοποιούν σε αλγορίθμους μηχανικής μάθησης, με συνηθέστερα παραδείγματα την παλινδρόμηση μερικών ελαχίστων τετραγώνων (PLSR), το μοντέλο ακραίας μηχανικής μάθησης (ELM) και των τυχαίων δασών (RF), (πίνακας 3).

#### Γραμμική παλινδρόμηση

Η γραμμική παλινδρόμηση αποτελεί μια από τις πιο δημοφιλείς μεθόδους μηχανικής μάθησης για την πραγματοποίηση προβλέψεων, διότι είναι αρκετά απλή, γεγονός που επιβεβαιώνει και η σύγκριση με τις προηγούμενες. Σε αυτήν την περίπτωση, η πρόβλεψη προκύπτει από τη γραμμική συνάρτηση που προσαρμόζεται καλύτερα στις παρατηρήσεις, οι οποίες δίνονται ως μεταβλητές εισόδου. Η διαδικασία αυτή περιλαμβάνει την εύρεση των βέλτιστων παραμέτρων (συντελεστές) της γραμμικής σχέσης, ώστε να ελαχιστοποιείται η απόκλιση μεταξύ των πραγματικών τιμών της εξαρτημένης μεταβλητής και των προβλέψεων που παράγονται από το μοντέλο. Η απόκλιση αυτή μπορεί να μετρηθεί με διάφορους τρόπους, όπως η τετραγωνική απόκλιση. Η γραμμική παλινδρόμηση είναι αρκετά χρήσιμη κυρίως λόγω της απλότητας και αποδοτικότητάς της. Μπορεί να υπολογίσει με ταχύτητα μεγάλο όγκο δεδομένων και γίνεται εύκολα κατανοητή στο χρήστη. Σημαντική είναι και η ευελιξία της, αφού έχει τη δυνατότητα εξέλιξης σε πιο σύνθετες μορφές, όπως η πολύπλοκη και η λογαριθμική παλινδρόμηση για την

αντιμετώπιση αντίστοιχων προβλημάτων. Η απλοϊκότητα της γενικής της μορφής, όμως, ευθύνεται για ορισμένες αδυναμίες που παρουσιάζει η μέθοδος, καθώς έχει περιορισμένη εκφραστική δύναμη, με αποτέλεσμα να μην εφαρμόζεται ικανοποιητικά σε πολύπλοκες ή μη γραμμικές σχέσεις. Επιπλέον, εφόσον λαμβάνει μια φορά δεδομένα εκπαίδευσης, προσαρμόζεται στα δεδομένα εισόδου και επηρεάζεται άμεσα από την κατανομή τους, γεγονός που είναι πιθανό να οδηγήσει σε υπερεκπαίδευση του μοντέλου και αδυναμία γενίκευσης, αλλά και αλλοίωση της πρόβλεψης λόγω της ευαισθησίας του μοντέλου στα outliers.

## PLSR

Η παλινδρόμηση μερικών ελαχίστων τετραγώνων (PLSR) είναι ένα αντιπροσωπευτικό μοντέλο των κλασικών μεθόδων φθίνουσας διάστασης (Wold, 1987). Η PLSR εισήχθη από τον Σουηδό στατιστικολόγο Wold (1987) και έχει χρησιμοποιηθεί ευρέως στη χημειομετρία και σε άλλους συναφείς επιστημονικούς τομείς. Είναι μια στατιστική μέθοδος που χρησιμοποιείται για την εύρεση της γραμμικής σχέσης μεταξύ δύο μεταβλητών. Συγκεκριμένα, η μέθοδος αυτή χρησιμοποιείται για να εκτιμήσει τα σημεία μιας γραμμικής καμπύλης που αντιστοιχούν σε μια σειρά παρατηρήσεων με μεταβλητές εξαρτημένες και ανεξάρτητες. Πρόκειται για μια εξέλιξη της παραδοσιακής παλινδρόμησης ελαχίστων τετραγώνων, η οποία χρησιμοποιείται όταν το πλήθος των ανεξάρτητων μεταβλητών υπερβαίνει το αντίστοιχο των διαθέσιμων παρατηρήσεων. Στην PLSR, οι ανεξάρτητες μεταβλητές χρησιμοποιούνται για να κατασκευάσουν μια νέα σειρά μεταβλητών που ονομάζεται συνιστώσα, με τη συνιστώσα να είναι γραμμικός συνδυασμός των αρχικών ανεξάρτητων μεταβλητών. Συνεπώς επιτρέπει την ανάλυση πολλαπλών ανεξάρτητων μεταβλητών, ενώ ταυτόχρονα μειώνει τον αριθμό των ανεξάρτητων μεταβλητών που χρησιμοποιούνται στην πρόβλεψη της εξαρτημένης μεταβλητής, με τον σκοπό της βελτίωσης της ικανότητας πρόβλεψης του μοντέλου. Συγκεκριμένα, η PLSR χωρίζει την εξήγηση της εξαρτημένης μεταβλητής σε διάφορες συνιστώσες, όπως και η PCA (Ανάλυση Κύριων Συνιστωσών).

Συχνά χρησιμοποιείται στη βελτίωση της ποιότητας της πρόβλεψης σε περιπτώσεις που οι δεδομένες ανεξάρτητες μεταβλητές είναι θορυβώδεις ή έχουν υψηλή συσχέτιση μεταξύ τους. Η PLSR επιλύει αυτό το πρόβλημα δημιουργώντας νέες μεταβλητές, γνωστές ως συνιστώσες PLS, που είναι συνδεδεμένες με τις αρχικές μεταβλητές και εξηγούν καλύτερα την εξαρτημένη μεταβλητή. Είναι επίσης χρήσιμη στην ανάλυση δεδομένων με μικρό αριθμό δειγμάτων, καθώς παρέχει αξιόπιστες προβλέψεις, ακόμη και όταν ο αριθμός των δειγμάτων είναι περιορισμένος.

Οι Castaldi et al., (2018) υπολόγισαν τα επίπεδα οργανικού άνθρακα στο έδαφος με τη χρήση PLSR. Οι χάρτες που δημιούργησαν έχουν ικανοποιητικά επίπεδα ακρίβειας με  $RMSE = 1.5-4.9 \text{ g}\cdot\text{kg}^{-1}$  και λόγο απόδοσης προς απόκλιση (RPD) = 1.4 -1.7. Οι Castaldi et al., (2019) χρησιμοποίησαν PLSR για τον υπολογισμό του SOC σε πέντε περιοχές με καλλιέργειες. Οι μετρικές που εξετάζουν την ακρίβεια της πρόβλεψης έχουν τις τιμές:  $RMSE = 3.7, 2.4, 4.7, 25.2, 1.9$  για τον Sentinel 2, 5.0, 3.2, 3.8, 33.7, 1.7 για τον αερομεταφερόμενο υπερφασματικό δέκτη, 8.3, 3.2, 3.7, 51.2, 1.7 για τον αερομεταφερόμενο υπερφασματικό δέκτη, προσαρμοσμένο στα κανάλια του Sentinel 2 και  $RPD = 1.7, 1.6, 1.0, 1.5, 1.1$  για τον Sentinel 2, 1.7, 1.4, 1.2, 1.1, 1.3 για τον αερομεταφερόμενο υπερφασματικό δέκτη, 1.0, 1.4, 1.2, 0.8, 1.3 για τον αερομεταφερόμενο υπερφασματικό δέκτη, προσαρμοσμένο στα κανάλια του Sentinel



2. Τα αποτελέσματα κρίνονται ως ικανοποιητικά για τις τέσσερις από τις πέντε περιοχές. Οι Long Guo et al., (2021) υπολόγισαν το SOC με δείκτες αξιολόγησης  $RMSE=0.18$ ,  $R^2=0.61$  και απόδοση στο διατεταρτημοριακό εύρος  $RPIQ=1.83$ .

## ELM

Η ELM κατασκευάστηκε από τους Huang et al. (2006) με βάση ένα νευρωνικό δίκτυο τροφοδοσίας ενός κρυφού στρώματος. Η επιλογή βαρών του κρυμμένου στρώματος της ELM μπορούν να τεθούν τυχαία για να βελτιώσουν την απόδοση γενίκευσης και την αποτελεσματικότητα της μάθησης. Η ταχύτητα της είναι χιλιάδες φορές μεγαλύτερη από τους παραδοσιακούς αλγορίθμους μάθησης δικτύου τροφοδότησης, απαιτεί ελάχιστη παρέμβαση από το χρήστη και η εκπαίδευση της γίνεται γρήγορα, καθώς τα βάρη επιλέγονται μια φορά και δεν απαιτείται περαιτέρω προσαρμογή. Σημαντικά πλεονεκτήματα της ELM είναι, επίσης, ότι παρέχει υψηλή ακρίβεια σε σύγκριση με άλλους αλγορίθμους, όπως η γραμμική παλινδρόμηση και υψηλή αντοχή στον θόρυβο.

Οι Long Guo et al. (2021) εφάρμοσαν ELM στη μελέτη τους. Ως συνάρτηση ενεργοποίησης χρησιμοποιήθηκε μια σιγμοειδής συνάρτηση και ο αριθμός των κρυφών κόμβων κυμαίνεται από 4 έως 100 με αύξηση 2. Η ELM δύναται να παράγει τυχαία βάρη και προκαταλήψεις ως παραμέτρους εισόδου. Κάθε μοντέλο επαναλήφθηκε 500 φορές για να επιλεγεί το βέλτιστο μοντέλο με το μικρότερο σφάλμα.

Οι ακρίβειες πρόβλεψης παρουσιάζονται με βάση τους παραδοσιακούς δείκτες αξιολόγησης του μέσου τετραγωνικού σφάλματος (RMSE), του συντελεστή προσδιορισμού ( $R^2$ ) και του RPIQ,  $RMSE = 0.16$ ,  $R^2 = 0.72$  και  $RPIQ = 2.03$

## SVMR

Η μηχανή παλινδρόμησης διανυσμάτων υποστήριξης (SVMR) είναι ένα μοντέλο μηχανικής μάθησης που χρησιμοποιείται για προβλήματα παλινδρόμησης, όπου ο στόχος είναι η πρόβλεψη μιας συνεχούς τιμής αντί για μια διακριτή κατηγορία. Το SVMR βασίζεται στην ίδια αρχή των διανυσμάτων υποστήριξης (support vectors) που χρησιμοποιούνται στη μηχανή διανυσμάτων υποστήριξης (SVM) για προβλήματα ταξινόμησης. Κατά την εκπαίδευση, το SVMR αναζητά ένα υπερεπίπεδο (ή πολυεπίπεδο) που προσεγγίζει τη σχέση μεταξύ των εισόδων και των εξόδων με τρόπο που ελαχιστοποιεί το σφάλμα πρόβλεψης. Τα διανύσματα υποστήριξης είναι τα δείγματα εισόδου που βρίσκονται κοντά σε αυτό το υπερεπίπεδο και συμβάλλουν στην πρόβλεψη των εξόδων. Στα πλεονεκτήματα του συγκεκριμένου μοντέλου είναι ότι παρουσιάζει ανθεκτικότητα στην υπερεκπαίδευση, ώστε να αποφεύγεται η υπερβολική προσαρμογή στα δεδομένα εκπαίδευσης. Το SVMR έχει εφαρμογές σε πολλούς τομείς, όπως προβλέψεις τιμών ακινήτων, χρηματιστηριακές αγορές, και προβλέψεις χρονοσειρών. Η ικανότητά του να αντιμετωπίσει προβλήματα παλινδρόμησης με συμπλέγματα δεδομένα το καθιστά ένα ισχυρό εργαλείο για την πρόβλεψη συνεχών μεταβλητών. Οι Gholizadeh et al., (2018) το χρησιμοποίησαν ως μέθοδο παλινδρόμησης για τη σύγκριση των δεδομένων τους. Τα αποτελέσματα έδειξαν ότι η ακρίβεια πρόβλεψης του SOC με βάση το εργαστηριακό φασματοσκόπιο, το αερομεταφερόμενο σύστημα και το Sentinel-2 στην πλειονότητα των περιοχών ήταν επαρκής,  $RMSE$  (Root Mean Square Error) για τις τέσσερις περιοχές να είναι για το φασματοσκόπιο: 0.14, 0.17, 0.12, 0.07, για τις

υπερφασματικές εικόνες: 0.12, 0.20, 0.13, 0.09 και για τις εικόνες Sentinel 2: 0.14, 0.24, 0.09, 0.08 και η σχετική απόκλιση μεταξύ των προβλεπόμενων τιμών RPD (Relative Percent Difference) να είναι για το φασματοσκόπιο: 1.77, 2.13, 1.29, 2.93, για τις υπερφασματικές εικόνες: 2.05, 1.80, 1.26, 1.78 και για τις εικόνες Sentinel 2: 1.60, 1.70, 1.73, 1.92.

## Τυχαία Δάση (Random Forest)

Ο Random Forest (RF) είναι ένας αλγόριθμος μηχανικής μάθησης που αναπτύχθηκε από τους Leo Breiman και Adele Cutler το 2001 και χρησιμοποιείται προβλέψεις ή ταξινόμηση, Breiman (2001). Αποτελείται από ένα σύνολο ασυσχέτιστων μεταξύ τους δέντρων αποφάσεων, καθένα από τα οποία δημιουργείται από τυχαία παραδείγματα των δοσμένων δεδομένων και δίνει μια εκτίμηση. Με αυτόν τον τρόπο μειώνεται η προσαρμογή του μοντέλου στο θόρυβο που υπάρχει στα δεδομένα. Ο αλγόριθμος επιλέγει τυχαία υποσύνολα του δείγματος εκπαίδευσης, τα οποία είναι διαφορετικά μεταξύ τους και συνεπώς στατιστικά ανεξάρτητα. Η τυχαία επιλογή των υποσυνόλων ενέχει την πιθανότητα να χρησιμοποιηθεί κάποιο τμήμα του δείγματος παραπάνω από μια φορά και ή κάποιο άλλο καμία. Επιπλέον παράγοντας είναι η τυχαία επιλογή των μεταβλητών (καναλιών), βάση των οποίων θα τεθούν τα φασματικά όρια διαχωρισμού σε κάθε εσωτερικό κόμβο (internal splitting node). Το πλήθος των καναλιών που χρησιμοποιούνται σε κάθε δέντρο απόφασης, καθορίζεται από το σύνολο των μεταβλητών ( $V$ ) που χρησιμοποιούνται στην ταξινόμηση και συνηθίζεται να είναι  $mtry = \sqrt{V}$ . Η επιλογή ενός υποσυνόλου εκπαίδευσης, επιδρά θετικά στη μείωση του υπολογιστικού χρόνου ενώ η επιλογή μερικώς ή εξ' ολοκλήρου διαφορετικού υποσυνόλου δεδομένων εκπαίδευσης για κάθε δέντρο, βοηθά στη δημιουργία διαφορετικών εκτιμητών, ικανών να προσαρμοστούν σε διαφορετικά είδη δεδομένων, για να βελτιωθεί έτσι η απόδοση του αλγορίθμου. Ακόμα, σύμφωνα με τους Hastie, et al. (2009), η επιλογή κάθε φορά  $mtry$  τυχαίων μεταβλητών σε κάθε εσωτερικό κόμβο διαχωρισμού (internal splitting node), βοηθά επίσης στη μείωση της συσχέτισης ( $\rho$ ) μεταξύ κάθε πιθανού ζεύγους δέντρων και κατ' επέκταση μείωση της μεγάλης αστάθειας (ύπαρξη τυχαίων σφαλμάτων) που παρουσιάζουν τα δενδροειδή μοντέλα από τα οποία απαρτίζεται το δάσος. Αφού επιλεγούν τα κανάλια και τα υποσύνολα των δεδομένων εκπαίδευσης που θα χρησιμοποιηθούν για την ανάπτυξη του κάθε δέντρου, σε κάθε εσωτερικό κόμβο επιλέγεται το σημείο διαχωρισμού μεταξύ των υποψηφίων φασματικών κλάσεων. Τα κριτήρια που θα καθορίσουν τη διαδρομή αυτού του εικονοστοιχείου μέχρι έναν τερματικό κόμβο (φύλλο), καθορίζονται από τα όρια διαχωρισμού. Αυτά προσδιορίζονται σε κάθε εσωτερικό κόμβο μέσα από μια επαναληπτική διαδικασία, κατά την οποία διάφορα όρια εξετάζονται και αξιολογούνται μέσω του υπολογισμού του δείκτη gini. Επιλέγεται το όριο με την χαμηλότερη τιμή του δείκτη, αφού με αυτό τον τρόπο επιτυγχάνεται η υπολογιστική ταχύτητα του αλγορίθμου, χωρίς να σπαταλιέται χρόνος που δε θα οδηγήσει το συντομότερο σε μία απόφαση. Η τελική απόφαση κατηγοριοποίησης του κάθε εικονοστοιχείου, λαμβάνεται βάση της ψήφου πλειοψηφίας στο σύνολο των αποφάσεων από τα δέντρα. Εκτός από τα δεδομένα εκπαίδευσης και τον αριθμό των μεταβλητών που επιλέγεται τυχαία για την εκπαίδευση του κάθε δέντρου, ακόμα ένας παράγοντας πρέπει να καθοριστεί για την κατασκευή του αλγορίθμου και αυτός είναι ο αριθμός των δέντρων απόφασης που απαρτίζουν το δάσος. Θεωρητικά, όσο το πλήθος των δέντρων αυξάνεται, τόσο πιο ομαλά είναι τα όρια διαχωρισμού και κατά συνέπεια καλύτερη η απόδοση του μοντέλου.

Για την πρόβλεψη του SOC είναι αναγκαία η αξιολόγηση της σημασίας του κάθε καναλιού, η οποία μπορεί να πραγματοποιηθεί με τη μέθοδο Relative Variable Importance (RVI). Με αυτόν τον τρόπο, υπολογίζεται το ποσοστό συνεισφοράς της κάθε μεταβλητής στη μείωση του σφάλματος του μοντέλου κατά την πρόβλεψη του αποτελέσματος.

Οι Castaldi et al. (2019) χρησιμοποιούν Random Forest (RF) για την πρόβλεψη του SOC. Η ακρίβεια των προβλέψεων τους υπολογίστηκε ανά περιοχή με τις μετρικές RMSE και RPD, με RMSE = 4.7, 2.5, 4.8, 18.6, 2.0 για τον Sentinel 2, 9.3, 3.3, 4.0, 18.4, 2.0 για τον αερομεταφερόμενο υπερφασματικό δέκτη, 1.0, 3.4, 4.1, 18.5, 2.0 για τον αερομεταφερόμενο υπερφασματικό δέκτη, προσαρμοσμένο στα κανάλια του Sentinel 2 και RPD= 1.3, 1.5, 1.0, 2.1, 1.0 για τον Sentinel 2, 1.0, 1.4, 1.1, 2.1, 1.1 για τον αερομεταφερόμενο υπερφασματικό δέκτη και 1.0, 1.3, 1.1, 2.1, 1.1 για τον αερομεταφερόμενο υπερφασματικό δέκτη, προσαρμοσμένο στα κανάλια του Sentinel 2, αντίστοιχα.

## Gradient Boosting Regressor

Ο αλγόριθμος Gradient Boosting Regressor είναι τεχνική παλινδρόμησης που χρησιμοποιείται για την πραγματοποίηση προβλέψεων και ανήκει στην κατηγορία των Boosting αλγορίθμων, δηλαδή των αλγορίθμων που συνδυάζουν πολλαπλά μοντέλα μηχανικής μάθησης για να πετύχουν μια καλύτερη επίδοση. Εφαρμόζει τη μεθοδολογία της σταδιακής μάθησης, χρησιμοποιώντας τη σύγκριση του αποτελέσματος κάθε προηγούμενου μοντέλου με τα πραγματικά δεδομένα, ώστε να υπολογίσει το σφάλμα που θα χρησιμοποιήσει κατά την εκπαίδευση του επόμενου μοντέλου. Με λίγα λόγια, λειτουργεί με παρόμοιο τρόπο με αυτόν του RF, δηλαδή με δέντρα αποφάσεων, με τη διαφορά ότι κάθε νέο δέντρο δεν είναι ανεξάρτητο από τα υπόλοιπα, αλλά αντίθετα εκπαιδεύεται με βάση το σφάλμα του προηγούμενου, ώστε να βελτιώσει την ακρίβεια της πρόβλεψης. Κύριος σκοπός του αλγορίθμου είναι η μείωση της τιμής του σφάλματος. Συγκεντρώνει σημαντικά πλεονεκτήματα, όπως η υψηλή ακρίβεια και η ευκολία στη χρήση. Ωστόσο, παρουσιάζει μειονεκτήματα όσον αφορά την ταχύτητα εκπαίδευσης του μοντέλου, την πολυπλοκότητα του που έχει ως συνέπεια και υψηλότερες απαιτήσεις σε υπολογιστική ισχύ, καθώς και την ευαισθησία του στις ακραίες τιμές (outliers), γεγονός που μπορεί να οδηγήσει σε υπερπροσαρμογή του μοντέλου στα δεδομένα εκπαίδευσης, αλλά είναι δυνατόν να αποφευχθεί με τη χρήση τεχνικών, όπως οι περιορισμοί του βάθους του δέντρου ή η ρύθμιση της στάθμης εκμάθησης.

## Δέντρο απόφασης

Η πρόβλεψη ενός αποτελέσματος μέσω ενός δέντρου απόφασης γίνεται με τη διαίρεση των δεδομένων σε υποσύνολα, σύμφωνα με τις τιμές τους, με στόχο τη μείωση της αβεβαιότητας του αποτελέσματος. Η κύρια διαφορά του DT από το RF είναι ότι πρόκειται για ένα μόνο δέντρο το οποίο εκπαιδεύεται στα αρχικά δεδομένα και άρα προσαρμόζεται σημαντικά σε αυτά, με αποτέλεσμα πιθανή αδυναμία γενίκευσης.

Το μοντέλο παλινδρόμησης Ridge χρησιμοποιείται για την αντιμετώπιση ακριβώς αυτού του προβλήματος. Στη μέθοδο Ridge γίνεται κανονικοποίηση του μοντέλου με τη μείωση του μεγέθους των συντελεστών, μειώνοντας τη μεταξύ τους διακύμανση, χωρίς ωστόσο να έχει τη δυνατότητα επιλογής βαρών για τις μεταβλητές.

Πίνακας 3. Μελέτες που αφορούν μοντέλα παλινδρόμησης

Authors	Title	Study Area	Data	Parameters	Crop	Method
Long Guo, Xiaoru Sun, Peng Fu, Tiezhu Shi, Lina Dang, Yiyun Chen, M. Linderman, Ganlin Zhang, Yu Zhang, Qinghu Jiang, Haitao Zhang, Chen Zeng	<b>Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas</b>	Iowa	hyperspectral image (Headwall-Hyperspec, 380–1700 nm), Sentinel-2, Landsat 8, soil samples	soil moisture, SBD soil bulk density, and SOCD, GNDVI, WDRVI, NDVI	corn, soybeans, oats, red clover, flaxseed	PLSR, ELM
Fabio Castaldi, Andreas Hueni, Sabine Chabrilat, Kathrin Ward, Gabriele Buttafuoco, Bart Bomans, Kristin Vreys, Maximilian Brell, Bas van Wesemael	<b>Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands</b>	Germany, Luxembourg and Belgium	Sentinel-2, Hyperspectral airborne data (APEX or HySpex), Airborne data (APEX or HySpex) resampled, soil samples	NDVI, SNR	In Germany: rotation of winter cereals, root crops (maize, sugar beet, potato), Grass-land and pasture. In Belgium cropland.	PLSR, RF
Sushil Lamichhane, Lalit Kumar, Brian Wilson	<b>Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review</b>	Mainly China, Australia, USA				MLR, RF, Cu/RT, NN, BRT, SVM, GWR, GLM, KED, GAM, KNN, LMM
Xiangtian Meng, Yilin Bao, Yiang Wang, Xinle Zhang, Huanjun Liu	<b>An advanced soil organic carbon content prediction model via fused temporal-spatial-spectral (TSS) information based on machine learning and deep learning algorithms</b>	China	Landsat, GaoFen-5 hyperspectral images, in situ, DEM, slope Curvature	discrete wavelet transform based on the regional energy weight (RW-DWT) and spectral band segmentation methods. NDMI	annual crops	PLSR, random forest (RF) and convolutional neural network (CNN)

E.Vaudour, J.M.Gilliot, L.Bel, J.Lefevre, K.Chehdi	<b>Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra</b>	France	Airborne AISA-Eagle data, SPOT, field measurements with spectroradiometer	NDVI	rapeseed, winter wheat, spring barley, maize and dry pea or fodder crops	bootstrap PLSR, KSS algorithm, conditioned Latin Hypercube sampling (cLHS)
Harm Bartholomeus, Lammert Kooistra, Antoine Stevens, Martin van Leeuwen, Bas van Wesemael, Eyal Ben-Dor, Bernard Tychon	<b>Soil Organic Carbon mapping of partially vegetated agricultural fields with imaging spectroscopy</b>	Belgium	AHS-160 sensor, lab	Reflectance --> Rmaize, Rsoil	maize and bare soil	PLSR, linear spectral unmixing using PROSAIL, Residual Spectral Unmixing (RSU)
Fabio Castaldi, Sabine Chabrilat, Arwyn Jones, Kristin Vreys, Bart Bomans and Bas Van Wesemael	<b>Soil Organic Carbon Estimation in Croplands by Hyperspectral Remote APEX Data Using the LUCAS Topsoil Database</b>	Belgium, Luxembourg	Airborne Prism Experiment (APEX), LUCAS			PLSR
Gholizadeh, Daniel Žižalaa, Mohammadmehdi Saberioon, Luboš Borůvka	<b>Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging</b>	Czech Republic	Sentinel-2, airborne hyperspectral (CASI/SASI sensors), lab ASD FieldSpec spectroradiometer measurements, soil samples	NDVI, TVI, EVI, SATVI, SAVI, MSI, GNDVI, GRVI, LSWI, TSAVI, MSAVI, MSAVI2, WDWI, BI, BI2, RI, CI, V		SVMR

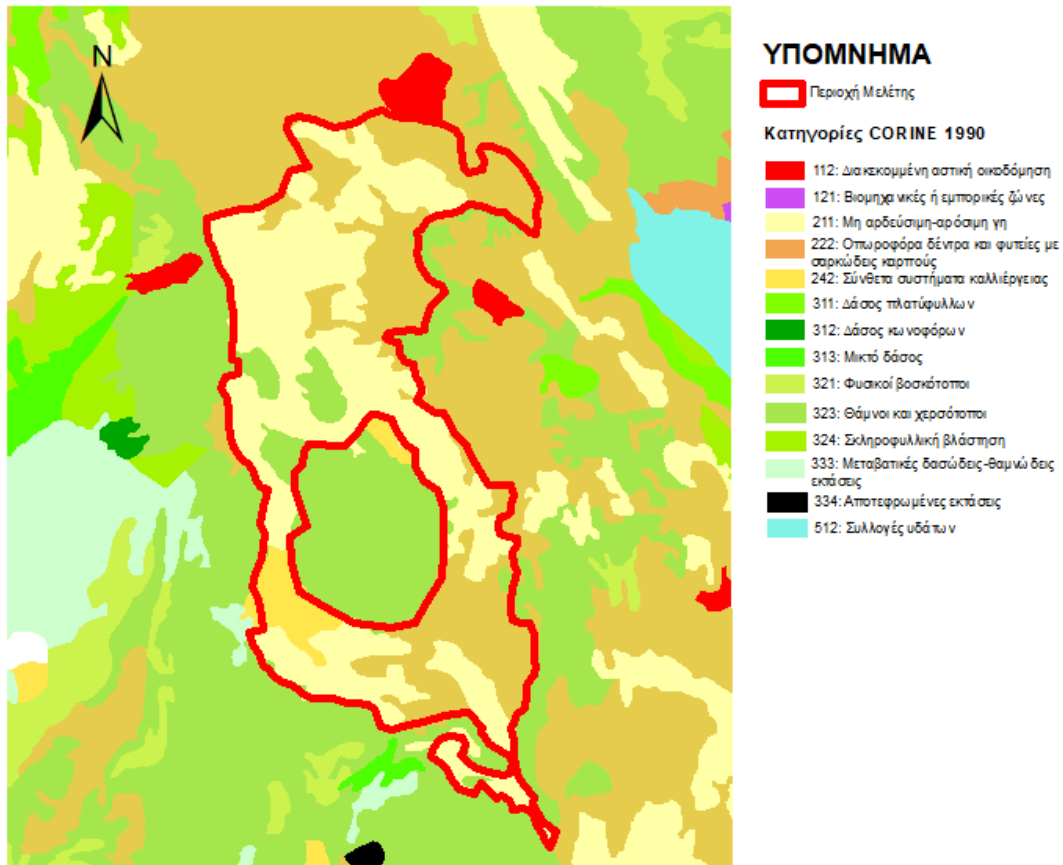
## Κεφάλαιο 3 – Μεθοδολογία

### 3.1. Η περιοχή μελέτης

Η περιοχή μελέτης είναι μια έκταση 33.974 στρεμμάτων στο υψίπεδο του Ξηρομέρου Αιτωλοακαρνανίας, στο νότιο τμήμα του Δήμου Ακτίου - Βόνιτσας της Περιφέρειας Δυτικής Ελλάδας. Κατά την χρονική περίοδο που εξετάζεται, η συνολική καλλιεργήσιμη έκταση είναι 23.690 στρέμματα και αποτελείται από μικρά αγροτεμάχια στα οποία καλλιεργούνται κυρίως καπνός, γρασίδια και χειμερινά σιτηρά, καθώς και πολυετείς καλλιέργειες με οπωροφόρα.



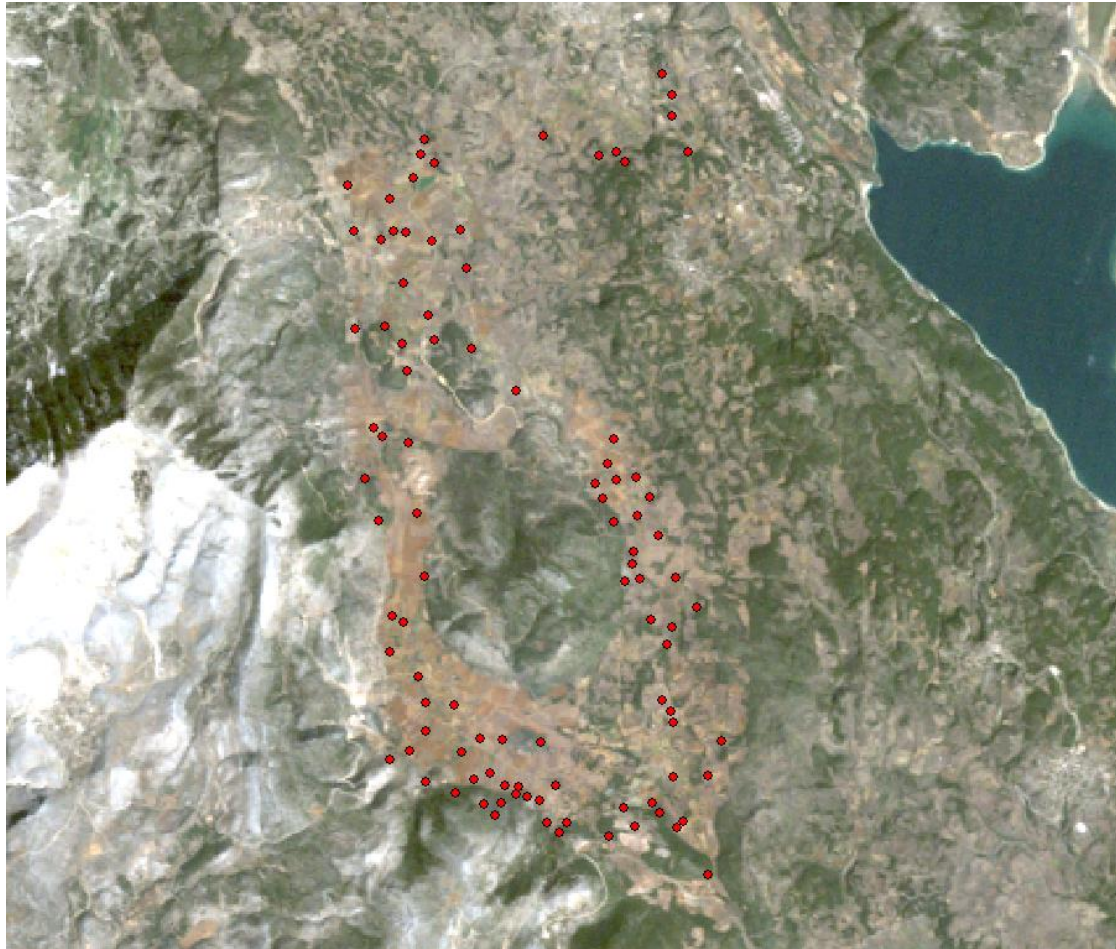
*Εικόνα 2. Η περιοχή μελέτης.*



Εικόνα 3. Χρήσεις γης - Κάλυψη γης σύμφωνα με το CORINE 1990.

### 3.2. Δεδομένα

Για την παρούσα μελέτη χρησιμοποιήθηκαν δορυφορικές εικόνες και δεδομένα εδάφους. Τα δεδομένα εδάφους δόθηκαν από το Ινστιτούτο Εδαφοϋδατικών πόρων και προέρχονται από την εδαφολογική μελέτη περιοχής Ξηρομέρου Ν. Αιτωλοακαρνανίας, του έτους 1995, του Ινστιτούτου Χαρτογράφησης και Ταξινόμησης Εδαφών Λάρισας, του Εθνικού Ιδρύματος Αγροτικής Έρευνας. Πρόκειται για 99 σημεία γνωστών συντεταγμένων για τα οποία είναι διαθέσιμες πληροφορίες για τα εδαφολογικά τους χαρακτηριστικά, καθώς και η τιμή της οργανικής ουσίας. Για το τμήμα της μελέτης για το οποίο υπάρχει η σχετική πληροφορία, τα δείγματα στο ανώτερο στρώμα του εδάφους (0 - 25cm) καταγράφονται στην πλειοψηφία τους ως οργανικά, αργιλώδη (C), πηλώδη (L), ή αργιλοπηλώδη (CL), ενώ σε λίγες περιπτώσεις καταγράφεται η παρουσία άμμου (S). Η οργανική ουσία παρουσιάζει τιμές από 0,5 έως 10%, με μέσο όρο 2,56 % και το μεγαλύτερο όγκο δεδομένων να συγκεντρώνεται στην περιοχή 2,1% έως 4,2% και επομένως τα εδάφη χαρακτηρίζονται ως μέσης περιεκτικότητας, γεγονός αναμενόμενο για την περιοχή και τη χρήση γης. Χρησιμοποιήθηκαν δορυφορικές εικόνες Landsat 5, ώστε να καλύπτουν την ίδια χρονική περίοδο με τα δείγματα εδάφους.



Εικόνα 4. Η κατανομή των δειγμάτων σε εικόνα Landsat 5 (έγχρωμο σύνθετο, κανάλια 3-2-1)

### 3.3. Μεθοδολογία

Η μέθοδος που εφαρμόστηκε βασίζεται στο Σύστημα γεωχωρικής ανίχνευσης εδάφους (GEOS3) (José Alexandre Melo Demattê et al - 2018). Πρόκειται για μια διαδικασία κατά την οποία δημιουργείται μια χρονοσειρά εικόνων, επιλέγονται οι περιοχές με γυμνό έδαφος για κάθε εικόνα και υπολογίζεται η φασματική ανάκλαση για κάθε εικονοστοιχείου, η οποία αναφέρεται ως χρονική φασματική ανάκλαση εδάφους (Temporal Soil Spectral Reflectance - TESS). Από τα TESS δημιουργείται μια συνθετική εικόνα (Synthetic Soil Image - SYSI), η οποία παρουσιάζει τις τιμές του γυμνού εδάφους στην περιοχή.

#### 3.3.1. Google Earth Engine

Στην παρούσα μελέτη, για την παραγωγή της συνθετικής εικόνας χρησιμοποιήθηκε η πλατφόρμα google earth engine (GEE). Η συγκεκριμένη πλατφόρμα είναι cloud-based και προσφέρει πλήθος ελεύθερων γεωχωρικών δεδομένων και εργαλείων για την παρατήρηση της γης και του περιβάλλοντος. Διαθέτει δεδομένα από δορυφόρους, όπως Landsat 1-5 MSS, Landsat 4-5 TM, Landsat 7, Landsat 8, Landsat 9, Sentinel 1, Sentinel 2, Sentinel 3, Sentinel 5P,



MODIS, κ.α., καθώς και αερομεταφερόμενους δέκτες, τα οποία αφορούν εικόνες πολύ υψηλής ανάλυσης, ατμοσφαιρικά και μετεωρολογικά δεδομένα, δεδομένα ραντάρ (SAR), ψηφιακά μοντέλα εδάφους, χάρτες κάλυψης και χρήσεων γης, κ.α. Η πρόσβαση στα δεδομένα είναι δωρεάν και παράλληλα παρέχεται η δυνατότητα επεξεργασίας τους σε προγραμματιστικό περιβάλλον JavaScript ή Python μέσω του Google Earth Engine API (Application Programming Interface). Το περιβάλλον του GEE είναι αρκετά φιλικό προς το χρήστη, αποτελείται από έναν χάρτη στον οποίο προβάλλονται τα δεδομένα, τα οποία εισάγει ή επεξεργάζεται ο χρήστης, ενώ παράλληλα έχει τη δυνατότητα να επιλέξει την περιοχή που αφορά τη μελέτη του. Περιέχει ένα παράθυρο στο οποίο γίνεται η επεξεργασία και η εκτέλεση κώδικα (code editor), ένα παράθυρο με καρτέλες που περιλαμβάνουν αρχεία αποθηκευμένου κώδικα (scripts), περιγραφή των έτοιμων εργαλείων (docs) και ένα πεδίο για εισαγωγή διανυσματικών αρχείων ή αρχείων εικόνων που αφορούν την εκάστοτε περιοχή μελέτης (assets). Τέλος, υπάρχει ένα παράθυρο με καρτέλες για την εποπτεία των διαφορετικών θεματικών επιπέδων (inspector), των δεδομένων εξόδου του κώδικα (console) και των αρχείων που δημιουργούνται (tasks). Η πλατφόρμα παρέχει τη δυνατότητα εξαγωγής των παραγόμενων αρχείων στο Google Drive. Τα σημαντικότερα πλεονεκτήματα του GEE, τα οποία οδήγησαν και στην επιλογή του για την παρούσα εργασία, είναι η δυνατότητα επεξεργασίας χρονοσειρών εικόνων, το γεγονός ότι είναι cloud-based και δεν απαιτεί αποθήκευση των αρχικών δεδομένων, επομένως δεν υπάρχουν απαιτήσεις σε υπολογιστική ισχύ και αποθηκευτικό χώρο, είναι εύχρηστο και δωρεάν.

Η επεξεργασία των δεδομένων πραγματοποιήθηκε με τη χρήση JavaScript, προκειμένου να εναρμονιστεί με τα έτοιμα κομμάτια κώδικα από το εγχειρίδιο χρήσης της πλατφόρμας. Αρχικά ορίστηκε χωρικά η περιοχή μελέτης και έγινε εισαγωγή του διανυσματικού επιπέδου με τα δείγματα εδάφους. Επιλέχθηκαν εικόνες Landsat 5, level-2, Tier 1, οι οποίες είναι ραδιομετρικά, ατμοσφαιρικά και γεωμετρικά διορθωμένες. Δημιουργήθηκαν χρονοσειρές εικόνων, στις οποίες εφαρμόστηκαν μάσκες για νεφοκάλυψη <10% - 20% (για τα διαφορετικά πειράματα) στο κάτω και στο ανώτερο μέρος της ατμόσφαιρας, για σκιές, χιόνι, θυσάνους (cirrus), καθώς και μάσκες για τις χρήσεις γης σύμφωνα με το CORINE, ώστε να απομονωθούν οι καλλιέργειες. Συγκεκριμένα χρησιμοποιήθηκαν μάσκες για τις κατηγορίες που αφορούν στο ανθρωπογενές περιβάλλον, τα δάση, τις περιοχές με βλάστηση, τις βαλτώδεις εκτάσεις και το νερό, όπως φαίνεται στον πίνακα 4.

Πίνακας 4. Οι κατηγορίες του CORINE, που εξαιρέθηκαν μέσω της μάσκας.

Επίπεδο 1	Επίπεδο 2	Επίπεδο 3
1. Τεχνητές επιφάνειες	1.1 Αστική οικοδόμηση	1.1.1 Συνεχής αστική οικοδόμηση
		1.1.2 Διακεκομμένη αστική οικοδόμηση
	1.2 Βιομηχανικές, εμπορικές ζώνες και δίκτυα επικοινωνίας	1.2.1 Βιομηχανικές ή εμπορικές ζώνες
		1.2.2 Οδικά και σιδηροδρομικά δίκτυα και γειτνιάζουσα γη
		1.2.3 Ζώνες λιμένων
		1.2.4 Αεροδρόμια
	1.3 Ορυχεία, χώροι απορρίψεως απορριμμάτων και χώροι οικοδόμησης	1.3.1 Χώροι εξόρυξης ορυκτών
	1.4 Τεχνητές, μη γεωργικές ζώνες πρασίνου	1.4.1 Περιοχές αστικού πρασίνου
		1.4.2 Εγκαταστάσεις αθλητισμού και αναψυχής
	3. Δάση και ημι-φυσικές περιοχές	3.1 Δάση
3.1.2 Δάσος κωνοφόρων		
3.1.3 Μικτό δάσος		
3.2 Συνδυασμοί θαμνώδους και/ή ποώδους βλάστησης		3.2.3 Σκληροφυλλική βλάστηση
		3.2.4 Μεταβατικές δασώδεις-θαμνώδεις εκτάσεις
4. Υγρές ζώνες	4.1 Εσωτερικές υγρές ζώνες	4.1.1 Βάλτοι στην ενδοχώρα
		4.1.2 Τυφώνες
	4.2 Παραθαλάσσιες υγρές ζώνες	4.2.1 Παραθαλάσσιοι βάλτοι
		4.2.2 Αλυκές
		4.2.3 Παλιρροιακά επίπεδα
5. Υδάτινες επιφάνειες	5.1 Χερσαία ύδατα	5.1.1 Ροές υδάτων
		5.1.2 Συλλογές υδάτων
	5.2 Θαλάσσια ύδατα	5.2.1 Παράκτιες λιμνοθάλασσες
		5.2.2 Εκβολές ποταμών
		5.2.3 Θάλασσα και ωκεανός

Στη συνέχεια εφαρμόστηκαν οι δείκτες NDVI, NBR2 και VNSIR (πίνακας 5) για το διαχωρισμό του γυμνού εδάφους. Ο κανονικοποιημένος δείκτης βλάστησης NDVI χρησιμοποιείται για την ανίχνευση του πρασίνου των φυτών και αποτελεί τον πιο ευρέως διαδεδομένο δείκτη

βλάστησης. Μπορεί να αξιοποιηθεί για τον προσδιορισμό της πυκνότητας της βλάστησης, την υγεία των φυτών και την παρακολούθηση αλλαγών σε αυτά. Υπολογίζεται ως η κανονικοποιημένη διαφορά της ανακλαστικότητας στο κόκκινο από το εγγύς υπέρυθρο τμήμα του φάσματος. Ο δείκτης λαμβάνει τιμές από -1 έως 1, όπου οι αρνητικές τιμές σηματοδοτούν την απουσία βλάστησης, συγκεκριμένα η τιμή -1 παρατηρείται σε υδάτινους όγκους, οι τιμές κοντά στο 0 παρατηρούνται σε γυμνό έδαφος ή περιοχές με ελάχιστη βλάστηση, ενώ οι τιμές κοντά στο 1 δηλώνουν την πολύ πυκνή, βλάστηση. Ο κανονικοποιημένος δείκτης καύσης 2 (NBR2) αποτελεί τη βελτιωμένη έκδοση του NBR και παρουσιάζει μεγαλύτερη ευαισθησία στην ανίχνευση υγρασίας στη βλάστηση. Εκφράζει τη ζημιά που έχει η υποστεί η βλάστηση σε μια περιοχή λόγω πυρκαγιάς ή άλλων παραγόντων. Υπολογίζεται ως η κανονικοποιημένη διαφορά της ανακλαστικότητας των καναλιών SWIR1 και SWIR2 και λαμβάνει τιμές από -1 έως 1. Οι αρνητικές τιμές αντιστοιχούν σε καταστροφή της βλάστησης, ενώ οι θετικές σε ανάκαμψή της. Ο εν λόγω δείκτης είναι χρήσιμος σε μελέτες αποκατάστασης του περιβάλλοντος από πυρκαγιές, αλλά και γενικότερα αξιολόγηση της κατάστασης της βλάστησης σε μια περιοχή. Οι José A. M. Demattê et al., (2020) υπολόγισαν τον δείκτη VNSIR, ο οποίος απεικονίζει την τάση από το ορατό, στο εγγύς και το μέσο υπέρυθρο εφαρμόζοντας μια γραμμική σχέση για τα κανάλια Red, Green, Blue, NIR και SWIR2, με σκοπό την ανάδειξη του ακάλυπτου εδάφους.

Η επιλογή των συγκεκριμένων δεικτών έγινε με σκοπό την παραγωγή μιας συνθετικής εικόνας (composite), η οποία να περιλαμβάνει μόνο εικονοστοιχεία με γυμνό έδαφος από τη χρονοσειρά των αρχικών εικόνων. Στις περιπτώσεις που υπάρχουν διαθέσιμα εικονοστοιχεία γυμνού εδάφους στην ίδια περιοχή από περισσότερες από μια εικόνες της χρονοσειράς, η τιμή του εικονοστοιχείου της συνθετικής εικόνας ταυτίζεται με την ενδιάμεση τιμή των διαθέσιμων εικονοστοιχείων. Με τον τρόπο αυτό επιτυγχάνεται η σύνθεση υπαρκτών τμημάτων του εδάφους για τη δημιουργία της εικόνας, αντί για την απόδοση με τη χρήση της μέσης τιμής που πιθανόν να είναι τεχνητή συνθήκη.

Οι José A. M. Demattê et al., (2020) εφάρμοσαν τους δείκτες ορίζοντας κατώφλια (thresholds) για τις τιμές που δέχονται ως γυμνό έδαφος για κάθε δείκτη. Οι τιμές των κατωφλιών που χρησιμοποίησαν, για τον NDVI από -0,25 έως 0,25, για τον NBR2 από -0,30 έως 0,10 και για τον VNSIR μικρότερο από 0,90. Οι Castaldi et al., (2019) εφάρμοσαν κατώφλια για τον NDVI έως 0,35 και για τον NBR2 πραγματοποίησαν πειράματα στο κάτω όριο με τις τιμές 0,025, 0,05, 0,075 και ως άνω όριο έθεσαν την τιμή 0,35. Οι Uta Heiden et. al (2022) επισήμαναν τη δυσκολία διαχωρισμού του γυμνού εδάφους από την ξηρή, μη φωτοσυνθετικά ενεργή βλάστηση, με τη χρήση πολυφασματικών εικόνων και την σύγχυση τους από τον NBR2. Οι Klara Dvorakova et al., (2023) έδειξαν ότι ένα κατώφλι με τιμή 0,05 για τον δείκτη NBR2 διαχωρίζει το γυμνό έδαφος και το χώμα σε συνθήκες σποράς, από τα υγρά εδάφη, τη βλάστηση και τα υπολείμματα καλλιεργειών.

Πίνακας 5. Εξισώσεις δεικτών

Δείκτης	Ορισμός	Εξίσωση σε Landsat 5 (TM)
NDVI	$\frac{NIR - Red}{NIR + Red}$	$\frac{B4 - B3}{B4 + B3}$
NBR2	$\frac{SWIR1 - SWIR2}{SWIR1 + SWIR2}$	$\frac{B5 - B7}{B5 + B7}$
VNSIR	$1 - [(2 * Red - Green - Blue) + 3(SWIR2 - NIR)]$	$1 - [(2 * B4 - B3 - B2) + 3(B7 - B4)]$

Εφαρμόστηκαν διαδοχικά έξι διαφορετικοί συνδυασμοί κατωφλιών για τους δείκτες NDVI, NBR2 και VNSIR (πίνακας 6) σε χρονοσειρές εικόνων Landsat 5 για τα έτη 1995, 1994 -1996, 1993 - 1997 και 1992 - 1998.

Πίνακας 6. Κατώφλια πειραμάτων

	NDVI upper	NDVI lower	NBR2 upper	NBR2 lower	VNSIR	cloud coverage
1	0,250	-0,250	0,100	-0,300	0,9	10%
2	0,250	-0,250	0,350	-0,300	0,9	10%
3	0,250	-0,250	0,200	-0,300	0,9	10%
4	0,350	-0,250	0,200	-0,300	0,9	10%
5	0,250	-0,250	0,050	0,000	0,9	10%
6	0,250	-0,250	0,150	0,000	0,9	20%

Από τους παραπάνω συνδυασμούς δημιουργήθηκαν για κάθε περίπτωση, μια συνθετική εικόνα η οποία περιλαμβάνει τις ενδιάμεσες τιμές όλων των διαθέσιμων εικονοστοιχείων, καθώς και ένας πίνακας με τις τιμές όλων των καναλιών και των δεικτών που χρησιμοποιήθηκαν για τα διαθέσιμα εικονοστοιχεία τα οποία συμπίπτουν με τις συντεταγμένες των δειγμάτων εδάφους. Στο σημείο αυτό πραγματοποιήθηκε μετατροπή των συντεταγμένων των δειγμάτων εδάφους, καθώς και της παραγόμενης εικόνας, ώστε να λυθεί η μεταξύ τους ασυμβατότητα. Τα δείγματα εδάφους δόθηκαν σε ένα διανυσματικό αρχείο (shapfile) που ανήκει στο ΕΓΣΑ '87, ενώ το προεπιλεγμένο σύστημα αναφοράς της εικόνας είναι το WGS 84. Για την επίλυση της συγκεκριμένης ασυμβατότητας με τις μικρότερες δυνατές παραμορφώσεις, έγινε μετατροπή και των δύο αρχείων σε WGS84 / UTM zone 34N, που βρίσκεται ανάμεσα στους μεσημβρινούς 18°E και 24°E και κεντρικές συντεταγμένες Χ:500000.0, Υ:4649776.22 και αποτελεί τη ζώνη του WGS84 που προσαρμόζεται καλύτερα στην Ελλάδα.

Επίσης, τα δεδομένα από τα δείγματα εδάφους αφορούσαν την οργανική ουσία (SOM) και πραγματοποιήθηκε η μετατροπή τους σε SOC, σύμφωνα με την εξίσωση του Jackson (1965):

$$SOM = SOC * 1.724$$

Για τα εικονοστοιχεία στα οποία ανιχνεύτηκαν ως γυμνό έδαφος και παράλληλα υπάρχει πληροφορία για τις τιμές του SOC από τα δείγματα, πραγματοποιήθηκε έλεγχος της γραμμής

εδάφους (soil line) για τους λόγους καναλιών B3 / B4 και B5 / B7 για την ανίχνευση μεταβολών στην επιφάνεια της γης, την υφή, την ποιότητα και την υγρασία του εδάφους.

### 3.3.2. Δείκτες

Στη συνέχεια υπολογίστηκαν περισσότεροι δείκτες οι οποίοι σχετίζονται με το έδαφος και βιβλιογραφικά χρησιμοποιούνται για την ανίχνευση οργανικού άνθρακα και πραγματοποιήθηκαν συγκρίσεις ανάμεσα στις τιμές τους και τις δοσμένες τιμές του SOC από τις μετρήσεις στο εργαστήριο. Οι εξισώσεις υπολογισμού τους παρουσιάζονται αναλυτικά στον πίνακα 7.

Ο ενισχυμένος δείκτης βλάστησης (EVI) είναι παρόμοιος με τον δείκτη κανονικοποιημένης διαφοράς βλάστησης (NDVI) και βελτιωμένος σε σύγκριση με τον NDVI, ιδιαίτερα σε περιοχές με υψηλή βιομάζα, καθώς λαμβάνει υπόψη τις διορθώσεις για τον ανακλαστικότητα του εδάφους και την ατμοσφαιρική διασπορά του φωτός. Υπολογίζεται ως συνδυασμός του μπλε, του κόκκινου και του εγγύς υπέρυθρου καναλιού. Η εξίσωση υπολογισμού του περιέχει τους συντελεστές C1 και C2 για τη διόρθωση της σκέδασης των αερολυμάτων που υπάρχουν στην ατμόσφαιρα και τον συντελεστή L για την προσαρμογή του υποβάθρου του εδάφους και της φυτοκάλυψης και μπορεί να λάβει τιμές από -1 έως 1. Ο EVI μπορεί να χρησιμοποιηθεί για την ανίχνευση της ποσότητας του πρασίνου της βλάστησης σε μια εικόνα, είναι χρήσιμος για παρακολούθηση καλλιεργειών, διαχείριση των υδάτων και ανίχνευση αλλαγών στο περιβάλλον.

Ο δείκτης βλάστησης προσαρμοσμένος στο έδαφος (SAVI) αποτελεί μια βελτιωμένη έκδοση του NDVI, προσαρμοσμένη σε περιοχές με χαμηλή φυτοκάλυψη. Η εξίσωση υπολογισμού του περιλαμβάνει τις τιμές της ανακλαστικότητας στο κόκκινο και το εγγύς υπέρυθρο και οι τιμές που μπορεί να λάβει είναι από -1 έως 1. Ο SAVI έχει τη δυνατότητα να διορθώσει την υπερεκτίμηση της χλωροφύλλης που μπορεί να προκύψει σε περιοχές με χαμηλή βλάστηση ή σε εδάφη με έντονη ανακλώμενη ακτινοβολία από το έδαφος με την προσθήκη ενός συντελεστή προσαρμογής εδάφους L στην εξίσωση του NDVI για να διορθώσει τις επιπτώσεις που τείνουν να επηρεάζουν τα αποτελέσματα. Αυτό το καθιστά ιδιαίτερα χρήσιμο για την εκτίμηση της βλάστησης σε περιοχές με μεγάλη διαφοροποίηση στις συνθήκες του εδάφους. Για τους δείκτες φωτεινότητας BI και BI2 χρησιμοποιήθηκαν οι εξισώσεις του Escadafal (1989), με σκοπό τον εντοπισμό διαφοροποιήσεων στην ανακλαστικότητα στο ορατό τμήμα του φάσματος, για τις περιοχές με γυμνό έδαφος. Βιβλιογραφικά οι δείκτες φωτεινότητας αναφέρονται ως συνδυασμός των καναλιών του ορατού τμήματος του οπτικού φάσματος, ή μέρους του ορατού και του εγγύς υπέρυθρου.

Ο δείκτης γυμνού εδάφους (BSI) χρησιμοποιείται για την ανίχνευση περιοχών με ακάλυπτο έδαφος. οι χαμηλές τιμές του σχετίζονται με πιο υψηλή ή ζωνρή βλάστηση, ενώ οι υψηλές τιμές, με γυμνό έδαφος, ή αποξηραμένη βλάστηση. Ο υπολογισμός του πραγματοποιείται με συνδυασμό του μπλε, κόκκινου, NIR και SWIR και μπορεί να λάβει τιμές από -1 έως 1. Ο συγκεκριμένος δείκτης παρέχει πληροφορίες για την κατάσταση των καλλιεργειών, την πρόσφατη αποψίλωση των δασών, τις μεταβολές στην υγρασία του εδάφους. Μπορεί να χρησιμοποιηθεί για την παρακολούθηση μεταβολών στις χρήσεις γης, την πρόβλεψη πυρκαγιών, την ανίχνευση αλλαγών στο περιβάλλον καθώς και την ανίχνευση κατολισθήσεων ή τον προσδιορισμό της έκτασης της διάβρωσης σε μη φυτεμένες περιοχές Δυστυχώς,

αναδεικνύει επίσης ορισμένα κτίρια, καθιστώντας δύσκολο τον διαχωρισμό του γυμνού εδάφους σε περιοχές που υπάρχει δόμηση.

Οι S. Mirzaee et al., (2016) χρησιμοποίησαν τους δείκτες αργίλου και ανθρακικών αλάτων. Η μέθοδος αναλογίας αργιλικών ορυκτών (Clay Index – CI) είναι ένας γεωλογικός δείκτης για τον εντοπισμό ορυκτών στοιχείων που περιέχουν άργιλο και αλουμίτη, χρησιμοποιώντας δύο υπέρυθρες ζώνες μικρού μήκους κύματος (SWIR). Η μέθοδος αυτή χρησιμοποιείται στη χαρτογράφηση σύνθετων ορυκτών. Επίσης, μετριάζει τις μεταβολές του φωτισμού που οφείλονται στο έδαφος, δεδομένου ότι είναι ένας λόγος. Ο λόγος ανθρακικών αλάτων (Carbonate Index) αφορά τις φασματικές ζώνες του πράσινου και του κόκκινου.

Πίνακας 7. Εξισώσεις υπολογισμού των δεικτών.

Δείκτης	Ορισμός	Εξίσωση σε Landsat 5 (TM)
EVI	$G \frac{(NIR - R)}{(NIR + C1 * R - C2 * B + L)}$	$2.5 * \frac{(B4 - B3)}{(B4 + 6 * B3 - 7.5 * B1 + 1)}$
SAVI	$\frac{(NIR - R)}{(NIR + R + L)} * (1 + L)$	$\frac{(B4 - B3)}{(B4 + B3 + 0.5)} * (1.5)$
BI	$\frac{\sqrt{R^2 + G^2}}{2}$	$\frac{\sqrt{(B3)^2 + (B2)^2}}{2}$
BI2	$\frac{\sqrt{R^2 + G^2 + NIR^2}}{3}$	$\frac{\sqrt{(B3)^2 + (B2)^2 + (B4)^2}}{3}$
BSI	$\frac{((R + SWIR) - (NIR + B))}{((R + SWIR) + (NIR + B))}$	$\frac{(B5 + B3) - (B4 + B1)}{(B5 + B3) + (B4 + B1)}$
CI	SWIR1 / SWIR2	B5 / B7
Cal	R / G	B3 / B2

\* Οι συντελεστές της εξίσωσης του EVI δίνονται βιβλιογραφικά: G = 2.5, C1 = 6, C2 = 7.5, L = 1 και ο συντελεστής του SAVI, L = 0.5

### 3.3.3. Μέθοδοι εκτίμησης SOC

Πραγματοποιήθηκε έλεγχος του βαθμού συσχέτισης ανάμεσα στις τιμές των εικονοστοιχείων ανά κανάλι ή δείκτη, με την τιμή του SOC στο σημείο αυτό. Για το σκοπό αυτό χρησιμοποιήθηκε, αρχικά, η γραμμική παλινδρόμηση για τις τιμές κάθε καναλιού και κάθε δείκτη ξεχωριστά σε σύγκριση με το SOC και σε δεύτερο χρόνο χρησιμοποιήθηκε ο ταξινομητής των τυχαίων δασών, με σκοπό τη συσχέτιση των τιμών της ανακλώμενης ακτινοβολίας και των δεικτών (ανεξάρτητες μεταβλητές) με τις τιμές του SOC (εξαρτημένη μεταβλητή). Τα δείγματα χωρίστηκαν σε αναλογία 80% train - 20% test και εφαρμόστηκε cross validation. Το CV χρησιμοποιήθηκε για την εξακρίβωση της ευαισθησίας του μοντέλου και της αξιοπιστίας των δεδομένων. Χρησιμοποιήθηκε η μέθοδος grid-search για την εύρεση των βέλτιστων παραμέτρων για το μοντέλο παλινδρόμησης Random Forest.

## Κεφάλαιο 4 – Αποτελέσματα και αξιολόγηση

### 4.1. Αποτελέσματα ανίχνευσης περιοχών γυμνου εδάφους

Το πλήθος των εικόνων από τις οποίες αποτελούνται οι χρονοσειρές παρουσιάζεται στον πίνακα 8.

Πίνακας 8. Πλήθος εικόνων ανά χρονοσειρά και ποσοστό νεφοκάλυψης.

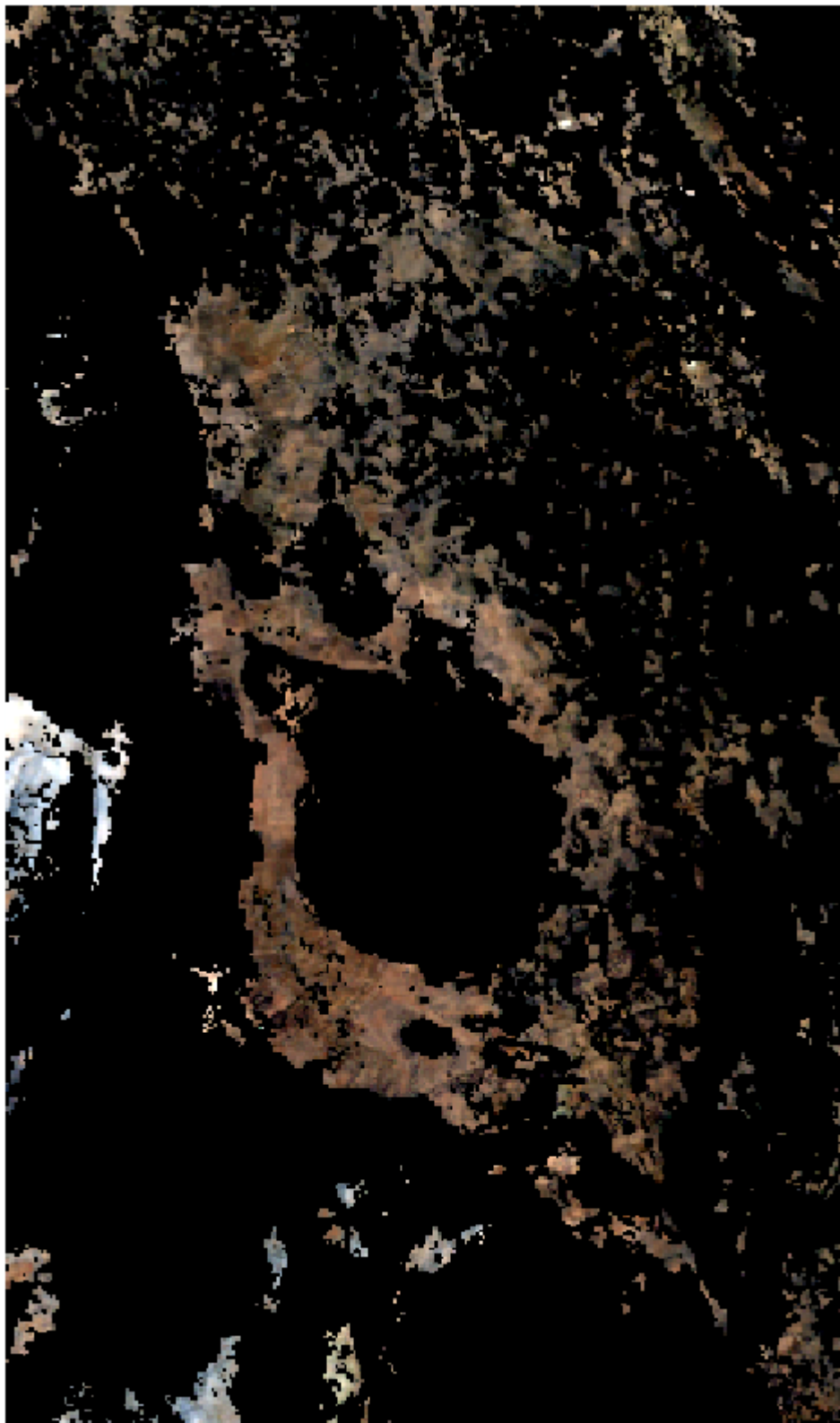
χρονοσειρά	1995	1994-1996	1993-1997	1992-1998
νεφοκάλυψη 10%	17	51	82	106
νεφοκάλυψη 20%	20	66	111	161

Στον πίνακα 9 παρουσιάζεται το πλήθος των σημείων των δειγμάτων εδάφους, τα οποία συμπίπτουν με εικονοστοιχεία γυμνού εδάφους για καθένα από τα πειράματα που πραγματοποιήθηκαν. Για τις περιπτώσεις, στις οποίες υπάρχουν διαθέσιμα εικονοστοιχεία από περισσότερες από μια εικόνες, το εικονοστοιχείο που εμπεριέχεται στην παραγόμενη εικόνα, είναι αυτό με την ενδιάμεση τιμή. Παρατηρείται ότι υπάρχει σημαντική διαφοροποίηση στα δείγματα που συμπεριλαμβάνονται στα πειράματα για τα πιο ελαστικά κατώφλια, σε σύγκριση με τα πιο αυστηρά, καθώς και στις χρονοσειρές που περιλαμβάνουν περισσότερα έτη. Όμως η διαφοροποίηση αυτή τείνει να εξαλειφθεί ανάμεσα στις χρονοσειρές 1993-1997 και 1992-1998.

Ο μέγιστος αριθμός δειγμάτων που αξιοποιήθηκε είναι 48, για το κατώφλι 4 και τις χρονοσειρές 1993-1997 και 1992-1998, ενώ το κατώφλι 5, το οποίο είναι και το αυστηρότερο του πειράματος, δεν αναγνώρισε κανένα εικονοστοιχείο εδάφους στις περιοχές των δειγμάτων για τις χρονοσειρές 1995 και 1994-1996 και μόλις 1 για τις σειρές 1993-1997 και 1992-1998, γεγονός που το καθιστά ακατάλληλο για σύγκριση, λόγω απουσίας επαρκών στοιχείων και συνεπώς απαγορευτικό για την εξαγωγή συμπερασμάτων. Αρκετά μικρά για ασφαλή σύγκριση είναι και τα δείγματα σε όλες τις χρονοσειρές για το κατώφλι 1. Παρατηρείται ότι τα κατώφλια 2 και 3 έδωσαν τα ίδια αποτελέσματα, γεγονός που οδηγεί στο συμπέρασμα ότι η διαφοροποίηση των τιμών μεταξύ των 2 κατωφλιών, δεν είναι σημαντική σε βαθμό που να επηρεάσει το αποτέλεσμα.

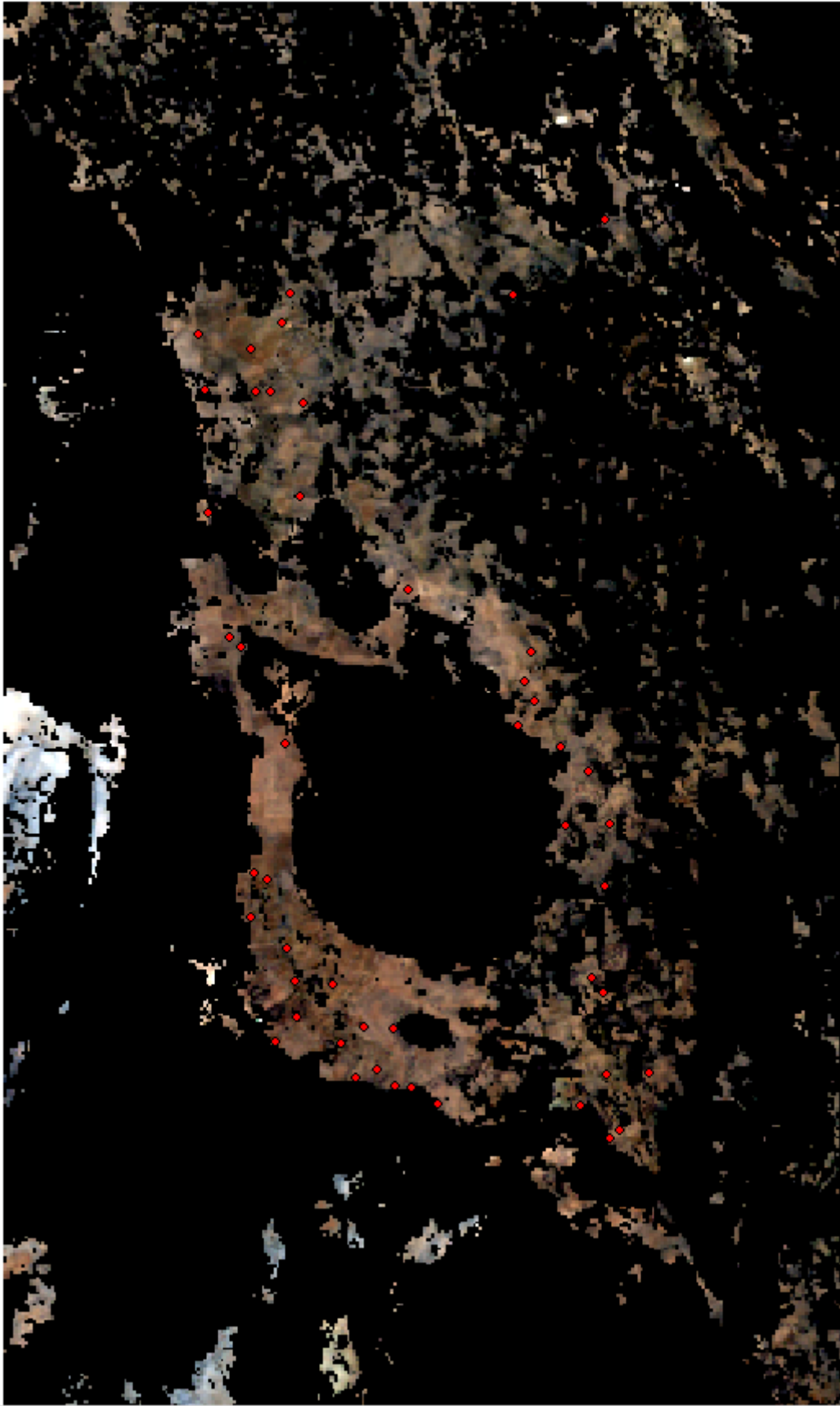
Πίνακας 9. Πλήθος δειγμάτων που αξιοποιήθηκαν ανά πείραμα

threshold	1995	1994 - 1996	1993 - 1997	1992 - 1998
1	7	9	13	16
2	26	43	44	44
3	26	43	44	44
4	41	46	48	48
5	0	0	1	1
6	22	31	37	39



Εικόνα 5. Περιοχές γυμνού εδάφους. Έγχρωμο σύνθετο 3-2-1 για τη χρονοσειρά 1992-1998, κατώφλι 4.





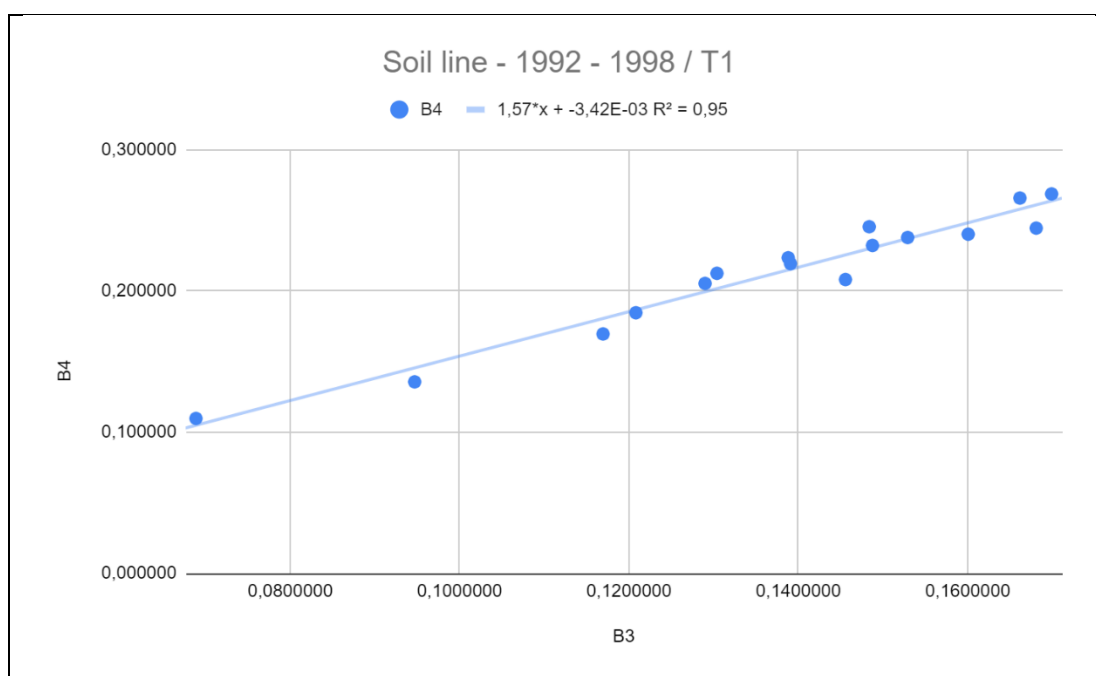
*Εικόνα 6. Τα ενεργά δείγματα για τη χρονοσειρά 1992-1998, κατώφλι 4, στο έγχρωμο σύνθετο 3-2-1.*

Η γραμμή εδάφους έδωσε ικανοποιητικά αποτελέσματα για την πλειοψηφία των πειραμάτων, τόσο μεταξύ των καναλιών B3 και B4, όσο και μεταξύ των B5 και B7. Οι συγκρίσεις που παρουσιάζονται παρακάτω, αφορούν τις χρονοσειρές 1994-1996, 1993-1997, 1992-1998 και τα κατώφλια 1, 2, 3, 4, 6, οι τιμές για το έτος 1995, καθώς και για το κατώφλι 5 δεν απεικονίζονται, διότι περιλαμβάνουν μικρό πλήθος σημείων, ανεπαρκές για την εξαγωγή συμπερασμάτων. Αναλυτικά, για τη σύγκριση B3 / B4, ο συντελεστής προσδιορισμού R<sup>2</sup>, ο οποίος εκφράζει την ακρίβεια (accuracy) του συγκεκριμένου λόγου είναι μεγαλύτερος από 0,75 σε όλες τις συγκρίσεις, αρκετά κοντά για τις διαφορετικές χρονοσειρές σε κάθε κατώφλι. Υψηλότερες τιμές παρατηρούνται στο κατώφλι 6 (0,961 για το 1993-1997) και στο κατώφλι 1 (0,959 για το 1993-1997) και ακολουθούν τα κατώφλια 2, 3 και 4.

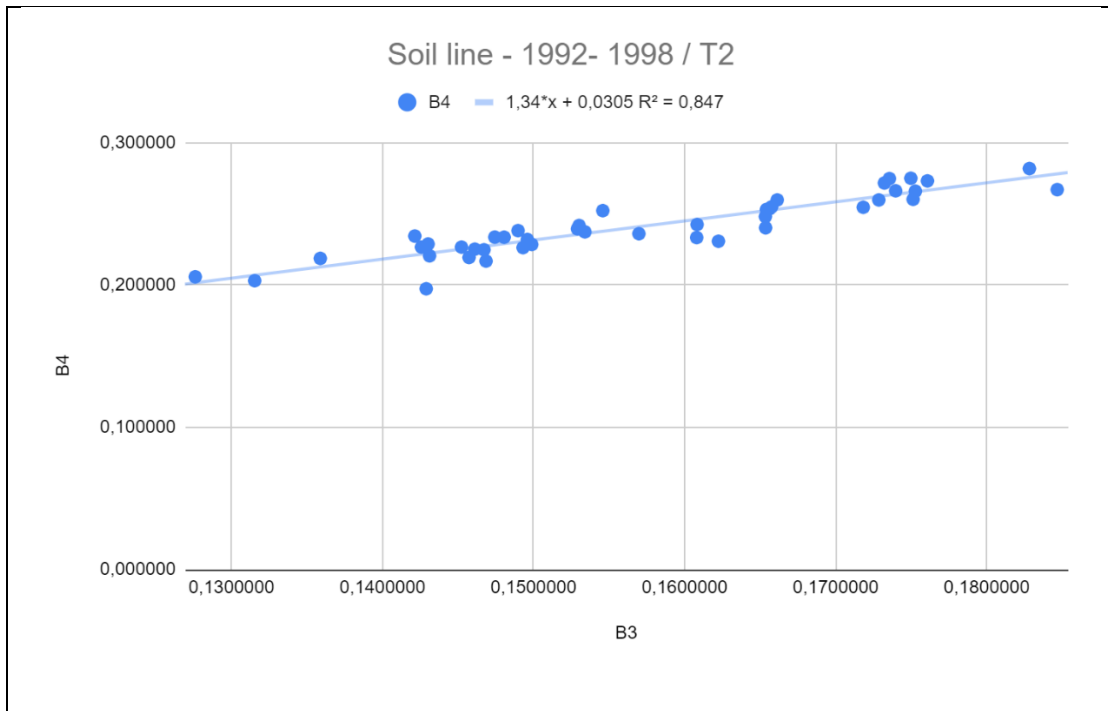
Πίνακας 10. Τιμές R<sup>2</sup> για τη σύγκριση B3 / B4.

	1994 - 1996	1993 - 1997	1992 - 1998
T1	0,938	0,959	0,950
T2	0,843	0,845	0,847
T3	0,841	0,828	0,832
T4	0,775	0,756	0,748
T6	0.929	0.961	0,950

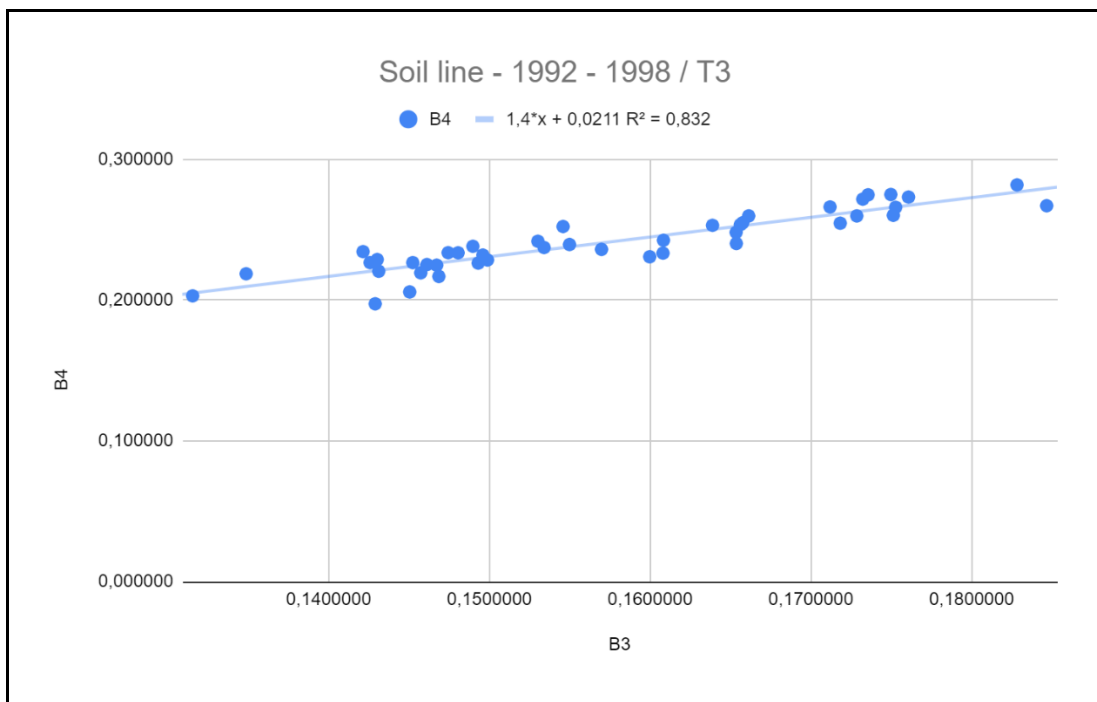
Διάγραμμα 1. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T1.



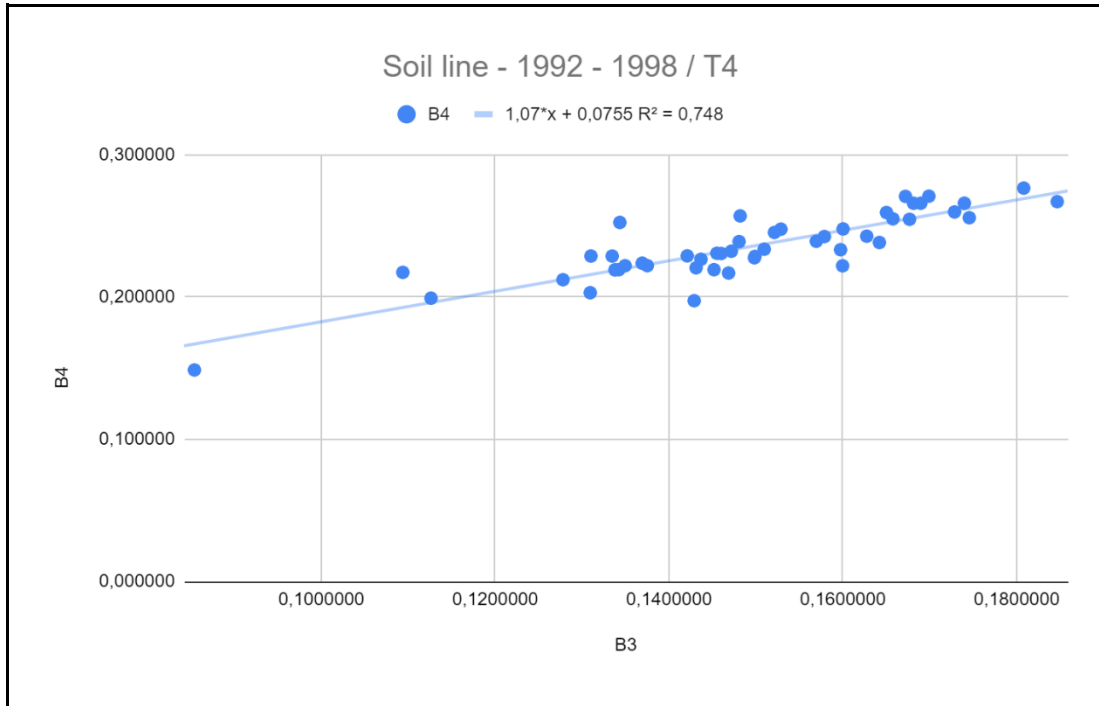
Διάγραμμα 2. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T2.



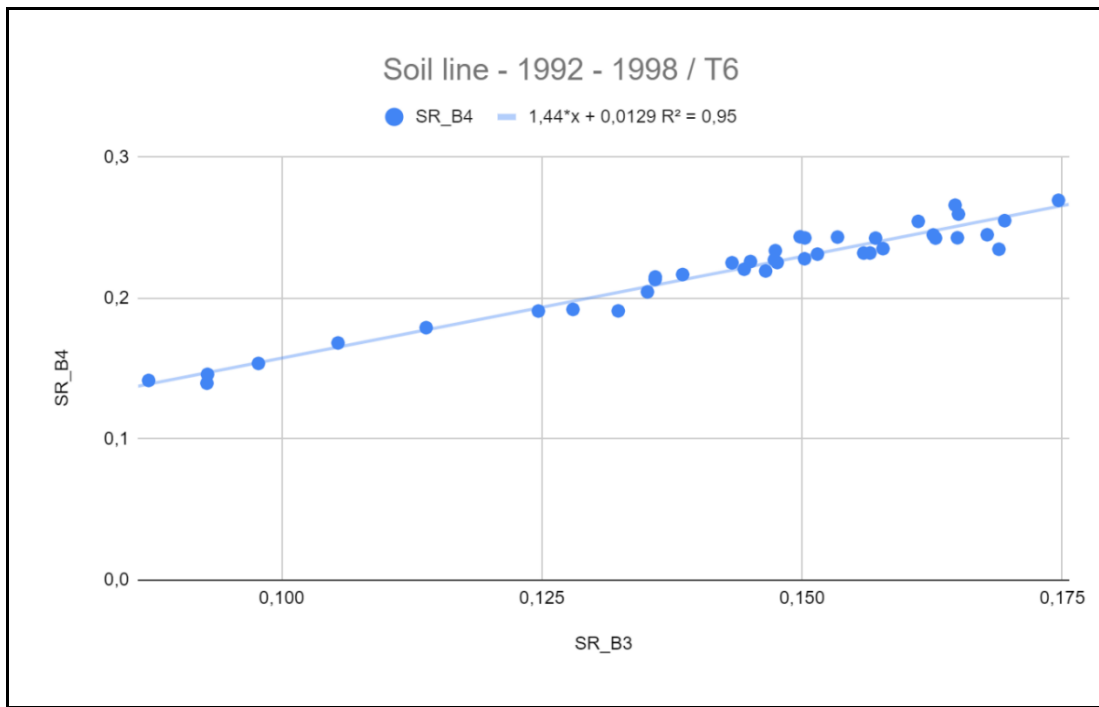
Διάγραμμα 3. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T3.



Διάγραμμα 4. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T4.



Διάγραμμα 5. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T6.

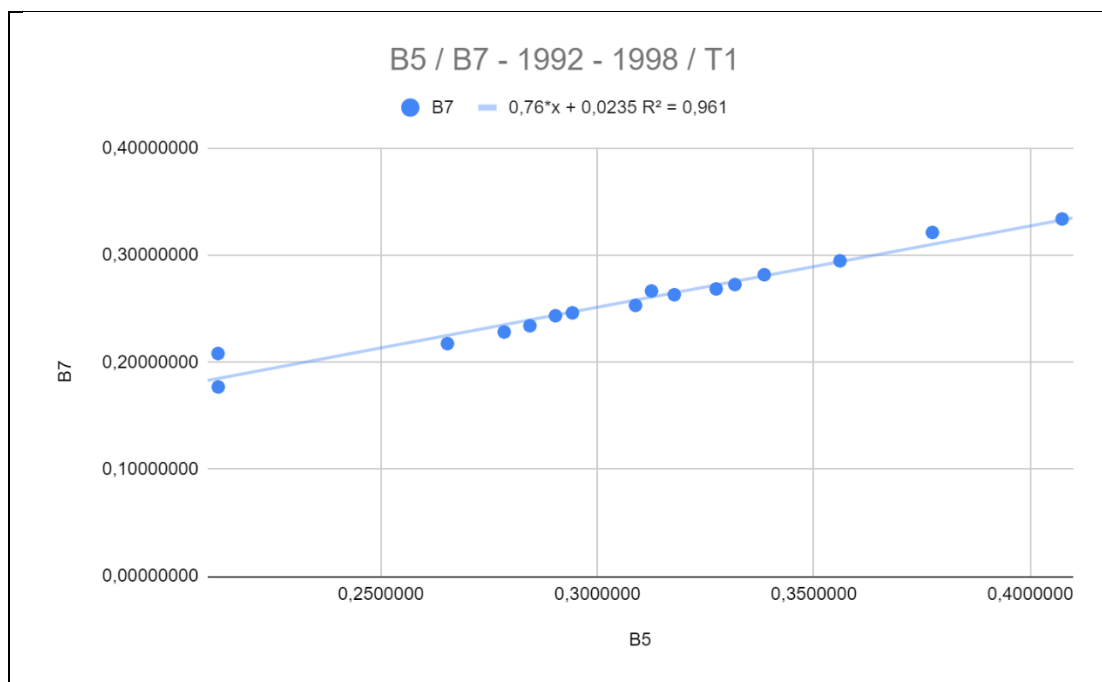


Για τη σύγκριση μεταξύ των SWIR καναλιών B5 και B7, ο συντελεστής  $R^2$  είναι ικανοποιητικός στις περισσότερες περιπτώσεις. Παρατηρούνται χαμηλά ποσοστά συσχέτισης στη χρονοσειρά 1994 - 1996 για τα thresholds 2, 3, 4, ενώ στις χρονοσειρές 1993 - 1997 και 1992 - 1998 τα αποτελέσματα είναι ικανοποιητικά, με ελαφρώς υψηλότερες τιμές στην 1992 - 1998 για τις περισσότερες περιπτώσεις. Τα thresholds 1 και 6 συγκεντρώνουν τιμές  $R^2$  μεγαλύτερες από 0,9 σε όλες τις περιπτώσεις.

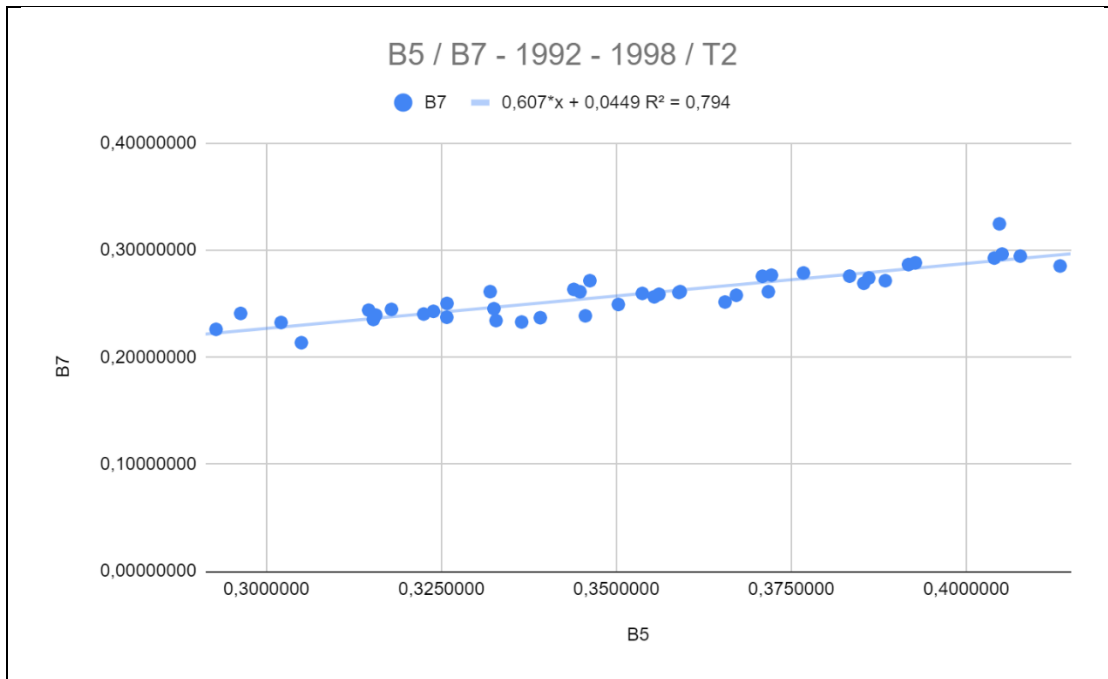
Πίνακας 11. Τιμές  $R^2$  για τη σύγκριση B5 / B7

	1994 - 1996	1993 - 1997	1992 - 1998
T1	0,992	0,987	0,961
T2	0,511	0,713	0,794
T3	0,511	0,726	0,797
T4	0,459	0,684	0,815
T6	0,971	0,957	0,959

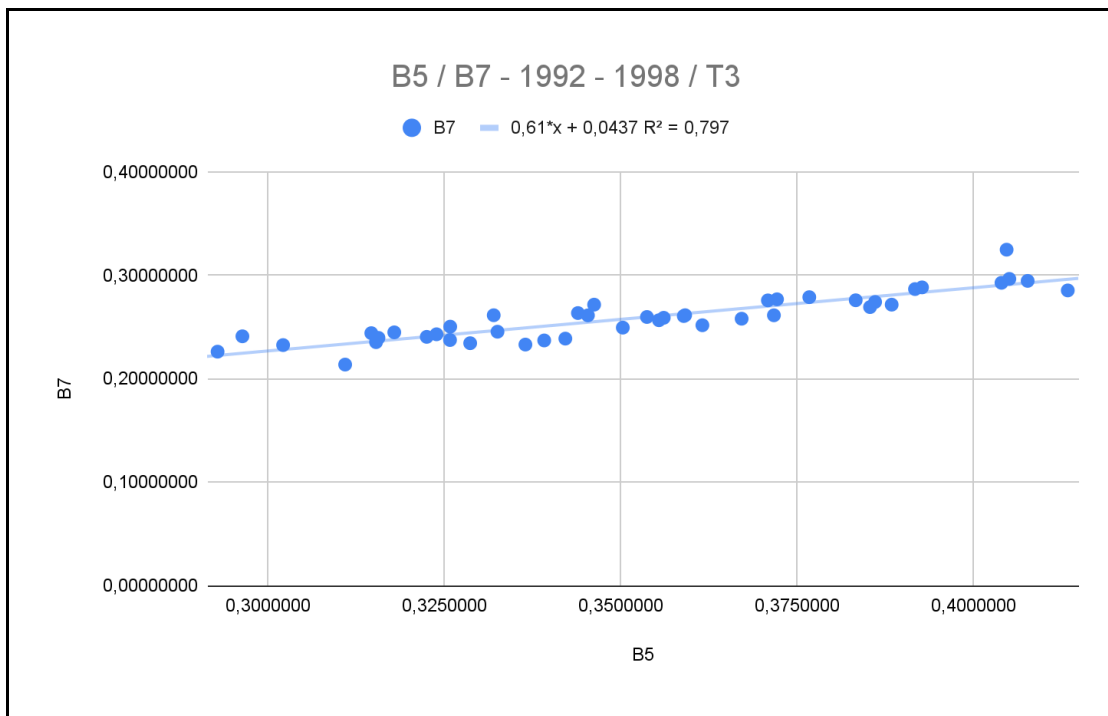
Διάγραμμα 6. B5 / B7 - 1992 - 1998 / T1



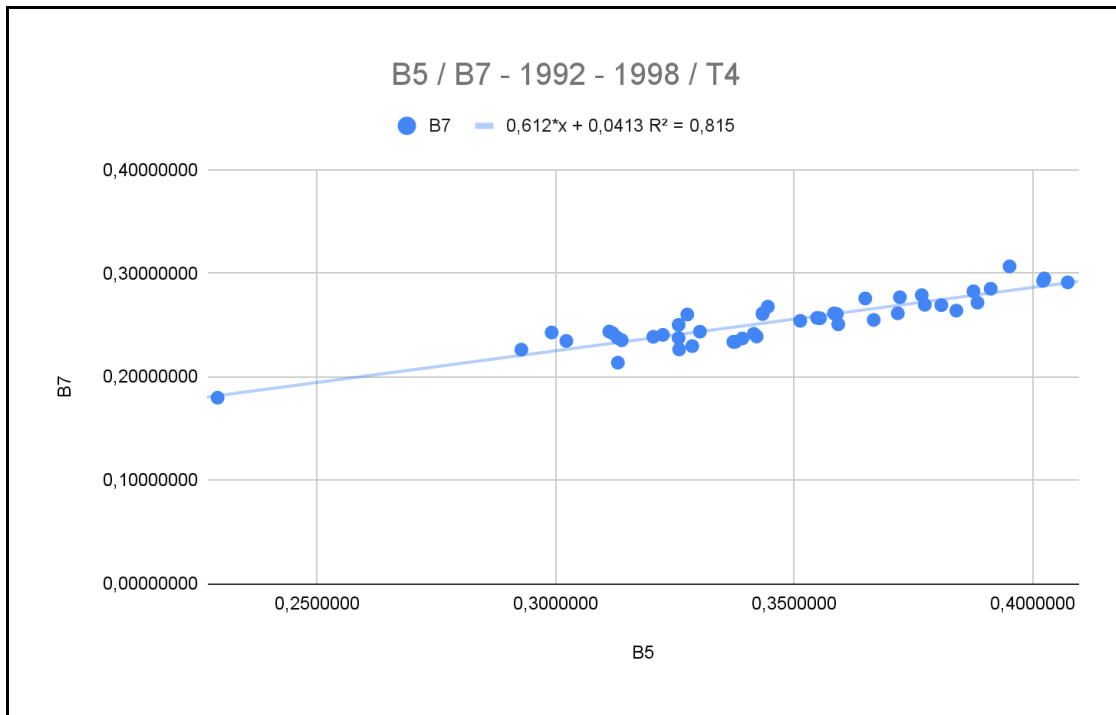
Διάγραμμα 7. B5 / B7 - 1992 - 1998 / T2



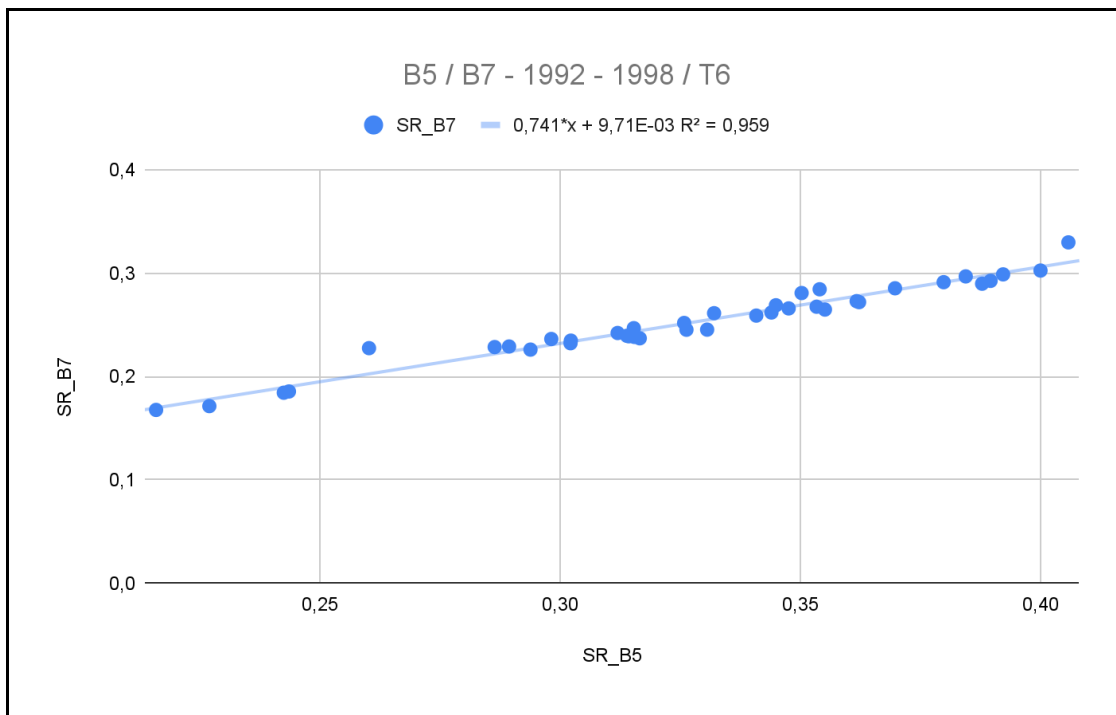
Διάγραμμα 8. B5 / B7 - 1992 - 1998 / T3



Διάγραμμα 9. B5 / B7 - 1992 - 1998 / T4



Διάγραμμα 10. B5 / B7 - 1992 - 1998 / T6



Από τον παραπάνω έλεγχο, προκύπτει ότι η ανίχνευση των pixel γυμνού εδάφους είναι επιτυχής και επιλέγονται τα thresholds 3, 4 και 6 στη χρονοσειρά 1992 - 1998 για σύγκριση με τις τιμές SOC των δειγμάτων. Το threshold 1 δεν χρησιμοποιήθηκε περεταίρω, επειδή περιλαμβάνει πολύ μικρό δείγμα και εξετάστηκε μέχρι αυτό το σημείο ενδεικτικά. Το threshold 2 δεν χρησιμοποιήθηκε επειδή οι τιμές του είναι παρεμφερείς με το T3 και δεν προκύπτει αξιοσημείωτη διαφορά στα μεταξύ τους αποτελέσματα. Ανάμεσα στα 3 πειράματα που χρησιμοποιήθηκαν για τη σύγκριση με το SOC, περισσότερο αξιόπιστο φαίνεται το T6, το οποίο παρουσιάζει συγκεντρωτικά το υψηλότερο  $R^2$ .

## 4.2. Μοντέλα εκτίμησης SOC

### 4.2.1 Γραμμική παλινδρόμηση

Αρχικά εξετάστηκε ο βαθμός σύγκλισης της συνθετικής εικόνας με τα δείγματα, με τη μέθοδο της γραμμικής παλινδρόμησης. Πραγματοποιήθηκαν συγκρίσεις ανάμεσα σε κάθε κανάλι και κάθε δείκτη ξεχωριστά με το SOC. Τα αποτελέσματα δεν εμφανίζουν σύνδεση ανάμεσα σε κανένα από τα κανάλια του Landsat 5 με το εργαστηριακά μετρημένο SOC. Για το κατώφλι 3, ο συντελεστής  $R^2$  είναι χαμηλότερος από 0,1 σε όλες τις συγκρίσεις, για το κατώφλι 4, χαμηλότερος από 0,06 και για το threshold σημειώνει, επίσης, χαμηλές τιμές με υψηλότερη την  $R^2 = 0,25$  για τον δείκτη SAVI. Οι παραπάνω τιμές υποδεικνύουν ότι δεν υπάρχει γραμμική συσχέτιση ανάμεσα στην ανακλώμενη ακτινοβολία των εικόνων Landsat 5 για τις περιοχές γυμνού εδάφους και τις τιμές του SOC. Επομένως, η συγκεκριμένη μέθοδος δεν μπορεί να οδηγήσει σε συμπεράσματα για την περιεκτικότητα μιας έκτασης σε οργανικό άνθρακα.

Πραγματοποιήθηκε σύγκριση ανάμεσα στα κανάλια B1, B2, B3, B4, B5, B7, τους δείκτες NBR2, NDVI, VNSIR, EVI, SAVI, BI, BI2, BSI, CI, Cal και τις τιμές του SOC. Για την αξιολόγηση των αποτελεσμάτων χρησιμοποιήθηκαν στατιστικοί δείκτες, ώστε η σύγκριση να επιτευχθεί μέσω της ποσοτικής διαφοράς τους. Συγκεκριμένα χρησιμοποιήθηκε ο συντελεστής προσδιορισμού  $R^2$  κάθε πειράματος, η στατιστική F για κάθε μεταβλητή και η στατιστική σημαντικότητα (P-Value) για κάθε μεταβλητή. Οι συγκρίσεις υποδεικνύουν πολύ χαμηλό βαθμό συσχέτισης ανάμεσα στις παραμέτρους που εξετάστηκαν και το SOC, σε όλες τις περιπτώσεις.

Για τη χρονοσειρά 1992 - 1998 και κατώφλι 3, ο συντελεστής  $R^2$  είναι 0,348 και ο δείκτης F-statistic ισούται με 0,9003. Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα από τη σύγκριση με συνήθη ελάχιστα τετράγωνα (OLS Regression), ο οποίος περιλαμβάνει και αναλυτικά στοιχεία για κάθε μεταβλητή. Υψηλότερη συσχέτιση, ανάμεσα στους συντελεστές, παρουσιάζει ο δείκτης SAVI και ακολουθούν τα κανάλια B3 και B4 και οι δείκτες BI και EVI, με χαμηλές τιμές ωστόσο.

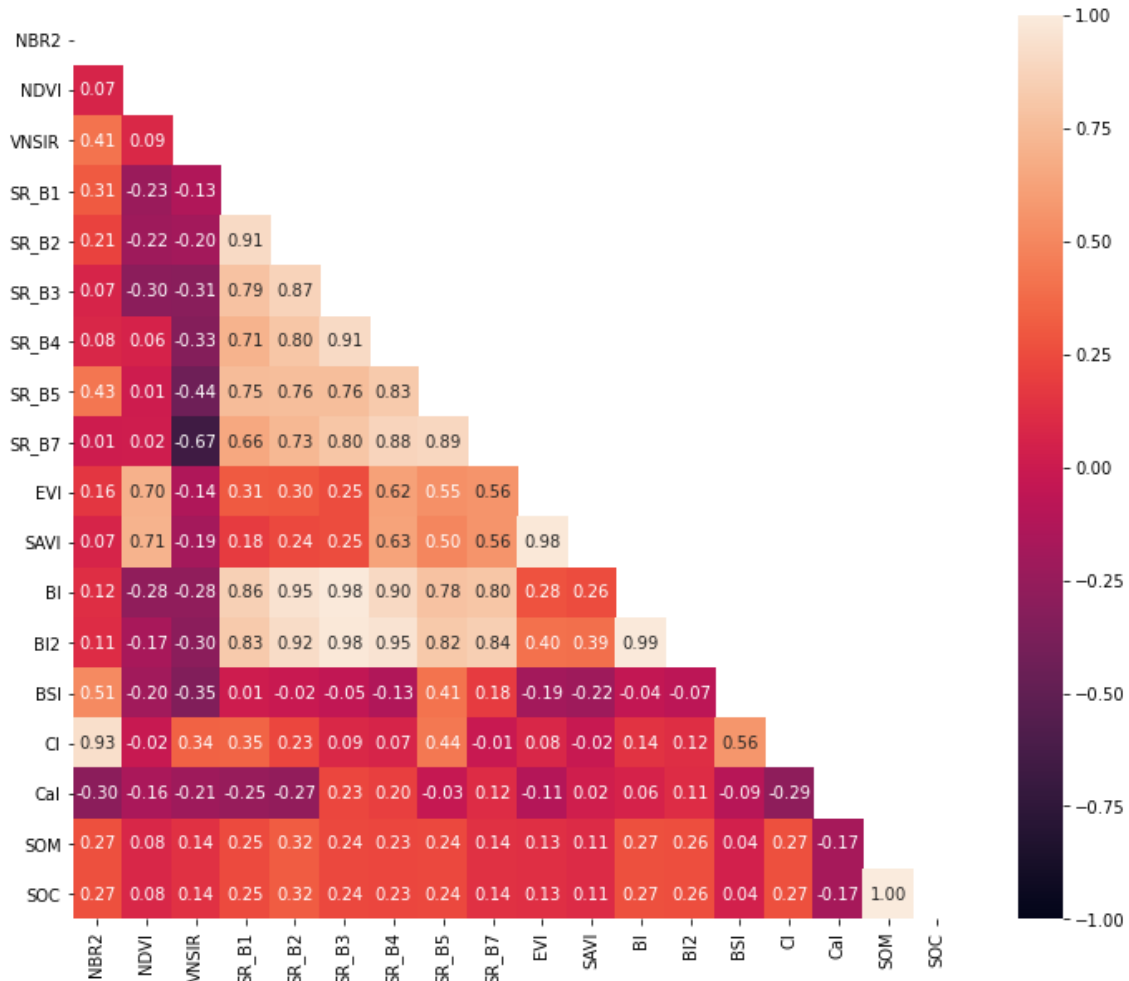


Πίνακας 12. OLS Regression Results 1992-1998 / T3

Dep. Variable:	SOC	R-squared:	0.348			
Model:	OLS	Adj. R-squared:	-0.039			
Method:	Least Squares	F-statistic:	0.9003			
Date:	Sun, 08 Oct 2023	Prob (F-statistic):	0.577			
Time:	16:26:43	Log-Likelihood:	15.238			
No. Observations:	44	AIC:	3.525			
Df Residuals:	27	BIC:	33.86			
Df Model:	16					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-3.8997	3.243	-1.203	0.240	-10.553	2.753
NBR2	-0.1538	0.550	-0.279	0.782	-1.283	0.976
NDVI	0.3477	0.330	1.053	0.302	-0.330	1.026
VNSIR	0.6110	0.689	0.887	0.383	-0.802	2.024
SR_B1	3.6314	2.542	1.429	0.165	-1.584	8.847
SR_B2	1.1640	20.853	0.056	0.956	-41.624	43.952
SR_B3	15.2559	40.529	0.376	0.710	-67.903	98.415
SR_B4	-15.6637	33.903	-0.462	0.648	-85.227	53.899
SR_B5	-3.6193	5.851	-0.619	0.541	-15.624	8.386
SR_B7	2.0621	4.128	0.500	0.621	-6.408	10.532
EVI	-13.4479	9.575	-1.404	0.172	-33.094	6.198
SAVI	20.4091	11.257	1.813	0.081	-2.688	43.506
BI	-13.2736	75.906	-0.175	0.862	-	142.473
					169.020	
BI2	8.4063	69.318	0.121	0.904	-	150.635
					133.822	
BSI	1.6445	2.943	0.559	0.581	-4.393	7.682
CI	0.9165	1.732	0.529	0.601	-2.637	4.470
Cal	-1.0109	2.285	-0.442	0.662	-5.699	3.678
Omnibus:	3.033	Durbin-Watson:	1.524			
Prob(Omnibus):	0.220	Jarque-Bera (JB):	2.168			
Skew:	0.531	Prob(JB):	0.338			
Kurtosis:	3.235	Cond. No.	7.47e+03			

Ο έλεγχος της συνδυαστικότητας των συντελεστών με τις τιμές του SOC δείχνει τεράστια διασπορά και συνεπώς υψηλό βαθμό τυχαιότητας, γεγονός που επιβεβαιώνει τις τιμές  $R^2$ , F-statistic και P-Values (διάγραμμα 11).

Παρακάτω παρουσιάζεται σε γραφική απεικόνιση ο βαθμός συσχέτισης μεταξύ των μεταβλητών (εικόνα 6). Παρατηρείται, επίσης χαμηλή συσχέτιση όλων των μεταβλητών με το SOC, μικρότερο του 0,25 κατ' απόλυτη τιμή.

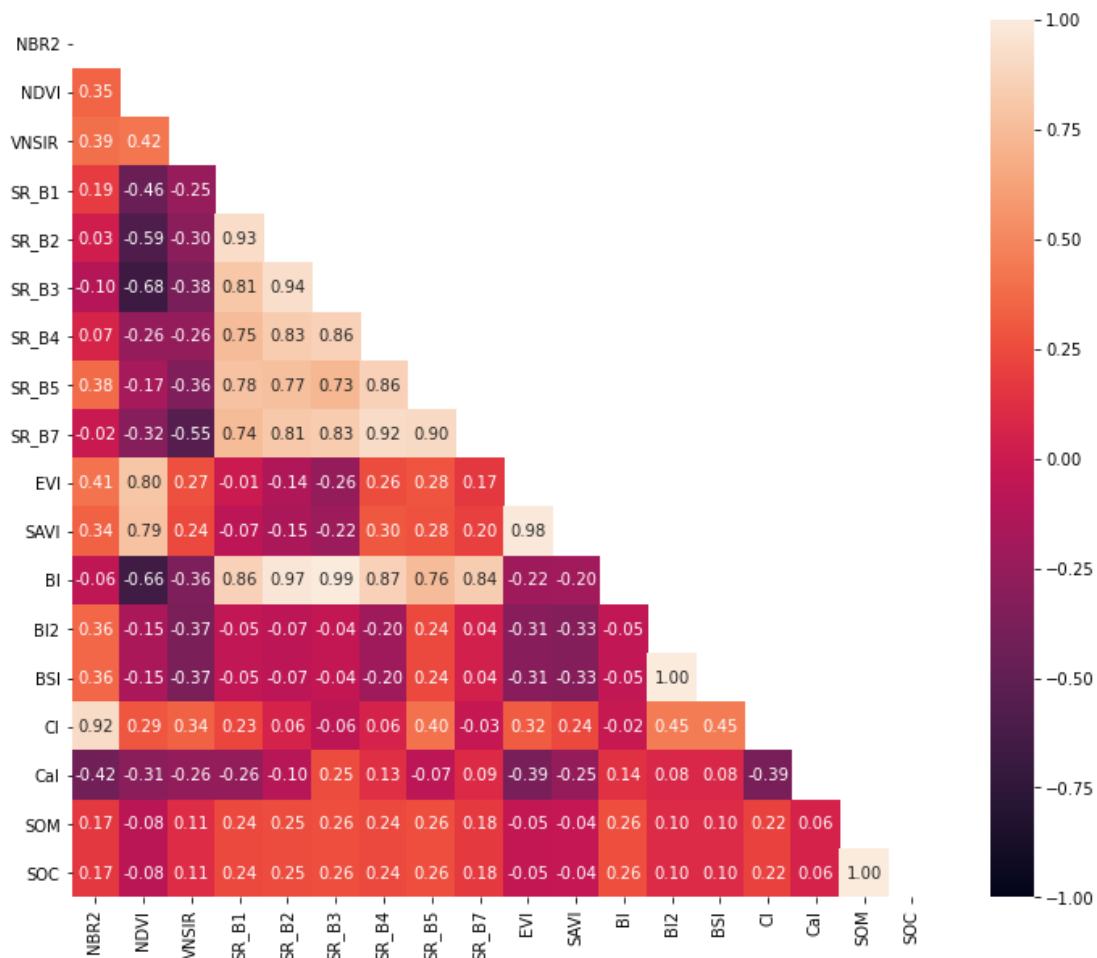


Εικόνα 7. Πίνακας συσχέτισης 1992 – 1998 / T3

Ακολουθούν τα αποτελέσματα για τη χρονοσειρά 1992 - 1998, με το threshold 4, στον πίνακα της γραμμικής παλινδρόμησης με συνήθη ελάχιστα τετράγωνα (πίνακας 12), τον πίνακα σύγκρισης (εικόνα 7), διάγραμμα (διάγραμμα 12). Το  $R^2$  είναι 0,343 και το F-statistic 1.115. Οι τιμές και των δύο μεγεθών είναι πολύ κοντά στην προηγούμενη περίπτωση, με μεταβλητές με την υψηλότερη συσχέτιση αυτή τη φορά τον δείκτη BI, τα κανάλια B3 και B2 και χαμηλότερα τον δείκτη SAVI.

Πίνακας 13. OLS Regression Results 1992-1998 - T4

<b>Dep. Variable:</b>	<b>SOC</b>	<b>R-squared:</b>	<b>0.343</b>			
<b>Model:</b>	OLS	Adj. R-squared:	0.035			
<b>Method:</b>	Least Squares	F-statistic:	1.115			
<b>Date:</b>	Sun, 08 Oct 2023	Prob (F-statistic):	0.383			
<b>Time:</b>	16:21:10	Log-Likelihood:	17.195			
<b>No. Observations:</b>	48	AIC:	-2.390			
<b>Df Residuals:</b>	32	BIC:	27.55			
<b>Df Model:</b>	15					
<b>Covariance Type:</b>	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
<b>const</b>	-5.2528	3.264	-1.609	0.117	-11.902	1.396
<b>NBR2</b>	-0.0988	0.380	-0.260	0.796	-0.872	0.675
<b>NDVI</b>	0.0867	0.468	0.185	0.854	-0.866	1.039
<b>VNSIR</b>	0.5919	0.341	1.734	0.093	-0.103	1.287
<b>SR_B1</b>	5.8089	3.455	1.681	0.102	-1.229	12.847
<b>SR_B2</b>	32.4299	36.130	0.898	0.376	-41.164	106.024
<b>SR_B3</b>	62.9998	66.941	0.941	0.354	-73.354	199.354
<b>SR_B4</b>	-1.4679	12.655	-0.116	0.908	-27.245	24.309
<b>SR_B5</b>	-7.9194	8.460	-0.936	0.356	-25.152	9.314
<b>SR_B7</b>	-1.1040	5.811	-0.190	0.851	-12.940	10.732
<b>EVI</b>	-7.0276	6.592	-1.066	0.294	-20.455	6.399
<b>SAVI</b>	13.1890	8.402	1.570	0.126	-3.925	30.303
<b>BI</b>	-88.2796	96.189	-0.918	0.366	-	107.650
					284.210	
<b>BI2</b>	2.2135	1.374	1.611	0.117	-0.584	5.011
<b>BSI</b>	2.2135	1.374	1.611	0.117	-0.584	5.011
<b>CI</b>	-0.4641	1.886	-0.246	0.807	-4.305	3.377
<b>Cal</b>	0.1091	1.721	0.063	0.950	-3.396	3.614
<b>Omnibus:</b>	4.250	Durbin-Watson:	1.577			
<b>Prob(Omnibus):</b>	0.119	Jarque-Bera (JB):	3.136			
<b>Skew:</b>	0.572	Prob(JB):	0.209			
<b>Kurtosis:</b>	3.511	Cond. No.	5.83e+18			



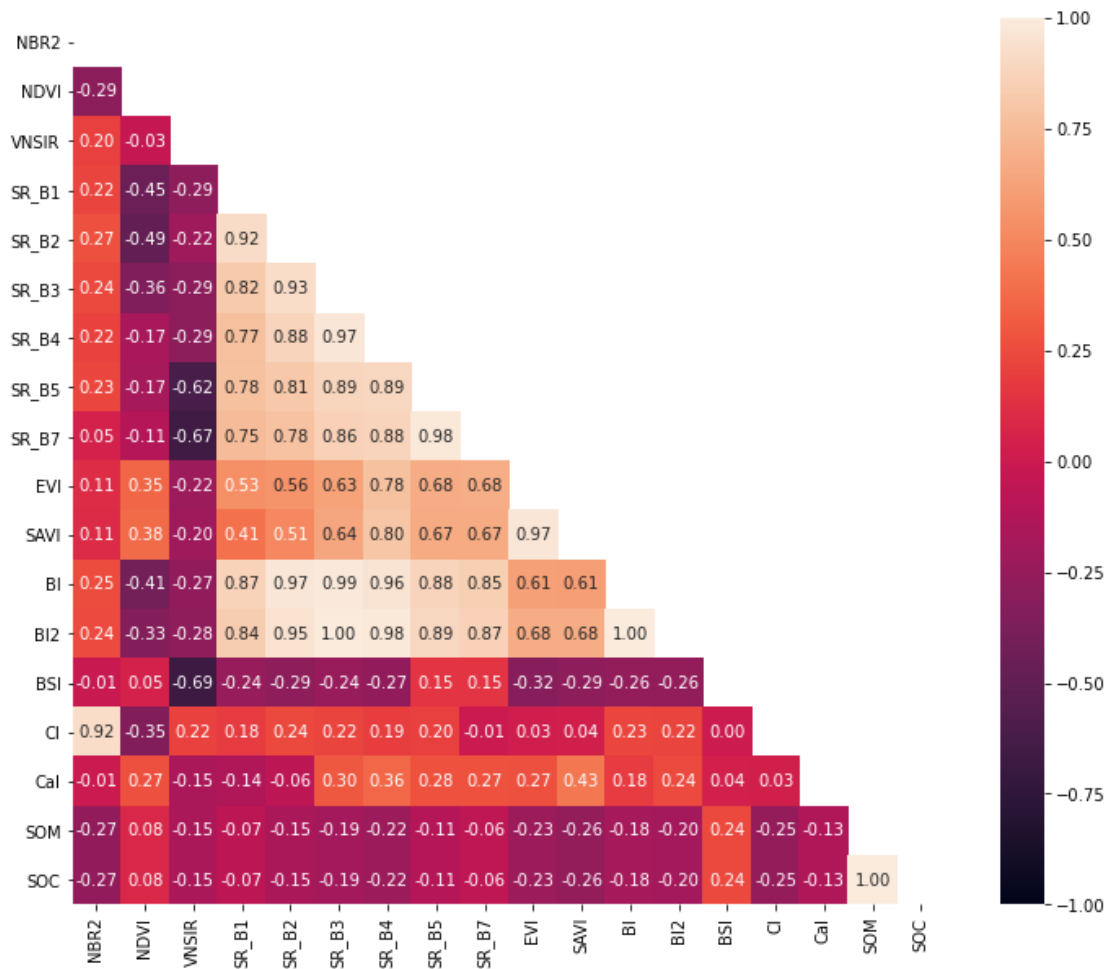
Εικόνα 8. Πίνακας συσχέτισης 1992 – 1998 / T4

Για τη χρονοσειρά 1992 - 1998, με το threshold 6, το  $R^2$  είναι 0,599, η συγκεκριμένη τιμή είναι σημαντικά υψηλότερη από τις προηγούμενες περιπτώσεις, αλλά είναι οριακή για την εξαγωγή ενός ασφαλούς συμπεράσματος για τη συσχέτιση των παραμέτρων που εξετάστηκαν με το SOC. Το F-statistic είναι ίσο με 2,055, τιμή επίσης υψηλότερη από τις προηγούμενες. Οι ψηλότερες συσχετίσεις παρουσιάζονται ανάμεσα στους δείκτες NBR2=-0,27, SAVI=-0,26, CI=-0,25, BSI=0,24, EVI=0,23 και το κανάλι B4=-0,22. Εξετάζοντας τα P-Values των συνιστωσών, σημαντικότερος στατιστικά φαίνεται να είναι ο δείκτης BI2, που περιλαμβάνει το πράσινο, το κόκκινο και το NIR, με πολύ υψηλή σύγκλιση, ίση με 0,008, γεγονός λογικό, αφού και το NIR έχει πολύ υψηλή σύγκλιση, ίση με 0,022. Σημαντικά υψηλή σύγκλιση έχει και ο δείκτης BI, ο οποίος εξαρτάται από το πράσινο και το κόκκινο κανάλι.

Πίνακας 14. OLS Regression 1992-1998 / T6

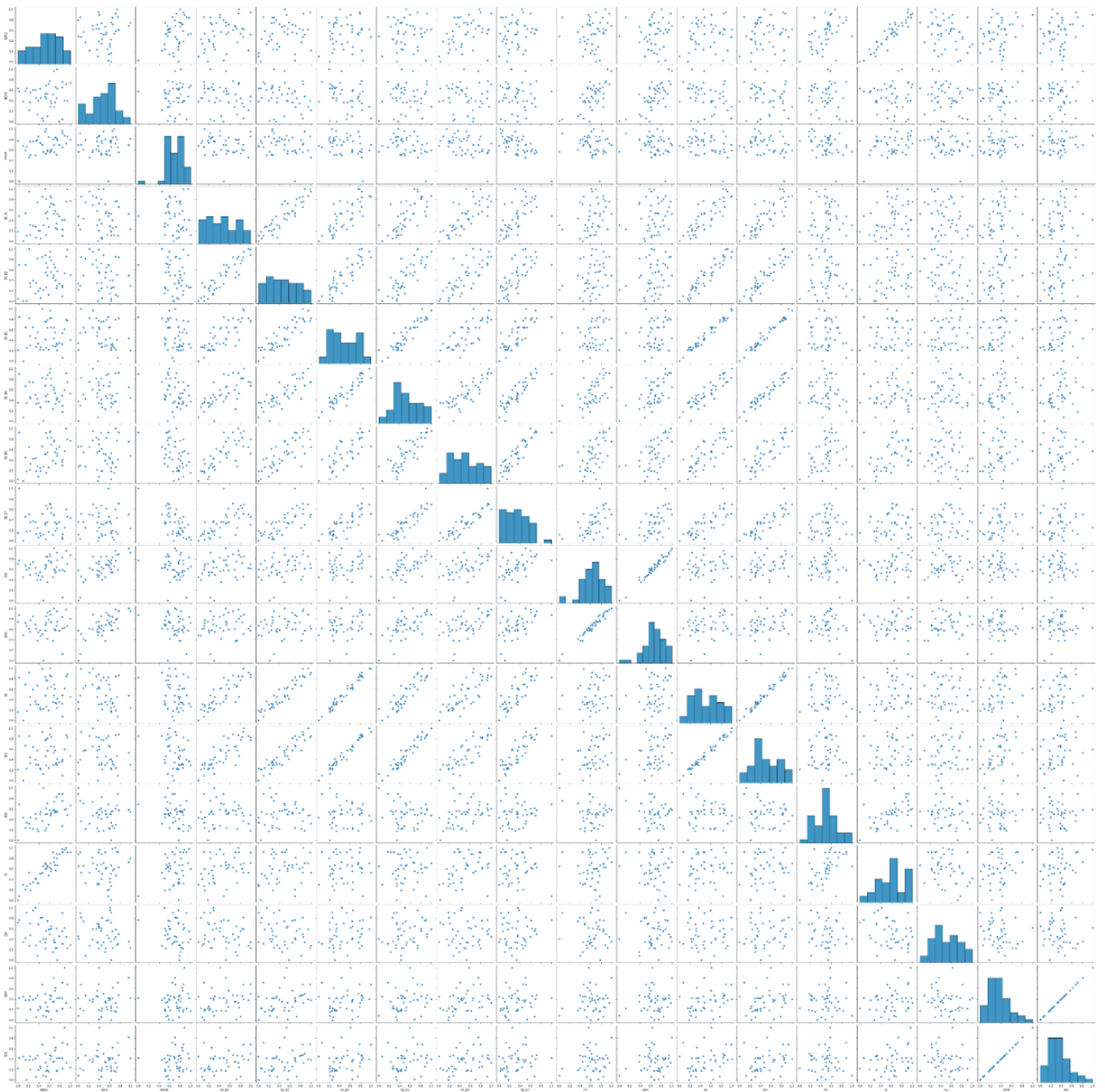
<b>Dep. Variable:</b>	<b>SOC</b>	<b>R-squared:</b>	<b>0.599</b>			
<b>Model:</b>	OLS	Adj. R-squared:	0.308			
<b>Method:</b>	Least Squares	F-statistic:	2.055			
<b>Date:</b>	Sun, 08 Oct 2023	Prob (F-statistic):	0.0586			
<b>Time:</b>	16:35:23	Log-Likelihood:	22.458			
<b>No. Observations:</b>	39	AIC:	-10.92			
<b>Df Residuals:</b>	22	BIC:	17.36			
<b>Df Model:</b>	16					
<b>Covariance Type:</b>	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
<b>const</b>	6.5528	3.853	1.701	0.103	-1.438	14.543
<b>NBR2</b>	0.4004	0.637	0.629	0.536	-0.920	1.721
<b>NDVI</b>	0.1683	0.381	0.441	0.663	-0.623	0.959
<b>VNSIR</b>	-0.2065	0.353	-0.584	0.565	-0.939	0.526
<b>SR_B1</b>	-4.1778	5.187	-0.805	0.429	-14.935	6.579
<b>SR_B2</b>	-6.5415	20.489	-0.319	0.753	-49.034	35.951
<b>SR_B3</b>	-9.1584	36.221	-0.253	0.803	-84.277	65.960
<b>SR_B4</b>	-98.3674	39.930	-2.464	0.022	-181.177	-15.558
<b>SR_B5</b>	-2.3041	7.157	-0.322	0.751	-17.146	12.538
<b>SR_B7</b>	0.5919	6.875	0.086	0.932	-13.667	14.851
<b>EVI</b>	11.7139	10.593	1.106	0.281	-10.255	33.683
<b>SAVI</b>	-13.1129	9.878	-1.328	0.198	-33.598	7.372
<b>BI</b>	-177.9023	96.845	-1.837	0.080	-378.747	22.942
<b>BI2</b>	291.2132	100.447	2.899	0.008	82.898	499.529
<b>BSI</b>	0.7045	1.595	0.442	0.663	-2.604	4.013
<b>CI</b>	-0.3965	1.998	-0.198	0.845	-4.541	3.747
<b>Cal</b>	-0.0607	0.765	-0.079	0.937	-1.647	1.526
<b>Omnibus:</b>	1.429	Durbin-Watson:	1.757			
<b>Prob(Omnibus):</b>	0.490	Jarque-Bera (JB):	1.015			
<b>Skew:</b>	-0.019	Prob(JB):	0.602			
<b>Kurtosis:</b>	2.211	Cond. No.	1.33e+04			

Οι χαμηλές τιμές επιβεβαιώνονται και από το διάγραμμα συνδυαστικότητας ανάμεσα σε κάθε μεταβλητή και τα νέφη που σχηματίζονται από τη διασπορά στη σύγκριση των μεταβλητών με το SOC.

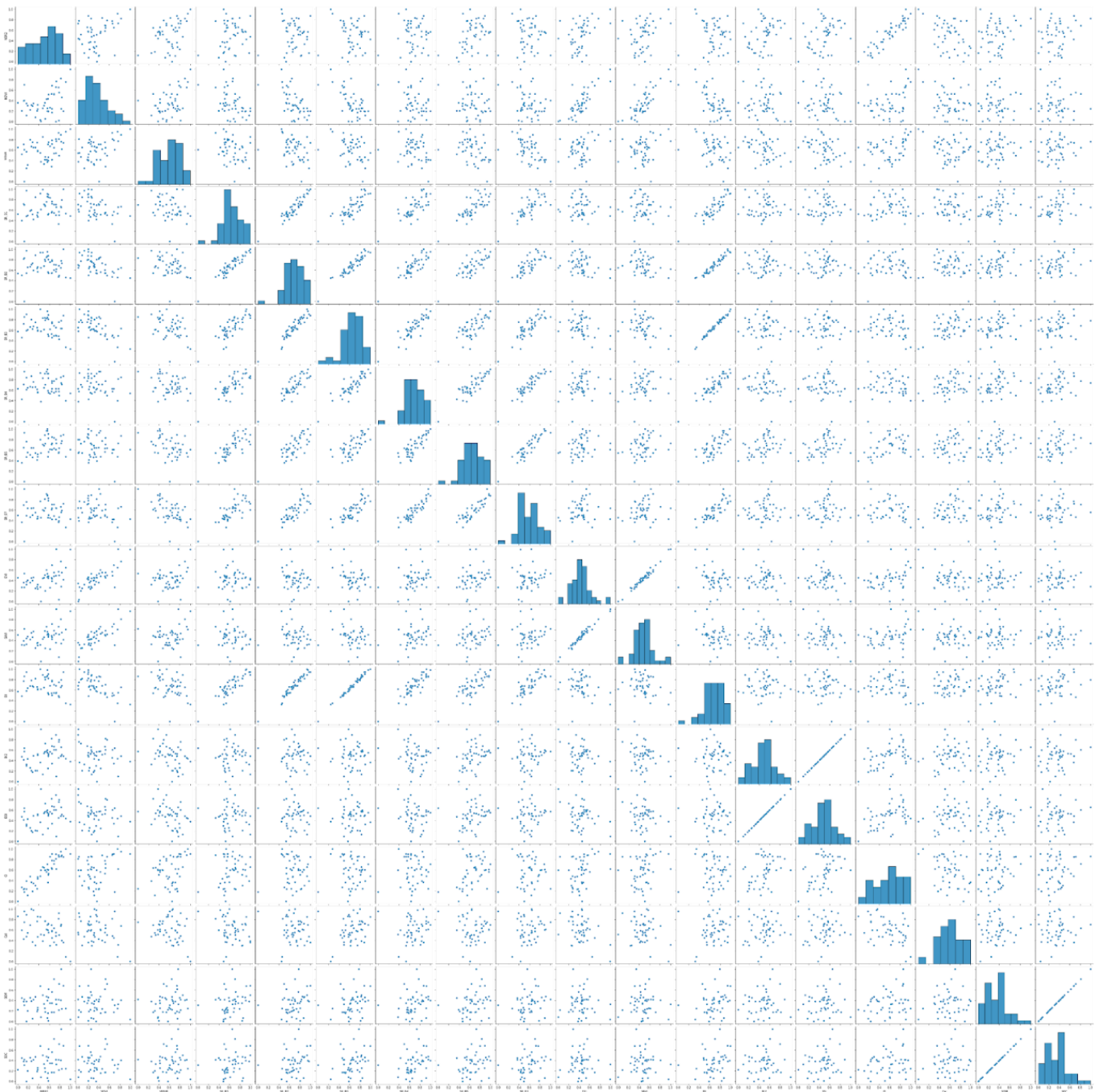


Εικόνα 9. Πίνακας συσχέτισης 1992 – 1998 / T6

Διάγραμμα 11. Συνδυαστικότητα (collinearity) 1992 - 1998 / T3

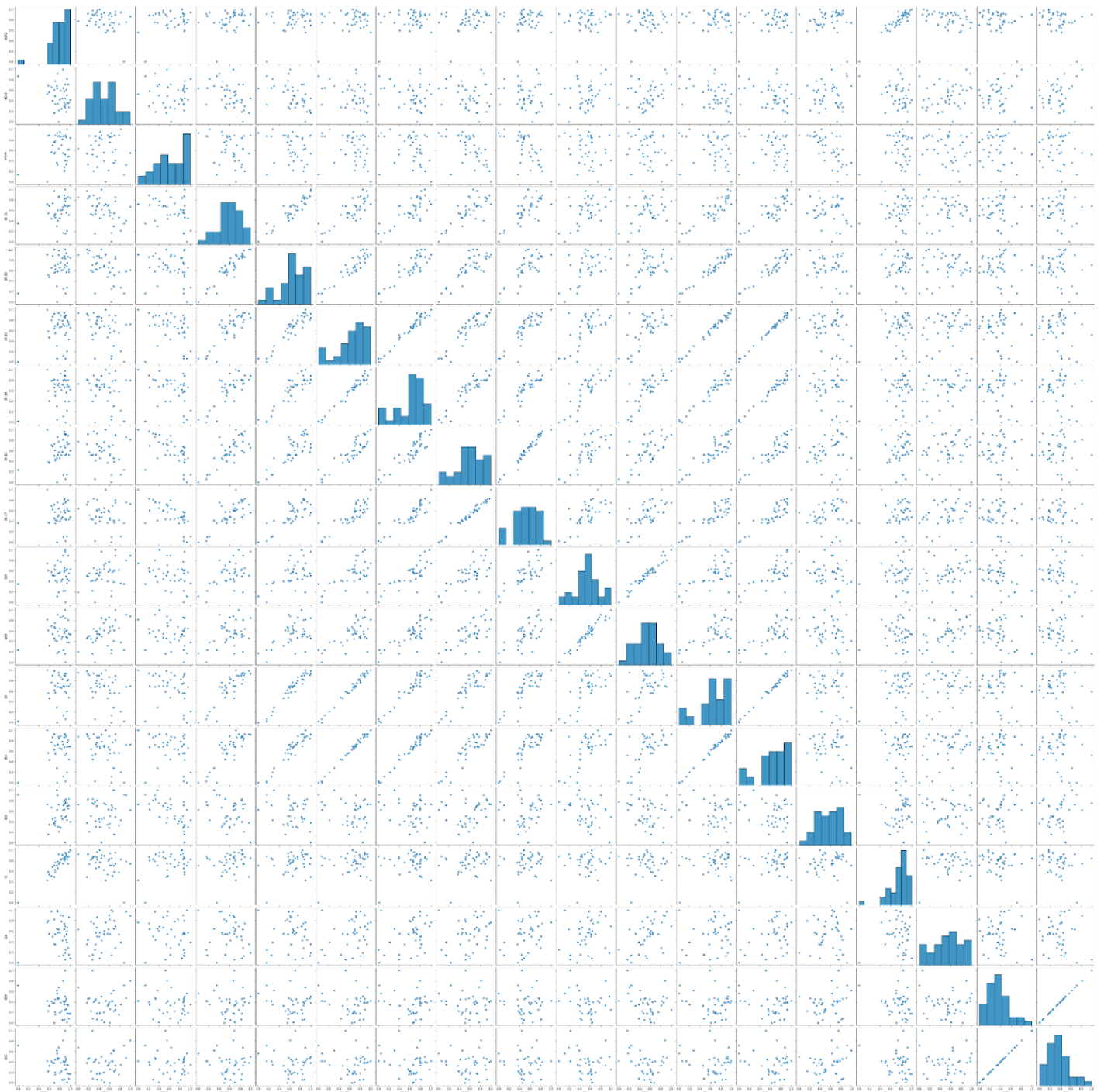


Διάγραμμα 12. Συνδυαστικότητα (collinearity) 1992 - 1998 / T4





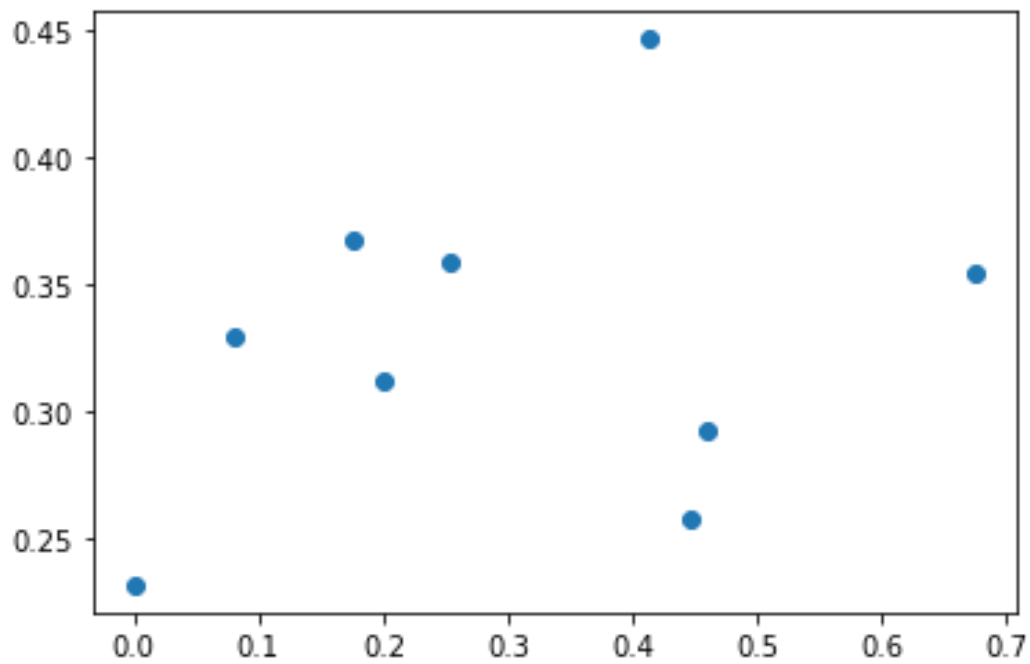
Διάγραμμα 13. Συνδυαστικότητα (collinearity) 1992 - 1998 / T6



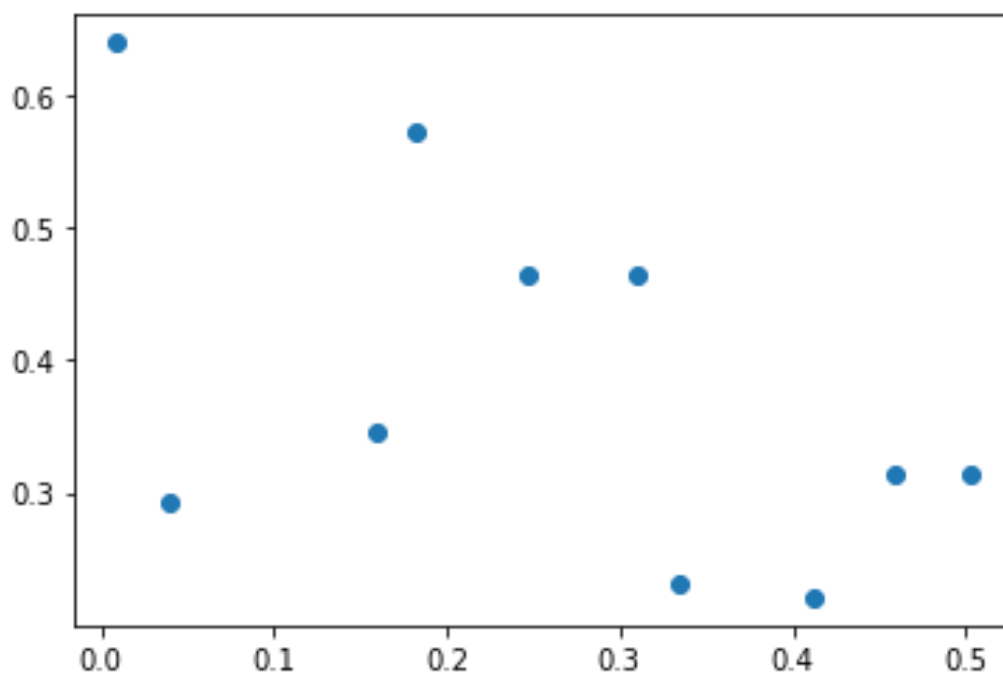
#### 4.2.2 Αποτελέσματα αλγορίθμου τυχαίων δασών (Random Forest)

Στη συνέχεια χρησιμοποιήθηκε ο αλγόριθμος random forest για την σύγκριση των τιμών των καναλιών και των δεικτών με το SOC. Η πρόβλεψη των τιμών του SOC, στο ποσοστό που χρησιμοποιήθηκε για test, επικυρώνει τα παραπάνω αποτελέσματα, δείχνοντας μεγάλες αποκλίσεις ανάμεσα στις προβλεπόμενες και τις πραγματικές τιμές του SOC και στις τρεις περιπτώσεις. Το μοντέλο δεν έχει καταφέρει να πλησιάσει τις πραγματικές τιμές του SOC.

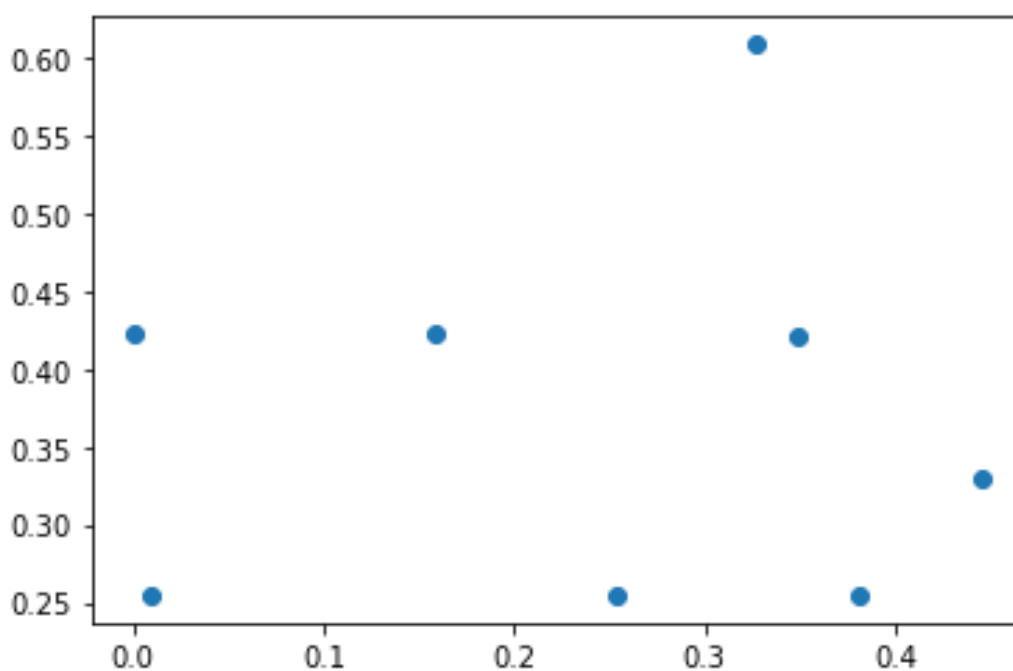
Διάγραμμα 14. Πρόβλεψη για το 1992 - 1998 / T3



Διάγραμμα 15. Πρόβλεψη για το 1992 - 1998 / T4



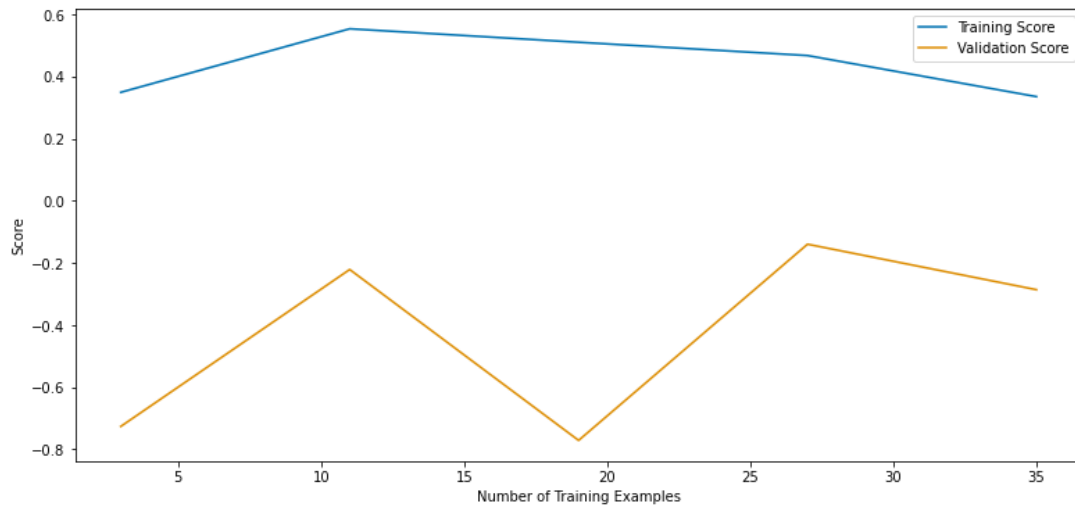
Διάγραμμα 16. Πρόβλεψη για το 1992 - 1998 / T6



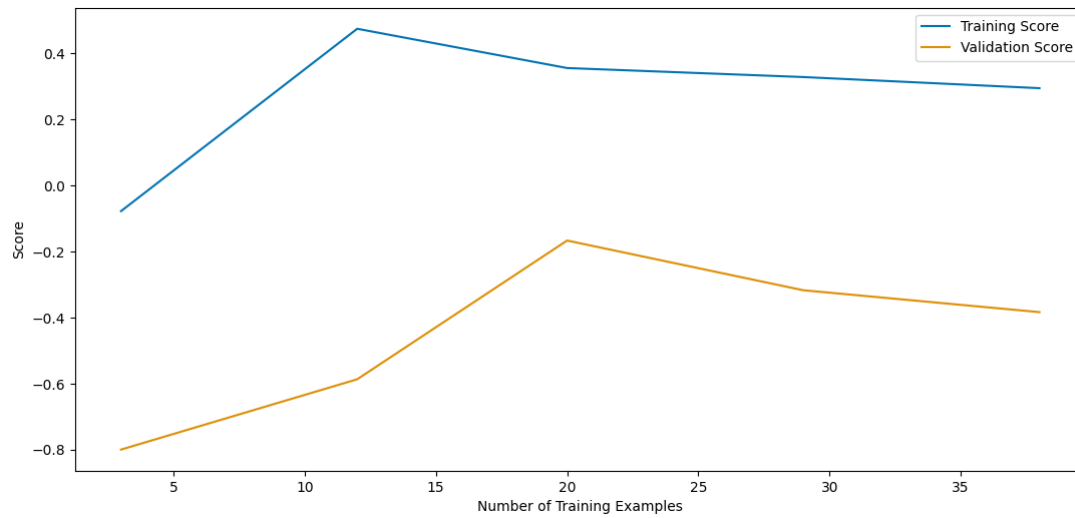
Τα παρακάτω διαγράμματα απεικονίζουν τις τιμές που λαμβάνει το SOC κατά την εκπαίδευση του μοντέλου και τις τιμές των δεδομένων ελέγχου. Παρατηρείται ότι και στις τρεις περιπτώσεις, το μοντέλο δεν έχει καταφέρει να ακολουθήσει τα δεδομένα ελέγχου, γεγονός που σημαίνει ότι αδυνατεί να γενικεύσει τα δεδομένα της εκπαίδευσης.



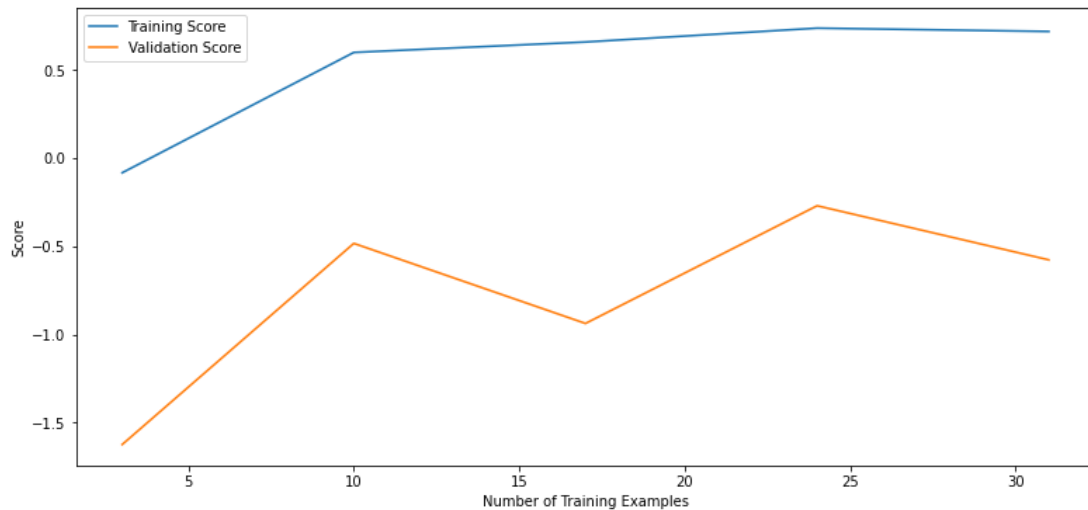
Διάγραμμα 17. Training score vs validation score, 1992 - 1998 / T3



Διάγραμμα 18. Training score vs validation score, 1992 - 1998 / T4



Διάγραμμα 19. Training score vs validation score, 1992 - 1998 / T6



# Κεφάλαιο 5 – Συμπεράσματα και προτάσεις

## 5.1. Συμπεράσματα

Κατά την παρούσα εργασία επιχειρήθηκε η εκτίμηση των επιπέδων του οργανικού άνθρακα στο έδαφος με την εφαρμογή δεικτών και τη χρήση αλγορίθμων μηχανικής μάθησης για πολυφασματικά δεδομένα του δορυφόρου Landsat 5, αξιοποιώντας την μέχρι τώρα γνώση για τη σύνδεση του οργανικού άνθρακα με συγκεκριμένα τμήματα του οπτικού φάσματος. Τα αποτελέσματα που προέκυψαν, μπορούν να σχολιαστούν σε δύο κατευθύνσεις: την επιτυχία στην αναγνώριση του γυμνού εδάφους στις εικόνες και την επιτυχία ανίχνευσης του οργανικού άνθρακα, τόσο από τις εικόνες, όσο και την καταλληλότητα των δεδομένων μεθόδων που χρησιμοποιήθηκαν.

Η αναγνώριση των περιοχών γυμνού εδάφους έγινε με τη χρήση αλγορίθμου, βασισμένου στο Σύστημα γεωχωρικής ανίχνευσης εδάφους (GEOS3), ο οποίος είναι σε θέση να παράγει μια συνθετική εικόνα αποτελούμενη από εικονοστοιχεία εδάφους. Από τα πειράματα που πραγματοποιήθηκαν στις χρονοσειρές και τις το εύρος τιμών των δεικτών, προέκυψε ότι η επιλογή δεδομένων περισσότερων ετών, συνέβαλε στη συγκέντρωση περισσότερων περιοχών γυμνού εδάφους, χωρίς να αλλοιώνεται η ποιότητα του αποτελέσματος. Όσον αφορά το εύρος των δεικτών, φαίνεται πως η καλύτερη επιλογή για τον NDVI είναι από -0,25 έως 0,25, αλλά και οι τιμές έως 0,35 λειτουργούν επαρκώς. Για τον NBR2 φαίνεται πως η διατήρηση των αρνητικών τιμών 0 έως -0,3 δεν προσδίδει πληροφορία γυμνού εδάφους. Αντίθετα, η επιλογή του άνω κατωφλίου είναι αρκετά σημαντική. Παρατηρήθηκε ότι όλες οι τιμές έως 0,350 είναι σε θέση να εντοπίσουν γυμνό έδαφος, δεν εμφανίστηκαν δηλαδή pixels με βλάστηση. Ωστόσο το αποτέλεσμα είναι ποιοτικά καλύτερο για τις τιμές έως 0,15. Αξίζει να σημειωθεί ότι τα αυστηρότερα κατώφλια περιορίζουν σε μεγάλο βαθμό τα αποτελέσματα, αποκλείοντας μεγάλο όγκο αξιοποιήσιμων δεδομένων.

Αξίζει να σημειωθεί ότι σε όλα τα πειράματα αξιοποιήθηκαν λιγότερα από τα μισά δείγματα εδάφους. Το γεγονός είναι πιθανόν να οφείλεται στο ότι η συλλογή αρκετών από τα δείγματα δεν αφορούσε μονοετείς καλλιέργειες, αλλά περιοχές με σταθερή βλάστηση, καθώς δεν εντοπίστηκε καθαρό έδαφος από το δορυφόρο σε διάστημα εξαετίας. Η σύνθεση του δείγματος περιορίσε σημαντικά την μετέπειτα εργασία, λόγω του μικρού αριθμού δειγμάτων που απέμειναν.

Όσον αφορά την επιλογή της μεθόδου σύγκρισης των μεταβλητών, η γραμμική παρεμβολή δεν κατάφερε να ανταποκριθεί, καθώς όπως προέκυψε από τα αποτελέσματα της, δεν υπάρχει γραμμική συσχέτιση ανάμεσα σε κάποιο από τα κανάλια του Landsat 5 και τον οργανικό άνθρακα του εδάφους, γεγονός που ισχύει και για τους δείκτες NBR2, NDVI, VNSIR, EVI, SAVI, BI, B12, BSI, CI και Cal. Συνεπώς, ο αλγόριθμος Random Forest είναι καταλληλότερος για τη διερεύνηση κάποιας πιο σύνθετης σύνδεσης ανάμεσα στα δεδομένα.

Από τα πειράματα που πραγματοποιήθηκαν με το Random Forest, η περίπτωση των κατωφλίων -0,25 έως 0,25 για τον NDVI και 0 έως 0,150 έδωσε τον υψηλότερο βαθμό συσχέτισης, ωστόσο δεν προέκυψε κάποια σαφής σχέση ανάμεσα στις τιμές της εικόνας Landsat 5 και τις αντίστοιχες τιμές του SOC. Ειδικότερα, σε αντίθεση με τη σχετική βιβλιογραφία, δεν παρατηρήθηκε σύνδεση του SOC με κανένα από τα SWIR κανάλια, B5, B7, σε καμία από τις τρεις περιπτώσεις, ούτε με τον δείκτη BI που αποτελεί λόγο των SWIR

καναλιών. Φαίνεται πως υπάρχει ορισμένη σύνδεση ανάμεσα στο ορατό τμήμα του οπτικού φάσματος και το SOC, όμως ο βαθμός συσχέτισης δεν είναι αρκετά ισχυρός ώστε να δώσει σαφή αποτελέσματα.

Παράγοντες που είναι πιθανό να επηρέασαν αρνητικά τη μελέτη είναι ο σχετικά μικρός αριθμός διαθέσιμων παρατηρήσεων, αλλά και η ποιότητα του εδάφους. Όπως αναφέρθηκε παραπάνω, η ανίχνευση του SOC από πολυφασματικές εικόνες είναι δυσκολότερη σε εδάφη με χαμηλότερη περιεκτικότητα. Γεγονός που μπορεί να σημαίνει ότι είναι δύσκολη η εκτίμηση του γενικότερα στον ελλαδικό χώρο και να δημιουργεί την ανάγκη χρήσης πιο σύνθετων δεκτών.

Συνοψίζοντας, οι δείκτες που εξετάστηκαν για τις εικόνες του δορυφόρου Landsat 5, δεν παρουσίασαν σαφή συσχέτιση, η οποία να οδηγήσει στην ανίχνευση του οργανικού άνθρακα στο έδαφος με αξιοπιστία.

## 5.2. Προτάσεις για μελλοντική έρευνα

Η δυσκολία της παρούσας μεθοδολογίας να εκτιμήσει τα επίπεδα του οργανικού άνθρακα αφήνει περιθώριο για περαιτέρω μελέτη του συγκεκριμένου ζητήματος. Το σχετικά μικρό πλήθος διαθέσιμων δεδομένων περιόρισε σημαντικά τις επιλογές σε πειραματισμό. Επόμενο βήμα στην κατεύθυνση αυτής της εργασίας θα μπορούσε να είναι η αξιοποίηση των δεδομένων του ευρωπαϊκού προγράμματος LUCAS. Ωστόσο, η λύση αυτή ενέχει άλλους κινδύνους κατά την προσπάθεια εκπαίδευσης, καθώς τα δεδομένα που παρέχει είναι αρκετά διασκορπισμένα και αφορούν διαφορετικές καλλιέργειες και διαφορετικές μεταξύ τους συνθήκες. Παρόλα αυτά αποτελεί μια πλούσια και χρήσιμη βάση δεδομένων, η οποία αξίζει να μελετηθεί.

# Βιβλιογραφία

- M.L. Jackson. "Soil Chemical Analysis", Englewood Cliffs, N.J., Prentice-Hall, USA (1965)
- B.E. Frazier. "Use of Landsat Thematic Mapper band ratios for soil investigations, Advances in Space" Research, Volume 9, Issue 1, 1989, Pages 155-158, ISSN 0273-1177, [https://doi.org/10.1016/0273-1177\(89\)90480-8](https://doi.org/10.1016/0273-1177(89)90480-8).
- A.R. Huete, K. Didan, T. Miura, E.P. Rodriguez, X. Gao, L.G. Ferreira. "Overview of the radiometric and biophysical performance of the MODIS vegetation indices." Remote Sens. Environ., 83 (2002), pp. 195-213
- R. Escadafal. "Remote sensing of arid soil surface color with Landsat Thematic Mapper" Adv. Space Res., 9 (1) (1989), pp. 159-163
- S. Mirzaee, S. Ghorbani-Dashtaki, J. Mohammadi, H. Asadi, F. Asadzadeh. "Spatial variability of soil organic matter using remote sensing data", CATENA, Volume 145, 2016, Pages 118-127, ISSN 0341-8162, <https://doi.org/10.1016/j.catena.2016.05.023>
- Ben-Dor, E., Y. Inbar, and Y. Chen. "The Reflectance Spectra of Organic Matter in the Visible Near-Infrared and Short Wave Infrared Region (400–2500 Nm) during a Controlled Decomposition Process." Remote Sensing of Environment 61, no. 1 (July 1, 1997): 1–15. [https://doi.org/10.1016/S0034-4257\(96\)00120-4](https://doi.org/10.1016/S0034-4257(96)00120-4).
- Breiman, Leo. "Random Forests." Machine Learning 45, no. 1 (October 1, 2001): 5–32. <https://doi.org/10.1023/A:1010933404324>.
- "Carbon Sequestration | Philosophical Transactions of the Royal Society B: Biological Sciences." Accessed October 9, 2023. <https://royalsocietypublishing.org/doi/10.1098/rstb.2007.2185>.
- Castaldi, Fabio, Sabine Chabrillat, Axel Don, and Bas van Wesemael. "Soil Organic Carbon Mapping Using LUCAS Topsoil Database and Sentinel-2 Data: An Approach to Reduce Soil Moisture and Crop Residue Effects." Remote Sensing 11, no. 18 (January 2019): 2121. <https://doi.org/10.3390/rs11182121>.
- Castaldi, Fabio, Andreas Hueni, Sabine Chabrillat, Kathrin Ward, Gabriele Buttafuoco, Bart Bomans, Kristin Vreys, Maximilian Brell, and Bas van Wesemael. "Evaluating the Capability of the Sentinel 2 Data for Soil Organic Carbon Prediction in Croplands." ISPRS Journal of Photogrammetry and Remote Sensing 147 (January 1, 2019): 267–82. <https://doi.org/10.1016/j.isprsjprs.2018.11.026>.
- Chapin, F. Stuart, Pamela A. Matson, and Peter M. Vitousek. Principles of Terrestrial Ecosystem Ecology. New York, NY: Springer New York, 2011. <https://doi.org/10.1007/978-1-4419-9504-9>.
- Demattê, José A. M., José Lucas Safanelli, Raul Roberto Poppiel, Rodnei Rizzo, Nélida Elizabet Quiñonez Silvero, Wanderson de Sousa Mendes, Benito Roberto Bonfatti, et al. "Bare



- Earth's Surface Spectra as a Proxy for Soil Resource Monitoring." *Scientific Reports* 10 (March 10, 2020): 4461. <https://doi.org/10.1038/s41598-020-61408-1>.
- Duchemin, Benoît, Philippe Maisongrande, Gilles Boulet, and Iskander Benhadj. "A Simple Algorithm for Yield Estimates: Evaluation for Semi-Arid Irrigated Winter Wheat Monitored with Green Leaf Area Index." *Environmental Modelling & Software* 23, no. 7 (July 1, 2008): 876–92. <https://doi.org/10.1016/j.envsoft.2007.10.003>.
- Gholizadeh, Asa, Daniel Žižala, Mohammadmehdi Saberioon, and Luboš Borůvka. "Soil Organic Carbon and Texture Retrieving and Mapping Using Proximal, Airborne and Sentinel-2 Spectral Imaging." *Remote Sensing of Environment* 218 (December 1, 2018): 89–103. <https://doi.org/10.1016/j.rse.2018.09.015>.
- Guo, L. B., and R. M. Gifford. "Soil Carbon Stocks and Land Use Change: A Meta Analysis." *Global Change Biology* 8, no. 4 (2002): 345–60. <https://doi.org/10.1046/j.1354-1013.2002.00486.x>.
- Guo, Long, Xiaoru Sun, Peng Fu, Tiezhu Shi, Lina Dang, Yiyun Chen, M. Linderman, et al. "Mapping Soil Organic Carbon Stock by Hyperspectral and Time-Series Multispectral Remote Sensing Images in Low-Relief Agricultural Areas." *Geoderma* 398 (September 15, 2021): 115118. <https://doi.org/10.1016/j.geoderma.2021.115118>.
- Hastie, Trevor, Jerome Friedman, and Robert Tibshirani. "Additive Models, Trees, and Related Methods." In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, edited by Trevor Hastie, Jerome Friedman, and Robert Tibshirani, 257–98. Springer Series in Statistics. New York, NY: Springer, 2001. [https://doi.org/10.1007/978-0-387-21606-5\\_9](https://doi.org/10.1007/978-0-387-21606-5_9).
- Heiden, Uta, Pablo d'Angelo, Peter Schwind, Paul Karlshöfer, Rupert Müller, Simone Zepp, Martin Wiesmeier, and Peter Reinartz. "Soil Reflectance Composites—Improved Thresholding and Performance Evaluation." *Remote Sensing* 14, no. 18 (September 10, 2022): 4526. <https://doi.org/10.3390/rs14184526>.
- Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme Learning Machine: Theory and Applications." *Neurocomputing, Neural Networks*, 70, no. 1 (December 1, 2006): 489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>.
- "Improving Soil Organic Carbon Predictions from a Sentinel–2 Soil Composite by Assessing Surface Conditions and Uncertainties - ScienceDirect." Accessed October 9, 2023. <https://www.sciencedirect.com/science/article/pii/S0016706122004359>.
- Jackson, Marion LeRoy. *Soil Chemical Analysis: Advanced Course*. UW-Madison Libraries Parallel Press, 2005.
- Jobbágy, Esteban G., and Robert B. Jackson. "The Vertical Distribution of Soil Organic Carbon and Its Relation to Climate and Vegetation." *Ecological Applications* 10, no. 2 (2000): 423–36. [https://doi.org/10.1890/1051-0761\(2000\)010\[0423:TVDOSO\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2).
- Marconato, Ulises, Roberto J. Fernández, and Gabriela Posse. "Cropland Net Ecosystem Exchange Estimation for the Inland Pampas (Argentina) Using EVI, Land Cover Maps, and Eddy Covariance Fluxes." *Frontiers in Soil Science* 2 (2022). <https://www.frontiersin.org/articles/10.3389/fsoil.2022.903544>.
- Nocita, Marco, Antoine Stevens, Bas van Wesemael, Matt Aitkenhead, Martin Bachmann, Bernard Barthès, Eyal Ben Dor, et al. "Chapter Four - Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring." In *Advances in Agronomy*, edited by Donald L.

- Sparks, 132:139–59. Academic Press, 2015.  
<https://doi.org/10.1016/bs.agron.2015.02.002>.
- “Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices - ScienceDirect.” Accessed October 9, 2023.  
<https://www.sciencedirect.com/science/article/abs/pii/S0034425702000962?via%3Dihub>.
- “Predicting the Spatial Variation of the Soil Organic Carbon Pool at a Regional Scale - Mishra - 2010 - Soil Science Society of America Journal - Wiley Online Library.” Accessed October 9, 2023.  
[https://access.onlinelibrary.wiley.com/doi/full/10.2136/sssaj2009.0158?saml\\_referre](https://access.onlinelibrary.wiley.com/doi/full/10.2136/sssaj2009.0158?saml_referre)r.
- “Remote Sensing | Free Full-Text | Combining High-Resolution Remote Sensing Products with a Crop Model to Estimate Carbon and Water Budget Components: Application to Sunflower.” Accessed October 9, 2023. <https://www.mdpi.com/2072-4292/12/18/2967>.
- Stockmann, Uta, Mark A. Adams, John W. Crawford, Damien J. Field, Nilusha Henakaarchchi, Meaghan Jenkins, Budiman Minasny, et al. “The Knowns, Known Unknowns and Unknowns of Sequestration of Soil Organic Carbon.” *Agriculture, Ecosystems & Environment* 164 (January 1, 2013): 80–99.  
<https://doi.org/10.1016/j.agee.2012.10.001>.
- Wibisono, Ari, Hanif Arief Wisesa, Novian Habibie, Aulia Arshad, Aditya Murdha, Wisnu Jatmiko, Ahmad Gamal, Indra Hermawan, and Siti Aminah. “Dataset of Short-Term Prediction of CO<sub>2</sub> Concentration Based on a Wireless Sensor Network.” *Data in Brief* 31 (August 1, 2020): 105924. <https://doi.org/10.1016/j.dib.2020.105924>.
- Wold, Herman O. A., ed. *Theoretical Empiricism: A General Rationale for Scientific Model-Building*. New York: Paragon House, 1987.
- Σινάνης, Κ. (2022). *Εδαφολογία. [Προπτυχιακό Εγχειρίδιο]. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις*. <http://dx.doi.org/10.57713/kallipos-14>
- Κοτρωνάκης, Μ (2018), Ο ρόλος της οργανικής ουσίας στη δομή και στην ποιότητα του εδάφους

## Διαγράμματα

Διάγραμμα 1. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T1. ....	33
Διάγραμμα 2. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T2. ....	34
Διάγραμμα 3. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T3. ....	34
Διάγραμμα 4. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T4. ....	35
Διάγραμμα 5. Γραμμή εδάφους ( B3 / B4 ) 1992 - 1998 για το T6. ....	35
Διάγραμμα 6. B5 / B7 - 1992 - 1998 / T1.....	36
Διάγραμμα 7. B5 / B7 - 1992 - 1998 / T2.....	37
Διάγραμμα 8. B5 / B7 - 1992 - 1998 / T3.....	37
Διάγραμμα 9. B5 / B7 - 1992 - 1998 / T4.....	38
Διάγραμμα 10. B5 / B7 - 1992 - 1998 / T6.....	38
Διάγραμμα 11. Συνδυαστικότητα (collinearity) 1992 - 1998 / T3.....	46
Διάγραμμα 12. Συνδυαστικότητα (collinearity) 1992 - 1998 / T4.....	47
Διάγραμμα 13. Συνδυαστικότητα (collinearity) 1992 - 1998 / T6.....	48
Διάγραμμα 14. Πρόβλεψη για το 1992 - 1998 / T3 .....	49
Διάγραμμα 15. Πρόβλεψη για το 1992 - 1998 / T4 .....	50
Διάγραμμα 16. Πρόβλεψη για το 1992 - 1998 / T6 .....	50
Διάγραμμα 17. Training score vs validation score, 1992 - 1998 / T3.....	52
Διάγραμμα 18. Training score vs validation score, 1992 - 1998 / T4.....	52
Διάγραμμα 19. Training score vs validation score, 1992 - 1998 / T6.....	52

## Παράρτημα εικόνων

Εικόνα 1. Ροές άνθρακα σε χερσαίο οικοσύστημα.....	5
Εικόνα 2. Η περιοχή μελέτης.....	21
Εικόνα 3. Χρήσεις γης - Κάλυψη γης σύμφωνα με το CORINE 1990. ....	22
Εικόνα 4. Η κατανομή των δειγμάτων σε εικόνα Landsat 5 (έγχρωμο σύνθετο, κανάλια 3-2-1) .....	23
Εικόνα 5. Περιοχές γυμνού εδάφους. Έγχρωμο σύνθετο 3-2-1 για τη χρονοσειρά 1992-1998, κατώφλι 4. ....	31
Εικόνα 6. Τα ενεργά δείγματα για τη χρονοσειρά 1992-1998, κατώφλι 4, στο έγχρωμο σύνθετο 3-2-1.....	32
Εικόνα 7. Πίνακας συσχέτισης 1992 – 1998 / T3 .....	41
Εικόνα 8. Πίνακας συσχέτισης 1992 – 1998 / T4 .....	43
Εικόνα 9. Πίνακας συσχέτισης 1992 – 1998 / T6 .....	45

## Παράρτημα πινάκων

Πίνακας 1. Τα κανάλια του Landsat 5 TM .....	10
Πίνακας 2. Μελέτες που αφορούν το μοντέλο SAFY .....	13
Πίνακας 3. Μελέτες που αφορούν μοντέλα παλινδρόμησης.....	19
Πίνακας 4. Οι κατηγορίες του CORINE, που εξαιρέθηκαν μέσω της μάσκας.....	25
Πίνακας 5. Εξισώσεις δεικτών .....	27
Πίνακας 6. Κατώφλια πειραμάτων .....	27
Πίνακας 7. Εξισώσεις υπολογισμού των δεικτών. ....	29
Πίνακας 8. Πλήθος εικόνων ανά χρονοσειρά και ποσοστό νεφοκάλυψης. ....	30
Πίνακας 9. Πλήθος δειγμάτων που αξιοποιήθηκαν ανά πείραμα .....	30
Πίνακας 10. Τιμές $R^2$ για τη σύγκριση B3 / B4. ....	33
Πίνακας 11. Τιμές $R^2$ για τη σύγκριση B5 / B7 .....	36
Πίνακας 12. OLS Regression Results 1992-1998 / T3 .....	40
Πίνακας 13. OLS Regression Results 1992-1998 - T4 .....	42
Πίνακας 14. OLS Regression 1992-1998 / T6 .....	44