



*Σύγχρονα χωρικά δεδομένα -
Μορφές, Διαχείριση και
Δυνατότητες*



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΕΙΟ
ΠΟΛΥΤΕΧΝΕΙΟ**

**ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ
ΜΑΘΗΜΑΤΙΚΩΝ &
ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΜΠΑΡΜΠΕΡΑΚΗΣ ΚΩΣΤΗΣ**

Φεβρουάριος 2020

Ευχαριστίες

Ευχαριστώ θερμά τον κ. Πέτρο Στεφανέα για το ενδιαφέρον που έδειξε από την πρώτη στιγμή ως προς την εκπόνηση της διπλωματικής μου εργασίας. Η συζήτηση μαζί του βοήθησε τα μέγιστα σε σχέση με τον προσανατολισμό σε ένα θέμα που θα έχει ιδιαίτερο ενδιαφέρον ως προς τις δικές μου αναζητήσεις και τη σύνδεση της επιστήμης των μαθηματικών με άλλα αντικείμενα που βρίσκονται στην πρώτη γραμμή της επικαιρότητας της παραγωγικής και επιστημονικής συγκυρίας. Η συνεργασία μαζί του υπήρξε άριστη.

Ευχαριστώ ιδιαίτερα τους φίλους και συναδέλφους (απόφοιτους του ΕΜΠ) Γιώργο Γκίμτσα και Ολύνα Γούναρη. Η τριβή, τόσο με τους ίδιους, όσο και με την επιστημονική τους ενασχόληση αποτέλεσε έμπνευση για τον ακριβή προσδιορισμό του κέντρου βάρους, καθώς και του πειραματικού σκέλους της παρούσας εργασίας. Η συζήτηση μαζί τους για θεωρητικές και τεχνικές πλευρές του παρόντος θέματος, καθώς και η απόλυτη προθυμία τους ως προς την αξιοποίηση της ακαδημαϊκής τους συνεισφοράς υπήρξε καταλυτική για την αναβάθμιση του ενδιαφέροντος και της τεκμηρίωσης της διπλωματικής μου εργασίας.

Ευχαριστώ πολύ την οικογένεια μου για τη στήριξη όλα αυτά τα χρόνια σε όλη αυτήν την πορεία με κάθε δυνατό τρόπο. Η εμπιστοσύνη και το ενδιαφέρον που έδειξαν αποτέλεσαν προωθητικό παράγοντα για την ολοκλήρωση αυτής της διαδρομής.

Ακόμα, ευχαριστώ την ΑΝ.ΑΡ.ΠΑ (φοιτητών ΣΕΜΦΕ) και κάθε συνάδελφο που αγωνιστήκαμε (και κερδίσαμε) όλα αυτά τα χρόνια μαζί για ένα ζωντανό φοιτητικό σύλλογο, στην πρώτη γραμμή των συλλογικών διεκδικήσεων, για μία καλύτερη σχολή και μία καλύτερη επαγγελματική προοπτική.

Τέλος, αφιερώνω την παρούσα εργασία στη μνήμη του προσφάτως εκλιπόντος (καθηγητή της σχολής) Αριστείδα Αραγεώργη. Ένας πραγματικός Δάσκαλος και αυθεντικός Αγωνιστής, απλός και ταπεινός, με σπάνια ηθική, επιστημονική και ιδεολογική αρτιότητα, η απουσία του οποίου αφήνει πολύ μεγάλο κενό στη σχολή μας (και την ακαδημαϊκή κοινότητα) σε όλα τα επίπεδα. Όσο πιο πολύ τέτοιοι Άνθρωποι σπανίζουν, τόσο περισσότερο η θύμηση του Άριθα αποτελεί τεράστια πηγή έμπνευσης για τους νέους ανθρώπους που επιλέγουν το μονοπάτι της αξιοπρέπειας και του συλλογικού αγώνα για την προσωπική και ακαδημαϊκή τους διαδρομή.

Πίνακας Περιεχομένων

Ευχαριστίες	2
ΚΕΦΑΛΑΙΟ 1 Το πλαίσιο του προβλήματος	5
1.1 Γενικά	5
1.2 ABSTRACT.....	6
1.3 Αντικείμενο και στόχοι.....	7
1.4 Κίνητρο και συνεισφορά.....	7
1.5 Δομή της εργασίας.....	7
ΚΕΦΑΛΑΙΟ 2 Χωρικά δεδομένα.....	8
2.1 Πώς παράγονται τα χωρικά δεδομένα;.....	8
2.2 Πώς συλλέγονται τα χωρικά δεδομένα;.....	11
2.2.1 Αναλογικοί Χάρτες	11
2.2.2 Δορυφορικές Εικόνες.....	12
2.2.3 Αεροφωτογραφίες	12
2.2.4 Τοπογραφικές Μετρήσεις.....	13
2.2.5 Άλλες Πηγές Γεωγραφικών Δεδομένων.....	13
2.3 Τύποι χωρικών δεδομένων.....	16
2.3.1 Αναπαράσταση Χωρικών Εννοιών.....	16
2.3.2 Μοντέλο Πεδίων	16
2.3.3 Μοντέλο Αντικειμένων	16
2.3.4 Δομές Γεωγραφικών Δεδομένων.....	16
2.3.5 Κανονικοποιημένη Δομή	17
2.3.6 Ιδιότητες Κανονικοποιημένης Δομής	17
2.3.7 Διανυσματική Δομή	18
2.3.8 Γεωμετρικά Αρχέτυπα.....	18
2.3.9 Δομές οργάνωσης της διανυσματικής δομής.....	19
2.4 Μορφότυποι χωρικών δεδομένων.....	22
2.4.1 Μορφότυποι διανυσματικών δεδομένων.....	22
2.4.2 Μορφότυποι ψηφιδωτών δεδομένων	24
2.5 Χωρικές Βάσεις Δεδομένων.....	27
2.5.1 Λίστα Χωρικών Βάσεων Δεδομένων.....	27
2.6 Ειδικοί τύποι χωρικών δεδομένων	29
2.6.1 Μεγάλα Χωρικά δεδομένα	29
2.6.2 Χωροχρονικά Δεδομένα.....	29
2.6.3 Χωρικά Blockchain	30
ΚΕΦΑΛΑΙΟ 3 Ανάλυση Χωρικών Δεδομένων	31
3.1 Δομικά ζητήματα της χωρικής ανάλυσης.....	33

3.2 Στατιστικοί Δείκτες Χωρικών Δεδομένων	34
3.2.1 Χωρικός Μέσος	34
3.2.2. Χωρική Τυπική Απόσταση.....	34
3.2.3 Χωρικός Αυτοσυσχετισμός	35
3.2.4 Χωρική Ομαδοποίηση.....	36
3.3 Προβλήματα Υπολογιστικής Γεωμετρίας	39
3.4 Εγγενή σφάλματα στη χωρική ανάλυση.....	43
3.5 Συστήματα Γεωγραφικών Πληροφοριών	44
3.5.1 Συστατικά στοιχεία ενός ΣΓΠ	44
3.5.2 Τεχνικές και τεχνολογίες.....	45
3.5.3 Αβεβαιότητα των ΣΓΠ	46
3.5.4 Λογισμικά ΣΓΠ.....	46
ΚΕΦΑΛΑΙΟ 4 ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΦΑΡΜΟΓΗΣ.....	48
4.1 Ταξινόμηση κάλυψης γης	48
4.1.1 Πρόγραμμα Landsat.....	48
4.1.2 Δορυφορικά δεδομένα	49
4.1.3 Πρόγραμμα COR.IN.E.....	53
4.2 Εφαρμογή Ταξινόμησης.....	54
4.2.1 Στόχοι της εφαρμογής	54
4.2.2 Περιγραφή εφαρμογής και δεδομένων	54
4.2.3 Προ-επεξεργασία εικόνων	56
4.2.4 Random Forest Classifier	58
4.2.5 Αποτελέσματα	59
4.2.6 Αξιολόγηση αποτελεσμάτων	64
4.3 Εφαρμογή Χωροθέτησης	71
4.3.1 Δεδομένα	71
4.3.2 Τεχνολογίες.....	71
4.3.3 Μεθοδολογία.....	72
4.3.4 Συμπεράσματα.....	77
ΚΕΦΑΛΑΙΟ 5 ΣΥΜΠΕΡΑΣΜΑΤΑ	78
Αντί επιλόγου.....	81
Κατάλογος Εικόνων	82
Κατάλογος Πινάκων.....	84
Βιβλιογραφία	85
Παράρτημα	86
OpenStreetMap	86
Ευρωπαϊκή Χωρική Υποδομή - Οδηγία INSPIRE	88

ΚΕΦΑΛΑΙΟ 1

Το πλαίσιο του προβλήματος

1.1 Γενικά

Την χρονική στιγμή της συγγραφής της παρούσας εργασίας, έχει παρέλθει ο χρόνος της έκρηξης των Μεγάλων Δεδομένων και πλέον είμαστε αντιμέτωποι με την ανάγκη για οργάνωση, αποθήκευση και διαχείριση μεγάλου όγκου πληροφορίας.

Από το 1944 ήδη, υπολογίζεται ότι ο όγκος των συγγραμμάτων διπλασιάζεται περίπου κάθε δεκαέξι χρόνια (Rider, 1944). Με την επικράτηση της ψηφιακής αποθήκευσης δεδομένων (καθώς αυτή έγινε οικονομικά προσιτή), η δυνατότητα αποθήκευσης δεδομένων από το 1986 έως το 2007 αυξάνεται κατά 25% ανά έτος (Hilbert & Lopez, 2011). Το 1999 παράχθηκαν σε όλο τον κόσμο περίπου 1.5 exabytes δεδομένων χωρίς να λαμβάνονται υπόψιν τα διπλότυπα (Lyman & Varian, 2000). Σε αντίστοιχη έρευνα για το έτος 2002 υπολογίζεται πως η παραγωγή των δεδομένων έφτασε τα 5 exabytes, με το 92% να βρίσκεται αποθηκευμένο σε μαγνητικά μέσα (δηλαδή σκληρούς δίσκους) ξεπερνώντας για πρώτη φορά άλλα μέσα όπως για παράδειγμα το χαρτί, το φιλμ ή οπτικά μέσα. Η παγκόσμια παραγωγή πληροφορίας για το έτος 2008 έφτασε συνολικά τα 14.7 exabytes (Bounie & Gille, 2012). (Siegel, 2013) και το μέγεθος αυτό αναμένεται να ανέλθει στα 1.7 MB ανά λεπτό, για κάθε άνθρωπο (Domo, 2018).

Μεγάλο κομμάτι αυτής της πληροφορίας έχει γεωχωρικό ή και γεωχωρο-χρονικό προσδιορισμό. Με τις σύγχρονες τεχνολογίες, με την ενσωμάτωση του GPS σε ολοένα και περισσότερους τύπους συσκευών και με την τάση για ανάπτυξη εφαρμογών που ενσωματώνουν υπηρεσίες που καταναλώνουν χωρικά δεδομένα (π.χ. social media) η αναλογία αυτή αυξάνεται και θα συνεχίσει να αυξάνεται ολοένα και περισσότερο.

Η τεράστια αυτή ποσότητα χωρικής πληροφορίας οδηγεί στην ανάπτυξη θεωρίας, τεχνικών και εργαλείων ανάλυσης και βρίσκει εφαρμογές στο σύνολο της ανθρώπινης κοινωνικο-οικονομικής δραστηριότητας. Η φύση όμως των χωρικών δεδομένων, ο κατακερματισμένος τρόπος απόκτησης - αποθήκευσης και διαχείρισης, η ad-hoc προσέγγιση στην ανάπτυξη των εργαλείων και των μεθόδων, οι καθιερωμένες πρακτικές στο χώρο και η αργοπορία στη δημιουργία πολιτικών επιδρούν αρνητικά στη δυναμική που αναπτύσσεται. Η δυναμική αυτή αποκρυσταλλώνεται στη δημιουργία εργαλείων που συστηματοποιούν την ανάλυση χωρικών δεδομένων και μπορούν να αξιοποιηθούν στο σχεδιασμό και την ερμηνεία πολύπλοκων κοινωνικο-οικονομικών δεδομένων με γεωχωρική υπόσταση. Η παρούσα εργασία επιχειρεί να παραθέσει αυτό το πλαίσιο, να αναδείξει τις προβληματικές και τις δυνατότητες που γεννώνται.

Λέξεις Κλειδιά: Χωρικά Δεδομένα, Ανάλυση Χωρικών Δεδομένων, Συστήματα Γεωγραφικών Πληροφοριών

1.2 ABSTRACT

At the time of writing this paper, Big Data explosion has already passed, and we are now faced with the need to organize, store and manage large volumes of information.

Since 1944, it has been estimated that the volume of textbooks has doubled every sixteen years (Rider, 1944). With the prevalence (as it became affordable) of digital data storage, data storage capacity from 1986 to 2007 increased by 25% per year (Hilbert & Lopez, 2011). In 1999, approximately 1.5 exabytes of data were produced worldwide without considering duplicate files (Lyman & Varian, 2000). A similar survey for 2002 estimates that data output reached 5 exabytes, with 92% being stored on magnetic media (i.e. hard drives) surpassing other media such as paper, film or optical media for the first time in history. Global output of information for the year 2008 totaled 14.7 exabytes (Bounie & Gille, 2012). From 2013 onwards, it is estimated that around 2.5 exabytes per day are produced (Siegel, 2013) and this size is expected to reach 1.7 MB per minute (Domo, 2018).

Much of this information is geospatial or geospatial temporal. With modern technologies, with the integration of GPS in more and more types of devices, with the tendency to develop applications that integrate spatial data services (e.g. social media) this proportion is increasing and will continue to increase.

This enormous amount of spatial information leads to the development of Theory, techniques and tools of analysis and finds applications in all human socio-economic activity. But the nature of spatial data, the fragmented way of acquisition - storage and management, the ad-hoc approach to the development of tools and methods, the established practices in the field and the delay in policy making have a negative impact on the dynamics that are being developed. This dynamic is can be seen in the creation of tools that systemize spatial data analysis and can be used in the design and interpretation of complex socio-economic data with geospatial correlation. The present paper attempts to set out this framework, highlighting the problems and opportunities that arise.

Keywords: *Spatial Data, Spatial Data Analysis, Geographic Information Systems*

1.3 Αντικείμενο και στόχοι

Αντικείμενο της παρούσας εργασίας είναι να περιγράψει την κατάσταση γύρω από τα χωρικά δεδομένα και τη ανάλυση χωρικών δεδομένων και να αναδείξει ότι η ύπαρξη των χωρικών δεδομένων, η ανάπτυξη της θεωρίας, η ανάπτυξη εργαλείων και τεχνικών ανάλυσης και η τεχνολογία υπολογιστών ανατροφοδοτούνται και ότι οι σωστές πρακτικές δημιουργούν ένα κύκλο θετικής ανατροφοδότησης με πολλαπλασιαστικά αποτελέσματα τόσο για τα ίδια τα τμήματα αυτού του “συστήματος” όσο και για τα πεδία εφαρμογής του.

1.4 Κίνητρο και συνεισφορά

Βασικός στόχος είναι η συνεισφορά στη συζήτηση η οποία έχει πολλά μέλη – την επιστημονική κοινότητα, το κράτος, τις εταιρείες του κλάδου, τους επιστήμονες & μηχανικούς και τελικά τον ίδιο τον πολίτη.

Η παραγόμενη αξία γύρω από τα χωρικά δεδομένα και την ανάλυσή τους, η πιθανή αξία που παραμένει αναξιοποίητη στο βαθμό που δεν ενσωματώνονται πλατιά στην έρευνα και στον κρατικό σχεδιασμό, τα ζητήματα ασφάλειας και προστασίας δεδομένων, η δυνατότητα απελευθέρωσης επιστημονικών – δημιουργικών δυνάμεων με θετικά αποτελέσματα για την κοινωνία, αναδεικνύουν τη σημασία αυτής της συζήτησης, την ανάγκη διεπιστημονικής της προσέγγισης και το ενδιαφέρον του συγγραφέα.

1.5 Δομή της εργασίας

Στο Κεφάλαιο 1 γίνεται η εισαγωγή στο πρόβλημα και παρατίθεται η θέση του συγγραφέα.

Στο Κεφάλαιο 2 περιγράφονται οι διαδικασίες παραγωγής και συλλογής χωρικών δεδομένων, η φύση, τα είδη, οι τύποι, οι μορφότυποί τους, καθώς και τα εργαλεία και οι τεχνικές αποθήκευσης και δημοσίευσής τους με στόχο να αναδειχθεί η πολυπλοκότητα του τεχνικού μέρους του προβλήματος.

Στο Κεφάλαιο 3 περιγράφεται η διαδικασία ανάλυσης χωρικών δεδομένων και γίνεται μια εισαγωγή στα Συστήματα Γεωγραφικών Πληροφοριών (ως το εργαλείο εκείνο μέσω του οποίου πραγματώνεται η διαχείριση, η απεικόνιση και η ανάλυση των χωρικών δεδομένων).

Στο Κεφάλαιο 4 παρουσιάζονται οι δύο εφαρμογές - υποθέσεις εργασίας - ανάλυσης χωρικών δεδομένων που αναπτύχθηκαν στα πλαίσια της παρούσας εργασίας αξιοποιώντας δημόσια διαθέσιμα δεδομένα και δημόσια διαθέσιμα εργαλεία πάνω στα εξής:

- Την ταξινόμηση κάλυψης γης στην περιοχή της ανατολικής Κρήτης με τη χρήση λογισμικού μηχανικής εκμάθησης.
- Την χωρική ανάλυση του οδικού δικτύου και τη χωροθέτηση υπηρεσιών στην Ιεράπετρα.

Στο Κεφάλαιο 5 αναπτύσσονται τα συμπεράσματα/θέσεις του συγγραφέα.

ΚΕΦΑΛΑΙΟ 2

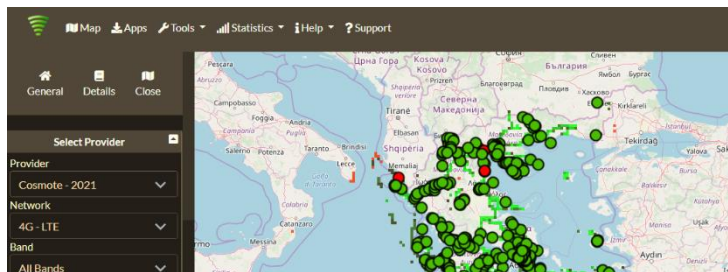
Χωρικά δεδομένα

Η προσοχή στην τοποθεσία, τη χωρική αλληλεπίδραση, τη δομή του χώρου και τις χωρικές διεργασίες βρίσκεται στην καρδιά της έρευνας σε όλο το φάσμα της ανθρώπινης δραστηριότητας, στις κοινωνικές επιστήμες, στις περιβαλλοντικές επιστήμες, στη μηχανική, στην οικονομία κτλ. Οι μελέτες σε αυτούς τους τομείς συνήθως χρησιμοποιούν δεδομένα για τα οποία τα χαρακτηριστικά τοποθεσίας ("το πού") αποτελούν σημαντική πηγή πληροφοριών. Αυτά τα δεδομένα τυπικά αποτελούνται από μία ή μερικές διασταυρώσεις παρατηρήσεων για τις υποομάδες, όπως νοικοκυριά, καλλιέργειες, υδάτινα στοιχεία ή για τις συνολικές χωρικές μονάδες, όπως ήπειροι, χώρες, διοικητικές διαιρέσεις ή όλος ο πλανήτης. Παρατηρήσεις όπως αυτές, για τις οποίες λαμβάνεται υπόψη η απόλυτη τοποθεσία ή/και η σχετική τοποθέτηση (χωροταξική διευθέτηση) αναφέρονται ως χωρικά δεδομένα. Τα χωρικά δεδομένα αποτελούν τον πυρήνα πολλών πεδίων και μελετώνται για να μοντελοποιήσουν τη χωρική δομή για μια σειρά κοινωνικοοικονομικών και μη μεταβλητών.

2.1 Πώς παράγονται τα χωρικά δεδομένα;

Χωρικά δεδομένα παράγονται σε όλο το φάσμα της σύγχρονης ανθρώπινης δραστηριότητας. Μια προσπάθεια κατηγοριοποίησης είναι σε γενικές γραμμές τα παρακάτω πλαίσια:

- **Στα πλαίσια ευρύτερου κρατικού σχεδιασμού**
Το εύρος της κρατικής δραστηριότητας που παράγει χωρικά δεδομένα είναι πραγματικά τεράστιο. Ο πολεοδομικός σχεδιασμός, ο περιβαλλοντικός σχεδιασμός, οι συγκοινωνίες, ο κοινωνικός σχεδιασμός, η ύδρευση, το δίκτυο ηλεκτρικής ενέργειας, η διαχείριση γης και οι επιδοτήσεις είναι μόνο μερικά από τα σημεία.
- **Στο πλαίσιο της σχέσης του πολίτη με το κράτος**
Αυτές οι σχέσεις συνήθως διαμεσολαβούνται από μηχανικούς και έχουν να κάνουν με δραστηριότητες όπως το κτηματολόγιο, την κατασκευή κτλ.
- **Τηλεπικοινωνίες**
Το κινητό και η κάθε συσκευή που συνδέεται στο διαδίκτυο ή στο δίκτυο κινητής τηλεφωνίας αποτελεί και μία μηχανή παραγωγής δεδομένων με χωρικό και χρονικό προσδιορισμό.



Εικόνα 1 - Παραγωγή χωρικών δεδομένων από τις τηλεπικοινωνίες (www.cellmapper.net)

Εικόνα 2 - Παραγωγή χωρικών δεδομένων από τη ναυσιπλοΐα (www.marinetraffic.com)
Εικόνα 3 - Παραγωγή χωρικών δεδομένων από τις τηλεπικοινωνίες (www.cellmapper.net)

- **Μεταφορές**

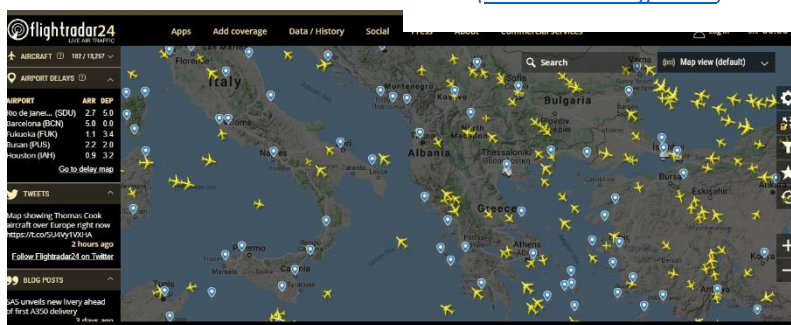
Ο σύγχρονος σχεδιασμός, ο συντονισμός και η παρακολούθηση μετακινήσεων και μεταφορών στον αέρα, στη γη, στη θάλασσα και στο διάστημα αποτελεί ένα σύνολο δραστηριοτήτων που στηρίζονται σε και ταυτόχρονα παράγουν χωρικά δεδομένα.



Εικόνα 4 - Παραγωγή χωρικών δεδομένων από τη ναυσιπλοΐα (www.marinetraffic.com)

Εικόνα 5 - Διαδρομή πτήσης των δορυφόρων LANDSAT (<https://earthnow.usgs.gov/observer/>)
Εικόνα 6 - Παραγωγή χωρικών δεδομένων από τη ναυσιπλοΐα (www.marinetraffic.com)

Εικόνα 7 - Παραγωγή χωρικών δεδομένων από την αεροπλοΐα (www.flightradar24.com)

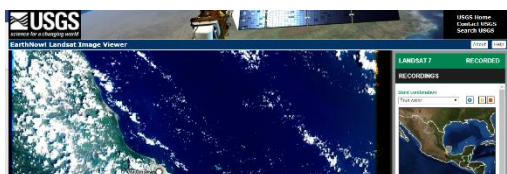


- **Στο πλαίσιο δραστηριοτήτων του ιδιωτικού τομέα**

Πολλές δραστηριότητες του ιδιωτικού τομέα απαιτούν τη συλλογή, επεξεργασία και διατήρηση χωρικών δεδομένων. Από την Google και τη Microsoft που διατηρούν (και ανανεώνουν) από τις μεγαλύτερες συλλογές χωρικών δεδομένων στον κόσμο και τα “οικοσυστήματα” εφαρμογών και υπηρεσιών που έχουν στηθεί τριγύρω τους μέχρι “μικρότερες” εφαρμογές όπως π.χ. οι αγωγοί και τα δίκτυα διανομής πετρελαίου και φυσικού αερίου.

- **Από τα διαστημικά προγράμματα**

Ίσως η ανθρώπινη δραστηριότητα που παράγει συστηματικά τα περισσότερα χωρικά δεδομένα.



Εικόνα 9 - Λήψη από βίντεο της συνεχούς απεικόνισης του Landsat 7 πάνω από το Μεξικό (<https://earthnow.usgs.gov/observer/>)

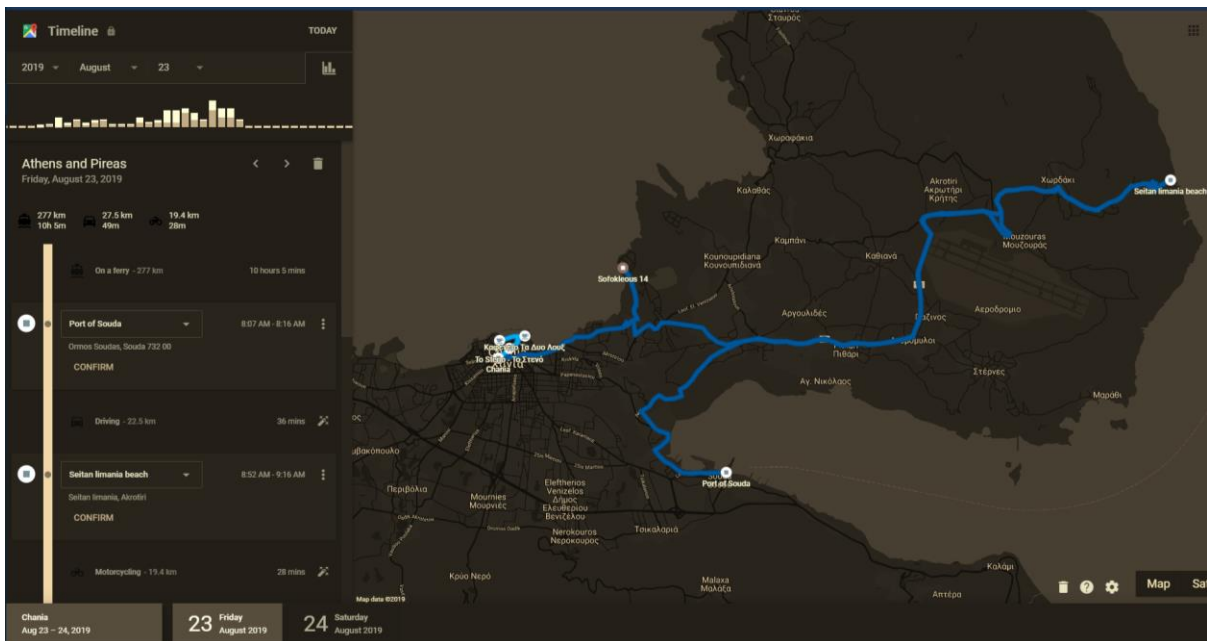


Εικόνα 8 - Διαδρομή πτήσης των δορυφόρων LANDSAT (<https://earthnow.usgs.gov/observer/>)

Πλήθος δορυφόρων από πολλές διαστημικές υπηρεσίες και ιδιωτικές πλέον εταιρείες παίρνουν εικόνες κατά μήκος της τροχιάς τους και τις διαθέτουν σε πραγματικό χρόνο .

- **Ως δευτερογενή προϊόντα εφαρμογών**

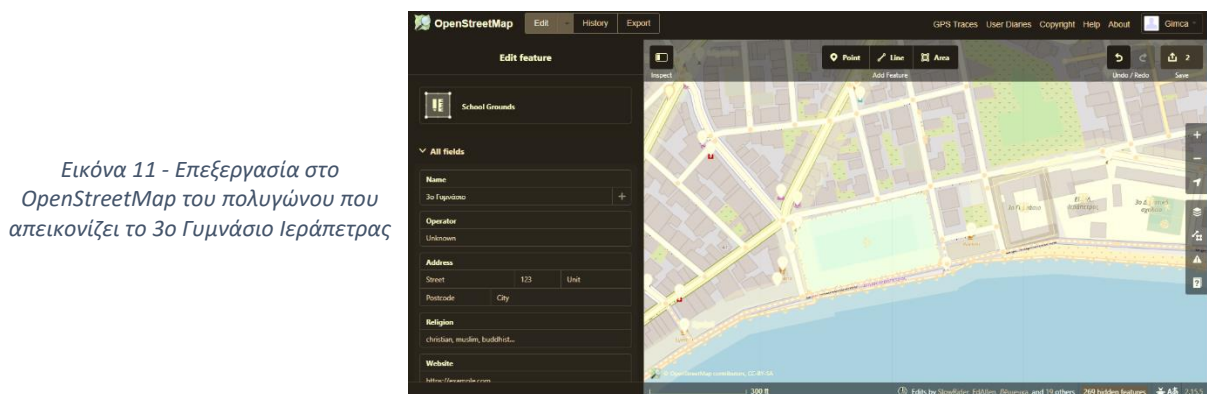
Χαρακτηριστικότερη ίσως περίπτωση σύνδεσης πρωτογενών χωρικών δεδομένων για την παραγωγή νέων ποιοτικά χωρικών δεδομένων είναι οι εφαρμογές δρομολόγησης οι οποίες συνδυάζουν το οδικό δίκτυο με τα δεδομένα της κινητής τηλεφωνίας για να υπολογίσουν την κίνηση τη δεδομένη στιγμή και το τελικό χωρικό προϊόν (διαδρομή) με υπηρεσίες (π.χ. προτάσεις για κατανάλωση)



Εικόνα 10 - Ιστορικό διαδρομών σε παραλίες και καταστήματα κατανάλωσης στα Χανιά, από χρήστη της Google

- **Από πληθοπορισμό (crowd sourcing)**

Την πλέον επιτυχημένη διαδικασία παραγωγής χωρικών δεδομένων από πληθοπορισμό αποτελεί το [OpenStreetMap](#). Με τον ιστότοπο ανοιχτό στον καθένα (ιδιώτη, κράτος, επιχείρηση) να προσθέσει, τροποποιήσει και αξιοποιήσει τα χωρικά δεδομένα του, αποτελεί τον πλέον ελκυστικό χάρτη και πάροχο χωρικής πληροφορίας κόντρα σε “κλειστές” αντίστοιχες λύσεις.



Εικόνα 11 - Επεξεργασία στο OpenStreetMap του πολυγώνου που απεικονίζει το 3ο Γυμνάσιο Ιεράπετρας

2.2 Πώς συλλέγονται τα χωρικά δεδομένα;

Οι πηγές των χωρικών δεδομένων διακρίνονται σε δύο τύπους:

- πηγές που παρέχουν πρωτογενή δεδομένα, δηλαδή δεδομένα που συλλέγονται άμεσα (π.χ. τοπογραφικές μετρήσεις)
- πηγές που παρέχουν δευτερογενή δεδομένα, δηλαδή δεδομένα που προέρχονται από υπάρχοντες χάρτες, προϋπάρχουσες βάσεις δεδομένων, πίνακες χαρακτηριστικών κλπ, ή που διατίθενται από οργανισμούς, κρατικούς φορείς, ιδιωτικές συλλογές κλπ.

Λόγω της πληθώρας των διαφορετικών πηγών γεωγραφικών δεδομένων, οι τεχνικές συλλογής που έχουν αναπτυχθεί προσαρμόστηκαν στην ιδιαίτερη φύση της κάθε πηγής. Γενικά, θα λέγαμε ότι οι κατηγορίες τεχνικών συλλογής γεωγραφικών δεδομένων κατηγοριοποιούνται σε αυτές που συλλέγουν πρωτογενή ή δευτερογενή δεδομένα και σε αυτές που αποθηκεύουν τα δεδομένα που συλλέχθηκαν σε κανονικοποιημένη ή σε διανυσματική δομή (ΣτΣ: Οι δομές δεδομένων αναλύονται στο Κεφάλαιο 3 του παρόντος).

	Δεδομένα κανονικοποιημένης δομής	Διανυσματικά δεδομένα
Πρωτογενή δεδομένα	Τηλεπισκόπηση - Φωτοερμηνεία	Μετρήσεις με GPS, Τοπογραφικές Μετρήσεις
Δευτερογενή δεδομένα	Σάρωση χαρτών και εικόνων	Ψηφιοποίηση χαρτών, Φωτογραμμετρία

*Πίνακας 1 - Κατηγοριοποίηση τεχνικών συλλογής
σε σχέση με τα χαρακτηριστικά των συλλεγμένων γεωγραφικών δεδομένων.*

2.2.1 Αναλογικοί Χάρτες

Από τις κυριότερες πηγές δευτερογενών δεδομένων παραμένουν οι αναλογικοί χάρτες. Για να είναι χρήσιμοι οι αναλογικοί χάρτες πρέπει να μετατραπούν σε ψηφιακή μορφή. Η μετατροπή των αναλογικών δεδομένων σε ψηφιακά ονομάζεται ψηφιοποίηση και αποτελεί δύσκολη, χρονοβόρα και επιρρεπή σε σφάλματα διαδικασία. Το κύριο μειονέκτημα της χρήσης των αναλογικών χαρτών ως κύρια πηγή δεδομένων είναι η περιορισμένη και εξαρτημένη από την κλίμακα ακρίβειά τους. Επιπλέον, οι αναλογικοί χάρτες δεν αποδίδουν με ακρίβεια τη θέση και το σχήμα των πραγματικών γεωγραφικών οντοτήτων, αφού αυτές έχουν υποστεί διαδικασίες γενίκευσης και συμβολισμού.

Η ποιότητα των δεδομένων δεν επηρεάζεται τόσο από τα σφάλματα των αρχικών μετρήσεων ή της φυσικής μεταβολής, αλλά από τα σφάλματα που έγιναν κατά την επεξεργασία των δεδομένων, δηλαδή κατά τη συλλογή και την αποθήκευση των δεδομένων. Τα σφάλματα που οφείλονται στην πεπερασμένη ακρίβεια του Η/Υ (αριθμητικά σφάλματα) επηρεάζουν όλες τις γεωμετρικές λειτουργίες, συμπεριλαμβανομένων και των τοπολογικών σχέσεων. Άλλη πιθανή αιτία σφαλμάτων είναι η χειροκίνητη καταχώρηση των δεδομένων που επηρεάζεται από την οπτική αντίληψη και τον συντονισμό ματιών-χειριού του χειριστή, την

υποκειμενικότητα στην επιλογή των σημείων κατά την ψηφιοποίηση καμπυλών και την κούραση μετά από μεγάλα διαστήματα εργασίας.

2.2.2 Δορυφορικές Εικόνες

Οι δορυφορικές εικόνες αποτελούν σημαντική πηγή δεδομένων διαθέσιμων σε ψηφιακή μορφή. Τα δεδομένα αυτά ακολουθούν την κανονικοποιημένη δομή δεδομένων. Η τιμή κάθε φαντίου της δομής αντιπροσωπεύει την ισχύ της ανακλώμενης ηλεκτρομαγνητικής ακτινοβολίας. Η αναγνώριση της τιμής των δεδομένων απαιτεί τη χρήση εργαλείων φωτοερμηνείας σε συνδυασμό με επίγειους ελέγχους για την εξαγωγή αξιόπιστων αποτελεσμάτων. Η γεωγραφική αναφορά των κανονικοποιημένων εικόνων προηγείται της συλλογής των δεδομένων. Οι δορυφορικές εικόνες χρησιμοποιούνται κυρίως σε θεματικές απεικονίσεις ή σε περιορισμένης έκτασης και μικρής κλίμακας τοπογραφικές απεικονίσεις.

2.2.3 Αεροφωτογραφίες

Οι αεροφωτογραφίες παρέχουν χωρικά οριζοντιογραφικά και υψομετρικά δεδομένα τα οποία συλλέγονται με διάφορες τεχνικές. Οι πρώτες λήψεις αεροφωτογραφιών πραγματοποιήθηκαν στη Γαλλία το 1858, από τον Tournaillon G. Από τότε, οι αεροφωτογραφίες χρησιμοποιήθηκαν για τη συλλογή τοπογραφικών πληροφοριών, την αποτύπωση όψεων ιστορικών και άλλων κτηρίων, τη χαρτογράφηση διαφόρων περιοχών αλλά και για στρατιωτικούς σκοπούς (π.χ. για την αναγνώριση και τη χαρτογράφηση εχθρικών εδαφών). Στις μέρες μας, η λήψη αεροφωτογραφιών βρίσκει εφαρμογή σε πολλούς τομείς, όπως στην τοπογραφία, τη χαρτογραφία, την εκπόνηση περιβαλλοντικών μελετών, την αρχιτεκτονική, τον πολεοδομικό σχεδιασμό, τη γεωλογία και τη δασολογία, κ.ά.

Η λήψη αεροφωτογραφιών πραγματοποιείται με ειδικές κάμερες οι οποίες διαθέτουν μεγάλη εστιακή απόσταση και εύρος πεδίου, καθώς και δυνατότητα συνεχούς λήψης. Η ποιότητα του οπτικού συστήματος των καμερών καθώς και ο τρόπος (ταχύτητα, κλίση, κλπ.) σάρωσης της επιφάνειας της Γης από το πτητικό μέσο επηρεάζουν τόσο τη χωρική όσο και τη φασματική ανάλυση των λαμβανόμενων αεροφωτογραφιών.

Οι αεροφωτογραφίες, σε σχέση με τα δεδομένα που συλλέγονται με επίγειες τεχνικές συλλογής δεδομένων, έχουν το πλεονέκτημα ότι παρέχουν τη δυνατότητα πολύ γρήγορης παραγωγής χαρτών και με σχετικά πολύ μικρότερο κόστος. Ακόμα, επιτρέπουν την άμεση προσπέλαση σε οποιαδήποτε περιοχή της γης, ανεξάρτητα από το ανάγλυφο ή την επικάλυψή της.

Παρότι οι αεροφωτογραφίες και οι δορυφορικές εικόνες παρουσιάζουν κοινά στοιχεία στον τρόπο λήψης τους, παραμένουν σημαντικές διαφορές στον τρόπο επεξεργασίας και ερμηνείας των δεδομένων. Παλαιότερα, η σπουδαιότερη διαφορά έγκειτο στο γεγονός ότι η λήψη αεροφωτογραφιών γινόταν με αναλογικές κάμερες και στη συνέχεια γινόταν η ψηφιοποίησή τους συνήθως με τη σάρωση των αρνητικών του φιλμ. Στις μέρες μας, η διαφορά αυτή τείνει να εκλείψει διότι η χρήση ψηφιακών καμερών για τη λήψη αεροφωτογραφιών επεκτείνεται ολοένα και περισσότερο.

2.2.4 Τοπογραφικές Μετρήσεις

Τα δεδομένα που συλλέγονται άμεσα από τοπογραφικές μετρήσεις, συμβατικές ή σύγχρονες, περιγράφουν άμεσα ή έμμεσα τη θέση γεωγραφικών οντοτήτων, είτε υπό τη μορφή συντεταγμένων, είτε με άλλες περιγραφικές ιδιότητες. Στην πραγματοποίηση τοπογραφικών μετρήσεων, σημαντική είναι επίσης και η συνεισφορά της τεχνολογίας του Παγκόσμιου Συστήματος Εντοπισμού (GPS), και ειδικότερα όσον αφορά τις εφαρμογές που συμβαίνουν σε πραγματικό χρόνο.

2.2.5 Άλλες Πηγές Γεωγραφικών Δεδομένων

Έτοιμα προς χρήση γεωγραφικά δεδομένα βρίσκονται επίσης διαθέσιμα σε διάφορες γεωγραφικές βάσεις δεδομένων και ΣΓΠ, σε πίνακες χαρακτηριστικών, ή διατίθενται δωρεάν ή με χρηματικό αντίτιμο από οργανισμούς, κρατικούς φορείς, ιδιωτικές συλλογές, κλπ.

Στη συνέχεια, επειδή ο αριθμός των υφιστάμενων εξωτερικών πηγών είναι πολύ μεγάλος, αναφέρονται μόνο κάποια ενδεικτικά παραδείγματα. Σε κάθε περίπτωση, πρέπει να ληφθεί υπόψη ότι τα χαρακτηριστικά και η διαθεσιμότητα των γεωγραφικών δεδομένων αλλάζουν συνεχώς. Ως γενική παρατήρηση, θα έλεγε κανείς ότι, ενώ στην Ευρώπη η διαθεσιμότητα των δεδομένων είναι ακόμα περιορισμένη, στη Βόρειο Αμερική τα ψηφιακά γεωγραφικά δεδομένα διατίθενται εύκολα και με σχετικά χαμηλό κόστος.

Όνομασία πηγής	Διατιθέμενα γεωγραφικά δεδομένα
United States Geological Survey (USGS)	Ψηφιακά χαρτογραφικά δεδομένα, κ.ά.
National Geodetic Survey (NGS)	Γεωδαιτικά/ τοπογραφικά δεδομένα (GPS δεδομένα, αεροφωτογραφίες, χάρτες ακτογραμμών, κ.ά.)
National Digital Orthophoto Programs (NDOP)	Αεροφωτογραφίες, ορθοφωτοχάρτες, κ.ά.
United State Department of Agriculture	Πύλη γεωχωρικών δεδομένων (ψηφιακές αεροφωτογραφίες, ψηφιακά μοντέλα εδάφους, χάρτες σύστασης εδάφους, κ.ά.)
National Geospatial Center of Excellence (NGCE)	Χάρτες κάλυψης γης, βάση δεδομένων απογραφής εθνικών πόρων (National Resources Inventory (NRI)), αεροφωτογραφίες, ψηφιακές ορθοφωτογραφίες, κ.ά.
National Aeronautics and Space Administration (NASA)	Δορυφορικές εικόνες, αεροφωτογραφίες, κ.ά.
United States Census Bureau Maps and Cartographic Resources	Ψηφιακά αρχεία γεωγραφικών δεδομένων, θεματικοί χάρτες μικρής και μεσαίας κλίμακας, κ.ά.
Ordnance Survey (UK)	Ψηφιακοί τοπογραφικοί και ιστορικοί χάρτες, κ.ά.
European Environmental Agency (EEA)	Δεδομένα κάλυψης γης του Ευρωπαϊκού προγράμματος CORINE Land Cover
OpenStreetMap	Ανοιχτού λογισμικού δωρεάν παγκόσμιος χάρτης με διανυσματικά και περιγραφικά δεδομένα

Πίνακας 2 - Παραδείγματα εξωτερικών πηγών γεωγραφικών δεδομένων στις ΗΠΑ και στην Ευρώπη

Στην Ελλάδα, οι κυριότερες πηγές γεωγραφικών δεδομένων ανήκουν στο δημόσιο τομέα. Υπάρχουν όμως και εταιρείες ιδιωτικού δικαίου που διαθέτουν και αυτές γεωγραφικά

δεδομένα σε ποικίλες μορφές (δορυφορικές εικόνες, αεροφωτογραφίες, κλπ.). Παρακάτω περιγράφονται παραδείγματα ελληνικών εξωτερικών πηγών γεωγραφικών δεδομένων.

- Ο κύριος σκοπός της [Γεωγραφικής Υπηρεσίας Στρατού \(ΓΥΣ\)](#) είναι να παράγει χαρτογραφικά προϊόντα για την υποστήριξη των Ενόπλων Δυνάμεων. Ακόμη, η ΓΥΣ παράγει, διαχειρίζεται και διανέμει γεωγραφικά δεδομένα που αποβλέπουν στην ανάπτυξη της χώρας και την ικανοποίηση των αναγκών του ευρύτερου κοινωνικού συνόλου. Τα γεωγραφικά δεδομένα διατίθενται προς τους πολίτες με χρηματικό αντίτιμο και περιλαμβάνουν τοπογραφικούς χάρτες (σε αναλογική ή ψηφιακή μορφή), υψομετρικές καμπύλες και ψηφιακά μοντέλα εδάφους, σχέδια πόλεων, οδικό δίκτυο, κλπ. Ακόμη, η ΓΥΣ παράγει χάρτες ανωμαλιών βαρύτητας και γεωμαγνητικής απόκλισης ενώ διατηρεί αρχείο ιστορικών χαρτών, λεξικό γεωγραφικών τοπωνυμίων, βιβλίο μετατροπής συντεταγμένων και ευρετήριο τοπογραφικών διαγραμμάτων σε κλίμακα 1:5000.
- Η [Ελληνική Στατιστική Αρχή \(ΕΛ.ΣΤΑΤ.\)](#) συγκεντρώνει, επεξεργάζεται και διαχέει προς τους κρατικούς φορείς και τους πολίτες, στοιχεία τα οποία προκύπτουν από στατιστικές έρευνες, απογραφές κι άλλες πηγές. Ακόμη, η υπηρεσία αυτή έχει προβεί σε κατάρτιση ψηφιακών χαρτογραφικών υποβάθρων που έχουν αποδοθεί σε αναλογικά διαγράμματα κλίμακας 1:5.000 και αφορούν σε 485 οικισμούς της Ελλάδος με πληθυσμό άνω των 2.000 κατοίκων και σε 125 οικισμούς με πληθυσμό κάτω των 2.000 κατοίκων. Οι παρεχόμενες πληροφορίες περιλαμβάνουν: 1) τους άξονες των δρόμων και την ονοματολογία τους, 2) τα περιγράμματα και την αρίθμηση των οικοδομικών τετραγώνων και 3) τα περιγράμματα και την αρίθμηση των κτιρίων στα πλαίσια των οικοδομικών τετραγώνων. Τα γεωγραφικά δεδομένα της ΕΛ.ΣΤΑΤ. διατίθενται έναντι χρηματικού αντιτίμου.
- Η [Υδρογραφική Υπηρεσία Πολεμικού Ναυτικού \(ΥΥΠΝ\)](#) συλλέγει, επεξεργάζεται, αξιοποιεί και διαθέτει υδρογραφικά, ωκεανογραφικά, χαρτογραφικά και ναυτιλιακά δεδομένα που αφορούν στον ελληνικό και στον ευρύτερο θαλάσσιο χώρο. Οι ναυτικοί χάρτες, τα διαγράμματα, τα ναυτιλιακά βοηθήματα, οι ωκεανογραφικές μελέτες και οι λοιπές εκδόσεις που εκδίδονται από την ΥΥΠΝ διατίθενται σε αναλογική ή ψηφιακή μορφή έναντι χρηματικού αντιτίμου.
- Το [Ινστιτούτο Γεωλογικών και Μεταλλευτικών Ερευνών \(ΙΓΜΕ\)](#) εκπονεί τη γεωλογική μελέτη της Ελλάδας σε συνδυασμό με έρευνες και αξιολογήσεις ορυκτών υλών και υπόγειων νερών. Αποτελεί θεσμοθετημένο τεχνικό σύμβουλο της πολιτείας σε θέματα γεωεπιστημών. Ειδικότερα, στις δραστηριότητες του ΙΓΜΕ περιλαμβάνεται η αποτύπωση της βασικής γεωλογικής δομής της χώρας (χαρτογραφήσεις, γεωφυσικές διασκοπήσεις, κλπ.). Το χαρτογραφικό υλικό καθώς και φωτοαντίγραφα των βιβλίων του ινστιτούτου διατίθενται από το ΙΓΜΕ με χρηματικό αντίτιμο.
- Η εταιρεία [Εθνικό Κτηματολόγιο και Χαρτογράφηση ΑΕ \(ΕΚΧΑ Α.Ε.\)](#) διαχειρίζεται το σύστημα του Εθνικού Κτηματολογίου στο οποίο καταγράφονται πληροφορίες (νομικού, τεχνικού, οικονομικού κ.ά. περιεχομένου) που αφορούν στα ακίνητα της χώρας. Μεταξύ των πληροφοριών αυτών είναι και η γεωγραφική περιγραφή των ακινήτων, δηλαδή το σχήμα, η θέση και οι διαστάσεις τους. Ακόμη, η ΕΚΧΑ Α.Ε. ανέπτυξε το σύστημα εντοπισμού Hellenic Positioning System (HEPOS) που επιτρέπει

τον προσδιορισμό της θέσης σημείων στο χώρο με υψηλή ακρίβεια, στηριζόμενο στο GPS. Τα δεδομένα του ΗΕΡΟΣ διατίθενται προς επαγγελματίες με ειδικότητα συναφή προς αυτήν του τοπογράφου και αγρονόμου μηχανικού (ιδιώτες μελετητές, κατασκευαστές έργων, πανεπιστημιακούς και ερευνητικούς φορείς, δημόσιες υπηρεσίες και οργανισμούς, κλπ.) έναντι εφάπαξ εγγραφής.

- Η διαδικτυακή πύλη www.geodata.gov.gr σχεδιάστηκε, αναπτύχθηκε και συντηρείται από το Ινστιτούτο Πληροφοριακών Συστημάτων του Ερευνητικού Κέντρου «Αθηνά» με σκοπό να αποτελέσει ένα κεντρικό σημείο συλλογής, αναζήτησης, διάθεσης και απεικόνισης της ανοικτής δημόσιας γεωχωρικής πληροφορίας. Ο διαδικτυακός τόπος ενημερώνεται διαρκώς με ανοικτά γεωχωρικά δεδομένα.

2.3 Τύποι χωρικών δεδομένων

2.3.1 Αναπαράσταση Χωρικών Εννοιών

Για την αναπαράσταση χωρικών εννοιών, τα κύρια μοντέλα δεδομένων (data models) που χρησιμοποιούνται είναι το μοντέλο πεδίων (field-based model) και το μοντέλο αντικειμένων (object-based model). Ανάλογα με το είδος των γεωγραφικών δεδομένων, προτιμάται άλλοτε το ένα και άλλοτε το άλλο μοντέλο.

2.3.2 Μοντέλο Πεδίων

Το μοντέλο πεδίων προσεγγίζει τον γεωγραφικό χώρο ως ένα συνεχές χωρικό μέσο το οποίο αποτελείται από επιμέρους υπο-περιοχές, τα πεδία. Ως πεδίο, ορίζεται η μοναδιαία υποδιαίρεση του γεωγραφικού χώρου στην οποία αποδίδεται η τιμή ενός χαρακτηριστικού ή μεγέθους, το οποίο παρουσιάζει συνεχή μεταβολή. Μπορούμε να παρομοιάσουμε τη διαδικασία απόδοσης τιμών στα πεδία με τον υπολογισμό των τιμών μιας συνάρτησης με μεταβλητή τη θέση. Για παράδειγμα, χωρίς να αποκλείεται η επιλογή του μοντέλου αντικειμένων, το μοντέλο πεδίων προτιμάται συνήθως για την αναπαράσταση της σύστασης του εδάφους μιας περιοχής, της υγρασίας, του υψομέτρου, κλπ., επειδή πρόκειται για μεγέθη των οποίων οι τιμές μεταβάλλονται συνεχώς από θέση σε θέση.

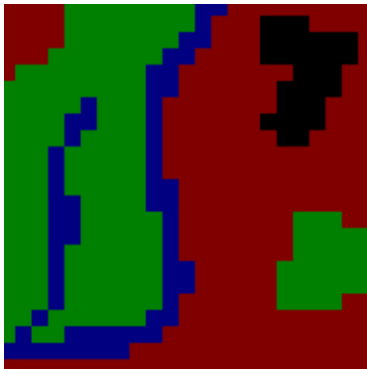
2.3.3 Μοντέλο Αντικειμένων

Βάσει του μοντέλου αντικειμένων, ο γεωγραφικός χώρος περιγράφεται ως αποτελούμενος από αντικείμενα (objects) που διαθέτουν γεωμετρικές ιδιότητες (σχήμα, θέση, κλπ.) και θεματικά χαρακτηριστικά. Προτιμάται έναντι του μοντέλου πεδίων για την αναπαράσταση γεωγραφικών αντικειμένων με διακριτά όρια, όπως π.χ. για την αναπαράσταση των διοικητικών διαιρέσεων μιας περιοχής, των υδρογραφικών και οδικών δικτύων, κλπ.

2.3.4 Δομές Γεωγραφικών Δεδομένων

Οι Δομές Γεωγραφικών Δεδομένων (ΔΓΔ) θεωρούνται ως ο συνδετικός κρίκος μεταξύ του μοντέλου αντίληψης και περιγραφής του γεωγραφικού χώρου και του χώρου που παρέχεται από ένα υπολογιστικό σύστημα. Οι ΔΓΔ ουσιαστικά υλοποιούν τα μοντέλα δεδομένων. Οι ΔΓΔ διακρίνονται στην κανονικοποιημένη δομή (raster structure) και στη διανυσματική δομή (vector structure). Η πρώτη αναπαριστά τα γεωγραφικά δεδομένα με τη βοήθεια πινάκων φαντίων ενώ η δεύτερη κάνει χρήση διανυσμάτων, όπως φανερώνει και το όνομά της. Οι δύο ΔΓΔ παρουσιάζουν ξεχωριστές ιδιότητες, πλεονεκτήματα και μειονεκτήματα. Ακόμη, οι ΔΓΔ διαχειρίζονται τη χωρική διάσταση των γεωγραφικών δεδομένων, έχοντας ως στόχο την όσο το δυνατόν μεγαλύτερη εξοικονόμηση χώρου κατά την αποθήκευσή τους και την ταχύτερη προσπέλαση, επεξεργασία και ανάκτησή τους.

2.3.5 Κανονικοποιημένη Δομή



Εικόνα 12 - Παράδειγμα απεικόνισης γεωγραφικών δεδομένων σύμφωνα με την κανονικοποιημένη δομή

Η κανονικοποιημένη δομή θεωρείται ως η πλέον κατάλληλη για την υλοποίηση του μοντέλου πεδίων, λόγω της δυνατότητάς της να αναπαριστά άμεσα τη μεταβολή της τιμής κάποιου γεωγραφικού χαρακτηριστικού/μεγέθους από θέση σε θέση.

Βάσει της δομής αυτής, τα γεωγραφικά δεδομένα απεικονίζονται σ' ένα χώρο αναπαράστασης ο οποίος αποτελείται από ισομεγέθη χωρία και παρομοιάζεται με κάναβο ή πίνακα φατνίων. Κάθε φατνίο ή εικονοστοιχείο (pixel) χαρακτηρίζεται από: 1) Ένα ζεύγος συντεταγμένων που προσδιορίζουν τη θέση του στον πίνακα (αριθμός γραμμής και αριθμός στήλης) και 2) μια τιμή που αντιπροσωπεύει το χαρακτηριστικό ή την ιδιότητα των γεωγραφικών δεδομένων στη θέση αυτή (π.χ. το είδος της εδαφοκάλυψης, των πετρωμάτων, της βλάστησης, κλπ.).

2.3.6 Ιδιότητες Κανονικοποιημένης Δομής

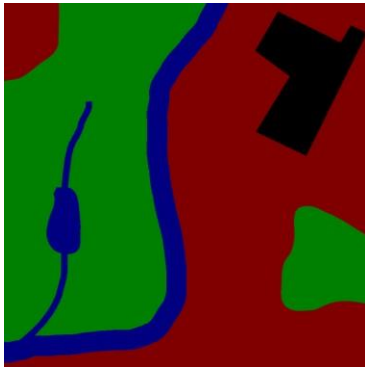
Ανάλογα με τα γεωμετρικά χαρακτηριστικά των αντικειμένων προς απεικόνιση, η κανονικοποιημένη δομή χρησιμοποιεί ένα ή περισσότερα φατνία για την αναπαράστασή τους. Στην απλοποιημένη τους μορφή, ιδεατά, τα σημειακά γεωγραφικά αντικείμενα αναπαρίστανται με τη βοήθεια ενός και μόνο φατνίου, τα δε γραμμικά με μια σειρά από διαδοχικά φατνία, και τα πολυγωνικά αντικείμενα από μια ομάδα γειτονικών φατνίων.

Τα φατνία επιλέγονται να έχουν συγκεκριμένο σχήμα και μέγεθος. Το ευρέως χρησιμοποιούμενο σχήμα φατνίου είναι το τετράγωνο και αυτό διότι είναι συμβατό με τις περισσότερες συσκευές συλλογής και απεικόνισης γεωγραφικών δεδομένων. Ωστόσο, υπάρχουν κανονικοποιημένες δομές που βασίζονται σε πίνακα ορθογώνιων, τριγωνικών και εξαγωνικών φατνίων.

Το μέγεθος των φατνίων ή η χωρική διακριτικότητα (spatial resolution) - όπως αλλιώς αναφέρεται - ποικίλει. Μπορεί να οριστεί μικρότερο του τετραγωνικού μέτρου μέχρι και πολλαπλάσιο του τετραγωνικού χιλιομέτρου. Επειδή η διακριτική ικανότητα του απεικονιζόμενου αποτελέσματος εξαρτάται από το μέγεθος αυτό, πρέπει να επιλέγεται με τρόπο ώστε να εναρμονίζεται με το μέγεθος των γεωγραφικών αντικειμένων. Εάν π.χ. επιλεγεί ένα πολύ μικρό μέγεθος φατνίου σε σχέση με το μέγεθος των απεικονιζόμενων αντικειμένων, η διακριτική ικανότητα θα είναι υψηλή, αλλά ο χώρος αποθήκευσης των δεδομένων θα αυξηθεί σημαντικά και ενδέχεται η επεξεργασία τους να γίνει ιδιαίτερα χρονοβόρα. Κατ' ανάλογο τρόπο, εάν το μέγεθος των φατνίων είναι πολύ μεγάλο σε σχέση με τα απεικονιζόμενα αντικείμενα, τότε θα μειωθούν ο χώρος αποθήκευσης και ο χρόνος επεξεργασίας, αλλά η διακριτική ικανότητα θα είναι πολύ χαμηλή, ενώ ενδέχεται να γίνουν και ασαφή τα όρια μεταξύ των απεικονιζόμενων αντικειμένων.

Λόγω του ότι σε κάθε φαντίο αποδίδεται η τιμή ενός μόνο χαρακτηριστικού ή μίας ιδιότητας δεδομένων, στην περίπτωση που η κανονικοποιημένη δομή χρησιμοποιηθεί για την αναπαράσταση περισσότερων χαρακτηριστικών ή ιδιοτήτων, δημιουργούνται πολλά επίπεδα αναπαράστασης (δηλαδή πολλοί πίνακες φαντίων για την αναπαράσταση της ίδιας περιοχής). Τα επίπεδα αυτά είναι γνωστά ως επιθέματα (layers).

2.3.7 Διανυσματική Δομή



Εικόνα 13 - Η διανυσματική εκδοχή της Εικόνας 8

Ενώ η κανονικοποιημένη δομή θεωρείται καταλληλότερη για την υλοποίηση του μοντέλου πεδίων, η διανυσματική δομή (Σχήμα 3.14) είναι στενά συνδεδεμένη με το μοντέλο αντικειμένων. Η διανυσματική δομή συχνά προτιμάται λόγω της ακρίβειας που προσφέρει κατά την αναπαράσταση γεωγραφικών δεδομένων, της οικονομίας σε χώρο αποθήκευσης, της ποιότητας της απεικόνισης που παρομοιάζεται με αυτήν του αναλογικού χάρτη, και της μεγάλης ποικιλίας λειτουργιών χωρικής ανάλυσης που παρέχονται συνήθως για τα γεωγραφικά δεδομένα της δομής.

Πλεονέκτημα της διανυσματικής δομής είναι ότι το γραφικό προϊόν μοιάζει περισσότερο με τους αναλογικούς χάρτες, αφού δεν εξαρτάται από τη χωρική ανάλυση όπως η κανονικοποιημένη δομή, με αποτέλεσμα κατά την αλλαγή κλίμακας (Σχήματα 3.15 και 3.16) (μεγέθυνση) να μην επηρεάζεται το απεικονιζόμενο αποτέλεσμα αλλά αντιθέτως να εμφανίζονται λεπτομέρειες που ενδεχομένως να μην είχαν γίνει αντιληπτές αρχικά.

Η διανυσματική δομή απεικονίζει τα γεωγραφικά δεδομένα, χρησιμοποιώντας ένα σύνολο γεωμετρικών αρχέτυπων. Όσον αφορά τον δισδιάστατο χώρο αναπαράστασης, τα αρχέτυπα αυτά είναι τα σημεία, οι γραμμές και τα πολύγωνα (ή επίπεδες επιφάνειες) (Σχήμα 3.17). Στον τρισδιάστατο χώρο αναπαράστασης, προστίθενται οι τρισδιάστατες επιφάνειες και οι όγκοι. Στη συνέχεια και για χάριν απλότητας, η περιγραφή της διανυσματικής δομής περιορίζεται στον χώρο δύο διαστάσεων.

2.3.8 Γεωμετρικά Αρχέτυπα

Τα γεωμετρικά αρχέτυπα που χρησιμοποιούνται για την αναπαράσταση διανυσματικών δεδομένων είναι τα εξής:

1. Τα σημεία. Θεωρούνται μηδενικής διάστασης. Αναπαριστούν είτε γεωγραφικές οντότητες με πολύ μικρό μέγεθος σε σχέση με την κλίμακα αναπαράστασης είτε πολύ συγκεκριμένες θέσεις στο χώρο. Στο δισδιάστατο χώρο, προσδιορίζονται μέσω ενός ζεύγους συντεταγμένων.
2. Οι γραμμές. Χρησιμοποιούνται κατά την αναπαράσταση γεωγραφικών οντοτήτων μιας διάστασης. Αποτελούνται από ένα σύνολο διαδοχικών τόξων. Κάθε τόξο

προσδιορίζεται από τις συντεταγμένες των δύο κορυφών του, δηλαδή του σημείου αρχής και του σημείου τέλους του. Μια γραμμή θεωρείται κλειστή εάν το σημείο τέλους του τελευταίου τόξου ταυτίζεται με το σημείο αρχής του πρώτου τόξου.

3. Το πολύγωνο. Θεωρείται αρχέτυπο δύο διαστάσεων. Προσδιορίζεται από την κλειστή τεθλασμένη γραμμή που σχηματίζει το περίγραμμά του. Δύναται να είναι συμπαγές ή να φέρει “τρύπες” στο εσωτερικό του. Μπορεί ακόμη να χαρακτηριστεί ως απλό ή σύνθετο και κυρτό ή μη κυρτό.

2.3.9 Δομές οργάνωσης της διανυσματικής δομής

Spaghetti

Η απλούστερη διανυσματική δομή γεωγραφικών δεδομένων είναι η δομή Spaghetti. Μπορεί να θεωρηθεί ως μία γραμμή-προς-γραμμή αντιστοίχιση δεδομένων αναλογικής μορφής (π.χ. ενός χάρτη) σε δεδομένα ψηφιακής μορφής. Κάθε γεωγραφικό αντικείμενο κωδικοποιείται ανεξάρτητα, δηλαδή χωρίς καμία σύνδεση ή συσχέτιση με τα υπόλοιπα, και προσδιορίζεται από μια σειρά σημείων.

Τα σημεία κωδικοποιούνται ψηφιακά με τη βοήθεια ζευγών συντεταγμένων (x, y) . Οι γραμμές κωδικοποιούνται ως σειρές σημείων που ορίζουν διαδοχικά τόξα. Τα πολύγωνα κωδικοποιούνται, όπως οι γραμμές, από σειρές σημείων, με τη διαφορά ότι το πρώτο και το τελευταίο σημείο ταυτίζονται.

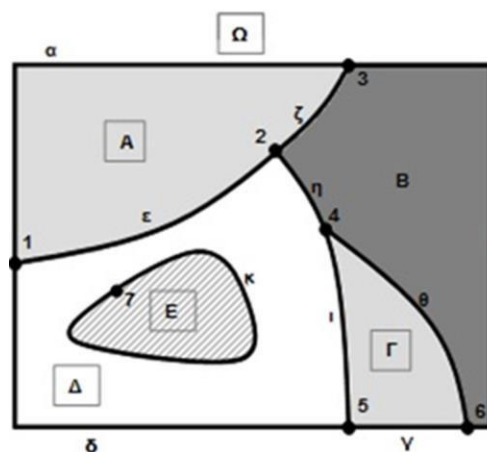
Στην ουσία, τα αντικείμενα κωδικοποιούνται βάσει των γραμμών που αποτελούν το περίγραμμά τους.

Γεωμετρία Αντικειμένου	Προσδιορισμός
Σημείο	(x,y)
Γραμμή	$((x_1,y_1), (x_2,y_2), \dots, (x_n,y_n))$
Πολύγωνο	$((x'_1,y'_1), (x'_2,y'_2), \dots, (x'_n,y'_n), (x'_1,y'_1))$
	$((x''_1,y''_1), (x''_2,y''_2), \dots, (x''_n,y''_n), (x''_1,y''_1))$

Τοπολογική Δομή

Η τοπολογική δομή δεν είναι παρά μια διανυσματική δομή στην οποία καταγράφονται και διαχειρίζονται οι τοπολογικές σχέσεις γεωγραφικών αντικειμένων.

Η καταγραφή των τοπολογικών σχέσεων που διέπουν μια συλλογή γεωγραφικών αντικειμένων αυξάνει τον απαιτούμενο χώρο αποθήκευσης στο υπολογιστικό σύστημα και προϋποθέτει την εκτέλεση πολύπλοκων και χρονοβόρων αλγορίθμων για τον προσδιορισμό τους.



Εικόνα 14 - Παράδειγμα τοπολογικής δομής

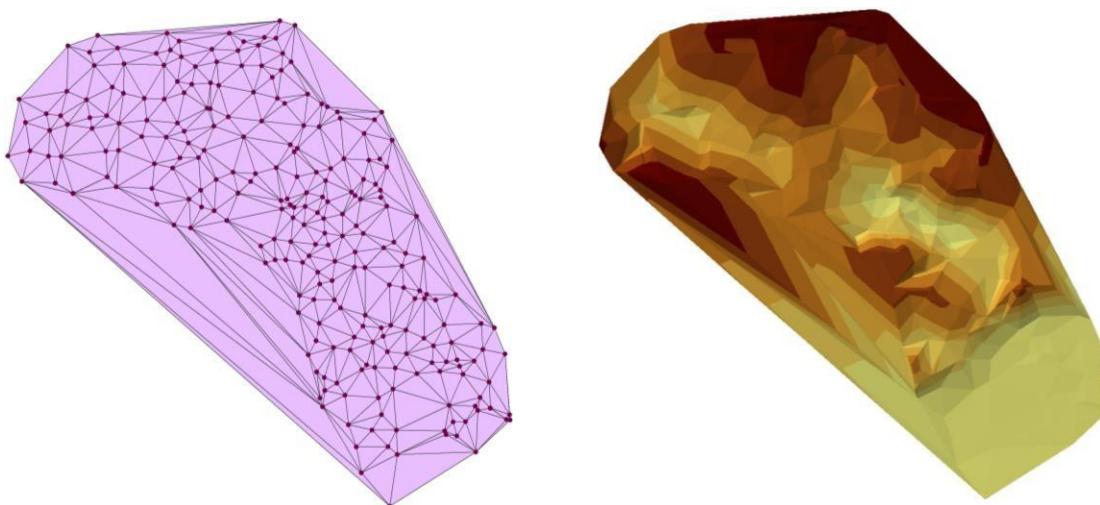
Ωστόσο, τα οφέλη από τη χρησιμοποίηση τοπολογικής δομής είναι σημαντικά. Για παράδειγμα, η τοπολογική δομή συμβάλλει στον εντοπισμό λαθών. Η τοπολογική δομή βοηθά στον έλεγχο της ορθότητας των δεδομένων επισημαίνοντας λάθη όπως: μη συνδεσιμότητα των αντικειμένων που απαρτίζουν ένα δίκτυο, επικάλυψη ή ύπαρξη κενού χώρου μεταξύ πολυγωνικών αντικειμένων που θα έπρεπε αντ' αυτού να εφάπτονται, παράληψη καταχώρησης στοιχείων για τα σημεία τομής γραμμικών αντικειμένων, πολλαπλή καταχώρηση ίδιων αντικειμένων, κλπ.

Δίκτυο Ακανόνιστων Τριγώνων (ΔΑΤ) - (Triangulated Irregular Network - TIN)

Πρόκειται για μια ειδική περίπτωση διανυσματικής τοπολογικής δομής. Το ΔΑΤ σχηματίζεται βάσει του τριγωνισμού Delaunay έτσι ώστε τα σημεία να ενώνονται μεταξύ τους με ευθύγραμμα τμήματα σχηματίζοντας μη επικαλυπτόμενα τρίγωνα.

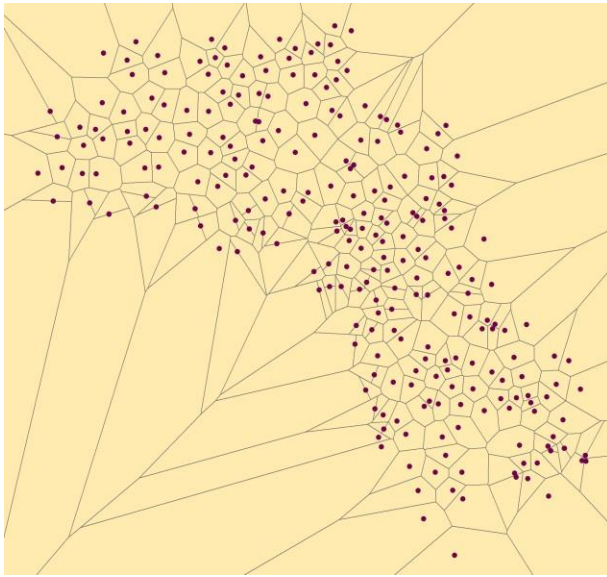
Ο τριγωνισμός Delaunay εφευρέθηκε από τον Ρώσο μαθηματικό Boris Delaunay (1934). Αρχικά, δημιουργούνται κύκλοι που διέρχονται από τρία σημεία του χώρου και δεν περικλείουν στο εσωτερικό τους κανένα άλλο σημείο. Κάθε τριάδα σημείων, που συμμετέχει στη δημιουργία ενός τέτοιου κύκλου, σχηματίζει ένα τρίγωνο. Το σύνολο των τριγώνων αυτών αποκαλείται τριγωνισμός Delaunay. Οι ιδιότητες τριγωνισμού Delaunay είναι: α) διαμορφώνονται όσο το δυνατόν πιο ισόπλευρα τρίγωνα, β) θεωρείται μοναδικός και γ) δίνει το ίδιο αποτέλεσμα ανεξάρτητα από το σημείο εκκίνησης δημιουργίας των τριγώνων.

Το ΔΑΤ δημιουργείται βάσει ενός συνόλου σημείων που χαρακτηρίζονται από ένα ζεύγος συντεταγμένων (x, y) στο επίπεδο και μια τιμή z στην κάθετη διάσταση. Χρησιμοποιείται για την αναπαράσταση επιφανειών 2.5 διαστάσεων. Ο όρος 2.5 διαστάσεις μπορεί να ξενίσει αλλά μια αληθινά τρισδιάστατη αναπαράσταση μοντελοποιεί την τιμή z ως τρίτη διάσταση του διανύσματος θέσης στον τρισδιάστατο χώρο (x,y,z) και όχι ως μια τιμή χαρακτηριστικού για το διάνυσμα θέσης στον δισδιάστατο χώρο (x,y) .



Εικόνα 15 - Δημιουργία τριγωνισμού Delaunay από σημειακά δεδομένα με πληροφορία υψομέτρου και σχηματισμός ΔΑΤ από υψομετρικά δεδομένα.

Πολύγωνα Thiessen



Εικόνα 16 - Τα πολύγωνα Thiessen που προκύπτουν από το δίκτυο σημείων της Εικόνας 11

Τα πολύγωνα Thiessen (ή διάγραμμα Voronoi) προτάθηκαν από τους Thiessen και Alter (1911) ως μέθοδος παρεμβολής για την εκτίμηση του ποσοστού βροχόπτωσης σε περιοχές γύρω από θέσεις εγκατεστημένων βαρομέτρων. Σχηματίζονται από την ένωση των κέντρων των ίδιων κύκλων που χρησιμοποιήθηκαν κατά τον τριγωνισμό Delaunay ή βάσει του τριγωνισμού Delaunay, με την χάραξη των μεσοκαθέτων των τριγώνων σταματώντας στα σημεία τομής τους. Υποστηρίζουν σύνθετες γεωμετρικές λειτουργίες (αναζήτηση του πλησιέστερου γείτονα, μεθόδους γενίκευσης δεδομένων κλπ.).

Είναι γνωστά και ως πολύγωνα εγγύτητας επειδή αυτός ο τύπος ανάλυσης χρησιμοποιείται ιδιαίτερα για τον προσδιορισμό της περιοχής που είναι πιο κοντά σ' ένα σημείο ενδιαφέροντος π.χ. νοσοκομείο ή σχολείο. Τα πολύγωνα Thiessen χρησιμοποιούνται ευρέως σε πλήθος πεδίων όπως στη γεωλογία, την οικολογία, τη βιολογία κ.ά.

2.4 Μορφότυποι χωρικών δεδομένων

Κομμάτι της πολυπλοκότητας του ζητήματος αποτελεί το πλήθος των μορφότυπων στους οποίους αποθηκεύονται τα χωρικά δεδομένα. Ο καθένας με την ιδιαίτερή του αρχιτεκτονική και θετικά και αρνητικά σε σχέση με τους άλλους. Παρακάτω παρατίθενται οι γνωστότεροι.

2.4.1 Μορφότυποι διανυσματικών δεδομένων

Esri Shapefile

Το shapefile της ESRI είναι από τους πιο συνηθισμένους τύπους αρχείων για την αποθήκευση χωρικών δεδομένων.

Απαιτεί ένα σύνολο τριών αρχείων για να λειτουργήσει:

- SHP - περιέχει τις γεωμετρίες
- SHX - περιέχει την αρίθμηση των γεωμετριών
- DBF - περιέχει την περιγραφική πληροφορία των γεωμετριών

Προαιρετικά περιέχει και τα παρακάτω αρχεία:

- PRJ - Μεταδεδομένα που περιγράφουν το σύστημα αναφοράς
- XML - Τα συνοδευτικά μεταδεδομένα
- SBN - Η χωρική δεικτοδότηση των γεωμετριών
- SBX - Βελτιστοποίηση χρόνων φόρτωσης

Geographic JavaScript Object Notation (GeoJSON)

Η μορφή GeoJSON χρησιμοποιείται ως επί το πλείστον για διαχείριση χωρικών δεδομένων μέσω διαδικτύου. Το GeoJSON αποθηκεύει τις συντεταγμένες ως κείμενο σε μορφή JavaScript (Object Notation). Αποθηκεύει τόσο γεωμετρίες όσο και περιγραφικά δεδομένα.

Το GeoJSON αποθηκεύει αντικείμενα μέσα σε {} και γενικά έχει μικρό κόστος σε υπολογιστική ισχύ. Έχει απλή σύνταξη που τροποποιείται σε οποιοδήποτε πρόγραμμα επεξεργασίας κειμένου. Οι φυλλομετρητές περιήγησης στο διαδίκτυο κατανοούν τη JavaScript, και άρα από προεπιλογή το GeoJSON είναι μια κοινή μορφή ιστού.

Geography Markup Language (GML)

Το GML επιτρέπει τη χρήση γεωγραφικών συντεταγμένων όντας επέκταση της XML. Όπως και η XML, είναι αναγνώσιμη τόσο από άνθρωπο όσο και από μηχανές.

Το GML αποθηκεύει γεωγραφικές οντότητες (χαρακτηριστικά) με τη μορφή κειμένου. Παρόμοια με το GeoJSON, το GML μπορεί να ενημερωθεί σε οποιοδήποτε πρόγραμμα

επεξεργασίας κειμένου. Κάθε αντικείμενο έχει έναν κατάλογο ιδιοτήτων, γεωμετρία (σημεία, γραμμές, καμπύλες, επιφάνειες και πολύγωνα) και χωρικό σύστημα αναφοράς. Χρησιμοποιείται κυρίως σε χωρικές εφαρμογές διαδικτύου και έχει μεγαλύτερες απαιτήσεις σε επικοινωνία δικτύου απ' ό,τι το GeoJSON.

Google Keyhole Markup Language (KML/KMZ)

Και αυτός ο μορφότυπος βασίζεται πάνω στην XML. Χρησιμοποιείται κυρίως για το Google Earth. Το KMZ (KML-Zipped) αντικατέστησε το KML ως την προεπιλεγμένη γεωχωρική μορφή του Google Earth, επειδή είναι μια συμπιεσμένη έκδοση του αρχείου. Το KML / KMZ έγινε το διεθνές πρότυπο της Open Geospatial Consortium το 2008.

Τα στοιχεία γεωγραφικού μήκους, γεωγραφικού πλάτους (δεκαδικοί βαθμοί) είναι όπως ορίζονται από το Παγκόσμιο Γεωδαιτικό Σύστημα του 1984 (WGS84). Η κατακόρυφη συνιστώσα (υψόμετρο) μετράται σε μέτρα από το κατακόρυφο γεωγραφικό σημείο γεωμετρίας EGG96 EGM96.

GPS eXchange Format (GPX)

Το GPX είναι και αυτό τύπου XML και χρησιμοποιείται για να περιγράψει σημεία και διαδρομές που αποτυπώνει ένας δέκτης GPS. Είναι μορφότυπος μεταφοράς δεδομένων και γι' αυτό χρησιμοποιείται ευρέως χωρίς να χρειάζεται κάποια μετατροπή. Η ελάχιστη πληροφορία που απαιτείται είναι ένα ζεύγος συντεταγμένων αλλά μπορεί να περιέχει και άλλη πληροφορία όπως χρόνος, υψόμετρο κτλ.

OpenStreetMap OSM XML

Ο μορφότυπος OSM είναι και αυτός τύπου XML και περιέχει διανυσματικά δεδομένα του OpenStreetMap. Υπάρχει και σε ελαφρύτερη μορφή (δυαδική) PBF.

Digital Line Graph (DLG)

Διανυσματικά αρχεία από ψηφιοποιήσεις αναλογικών τοπογραφικών χαρτών πάνω στα οποία χτίστηκαν τα δεδομένα της υπηρεσίας απογραφής των ΗΠΑ.

MapInfo TAB

Αρχεία τύπου ASCII τα οποία χρησιμοποιεί το πρόγραμμα MapInfo. Και αυτά όπως το shapefile αποτελούνται από μια συλλογή αρχείων.

- TAB - Το αρχείο που συνδέει όλα τα άλλα αρχεία
- DAT - περιέχει τα περιγραφικά δεδομένα
- ID - συσχετίζει τις γεωμετρίες με τα περιγραφικά δεδομένα

- MAP - περιέχει τις γεωμετρίες
- IND - δεικτοδοτεί τα περιγραφικά δεδομένα

Geographic Base File-Dual Independent Mask Encoding (GBF-DIME)

Η μορφή αρχείου GPF-DIME αναπτύχθηκε από το Γραφείο Απογραφής των ΗΠΑ στα τέλη της δεκαετίας του 1960 ως μία από τις πρώτες μορφές δεδομένων GIS που υπάρχουν. Χρησιμοποιήθηκε για την αποθήκευση του οδικού δικτύου των ΗΠΑ για μεγάλες αστικές περιοχές, που αποτελεί βασικό παράγοντα για την απογραφή των απογραφών.

Το DIME αποτελεί βασικό στοιχείο του τρέχοντος συστήματος TIGER (Topologically Integrated Geographic Encoding and Reference), το οποίο δημιουργήθηκε από το Αμερικανικό Γραφείο Απογραφής.

2.4.2 Μορφότυποι ψηφιδωτών δεδομένων

Τα ψηφιδωτά δεδομένα είναι δυνατόν να απεικονιστούν σχεδόν από οποιοδήποτε μορφότυπο ψηφιακής εικόνας. Οι πιο γνωστοί εξειδικευμένοι μορφότυποι είναι:

ERDAS Imagine (IMG)

Το ERDAS Imagine IMG αρχεία είναι μια ιδιόκτητη μορφή αρχείου που αναπτύχθηκε από το Hexagon Geospatial. Τα αρχεία IMG χρησιμοποιούνται συνήθως για δεδομένα ράστερ για την αποθήκευση ενιαίων και πολλαπλών ζωνών δορυφορικών δεδομένων.

Τα αρχεία IMG χρησιμοποιούν μια ιεραρχική μορφή (HFA) που είναι προαιρετικά για την αποθήκευση βασικών πληροφοριών σχετικά με το αρχείο. Για παράδειγμα, αυτό μπορεί να περιλαμβάνει πληροφορίες αρχείου, σημεία ελέγχου εδάφους και τύπο αισθητήρα.

Κάθε στρώμα raster ως μέρος ενός αρχείου IMG περιέχει πληροφορίες σχετικά με τις τιμές δεδομένων του. Για παράδειγμα, αυτό περιλαμβάνει την προβολή, τα στατιστικά στοιχεία, τις ιδιότητες, τις πυραμίδες και αν είναι ή όχι ένας συνεχής ή διακριτός τύπος ράστερ.

American Standard Code for Information Interchange ASCII Grid

Το ASCII χρησιμοποιεί ένα σύνολο αριθμών (συμπεριλαμβανομένων των πλωτών) μεταξύ 0 και 255 για αποθήκευση και επεξεργασία πληροφοριών. Περιέχουν επίσης πληροφορίες κεφαλίδας με ένα σύνολο λέξεων-κλειδιών.

Στη φυσική τους μορφή, τα αρχεία κειμένου ASCII αποθηκεύουν δεδομένα GIS σε μια οριοθετημένη μορφή. Αυτό θα μπορούσε να είναι κόμμα, διαστήματος ή tab-delimited μορφή. Μεταβαίνοντας από μη χωρικά σε χωρικά δεδομένα, μπορείτε να εκτελέσετε ένα εργαλείο διαδικασίας μετατροπής, όπως το ASCII, για να δημιουργήσετε raster.

GeoTIFF

Το GeoTIFF έχει γίνει ένα πρότυπο αρχείο εικόνας για εφαρμογές GIS και δορυφορικής τηλεπισκόπησης. Τα GeoTIFFs μπορούν να συνοδεύονται από άλλα αρχεία:

- TFW - αρχείο γεωεντοπισμού
- XML - τα μεταδεδομένα
- AUX - βοηθητικό αρχείο που αποθηκεύει προβολές και άλλες πληροφορίες
- OVR - αρχεία πυραμίδας που βελτιώνουν την ταχύτητα απόδοσης της πληροφορίας

ER Mapper Enhanced Compression Wavelet (ECW)

Το ECW είναι μια μορφή συμπιεσμένης εικόνας που χρησιμοποιείται συνήθως για εναέρια και δορυφορική απεικόνιση. Αυτός ο τύπος αρχείου είναι γνωστός για τις υψηλές αναλογίες συμπίεσης διατηρώντας παράλληλα την ποιότητα αντίθεση στις εικόνες.

Η μορφή ECW αναπτύχθηκε από την ER Mapper, αλλά τώρα ανήκει στο Hexagon Geospatial.

Envi RAW Raster

Τα αρχεία Interleaved Band είναι μια επέκταση αποθήκευσης ράστερ για εναέρια και δορυφορικές εικόνες μονής / πολλαπλής ζώνης.

- Το Band Interleaved for Line (BIL) αποθηκεύει πληροφορίες εικονοστοιχείων βασισμένες σε σειρές για όλες τις ζώνες μιας εικόνας.
- Το Band Interleaved by pixel (BIP) εκχωρεί τις τιμές των εικονοστοιχείων για κάθε ζώνη κατά σειρές.
- Τέλος, η ακολουθία μορφής Band (BSQ) αποθηκεύει ξεχωριστές ζώνες ανά σειρές.

Τα αρχεία BIL αποτελούνται από ένα αρχείο κεφαλίδας (HDR) που περιγράφει τον αριθμό των στηλών, των γραμμών, των ζωνών, το βάθος των δυαδικών ψηφίων και τη διάταξη σε μια εικόνα.

PCI Geomatics Database File (PCIDSK)

Τα αρχεία PIX είναι στρώματα αποθήκευσης ράστερ που αναπτύχθηκαν από την PCI Geomatics. Είναι ένας ευέλικτος τύπος αρχείου που αποθηκεύει όλα τα δεδομένα εικόνας και βοηθητικά δεδομένα που ονομάζονται "τμήματα" σε ένα αυτόνομο αρχείο. Για παράδειγμα, τα τμήματα μπορούν να περιλαμβάνουν κανάλια εικόνας, τοποθεσία εκπαίδευσης και πληροφορίες ιστόγραμμα.

Ως αρχείο βάσης δεδομένων, τα αρχεία PIX μπορούν να χωρέσουν κανάλια ράστερ με διαφορετικά βάθη bit. Μπορούν επίσης να αποθηκεύουν προβολές, πληροφορίες χαρακτηριστικών, μεταδεδομένα και εικόνες / φορείς.

Esri Grid

Τα αρχεία Grid αποτελούν ιδιόκτητο σχήμα που αναπτύχθηκε από την Esri. Τα Grid δεν έχουν επέκταση και είναι μοναδικά επειδή μπορούν να διατηρούν επιπλέον (αριθμητικά) δεδομένα σε ένα αρχείο ράστερ. Αυτά αποθηκεύονται σε πίνακες - ένας για κάθε μοναδική τιμή στο πλέγμα και ο αριθμός που αντιπροσωπεύει τον αριθμό των κελιών.

2.5 Χωρικές Βάσεις Δεδομένων

Μια βάση δεδομένων (ΒΔ) μπορεί να ορισθεί ως μια οργανωμένη συλλογή συσχετιζόμενων δεδομένων. Ειδικότερα, μια βάση δεδομένων ονομάζεται βάση γεωγραφικών δεδομένων (ΒΓΔ) όταν τα κύρια δεδομένα της συνδέονται με μια συγκεκριμένη θέση στον γεωγραφικό χώρο μέσω ενός συστήματος συντεταγμένων. Για παράδειγμα, τα δεδομένα μιας βάσης μπορεί να σχετίζονται με την περιγραφή κάποιας γεωγραφικής περιοχής ή ενός γεωγραφικού φαινομένου. Μια ΒΓΔ διαφέρει από μια βάση χωρικών δεδομένων στο ότι τα δεδομένα της δεύτερης μπορούν να αναφέρονται και σε χώρο μη γεωγραφικής κλίμακας.

Οι ΒΓΔ σχεδιάζονται, δημιουργούνται και συντηρούνται με τρόπο ώστε να μπορούν να στηρίζουν τις διαδικασίες ανάλυσης, εξαγωγής συμπερασμάτων και λήψης αποφάσεων που συντελούνται στο πλαίσιο των ΣΓΠ. Σε σχέση με τα απλά αρχεία δεδομένων, οι ΒΓΔ προσφέρουν πολλά πλεονεκτήματα, όπως:

- εφαρμογή ενός αποτελεσματικότερου τρόπου οργάνωσης των δεδομένων, αποφεύγοντας την αποθήκευση περιττών πληροφοριών
- γρηγορότερη και ευκολότερη ενημέρωση και αναζήτηση δεδομένων
- ταυτόχρονη προσπέλαση της ίδιας ΒΓΔ από πολλούς χρήστες, εξοικονομώντας χρόνο και διευκολύνοντας τη διαχείριση των δεδομένων

2.5.1 Λίστα Χωρικών Βάσεων Δεδομένων

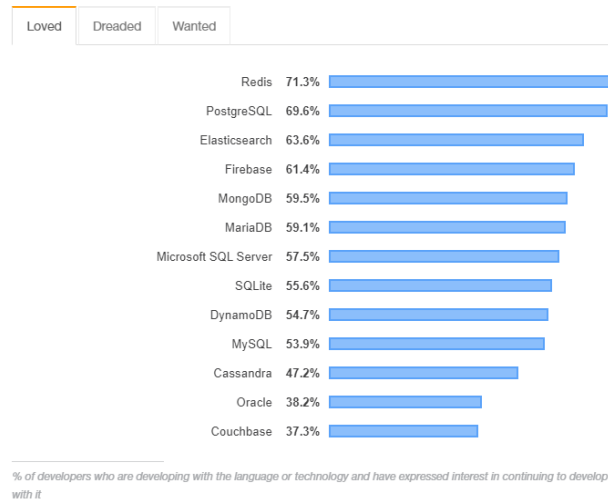
Βάσεις Δεδομένων σε ένα αρχείο:

- Esri File Geodatabase
- Esri Personal Geodatabase
- OGC GeoPackage
- Mapbox MBTiles
- GE Smallworld Version Managed Data Store
- SQLite/Spatialite

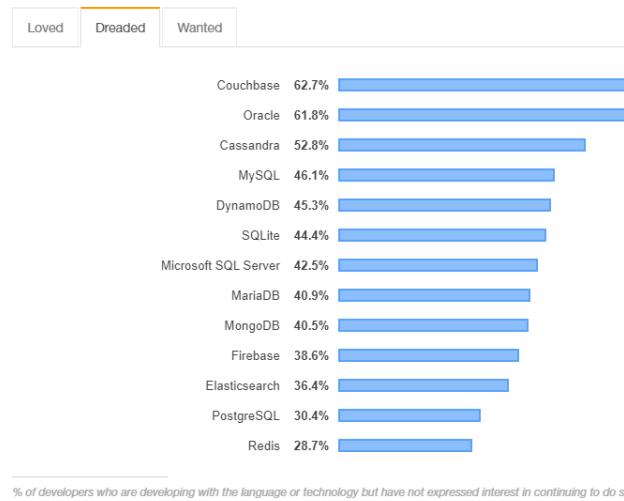
Σχεσιακές βάσεις δεδομένων

- PostgreSQL/PostGIS
- ORACLE Spatial
- SQL SERVER

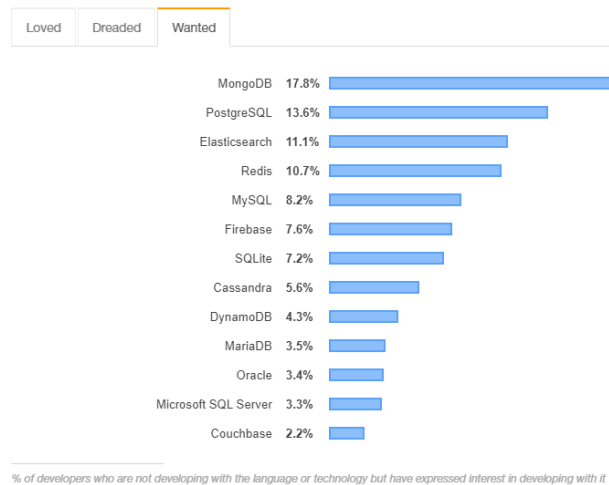
Most Loved, Dreaded, and Wanted Databases



Most Loved, Dreaded, and Wanted Databases



Most Loved, Dreaded, and Wanted Databases



Εικόνα 17 - Ιστογράμματα με τις πιο αγαπητές, μισητές και "ποθητές" βάσεις, βάσει του ερωτηματολογίου του Stack Overflow για το2019

2.6 Ειδικοί τύποι χωρικών δεδομένων

2.6.1 Μεγάλα Χωρικά δεδομένα

Το 2001 ο αναλυτής Doug Laney, αναφέρει για πρώτη φορά, ως χαρακτηριστικά των δεδομένων, τις λέξεις “Volume, Velocity και Variety”. Περίπου μια δεκαετία αργότερα, οι λέξεις αυτές θα γίνουν κοινώς αποδεκτές ώστε να περιγραφούν οι τρεις άξονες των Μεγάλων Δεδομένων. Οι άξονες αυτοί περιγράφουν τόσο τις ευκαιρίες όσο και τις προκλήσεις με τις οποίες ερχόμαστε αντιμέτωποι και αποτελούν τα διάσημα ‘3V’s’ που σήμερα αναφέρονται κατά τον ορισμό των Μεγάλων Δεδομένων, παρ’ ότι ο όρος αυτός δεν υπήρχε αυτοτελώς στο άρθρο του Doug Laney. Πιο συγκεκριμένα, ο όρος “Volume” αναφέρεται στον όγκο των δεδομένων, ο όρος ‘Velocity’ αναφέρεται στην συχνότητα με την οποία παράγεται κάποια πληροφορία, για παράδειγμα ωριαία, ημερήσια κλπ. και ο όρος “Variety” αναφέρεται στην ποικιλία των δεδομένων όπως αυτά προκύπτουν από διαφορετικές πηγές π.χ. αισθητήρες, μέσα κοινωνικής δικτύωσης κ.α. Κατά διαστήματα έχουν προστεθεί επιπλέον όροι όπως για παράδειγμα “Veracity, Variability, Validity, Volatility, Visibility, Value και Visualization”, δεν χρησιμοποιούνται εν τούτοις σταθερά διότι δεν εκφράζουν ιδιότητες του μεγέθους και επίσης κάθε επιμέρους χαρακτηριστικό μπορεί να αφορά περισσότερο κάποιες συγκεκριμένες εφαρμογές σε σχέση με κάποιες άλλες.

Ο ορισμός των Μεγάλων Δεδομένων τείνει να διαφοροποιείται ελαφρώς από τομέα σε τομέα και ανάλογα με τις διάφορες οπτικές, π.χ. ακαδημαϊκά, βιομηχανικά, τεχνολογικά κ.α. Αυτός είναι και λόγος που δεν υπάρχει ένας καθολικός ορισμός. Σε γενικές γραμμές, Μεγάλα Δεδομένα (Big Data) μπορούν να χαρακτηριστούν τα δομημένα και μη - δομημένα δεδομένα μεγάλου όγκου που δεν είναι δυνατόν να αποκτηθούν, αποθηκευτούν, διαχειριστούν και να αναλυθούν από τα παραδοσιακά hardware, software και τις βάσεις δεδομένων.

Όταν τα Μεγάλα δεδομένα έχουν χωρική διάσταση τότε αναφερόμαστε στα Μεγάλα Χωρικά Δεδομένα.

2.6.2 Χωροχρονικά Δεδομένα

Η συνεχής συλλογή γεωχωρικών δεδομένων έχει ως αποτέλεσμα την δημιουργία ενός συνόλου πληροφορίας η οποία εκτείνεται, εκτός από τον χώρο, και στον χρόνο. Η πληροφορία αυτή είναι γνωστή ως Χωροχρονικά Δεδομένα (Spatiotemporal Data). Στην πραγματικότητα, το μεγαλύτερο ποσοστό όλων των δεδομένων που μπορούν να συλλεχθούν είναι χωροχρονικά δεδομένα, διότι το πιθανότερο είναι ότι έχουν τόσο χωρική όσο και χρονική υπόσταση, είτε έχουν διάρκεια, είτε επαναλαμβάνονται, είτε συνέβησαν σε μια συγκεκριμένη χρονική στιγμή και τελειώσαν. Επομένως, τα Μεγάλα Χωροχρονικά Δεδομένα (Spatiotemporal Big Data) είναι συνήθως η πιο ολοκληρωμένη φύση μιας πληροφορίας και αποτελούν τα καταλληλότερα δεδομένα για μια χωρο-χρονική ανάλυση.

Η διάσταση του χρόνου ενεργοποιεί την δυνατότητα πραγματοποίησης ακόμα πιο σύνθετων ερωτημάτων σε σχέση με τα χωρικά δεδομένα, αλλά αυξάνει την δυσκολία διαχείρισης και μοντελοποίησης της πληροφορίας. Υπάρχουν διάφορες προσεγγίσεις για τον τύπο

μοντελοποίησης των χωροχρονικών δεδομένων: τα στιγμιότυπα (Snapshots), η αποτύπωση του χρόνου (Time-Stamping), το ιστορικό γράφημα (History Graph), τα τρία πεδία (Tree-Domain), τα κινούμενα αντικείμενα (Moving Objects) και το αντικειμενοστρεφές μοντέλο (Object-oriented Model).

2.6.3 Χωρικά Blockchain

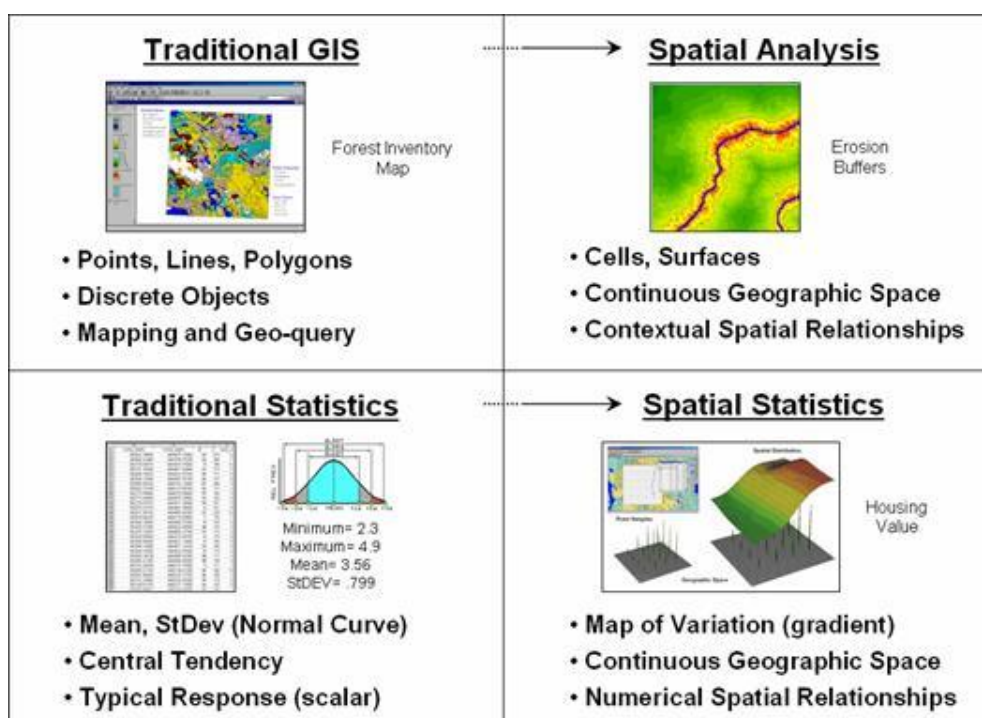
Η τεχνολογία του Blockchain εμφανίστηκε με το κρυπτονόμισμα Bitcoin και επέτρεψε ασφαλείς και αμετάβλητες συναλλαγές μεταξύ των μελών (peer-to-peer) χωρίς μεσάζοντες για να εγγυηθούν την συναλλαγή (trusted authorities). Ουσιαστικά αποτελεί μια κατακευματισμένη βάση δεδομένων όπου αποθηκεύεται κρυπτογραφημένη πληροφορία με τρόπο που να δημιουργεί μια διαδοχή δεδομένων. Εγγυητές μιας συναλλαγής είναι οι χρήστες, οι οποίοι αποτελούν τους κόμβους της κατακευματισμένης βάσης, και άρα οι συναλλαγές είναι δημόσια προσβάσιμες. Η τεχνολογία αυτή βρίσκει σταδιακά τις εφαρμογές της πάνω σε χωρικά δεδομένα.

ΚΕΦΑΛΑΙΟ 3

Ανάλυση Χωρικών Δεδομένων

Η ανάλυση χωρικών δεδομένων ή “χωρική στατιστική” περιλαμβάνει το σύνολο εκείνων των τεχνικών που εφαρμόζονται στη μελέτη οντοτήτων χρησιμοποιώντας τις τοπολογικές, γεωμετρικές ή γεωγραφικές τους ιδιότητες. Η χωρική ανάλυση περιλαμβάνει μια ποικιλία τεχνικών, πολλές ακόμα στην πρώιμη ανάπτυξή τους, ενσωματώνει διαφορετικές αναλυτικές προσεγγίσεις και εφαρμόζεται σε ποικίλους τομείς. Υπό μία έννοια, η χωρική ανάλυση είναι οι τεχνικές που εφαρμόζονται στην ανάλυση των γεωγραφικών δεδομένων σε όλο το εύρος της ανθρώπινης δραστηριότητας,

Σύνθετα ζητήματα προκύπτουν από τη χωρική ανάλυση, πολλά από τα οποία ούτε είναι σαφώς καθορισμένα ούτε έχουν επιλυθεί πλήρως, αλλά αποτελούν αντικείμενο σύγχρονης διεπιστημονικής έρευνας. Το πιο θεμελιώδες από αυτά είναι το πρόβλημα του καθορισμού της χωρικής θέσης των φορέων που μελετώνται.



Εικόνα 18- Η Χωρική Ανάλυση και Χωρική στατιστική *Beyond Mapping III - innovativegis*

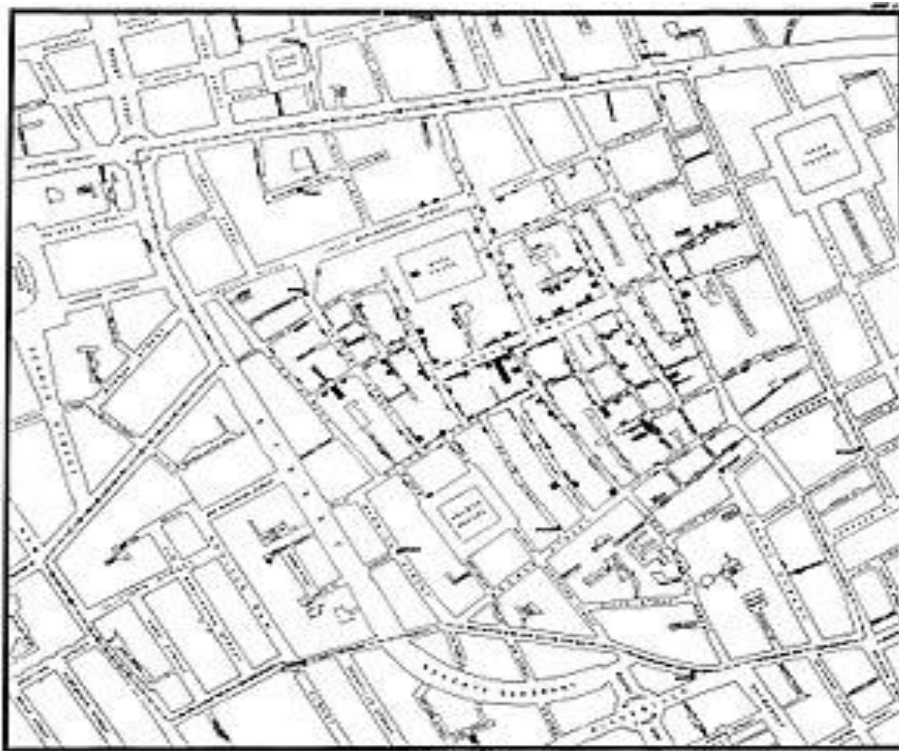
Η ταξινόμηση των τεχνικών χωρικής ανάλυσης είναι δύσκολη λόγω του μεγάλου αριθμού διαφορετικών πεδίων έρευνας, των διαφορετικών θεμελιωδών προσεγγίσεων που μπορούν να επιλεγούν και των πολλών μορφών που μπορούν να λάβουν τα δεδομένα.

Η χωρική ανάλυση μπορεί ίσως να θεωρηθεί ότι προέκυψε με πρώτες προσπάθειες χαρτογράφησης και τοπογραφίας, αλλά η αλήθεια είναι πως πολλοί επιστημονικοί τομείς συνέβαλαν στην εξέλιξή της στη σύγχρονη της μορφή και στη διεύρυνση της χρήσης της ως

εργαλείο. Τα μαθηματικά παρέχουν τα θεμελιώδη εργαλεία ανάλυσης, η επιστήμη των υπολογιστών συμβάλει εκτενώς στη μελέτη αλγορίθμων, κυρίως στην υπολογιστική γεωμετρία, η τηλεπισκόπηση συνεισφέρει με τη μορφομετρική ανάλυση και την ανάλυση συστάδων, και τα GIS (Γεωγραφικά Συστήματα Πληροφοριών) αποτελούν τον τεχνικό/τεχνολογικό ιστό που “δένει” όλα τα παραπάνω και αποτελεί το λογισμικό διαχείρισης χωρικών δεδομένων και τη σύγχρονη εργαλειοθήκη αναλυτικών εργαλείων.

Μια από τις πλέον χαρακτηριστικές εφαρμογές της χωρικής ανάλυσης με αποτελέσματα για τον άνθρωπο σε πλανητικό επίπεδο αποτελεί η επιδημιολογία, στο έργο του John Snow για τη χαρτογράφηση μιας επιδημίας χολέρας, με την έρευνα και τη χαρτογράφηση της εξάπλωσης της νόσου και με τις μελέτες τοποθεσίας για την παροχή υπηρεσιών υγείας.

Ο Snow χρησιμοποίησε έναν χάρτη συγκέντρωσης για να απεικονίσει με κουκίδες τη συγκέντρωση περιπτώσεων χολέρας γύρω από αντλία νερού. Χρησιμοποίησε επίσης στατιστικές για να δείξει τη σχέση μεταξύ της ποιότητας, της πηγής νερού και των περιπτώσεων χολέρας. Έδειξε ότι η Southwark και η εταιρεία Waterworks της Vauxhall έπαιρναν νερό από τμήματα του Τάμεση που είχαν μολυνθεί από τα λύματα και έδιναν νερό στα σπίτια οδηγώντας σε αυξημένη συχνότητα εμφάνισης χολέρας. Η μελέτη του Snow ήταν ένα σημαντικό γεγονός στην ιστορία της δημόσιας υγείας και της γεωγραφίας. Θεωρείται ως το ιδρυτικό γεγονός της επιστήμης της επιδημιολογίας



Εικόνα 19 - Αρχικός χάρτης από τον John Snow που δείχνει τις συστάδες περιπτώσεων χολέρας στην επιδημία του Λονδίνου του 1854, όπως σχεδιάστηκε και λιθογραφήθηκε από τον Charles Cheffins.

3.1 Δομικά ζητήματα της χωρικής ανάλυσης

Η χωρική ανάλυση αντιμετωπίζει πολλά δομικά ζητήματα στον ορισμό των αντικειμένων που μελετά, στην κατασκευή των αναλυτικών λειτουργιών που πρόκειται να χρησιμοποιηθούν, στη χρήση υπολογιστών για ανάλυση, στους περιορισμούς και τις ιδιαιτερότητες των αναλύσεων που είναι γνωστές και στην παρουσίαση των αναλυτικών αποτελεσμάτων.

Οι μέθοδοι χωρικής ανάλυσης περιλαμβάνουν λειτουργίες για:

- τον υπολογισμό στατιστικών δεικτών που αφορούν σε χωρικά δεδομένα, όπως ο χωρικός μέσος, η τυπική απόσταση και ο χωρικός αυτοσυσχετισμός,
- την επίλυση προβλημάτων υπολογιστικής γεωμετρίας,
- την επίλυση προβλημάτων που βασίζονται στον υπολογισμό αποστάσεων,
- την εκτίμηση της μορφής επιφανειών βάσει ενός συνόλου παρατηρήσεων,
- την ανάλυση του πεδίου ως προς τα πρωτογενή και τα δευτερογενή χαρακτηριστικά του,
- τον συνδυασμό πολλών χαρτογραφικών επιθεμάτων (χαρτογραφική επίθεση).

Τα αποτελέσματα αυτών των μεθόδων μεταβάλλονται με την αλλαγή τουλάχιστον ενός εκ των ακόλουθων χαρακτηριστικών:

- 1) της γεωγραφικής θέσης
- 2) των τοπολογικών σχέσεων και
- 3) της γεωμετρίας των χωρικών δεδομένων.

Στην περίπτωση που η οποιαδήποτε μεταβολή των χαρακτηριστικών αυτών δεν επιφέρει αλλαγή στα αποτελέσματα, τότε οι εφαρμοζόμενες μέθοδοι δεν είναι μέθοδοι χωρικής ανάλυσης.

3.2 Στατιστικοί Δείκτες Χωρικών Δεδομένων

3.2.1 Χωρικός Μέσος

Για να δώσει κανείς μια σύντομη μαθηματική περιγραφή ενός χρονικά μεταβαλλόμενου μεγέθους, μπορεί να υπολογίσει τον μέσο όρο των τιμών του, κατά τη διάρκεια ενός συγκεκριμένου χρονικού διαστήματος

Κατ' ανάλογο τρόπο, τα μεγέθη, τα οποία μεταβάλλονται χωρικά, περιγράφονται συνοπτικά από τον χωρικό μέσο (spatial median). Για παράδειγμα, η πολεοδομική ανάπτυξη μιας πόλης, με το πέρασμα του χρόνου, γίνεται συνήθως προς την κατεύθυνση της γειτονικής πόλης με την οποία πραγματοποιούνται οι περισσότερες οικονομικές και πολιτικές συναλλαγές. Στην περίπτωση αυτή, η μεταβολή του χωρικού μέσου περιγράφει συνοπτικά τη μετακίνηση που υφίσταται το πολεοδομικό κέντρο βάρους. Στον χώρο δύο διαστάσεων, ο χωρικός μέσος ενός συνόλου n σημείων υπολογίζεται από τους τύπους:

$$\bar{x} = \frac{\sum_i^n w_i x_i}{\sum_i^n w_i}$$

$$\bar{y} = \frac{\sum_i^n w_i y_i}{\sum_i^n w_i}$$

Όπου (\bar{x}, \bar{y}) , οι συντεταγμένες του χωρικού μέσου, (x_i, y_i) ($i=1..n$), οι συντεταγμένες των n σημείων του χώρου, και w_i , συντελεστές που αποδίδουν ειδική βαρύτητα σε κάθε σημείο που συμμετέχει στον προσδιορισμό του χωρικού μέσου.

Εάν δεν χρησιμοποιηθούν συντελεστές απόδοσης ειδικής βαρύτητας, ο χωρικός μέσος ονομάζεται κεντροειδής (centroid) ή γεωγραφικός μέσος (geographic mean) και η θέση του προσδιορίζεται από τους τύπους:

$$\bar{x} = \frac{\sum_i^n x_i}{n}$$

$$\bar{y} = \frac{\sum_i^n y_i}{n}$$

Όπου (\bar{x}, \bar{y}) , οι συντεταγμένες του κεντροειδούς, (x_i, y_i) ($i=1..n$) οι συντεταγμένες των n σημείων του χώρου.

3.2.2. Χωρική Τυπική Απόσταση

Η τυπική απόσταση (standard distance) ορίζεται ως ο μέσος όρος των αποστάσεων ενός συνόλου παρατηρήσεων από τον αντίστοιχο χωρικό μέσο. Η τυπική απόσταση δίνει ένα μέτρο της διασποράς που υπάρχει σ' ένα σύνολο παρατηρήσεων. Για παράδειγμα, η τυπική απόσταση μπορεί να χρησιμοποιηθεί για την εκτίμηση της διασποράς (dispersion) των

οικιών μαθητών σε σχέση με τη θέση του σχολείου τους. Ο τύπος υπολογισμού της χωρικής τυπικής απόστασης έχει ως ακολούθως:

$$S = \sqrt{\frac{\sum_i^n di^2}{n}}$$

Όπου d_i η απόσταση της i ης παρατήρησης από τον χωρικό μέσο και n ο αριθμός των παρατηρήσεων. Με την ενσωμάτωση των συντελεστών w_i που αποδίδουν ειδική βαρύτητα σε κάθε παρατήρηση i , ο τύπος διαμορφώνεται ως εξής:

$$S = \sqrt{\frac{\sum_i^n w_i di^2}{\sum_i^n w_i}}$$

3.2.3 Χωρικός Αυτοσυσχετισμός

Εάν καταγράψει κανείς παρατηρήσεις σχετικά με την εξέλιξη κάποιου φαινομένου στον χρόνο, μπορεί να δώσει μια εκτίμηση του πώς θα εξελιχθεί το φαινόμενο αυτό στο μέλλον. Αυτή η αλληλεξάρτηση των τιμών ενός φαινομένου στον χρόνο παρατηρείται και στην περίπτωση ενός χωρικά μεταβαλλόμενου μεγέθους και ονομάζεται χωρικός αυτοσυσχετισμός (spatial autocorrelation).

Στον θεωρητικό πυρήνα του χωρικού αυτοσυσχετισμού βρίσκεται η παρατήρηση του Waldo Tobler (1970), η οποία είναι γνωστή ως ο πρώτος νόμος της γεωγραφίας: “Everything is related to everything else, but near things are more related to each other” (Τα πάντα σχετίζονται μεταξύ τους αλλά τα γειτονικά πράγματα ακόμη περισσότερο).

Ο χωρικός αυτοσυσχετισμός είναι μετρήσιμος. Οι δείκτες του Moran και του Geary, που έχουν προταθεί για τον σκοπό αυτό, εφαρμόζονται σε πολυγωνικές ζώνες στις οποίες έχουν καταγραφεί δειγματοληπτικά παρατηρήσεις. Ο δείκτης του Moran συμβολίζεται με I και υπολογίζεται σύμφωνα με τους τύπους:

$$I = \frac{1}{p} \frac{\sum_i \sum_j w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_i (z_i - \bar{z})^2}, \text{ όπου}$$

$$p = \frac{\sum_i \sum_j w_{ij}}{n}$$

Ενώ, ο δείκτης του Geary συμβολίζεται με c και υπολογίζεται από τους τύπους:

$$c = \frac{1}{p} \frac{\sum_i \sum_j w_{ij} (z_i - z_j)^2}{\sum_i (z_i - \bar{z})^2}, \text{ όπου}$$

$$p = 2 \frac{\sum_i \sum_j w_{ij}}{n-1}$$

και z_i και z_j , ($i, j = 1..n$), οι τιμές των παρατηρήσεων i και j , \bar{z} , ο μέσος όρος τους, και w_{ij} , συντελεστές των οποίων η τιμή εξαρτάται από την εγγύτητα των ζωνών όπου έγιναν οι παρατηρήσεις i και j . Συνήθως, οι τιμές που αποδίδονται στους συντελεστές w_{ij} , είναι ίσες με 1 όταν οι παρατηρήσεις i και j βρίσκονται σε εφαιπτόμενες περιοχές, και ίσες με 0 σε διαφορετική περίπτωση.

Όσο περισσότερο η τιμή του I πλησιάζει το 0, τόσο μικρότερος είναι ο χωρικός αυτοσυσχετισμός, όσο πλησιάζει τις ακραίες τιμές +1 ή -1, τόσο αυξάνεται ο θετικός ή ο αρνητικός χωρικός αυτοσυσχετισμός αντίστοιχα. Για τον δείκτη c , ισχύει ότι, εάν βρεθεί ίσος με 1, δεν υπάρχει αυτοσυσχετισμός, ενώ, αν έχει τιμή μεγαλύτερη από 1, υπάρχει χωρικός αυτοσυσχετισμός.

3.2.4 Χωρική Ομαδοποίηση

Οι περισσότερες τεχνικές χωρικής ομαδοποίησης (spatial clustering) που έχουν προταθεί, εφαρμόζονται σε σημειακά αντικείμενα. Έτσι, γίνεται η υπόθεση ότι τα δεδομένα προβάλλονται υπό τη μορφή σημείων σ' έναν χώρο απεικόνισης ή ότι τα γραμμικά και πολυγωνικά αντικείμενα αναπαρίστανται από ένα σύνολο αντιπροσωπευτικών σημείων, όπως το κεντροειδές τους, οι κορυφές τους ή οι κορυφές του ελάχιστου κυρτού πολυγώνου που τα περιέχει.

Η χωρική ομαδοποίηση έχει ως στόχο τον εντοπισμό ομάδων χωρικών δεδομένων, από τη μελέτη των αναμεταξύ τους αποστάσεων, για την ανακάλυψη σημαντικών νοημάτων και σχέσεων (εξόρυξη γνώσης - data mining). Ως ομάδα (cluster) ορίζεται το σύνολο των σημείων του χώρου απεικόνισης που έχουν πολύ μικρή απόσταση μεταξύ τους σε σχέση με τα σημεία του υπόλοιπου χώρου. Ο σχηματισμός ομάδων σ' έναν χώρο απεικόνισης αποδίδεται σε δυο είδη διαδικασιών:

- Σε διαδικασίες πρώτης τάξης (first order processes) σύμφωνα με τις οποίες οι ομάδες σχηματίζονται ανεξάρτητα από τη σχέση μεταξύ των σημείων. Για παράδειγμα, κάποιοι οργανισμοί σχηματίζουν πυκνές ομάδες πληθυσμού σε ορισμένες περιοχές λόγω των ιδιαίτερων κλιματολογικών και εδαφικών συνθηκών που επικρατούν εκεί, και όχι λόγω των σχέσεων με τους υπόλοιπους οργανισμούς του είδους.
- Σε διαδικασίες δεύτερης τάξης (second order processes) κατά τις οποίες ο σχηματισμός των ομάδων οφείλεται στις σχέσεις που διέπουν τα σημεία. Οι σχέσεις αυτές μπορεί να είναι ελκτικές ή ανταγωνιστικές.

Διαιρετικές Τεχνικές Ομαδοποίησης

Οι διαιρετικές τεχνικές ομαδοποίησης (partitioning clustering methods) επιδιώκουν να χωρίσουν ένα σύνολο σημείων του χώρου σε k ομάδες, με τρόπο που να ικανοποιείται κάποιο χωρικό κριτήριο. Η πιο γνωστή εξ' αυτών, η τεχνική k -means, επιδιώκει να διαμορφώσει ομάδες με κριτήριο η μέση τιμή των αποστάσεων των σημείων από το κεντροειδές της ομάδας τους να ελαχιστοποιείται. Τα βήματα που ακολουθούνται κατά την εκτέλεση ενός διαιρετικού αλγορίθμου, είναι τα ακόλουθα:

1. Ορίζεται ο αριθμός των ομάδων και η θέση των κεντροειδών τους.
2. Υπολογίζονται οι αποστάσεις των σημείων από τα κεντροειδή.
3. Αποδίδεται κάθε σημείο στην πλησιέστερη ομάδα.
4. Επαναπροσδιορίζονται τα κεντροειδή των ομάδων, βάσει των σημείων που τους αποδόθηκαν.
5. Συγκρίνεται η μεταβολή της θέσης των κεντροειδών με συγκεκριμένη τιμή κατωφλίου. Εάν βρεθεί μεγαλύτερη, βγαίνει το συμπέρασμα ότι δεν έχουν «φτάσει» τα κεντροειδή στη βέλτιστη θέση και ο αλγόριθμος συνεχίζεται από το δεύτερο βήμα. Εάν βρεθεί μικρότερη, ο αλγόριθμος σταματά.

Τα μειονεκτήματα της μεθόδου αυτής είναι αρκετά καθώς πρέπει να είναι εκ των προτέρων γνωστός ο αριθμός των ομάδων. Επίσης, το τελικό αποτέλεσμα εξαρτάται από την αρχική τοποθέτηση των κεντροειδών των ομάδων. Τέλος, ορισμένα σημεία αποδίδονται λανθασμένα σε κάποιες ομάδες παρότι βρίσκονται κοντά στο κεντροειδές τους. Αυτό συμβαίνει κυρίως όταν τα μεγέθη των ομάδων είναι πολύ δυσανάλογα ή όταν υπάρχουν ομάδες με κυρτό σχήμα.

Ιεραρχικές Τεχνικές Ομαδοποίησης

Από τις πλέον γνωστές τεχνικές ομαδοποίησης, οι ιεραρχικές (hierarchical clustering methods), έκαναν την εμφάνισή τους από τους Florek κ.ά. (1951). Από τότε, έχουν δημιουργηθεί αρκετές παραλλαγές οι οποίες όμως στηρίζονται πιστά στην πρώτη εκδοχή.

Οι ιεραρχικές τεχνικές ομαδοποίησης χωρίζονται σε δυο βασικούς τύπους: τις αθροιστικές (agglomerative methods) και τις διαιρετικές (divisive methods) τεχνικές, με δημοφιλέστερες τις πρώτες. Η αθροιστική ιεραρχική ομαδοποίηση συνίσταται στη δημιουργία μίας και μοναδικής ομάδας από έναν αρχικό αριθμό ομάδων ίσο με το πλήθος των σημείων. Στη διαιρετική ιεραρχική ομαδοποίηση, γίνεται ακριβώς το αντίθετο: στην αρχή, θεωρείται ότι υπάρχει μία και μοναδική ομάδα, η οποία περιέχει όλα τα σημεία. Σε κάθε βήμα εκτέλεσης του αλγορίθμου, η ομάδα αυτή διαιρείται μέχρις ότου δημιουργηθούν τόσες ομάδες, όσες και τα σημεία του χώρου.

Τα βήματα που ακολουθούνται κατά την ιεραρχική ομαδοποίηση απεικονίζονται συνήθως με τη βοήθεια μιας δενδροειδούς δομής, το δενδρόγραμμα. Όμως, πρέπει να σημειωθεί ότι το δενδρόγραμμα, δεν δίνει ευκρινή αποτελέσματα στην περίπτωση πολύ μεγάλου πλήθους σημείων. Άλλο μειονέκτημα της ιεραρχικής τεχνικής ομαδοποίησης θεωρείται το γεγονός ότι ένα σημείο δεν μπορεί να αλλάξει ομάδα, εφόσον έχει καταχωρηθεί σε άλλη κατά τη διάρκεια της εκτέλεσης του αλγορίθμου.

Ωστόσο, η τεχνική αυτή δεν παύει να αποτελεί έναν φυσικό και απλό τρόπο ομαδοποίησης σημείων. Οι επιδόσεις της στην εύρεση ομάδων ποικίλων σχημάτων είναι αναμφισβήτητες. Το πιο σημαντικό πλεονέκτημα της μεθόδου αυτής, έγκειται στο ότι δεν απαιτείται η εκ των προτέρων γνώση του αριθμού των ομάδων που θα σχηματιστούν. Τέλος, οι ιεραρχικές τεχνικές χρησιμοποιούνται με σκοπό να βελτιώσουν σημαντικά τα αποτελέσματα των

διακριτικών τεχνικών. Για το λόγο αυτό, αρκετοί επιστήμονες θεωρούν τις μεθόδους αυτές συμπληρωματικές και όχι ανταγωνιστικές.

Τεχνικές Ομαδοποίησης βάσει Πυκνότητας

Η βασική ιδέα των τεχνικών ομαδοποίησης σημείων βάσει της πυκνότητάς τους (density based clustering methods), είναι η δημιουργία ομάδων με τρόπο ώστε η πυκνότητα των σημείων εντός μιας ομάδας να είναι μεγαλύτερη από εκείνη εκτός της ομάδας. Στη συνέχεια, περιγράφεται ένα τυπικό παράδειγμα τεχνικής αυτού του τύπου, η τεχνική DENCLUE (Density based CLUstEring).

Η τεχνική DENCLUE βασίζεται στην παραδοχή ότι κάθε απεικονιζόμενο σημείο ασκεί επιρροή στη γειτονιά του, η οποία μπορεί να μοντελοποιηθεί μαθηματικά από μια συνάρτηση που ονομάζεται συνάρτηση επιρροής (impact function). Η γραφική παράσταση του αθροίσματος των συναρτήσεων επιρροής είναι μια επιφάνεια που αποδίδει τη μεταβολή της πυκνότητας των σημείων στον χώρο. Τα τοπικά μέγιστα της επιφάνειας, ή ελκυστές πυκνότητας (density-attractors), προσδιορίζουν τα κέντρα των ομάδων. Οι τεχνικές ομαδοποίησης βάσει της πυκνότητας παρέχουν τα ακόλουθα πλεονεκτήματα:

- Είναι κατάλληλες για τον εντοπισμό ομάδων οποιουδήποτε σχήματος.
- Έχουν καλή απόδοση ακόμα και όταν εφαρμόζονται σε δεδομένα με πολύ θόρυβο.

3.3 Προβλήματα Υπολογιστικής Γεωμετρίας

Οι λειτουργίες χωρικής ανάλυσης που έχουν ως στόχο την επίλυση προβλημάτων υπολογιστικής γεωμετρίας (computational geometry) αφορούν στον υπολογισμό μεγεθών όπως το μήκος και το εμβαδόν, δίνουν απάντηση στο εάν ένα αντικείμενο περιέχεται σε κάποιο άλλο, προσδιορίζουν την τομή ή το ελάχιστο κυρτό περίγραμμα διαφόρων γεωγραφικών αντικειμένων, κλπ.

Οι διάφοροι υπολογιστικοί τύποι που χρησιμοποιούνται εξαρτώνται από το εκάστοτε σύστημα αναφοράς. Εάν τα γεωγραφικά αντικείμενα είναι σχετικά μικρά (εκτείνονται σε περιοχή κάτω των 100τ.χλμ) ή στην περίπτωση που έχουν προβληθεί στο επίπεδο, θεωρείται ότι μπορεί να χρησιμοποιηθεί για τον προσδιορισμό της θέσης τους το καρτεσιανό σύστημα συντεταγμένων και ότι οι μετρήσεις μπορούν να βασιστούν στην Ευκλείδεια γεωμετρία. Εάν όμως τα γεωγραφικά αντικείμενα είναι τόσο μεγάλα ώστε να πρέπει να ληφθεί υπόψη η καμπυλότητα της γης, χρησιμοποιείται το σύστημα γεωγραφικών συντεταγμένων.

Παρακάτω περιγράφονται τα κυριότερα εξ αυτών των προβλημάτων:

Υπολογισμός Μήκους

Το μήκος των γραμμικών αντικειμένων αντιστοιχεί στο άθροισμα των μηκών των διαδοχικών τόξων (κατά προσέγγιση των ευθυγράμμων τμημάτων) που τα απαρτίζουν. Στο επίπεδο, το μήκος κάθε ευθύγραμμου τμήματος με άκρα τα σημεία $S_1(x_1, y_1)$ και $S_2(x_2, y_2)$ υπολογίζεται από τον τύπο της Ευκλείδειας απόστασης:

$$d_{(S_1, S_2)} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Αντίστοιχα, στο σύστημα γεωγραφικών συντεταγμένων, το μήκος κάθε ευθύγραμμου τμήματος με άκρα τα σημεία $S_1(\phi_1, \lambda_1)$ και $S_2(\phi_2, \lambda_2)$ υπολογίζεται από τον τύπο:

$$d_{(S_1, S_2)} = R^{-1} \cos [\sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos(\lambda_1 - \lambda_2)]$$

όπου R η ακτίνα της Γης, ϕ_1 και ϕ_2 , τα γεωγραφικά πλάτη των σημείων, και λ_1 και λ_2 τα γεωγραφικά μήκη των σημείων.

Εμβαδόν

Ο υπολογισμός του εμβαδού πολυγωνικών αντικειμένων βασίζεται στη μέθοδο του τραπεζίου. Πιο συγκεκριμένα, καθώς διατρέχονται με τη φορά των δεικτών του ρολογιού, τα τόξα που απαρτίζουν το περίγραμμα του πολυγώνου, υπολογίζονται τα εμβαδά των τραπεζίων που σχηματίζονται μεταξύ των τόξων και ενός εκ των κάθετων αξόνων του συστήματος αναφοράς. Τα εμβαδά των τραπεζίων των οποίων τα τόξα δείχνουν προς μια συγκεκριμένη κατεύθυνση προστίθενται στο ζητούμενο εμβαδόν ενώ τα εμβαδά των τόξων που δείχνουν προς την αντίθετη κατεύθυνση αφαιρούνται.

Σημείο Εντός Πολυγώνου

Το πρόβλημα χωρικής ανάλυσης γνωστό ως σημείο-εντός-πολυγώνου αφορά στο εάν ένα σημείο Σ βρίσκεται εντός των ορίων που ορίζει το περίγραμμα ενός πολυγωνικού αντικειμένου Π . Η επίλυση του προβλήματος γίνεται σύμφωνα με τα παρακάτω βήματα:

1. Σχηματίζουμε την ημιευθεία παράλληλη προς τον άξονα των τετμημένων, με αρχή το σημείο Σ και με κατεύθυνση προς το $+\infty$.
2. Προσδιορίζουμε τον αριθμό n των σημείων τομής της ημιευθείας με το περίγραμμα του πολυγώνου. Εάν η ημιευθεία τέμνει μία πλευρά του πολυγώνου στην κορυφή της, τότε το n προσαυξάνεται κατά 1 όταν η άλλη κορυφή της πλευράς κείται κάτω από την ημιευθεία και κατά 0 όταν κείται πάνω ή επί της ημιευθείας.
3. Εάν τελικά το n βρεθεί να είναι περιττός αριθμός, βγαίνει το συμπέρασμα ότι το Σ κείται εντός του Π , διαφορετικά, το Σ βρίσκεται εκτός του Π .

Προσδιορισμός του Ελάχιστου Κυρτού Πολυγώνου

Ο προσδιορισμός του ελάχιστου κυρτού πολυγώνου (ΕΚΠ) είναι συνήθης λειτουργία χωρικής ανάλυσης. Η περίπτωση προσδιορισμού του ΕΚΠ ενός συνόλου γραμμικών ή πολυγωνικών αντικειμένων ανάγεται στην περίπτωση προσδιορισμού του ΕΚΠ για το σύνολο των σημείων που ορίζουν τα τόξα των περιγραμμάτων τους. Συνεπώς η περιγραφή που ακολουθεί περιορίζεται στον τρόπο προσδιορισμού του ΕΚΠ για ένα σύνολο σημείων.

Ο προσδιορισμός του ΕΚΠ ξεκινά με την εύρεση του σημείου με τη μικρότερη τεταγμένη, έστω Σ_1 , για το οποίο είναι βέβαιο ότι αποτελεί κορυφή του ζητούμενου πολυγώνου. Σ' ένα πρώτο βήμα, σχηματίζουμε την ημιευθεία, παράλληλη προς τον άξονα των τετμημένων, με αρχή το Σ_1 . Περιστρέφοντας αριστερόστροφα την ημιευθεία, συναντάμε το επόμενο σημείο που θα αποτελέσει κορυφή του ΕΚΠ, έστω Σ_2 . Ακολουθούμε την ίδια διαδικασία, ξεκινώντας αυτή τη φορά από το σημείο Σ_2 . Στη συνέχεια, η διαδικασία επαναλαμβάνεται έως ότου συναντήσουμε πάλι το Σ_1 , δηλαδή το σημείο με τη μικρότερη τεταγμένη που προσδιορίσαμε αρχικά.

Τομή Γραμμικών Αντικειμένων

Ο προσδιορισμός της τομής δυο γραμμικών αντικειμένων ανάγεται στον προσδιορισμό της τομής όλων των ζευγών των τόξων ή ευθυγράμμων τμημάτων που τα απαρτίζουν.

Έστω ένα ευθύγραμμο τμήμα που ορίζεται από τα σημεία $\Sigma_1 (x_1, y_1)$ και $\Sigma_2 (x_2, y_2)$, και ένα δεύτερο ευθύγραμμο τμήμα που ορίζεται από τα σημεία $\Sigma_3 (x_3, y_3)$ και $\Sigma_4 (x_4, y_4)$. Τότε, οι υποκείμενες ευθείες ορίζονται αντίστοιχα από τις εξισώσεις:

$$y = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

$$y = y_3 + \frac{y_4 - y_3}{x_4 - x_3} (x - x_3)$$

Η επίλυση του συστήματος των εξισώσεων αυτών θα δώσει απάντηση στο αν τέμνονται τα ευθύγραμμα τμήματα και θα προσδιορίσει το σημείο τομής, έστω T.

Τομή Πολυγωνικών Αντικειμένων

Ο έλεγχος της τομής δυο πολυγωνικών αντικειμένων βασίζεται στον έλεγχο δυο καταστάσεων, εάν υπάρχει τομή μεταξύ των ευθυγράμμων τμημάτων που απαρτίζουν τα περιγράμματά τους και εάν υπάρχει κορυφή του ενός που να βρίσκεται εντός του άλλου; Εφόσον ισχύει κάποια από τις δύο προαναφερθείσες καταστάσεις, βγαίνει το συμπέρασμα ότι τα πολυγωνικά αντικείμενα τέμνονται.

Κέντρο Γραμμικού Αντικειμένου

Το κέντρο ενός γραμμικού αντικειμένου, είτε πρόκειται για ένα απλό τόξο, είτε για μια ακολουθία διαδοχικών τόξων, θεωρείται το κέντρο του ευθύγραμμου τμήματος που ενώνει την αρχή και το τέλος του αντικειμένου. Συνεπώς, το κέντρο δεν εντοπίζεται πάντα κατά μήκος του αντικειμένου αλλά μπορεί να βρίσκεται στον περιβάλλοντα χώρο.

Κέντρο Πολυγωνικού Αντικειμένου

Ο προσδιορισμός του κέντρου πολυγωνικών αντικειμένων είναι αρκετά σημαντικός διότι βοηθά στην πραγματοποίηση αρκετών λειτουργιών. Το κέντρο μπορεί να αποτελέσει το σημείο απ' όπου ο χρήστης επιλέγει, μετακινεί ή περιστρέφει ένα πολυγωνικό αντικείμενο. Σε ορισμένες περιπτώσεις, όπου το μέγεθος ενός πολυγώνου είναι σχετικά μικρό, το κέντρο θεωρείται ως το ιδανικό σημείο για την τοποθέτηση μιας αντιπροσωπευτικής επιγραφής. Ακόμη, το κέντρο χρησιμοποιείται και ως σημείο αναφοράς κατά τον υπολογισμό της απόστασης που χωρίζει ένα πολυγωνικό αντικείμενο από άλλα.

$$x_k = \frac{\sum_i^n x_i}{n} \text{ και}$$

$$y_k = \frac{\sum_i^n y_i}{n}$$

Δημιουργία Ζωνών Επιρροής

Η δημιουργία ζωνών επιρροής (buffering) είναι η λειτουργία της χωρικής ανάλυσης που αφορά στην οριοθέτηση μιας ή περισσότερων περιοχών γύρω από γεωγραφικά αντικείμενα και σε προκαθορισμένη απόσταση από αυτά. Οι περιοχές αυτές, που ονομάζονται ζώνες, σχηματίζονται βάσει του ελάχιστου κυρτού πολυγώνου που περιέχει τα αντικείμενα, δημιουργώντας κύκλους με κέντρο τις κορυφές του και στη συνέχεια, ενώνοντας τους κύκλους με εφαπτόμενες γραμμές.

Υπολογισμός Απόστασης

Ο υπολογισμός της απόστασης μεταξύ γεωγραφικών θέσεων ή αντικείμενων, είναι μια θεμελιώδης λειτουργία της χωρικής ανάλυσης. Εκτελείται είτε αυτόνομα είτε στο πλαίσιο πιο σύνθετων λειτουργιών (π.χ. κατά τη δημιουργία ζωνών, κατά την χωρική ομαδοποίηση, κλπ.).

Η τιμή της πραγματικής απόστασης μεταξύ γεωγραφικών αντικειμένων εξαρτάται από την καμπυλότητα της γης και από το γήινο ανάγλυφο. Στην πράξη, η απόσταση υπολογίζεται προσεγγιστικά, άλλοτε στο επίπεδο, κατά μήκος μιας ευθείας γραμμής ή διαδοχικών ευθυγράμμων τμημάτων, κι άλλοτε κατά μήκος καμπυλών, θεωρώντας τη γη σφαιρική ή σε σχήμα ελλειψοειδούς.

Για πιο απομακρυσμένα σημεία, με γεωγραφικά πλάτη, και γεωγραφικά μήκη, χρησιμοποιούνται τύποι υπολογισμού της απόστασης που λαμβάνουν υπόψη την καμπυλότητα της Γης, όπως η σφαιρική ή η ελλειψοειδής απόσταση.

Εύρεση του Συντομότερου Μονοπατιού

Σε ένα δίκτυο, το πρόβλημα της εύρεσης του συντομότερου μονοπατιού (shortest path) μεταξύ ενός συγκεκριμένου κόμβου και των υπολοίπων, έγκειται στον εντοπισμό της ακολουθίας των διαδοχικών ακμών η οποία ελαχιστοποιεί το άθροισμα των δικτυακών αποστάσεων ή το άθροισμα των βαρών που αποδίδονται στις ακμές. Ο αλγόριθμος του Dijkstra, που προτάθηκε από τον ολλανδό επιστήμονα Edsger Dijkstra (1959), επιλύει το πρόβλημα αυτό και περιγράφεται στη συνέχεια.

Ξεκινώντας από έναν κόμβο-αρχή, έστω K_0 , προσδιορίζουμε τον κόμβο K_i που συνδέεται με τον K_0 σύμφωνα με το μικρότερο κόστος σύνδεσης (μικρότερη απόσταση ή μικρότερο βάρος) και ορίζουμε το συντομότερο μονοπάτι να περιλάβει την ακμή $K_0 K_i$. Στη συνέχεια, σε κάθε βήμα εκτέλεσης του αλγορίθμου, αναζητούμε τον κόμβο K_{i+1} από τους υπόλοιπους κόμβους που δεν έχουμε ακόμα συμπεριλάβει στο συντομότερο μονοπάτι, ο οποίος συνδέεται με το K_i με τρόπον ώστε το συνολικό κόστος (από τον κόμβο- αρχή μέχρι τον κόμβο K_{i+1}) να ελαχιστοποιείται. Η διαδικασία συνεχίζεται μέχρις ότου να «επισκεφθεί» ο αλγόριθμος όλους τους κόμβους.

3.4 Εγγενή σφάλματα στη χωρική ανάλυση

Συχνά προκύπτουν κοινά σφάλματα στη χωρική ανάλυση, μερικά λόγω των μαθηματικών του χώρου, μερικά λόγω των ιδιαίτερων τρόπων που τα δεδομένα παρουσιάζονται χωρικά, μερικά λόγω των εργαλείων που είναι διαθέσιμα. Χαρακτηριστική περίπτωση αποτελεί η φράκταλ φύση της ακτογραμμής κάνει ακριβείς μετρήσεις του μήκους της δύσκολες, αν όχι αδύνατες. Ένα λογισμικό υπολογιστή που τοποθετεί ευθείες γραμμές στην καμπύλη μιας ακτογραμμής, μπορεί εύκολα να υπολογίσει τα μήκη των γραμμών που ορίζει. Ωστόσο, αυτές οι ευθείες γραμμές μπορεί να μην έχουν εγγενή σημασία στον πραγματικό κόσμο.

Τα σφάλματα που οφείλονται στην πεπερασμένη ακρίβεια του Η/Υ (αριθμητικά σφάλματα) επηρεάζουν όλες τις γεωμετρικές λειτουργίες, συμπεριλαμβανομένων και των τοπολογικών σχέσεων. Χαρακτηριστικό παράδειγμα αποτελούν κατά την χωρική ανάλυση σε υπολογιστή τα προϊόντα διαιρέσεων τα οποία μπορεί να είναι διαφορετικής ακρίβειας από τις συντεταγμένες που περιγράφουν. Αν η επιλεγμένη ακρίβεια των εργασιών είναι μεγαλύτερη τότε αντικείμενα που π.χ. εφάπτονται στην πραγματικότητα ο υπολογιστής θα βρίσκει ή ότι επικαλύπτονται ή ότι δεν έχουν κάποια τοπολογική σχέση.

Τα προβλήματα αυτά αντιπροσωπεύουν μια πρόκληση στη χωρική ανάλυση λόγω της δύναμης των χαρτών ως μέσα παρουσίασης. Όταν τα αποτελέσματα παρουσιάζονται ως χάρτες, η παρουσίαση συνδυάζει χωρικά δεδομένα τα οποία είναι γενικά ακριβή με αναλυτικά αποτελέσματα που μπορεί να είναι ανακριβή, οδηγώντας σε μια εντύπωση ότι τα αναλυτικά αποτελέσματα είναι ακριβέστερα από τα δεδομένα που θα έδειχναν.

Ο ορισμός της χωρικής παρουσίας μιας οντότητας περιορίζει την πιθανή ανάλυση που μπορεί να εφαρμοστεί σε αυτήν την οντότητα και επηρεάζει τα τελικά συμπεράσματα που μπορούν να επιτευχθούν. Ενώ αυτή η ιδιότητα είναι θεμελιωδώς αληθής για όλες τις αναλύσεις, είναι ιδιαίτερα σημαντική στη χωρική ανάλυση, επειδή τα εργαλεία για τον ορισμό και τη μελέτη οντοτήτων ευνοούν τους συγκεκριμένους χαρακτηρισμούς των οντοτήτων που μελετώνται. Οι στατιστικές τεχνικές ευνοούν τον χωρικό ορισμό αντικειμένων ως σημείων επειδή υπάρχουν πολύ λίγες στατιστικές τεχνικές οι οποίες λειτουργούν απευθείας σε στοιχεία γραμμής, περιοχής ή όγκου. Τα εργαλεία υπολογιστών ευνοούν τον χωρικό ορισμό των αντικειμένων ως ομοιογενή και ξεχωριστά στοιχεία λόγω του περιορισμένου αριθμού στοιχείων βάσης δεδομένων και των διαθέσιμων υπολογιστικών δομών και της ευκολίας με την οποία μπορούν να δημιουργηθούν αυτές οι πρωτόγονες δομές.

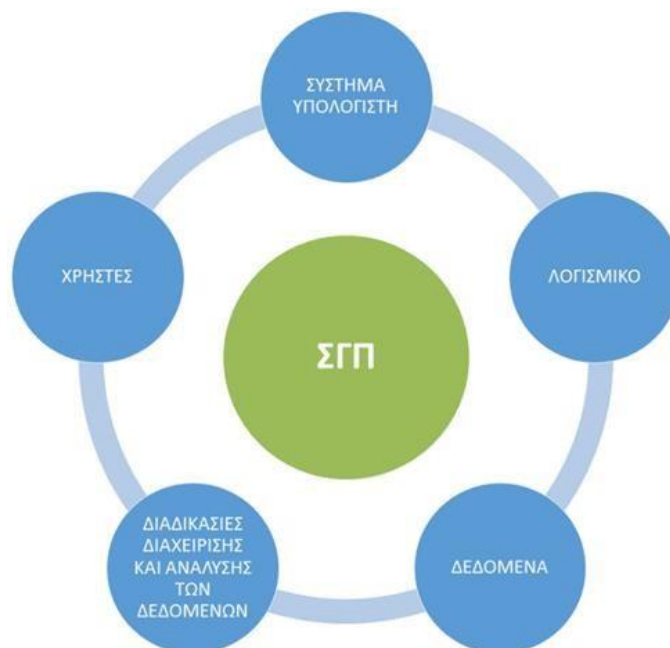
3.5 Συστήματα Γεωγραφικών Πληροφοριών

Αν η χωρική ανάλυση απαντάει στο “πώς” μελετώνται τα χωρικά δεδομένα τότε τα Συστήματα Γεωγραφικών Πληροφοριών είναι η απάντηση στο “που” γίνεται η χωρική ανάλυση. Ένα σύστημα γεωγραφικών πληροφοριών (GIS) είναι ένα σύστημα σχεδιασμένο για να συλλαμβάνει, να αποθηκεύει, να χειρίζεται, να αναλύει, να διαχειρίζεται και να παρουσιάζει χωρικά ή γεωγραφικά δεδομένα. Τα GIS είναι εργαλεία που επιτρέπουν στους χρήστες να δημιουργούν διαδραστικές ερωτήσεις (αναζητήσεις που δημιουργούνται από χρήστες), να αναλύουν χωρικές πληροφορίες, να επεξεργάζονται δεδομένα σε χάρτες και να παρουσιάζουν τα αποτελέσματα όλων αυτών των εργασιών. Ο όρος GIS αναφέρεται σε διάφορες τεχνολογίες, διαδικασίες, τεχνικές και μεθόδους. Είναι συνδεδεμένος με πολλές λειτουργίες και έχει πολλές εφαρμογές σχετικές με τη μηχανική, τον προγραμματισμό, τη διαχείριση γης, τις μεταφορές, τις τηλεπικοινωνίες. Οι εφαρμογές GIS και εντοπισμού θέσης αποτελούν το θεμέλιο για πολλές υπηρεσίες που σχετίζονται με την τοποθεσία και βασίζονται στην ανάλυση και την απεικόνιση.

3.5.1 Συστατικά στοιχεία ενός ΣΓΠ

Ένα ΣΓΠ δεν είναι απλώς ένα πακέτο λογισμικού που εκτελεί κάποιες λειτουργίες, αλλά ένα σύνθετο σύστημα που αποτελείται από επιμέρους συστατικά στοιχεία:

- το σύστημα του υπολογιστή, δηλαδή το υλικό (hardware) και το λειτουργικό σύστημα (operating system),
- το λογισμικό (software),
- τα χωρικά δεδομένα,
- τις διαδικασίες διαχείρισης και ανάλυσης των δεδομένων,
- τους χρήστες



Εικόνα 20 - Συστατικά στοιχεία ενός ΣΓΠ

Το λογισμικό επιτρέπει στους χρήστες την εισαγωγή, αποθήκευση, διαχείριση, επεξεργασία, ανάλυση και απεικόνιση των δεδομένων.

Τα δεδομένα των ΣΓΠ είναι κυρίως χωρικά, δηλαδή δεδομένα που αφορούν τη θέση των γεωγραφικών οντοτήτων αλλά και τις χωρικές τους σχέσεις με άλλες οντότητες. Επίσης τα ΣΓΠ διαχειρίζονται και μη-χωρικά δεδομένα τα οποία συνήθως αφορούν μη-χωρικά χαρακτηριστικά των γεωγραφικών οντοτήτων (π.χ., πληθυσμός των απογραφικών μονάδων, μέση ετήσια θερμοκρασία μιας περιοχής, κ.ά.).

Ο κλασικός τρόπος αναπαράστασης του γεωγραφικού χώρου είναι μέσω θεματικών επιπέδων πληροφορίας. Κάθε θεματικό επίπεδο αντιπροσωπεύει μια θεματική προσέγγιση για ένα συγκεκριμένο σκοπό και μπορεί να περιλαμβάνει ένα ή περισσότερα διαφορετικά είδη πληροφοριών.

Μια εναλλακτική προσέγγιση δόμησης των χωρικών δεδομένων θεωρεί τον γεωγραφικό χώρο ως μια ενιαία περιοχή όπου περιέχονται διάφορα αντικείμενα. Αυτά τα αντικείμενα θεωρούνται πως βρίσκονται σε ένα μόνο θεματικό επίπεδο, επιτρέποντας με αυτόν τον τρόπο την απεικόνιση της τρίτης διάστασης, όπως και την καθιέρωση ενός ενιαίου συστήματος αναφοράς. Αυτή η προσέγγιση σχετίζεται με την ικανότητα αναγνώρισης των αντικειμένων σύμφωνα με την ανθρώπινη λογική γι' αυτό ονομάζεται προσέγγιση των αντικειμένων.

3.5.2 Τεχνικές και τεχνολογίες

Οι σύγχρονες τεχνολογίες ΣΓΠ χρησιμοποιούν ψηφιακές πληροφορίες, για τις οποίες χρησιμοποιούνται διάφορες μέθοδοι δημιουργίας ψηφιακών δεδομένων. Η πιο συνηθισμένη μέθοδος δημιουργίας δεδομένων είναι η ψηφιοποίηση, όπου ένας χάρτης ή ένα σχέδιο μεταφέρονται σε ένα ψηφιακό μέσο μέσω της χρήσης ενός προγράμματος CAD και των δυνατοτήτων γεωαναφοράς. Με την ευρεία διαθεσιμότητα ορθο-διορθωμένων εικόνων (από δορυφόρους, αεροσκάφη, Helikites και UAV), η ψηφιοποίηση γίνεται η κύρια μέθοδος μέσω της οποίας εξάγονται τα γεωγραφικά δεδομένα. Η αρχική ψηφιοποίηση περιλαμβάνει την ανίχνευση γεωγραφικών δεδομένων απευθείας πάνω από την αεροφωτογραφία.

Η **γεωεπεξεργασία** είναι οι λειτουργίες ΣΓΠ που χρησιμοποιούνται για τον χειρισμό χωρικών δεδομένων. Μια τυπική διαδικασία γεωεπεξεργασίας παίρνει ένα σύνολο δεδομένων εισόδου, εκτελεί μια λειτουργία σε αυτό το σύνολο δεδομένων και επιστρέφει το αποτέλεσμα της λειτουργίας ως σύνολο δεδομένων εξόδου. Τέτοιες λειτουργίες αποτελούν επικάλυψη γεωγραφικών χαρακτηριστικών, η επιλογή και ανάλυση χαρακτηριστικών, η επεξεργασία τοπολογίας, η επεξεργασία ψηφιδωτών δεδομένων και μετατροπή δεδομένων. Η γεωεπεξεργασία επιτρέπει τον ορισμό, τη διαχείριση και την ανάλυση των πληροφοριών που προκύπτουν από τα χωρικά δεδομένα και χρησιμοποιούνται για τη λήψη αποφάσεων.

3.5.3 Αβεβαιότητα των ΣΓΠ

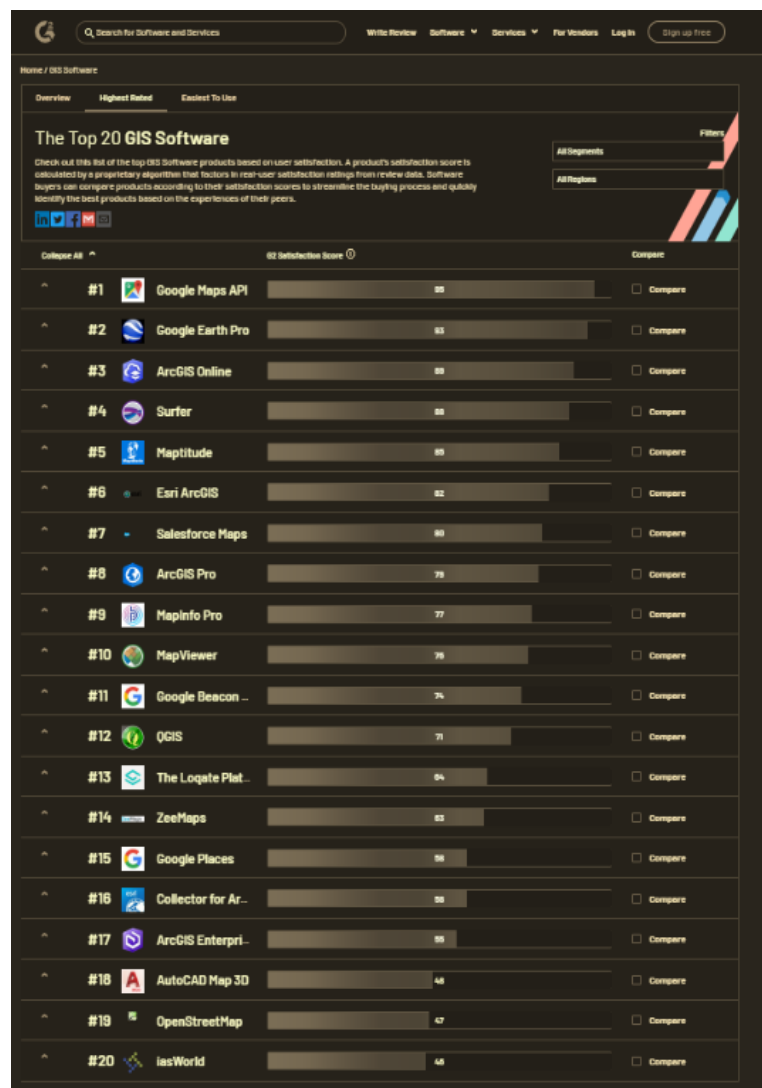
Η ακρίβεια των GIS εξαρτάται από τα δεδομένα εισόδου και από τον τρόπο με τον οποίο κωδικοποιούνται τα συγκεκριμένα δεδομένα. Υπάρχει η δυνατότητα λήψης δεδομένων με πολύ υψηλό επίπεδο ακρίβειας θέσης χρησιμοποιώντας τις θέσεις που προέρχονται από το GPS. Ψηφιακές εικόνες εδάφους και αεροφωτογραφίες υψηλής ανάλυσης, οι ισχυροί υπολογιστές και η τεχνολογία Ιστού αλλάζουν την ποιότητα, τη χρησιμότητα, αλλά υπάρχουν και άλλα δεδομένα πηγών που επηρεάζουν αρνητικά τη συνολική ακρίβεια των GIS όπως οι αναλογικοί χάρτες.

3.5.4 Λογισμικά ΣΓΠ

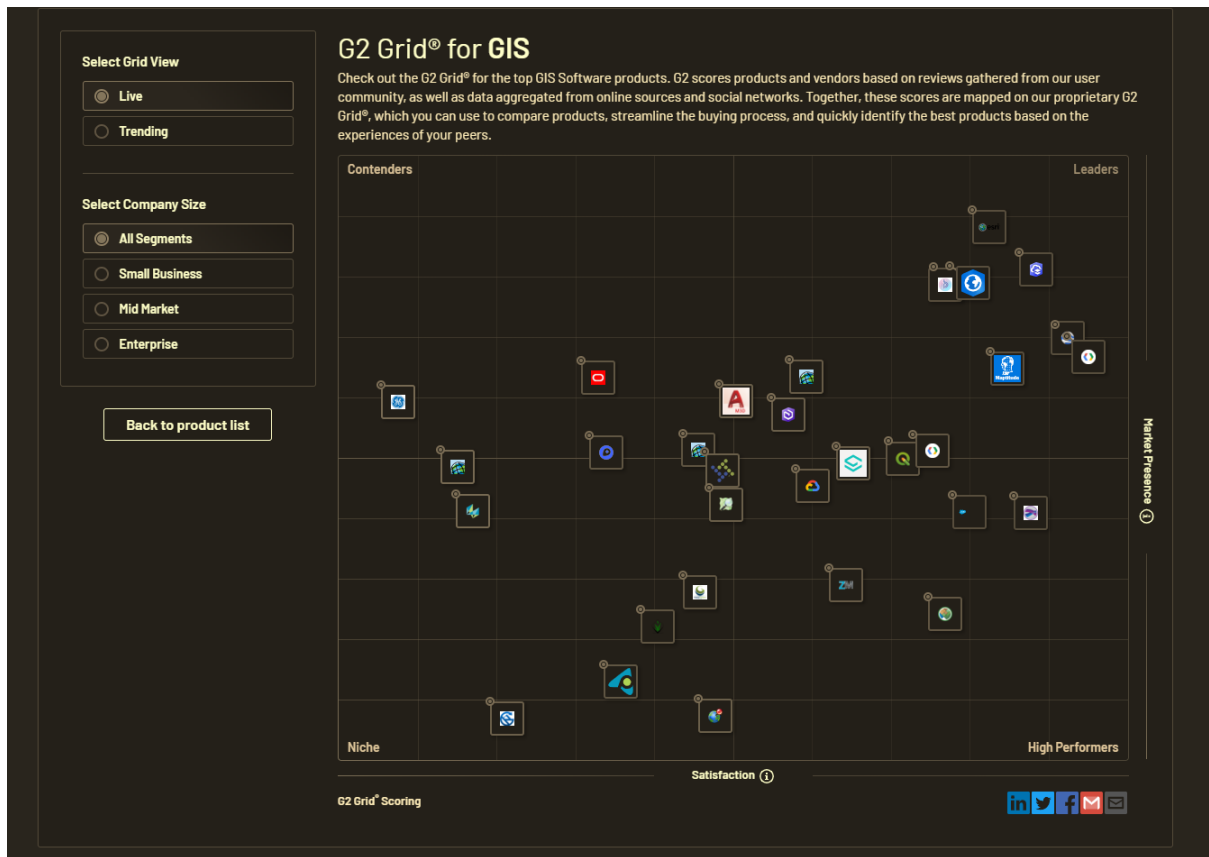
Η ανάλυση και η σύγκριση συγκεκριμένων λογισμικών ΣΓΠ είναι εκτός του πεδίου που η παρούσα εργασία εστιάζει. Παρ' όλα αυτά γίνεται μια αναφορά στο πλήθος των λογισμικών που καταγράφονται (με πολλά από αυτά να είναι αρκετά εξειδικευμένα)

Στον ιστότοπο [G2 \(ΣτΣ: go to\)](#) καταγράφονται **215 λογισμικά ΣΓΠ** που χρησιμοποιούνται σε ερευνητικό, εκπαιδευτικό και επαγγελματικό επίπεδο.

Αυτό σημαίνει ότι υπάρχουν εκατοντάδες εφαρμογές ανάλυσης χωρικών δεδομένων, η κάθε μία με δικά της υπέρ και κατά, με δικά της οπτική πάνω στις εργασίες ανάλυσης με τις δικές τις ιδιαιτερότητες στο χειρισμό, αποθήκευση και ανάλυση χωρικών δεδομένων.



Εικόνα 21 - Ιστόγραμμα Δημοφιλίας λογισμικών ΓΣΠ από το G2 Crowd (ιστότοπος αξιολόγησης λογισμικών από το κοινό)



Εικόνα 22 - Διάγραμμα διασποράς δημοφιλίας GIS λογισμικών από το G2 Crowd (ιστότοπος αξιολόγησης λογισμικών από το κοινό)

ΚΕΦΑΛΑΙΟ 4

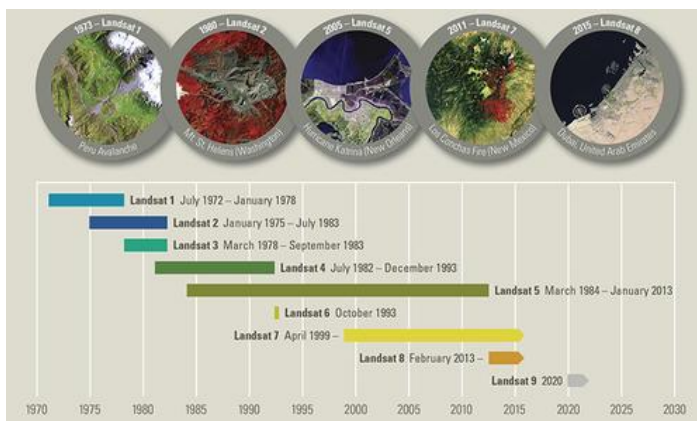
ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΦΑΡΜΟΓΗΣ

4.1 Ταξινόμηση κάλυψης γης

4.1.1 Πρόγραμμα Landsat

Σήμερα, το πρόγραμμα Landsat συμπληρώνει σχεδόν πενήντα χρόνια από την πρώτη εκτόξευση δορυφόρου. Η ιδέα ξεκίνησε το 1960 από το υπουργείο εσωτερικών, το υπουργείο γεωργίας των ΗΠΑ και την Αμερικάνικη Διαστημική Υπηρεσία (NASA). Ο δορυφόρος Landsat 1 αποτέλεσε τον πρώτο χαρτογραφικό δορυφόρο με πολυφασματικό σαρωτή και τέθηκε επιτυχώς σε τροχιά το 1972 στο πλαίσιο μιας συνεργασίας με συμβαλλόμενα μέρη την NASA και την Αμερικάνικη Γεωλογική Υπηρεσία (United States Geological Survey - USGS). Ακολούθησαν άλλες επτά διαδοχικές εκτοξεύσεις, εκ' των οποίων η μια απέτυχε να μπει σε τροχιά (Landsat 6). Σήμερα, είναι ενεργές οι αποστολές Landsat 7 και Landsat 8, και ήδη έχει ανακοινωθεί η αποστολή Landsat 9 η οποία αναμένεται να ξεκινήσει το 2023. Όλες οι αποστολές έχουν ως στόχο την συλλογή πολυφασματικών δεδομένων σε παγκόσμια χωρική κλίμακα. Αποτελούν τον ακρογωνιαίο λίθο για εφαρμογές που σχετίζονται με την παρατήρηση κάλυψης γης (land cover) και πιο συγκεκριμένα είναι σημαντική η συμβολή τους σε εφαρμογές που σχετίζονται με:

- Χαρτογραφία
- Χωροταξία
- Γεωργία
- Διαχείριση φυσικών πόρων
- Υδρολογία
- Παρακολούθηση θαλάσσιων υδάτων
- Κλιματική αλλαγή
- Παρακολούθηση δασών
- Γεωλογία
- Ανθρώπινες δραστηριότητες
- Χρήσεις γης
- Πολιτική
- Οικονομία



Εικόνα 23 - Διαστημικό Πρόγραμμα LANDSAT
(<https://landsat.gsfc.nasa.gov>)

4.1.2 Δορυφορικά δεδομένα

Η NASA και η USGS από το 2008 διανέμουν ελεύθερα τα δορυφορικά δεδομένα από την επίσημη ιστοσελίδα <https://earthexplorer.usgs.gov>. Κάθε χρήστης μπορεί να έχει πρόσβαση σε δεδομένα τα οποία έχουν υποστεί συγκεκριμένη προ-επεξεργασία (γεωμετρική και ραδιομετρική διόρθωση) ή και περαιτέρω προ-επεξεργασία σε διάφορα επίπεδα (πχ ατμοσφαιρικές διορθώσεις) προς διευκόλυνσή του χρήστη. Ανάλογα με τις ανάγκες και τις απαιτήσεις κάποιας εργασίας (και άρα την αρχική πληροφορία η οποία απαιτείται) μπορεί να επιλέξει το αντίστοιχο σύνολο δεδομένων. Τα δεδομένα τα οποία αξιοποιήθηκαν για την παρούσα εφαρμογή, αποκτήθηκαν από τον συγκεκριμένο ιστότοπο και έχουν υποστεί γεωμετρική και ραδιομετρική διόρθωση, αλλά όχι ατμοσφαιρική. Συγκεκριμένα, προέρχονται από τους δορυφόρους Landsat 5 και Landsat 8. Δεν επιλέχθηκε να χρησιμοποιηθούν λήψεις του Landsat 7, διότι ο συγκεκριμένος δορυφόρος, στις 31-05-2003 υπέστη ένα πρόβλημα στο διορθωτή της γραμμής σάρωσης (Scan Line Corrector - SLC) ο οποίος αντιστάθμιζε την εμπρόσθια κίνηση του δορυφόρου. Τα δεδομένα του δορυφόρου είναι αξιοποιήσιμα, απλά προϋποθέτουν μια επιπρόσθετη διεργασία κατά το στάδιο της προ-επεξεργασίας.

Δορυφόρος Αισθητήρας	Κανάλι Μήκος Κύματος	Ανάλυση (m)	Ενδείκνυται για:
LANDSAT_5 TM (Thematic Mapper)	B1 blue 0.45-0.52	30	Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation
	B2 green 0.52-0.60	30	Emphasizes peak vegetation, which is useful for assessing plant vigor
	B3 red 0.63-0.69	30	Discriminates vegetation slopes
	B4 Near Infrared (NIR) 0.77-0.90	30	Emphasizes biomass content and shorelines
	B5 Short-wave Infrared (SWIR) 1.55-1.75	30	Discriminates moisture content of soil and vegetation; penetrates thin clouds
	B6 Thermal Infrared 10.40-12.50	120	Thermal mapping and estimated soil moisture
	B7 Short-wave Infrared 2.09-2.35	30	Hydrothermally altered rocks associated with mineral deposits

Πίνακας 3 - Κανάλια δορυφόρου Landsat 5 (www.usgs.gov)

Τόσο η ποιότητα όσο και η ακρίβεια των δεδομένων, εξελίσσεται σταδιακά με την πρόοδο της διαστημικής τεχνολογίας και των αισθητήριων οργάνων του δορυφόρου. Έτσι, τα δεδομένα ανά αποστολή χαρακτηρίζονται από ορισμένες διαφοροποιήσεις. Στον πίνακα

φαίνονται τα κανάλια του κάθε δέκτη, τα μήκη κύματος στα οποία καταγράφει, η ανάλυση και η αντίστοιχη χρησιμότητα του καναλιού. Όπως και με την επιλογή του κατάλληλου συνόλου δεδομένων, έτσι και με τα κανάλια, ο αναλυτής καλείται να επιλέξει τα κατάλληλα (ή συνδυασμό καναλιών) για την εφαρμογή του. Σε κάθε κανάλι, η ποσότητα της φασματικής απόκρισης που καταγράφεται, μπορεί να περιγράψει εμμέσως κάποια συγκεκριμένη χημική ιδιότητα του αντικειμένου.

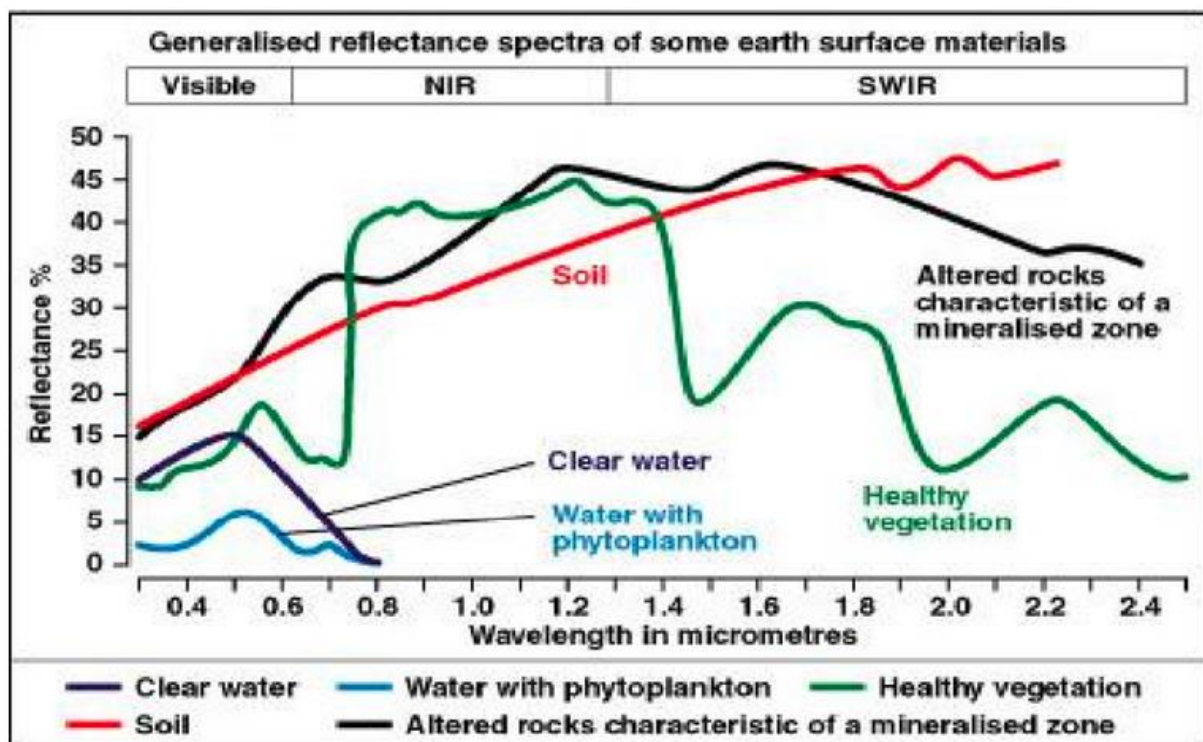
Δορυφόρος Αισθητήρας	Κανάλι Μήκος Κύματος	Ανάλυση (m)	Ενδείκνυται για:
LANDSAT_8 OLI (Operational Land Image) & TIRS (Thermal Infrared Sensor)	B1 coastal aerosol 0.43-0.45	30	Coastal and aerosol studies
	B2 blue 0.45-0.51	30	Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation
	B3 green 0.53-0.59	30	Emphasizes peak vegetation, which is useful for assessing plant vigor
	B4 red 0.64-0.67	30	Discriminates vegetation slopes
	B5 Near Infrared (NIR) 0.85-0.88	30	Emphasizes biomass content and shorelines
	B6 Short-wave Infrared (SWIR) 1 1.57-1.65	30	Discriminates moisture content of soil and vegetation; penetrates thin clouds
	B7 Short-wave Infrared (SWIR) 2 2.11-2.29	30	Improved moisture content of soil and vegetation; penetrates thin clouds
	B8 Panchromatic 0.50-0.68	15	15-meter resolution, sharper image definition
	B9 Cirrus 1.36-1.38	30	Improved detection of cirrus cloud contamination
	B10 TIRS 1 10.60-11.19	100*(30)	100-meter resolution, thermal mapping and estimated soil moisture

	B11 TIRS 2 11.50-12.51	100*(30)	100-meter resolution, improved thermal mapping and estimated soil moisture
--	---------------------------	----------	--

Πίνακας 4 - Κανάλια δορυφόρου Landsat 8 (www.usgs.gov)

Από τους παραπάνω πίνακες είναι δυνατό να προκύψουν οι διαφορές μεταξύ των δύο δορυφόρων που χρησιμοποιήθηκαν στο πλαίσιο της παρούσας εφαρμογής. Μια σημαντική διαφορά αποτελεί το κανάλι B8 που φαίνεται ότι υπάρχει μόνο στον Landsat 8, και έχει υψηλότερη χωρική ανάλυση (15 m) σε σχέση με τα υπόλοιπα κανάλια των δύο δορυφόρων (30m). Πρόκειται για το πανχρωματικό κανάλι, το οποίο ξεκίνησε να ενσωματώνεται στους αισθητήρες του προγράμματος Landsat, από τον Landsat 7 και μετά. Το εύρος στο οποίο αντιστοιχεί εκτείνεται από το ορατό μέχρι και το εγγύς υπέρυθρο και προσφέρει την δυνατότητα συγχώνευσης των πολυφασματικών εικόνων χαμηλής ανάλυσης με την πανχρωματική εικόνα υψηλής ανάλυσης. Το προϊόν που προκύπτει είναι εικόνες που συνδυάζουν την ραδιομετρική ανάλυση της πολυφασματικής (bits που καταγράφονται για κάθε ψηφιακή τιμή από τον αισθητήρα), με την χωρική ανάλυση της πανχρωματικής. Η διαδικασία αυτή (pansharpening) εφαρμόζεται για να ενισχυθούν κάποια χαρακτηριστικά που δεν εμφανίζονται στις επί μέρους εικόνες, για να εντοπισθούν ενδεχόμενες αλλαγές σε διαχρονικά δεδομένα, για να ενισχυθεί η γεωμετρική διόρθωση και να βελτιωθεί η ταξινόμηση (classification). Στην παρούσα εφαρμογή δεν πραγματοποιήθηκε pansharpening για λόγους σύγκρισης των δεδομένων από τους δυο διαφορετικούς δορυφόρους. Έτσι, η χωρική ανάλυση των εικόνων παραμένει στα 30 μέτρα.

Φασματική Απόκριση (Spectral Response)



Εικόνα 24 - Φασματική απόκριση πέντε βασικών επιφανειακών γήινων στοιχείων (www.researchgate.net)

Συνολικά, ένα συγκεκριμένο υλικό, όπως για παράδειγμα το νερό, τείνει να υιοθετεί ένα πρότυπο συμπεριφοράς. Το πρότυπο αυτό ονομάζεται φασματική καμπύλη και αποτελείται από τις φασματικές τιμές (ανακλαστικότητα - reflectance) που προκύπτουν μετά από συγκεκριμένη επεξεργασία της αντανάκλασης της ηλεκτρομαγνητικής ακτινοβολίας (radiance) που κατέγραψε ο δέκτης για αυτό το υλικό, σε κάθε κανάλι. Οι φασματικές καμπύλες μεταξύ διαφορετικών υλικών μπορούν να έχουν τέτοια διαφοροποίηση, έτσι ώστε να αποτελούν την φασματική υπογραφή αυτού του υλικού, και με αυτό το τρόπο να το καθιστούν αναγνωρίσιμο. Στην παραπάνω εικόνα φαίνονται οι φασματικές υπογραφές για τις πέντε βασικές κατηγορίες που είναι δυνατόν να βρεθούν στην επιφάνεια της γης. Αυτές είναι το καθαρό νερό, το νερό που περιέχει φυτοπλαγκτόν, η υγιής βλάστηση, το χώμα και η πέτρα (ή άμμος).

Για κάθε μια από τις παραπάνω κατηγορίες, οποιαδήποτε χημική μεταβολή, θα έχει ως αποτέλεσμα την μεταβολή της φασματικής απόκρισης, με συγκεκριμένο κάθε φορά τρόπο. Από την πρόταση αυτή, προκύπτει το συμπέρασμα ότι η φασματική απόκριση μιας καλλιέργειας καλαμποκιού θα είναι διαφορετική από αυτή μιας καλλιέργειας βαμβακιού κοκ. Συγχρόνως, η φασματική υπογραφή μεταβάλλεται και στο χρόνο. Για παράδειγμα, η φασματική υπογραφή του καλαμποκιού δεν είναι η ίδια τον Απρίλιο που το φυτό βλασταίνει και τον Αύγουστο που το φυτό βρίσκεται σε πλήρη ανάπτυξη. Εντούτοις, παρουσιάζει μια περιοδικότητα η οποία μπορεί να παρατηρηθεί και είναι σχετικά σταθερή διότι βασίζεται στο φαινολογικό κύκλο του φυτού. Σε αντίθεση με την βλάστηση, οι κατηγορίες όπως πέτρα, άμμος, τσιμέντο κοκ δεν παρουσιάζουν μεταβολές στο χρόνο, αλλά χαρακτηρίζονται από μεταξύ τους διαφορές.

Υπερφασματικά δεδομένα

Όσον αφορά την εξαγωγή θεματικών υποβάθρων κάλυψης γης με τις παραπάνω μεθόδους, πολλές φορές χρησιμοποιούνται και υπερφασματικά δεδομένα, τα οποία προκύπτουν από δορυφόρους που είναι εξοπλισμένοι με αισθητήρες αντίστοιχης τεχνολογίας. Συγκεκριμένα, τα υπερφασματικά δεδομένα είναι αντίστοιχα των πολυφασματικών, με την διαφορά ότι ο αισθητήρας καταγράφει μεγαλύτερο αριθμό καναλιών (πχ 200), σε φασματικές ζώνες στενότερου εύρους σε σχέση με αυτές των πολυφασματικών. Τα δεδομένα αυτά προσφέρουν την δυνατότητα διάκρισης ανεπαίσθητων χαρακτηριστικών και διαφορών μεταξύ των ιδιομορφιών της επιφάνειας της Γης.

Παρ' όλα αυτά, οι δυνατότητες μόνο των πολύφασματικών ή/και υπερφασματικών δεδομένων μπορούν να έχουν ως αποτέλεσμα την εξαγωγή θεματικών υποβάθρων κάλυψης γης, και δεν είναι δυνατόν να εντοπίσουν ανεπαίσθητες διαφοροποιήσεις στην χρήση γης. Για παράδειγμα, είναι δυνατόν να εντοπιστεί ένα αγροτεμάχιο στο οποίο κάποια χρονιά δεν καλλιεργήθηκε κάποιο αγροτικό προϊόν διότι πραγματοποιήθηκε αγρανάπαυση. Όπως επίσης μπορεί να εντοπιστεί κάποιο κτίσμα το οποίο χρησιμοποιείται ως αποθήκη αγροτικών εργαλείων, αξιοποιώντας μεθόδους με προηγμένους αλγόριθμους ταξινόμησης (πχ k-Nearest Centroid Neighbor - k-NCN) που βασίζονται όχι μόνο στην φασματική πληροφορία, αλλά και στην εγγύτητα και την χωρική συσχέτιση μεταξύ διαφορετικών εικονοστοιχείων

(pixels). Η ολοκλήρωση (integration) των φασματικών δεδομένων με άλλη χωρική πληροφορία η οποία μπορεί να συμβάλει συμπληρωματικά αποτελεί ένα ισχυρό εργαλείο.

4.1.3 Πρόγραμμα COR.IN.E.

Το COR.IN.E. (Coordination of Information on the Environment) είναι πρόγραμμα συντονισμού πληροφοριών για το περιβάλλον, το οποίο θεσπίστηκε το 1985. Κύριος δημόσιος φορέας στην Ευρώπη είναι ο Ευρωπαϊκός Οργανισμός Περιβάλλοντος (European Environment Agency), ο οποίος είναι αρμόδιος για την παροχή τεκμηριωμένων και αντικειμενικών πληροφοριών του περιβάλλοντος. Ο ΕΟΠ, ο οποίος λειτουργεί από το 1994 με έδρα την Κοπεγχάγη, είναι ο κεντρικός κόμβος του Ευρωπαϊκού Δικτύου Πληροφοριών και Παρατηρήσεων για το Περιβάλλον (EIONET), ενός δικτύου που συνδέει 300 περίπου φορείς σε ολόκληρη την Ευρώπη, συλλέγοντας και διαδίδοντας πληροφορίες και στοιχεία για το περιβάλλον. Σήμερα 33 κράτη είναι μέλη του ΕΟΠ, 28 χώρες της ευρωπαϊκής ένωσης και η Ελβετία, η Ισλανδία, το Λιχτενστάιν, η Νορβηγία και η Τουρκία.

Το COR.IN.E. συνίσταται από ψηφιακούς χάρτες που δείχνουν τις πολλαπλές μεταβολές που υπέστησαν τα ευρωπαϊκά τοπία, ενώ συμπεριλαμβάνει κατηγορίες που θα χαρακτηριζόταν όχι μόνο σαν κάλυψη αλλά και σαν χρήση γης. Ο χάρτης αυτός δίνει τη δυνατότητα στους αρμοδίους για τη χάραξη πολιτικής να εξάγουν συμπεράσματα σε τομείς όπως της γεωργίας και των μεταφορών σχετικά με τις επιπτώσεις των ληφθέντων αποφάσεων στους περιορισμένους εδαφικούς πόρους μιας περιοχής και στο ευρύτερο περιβάλλον. Χρησιμοποιείται, μεταξύ άλλων, για την εξαγωγή πορισμάτων σχετικά με την εξάπλωση των αστικών περιοχών, για την υποστήριξη της προστασίας των οικοσυστημάτων, της αναστολής της υποβάθμισης της βιοποικιλότητας και της καταγραφής των συνεπειών της μεταβολής του κλίματος. Τα προϊόντα είναι δημοσιά διαθέσιμα, χωρίς κόστος στο δικτυακό τόπο του ΕΟΠ.

4.2 Εφαρμογή Ταξινόμησης

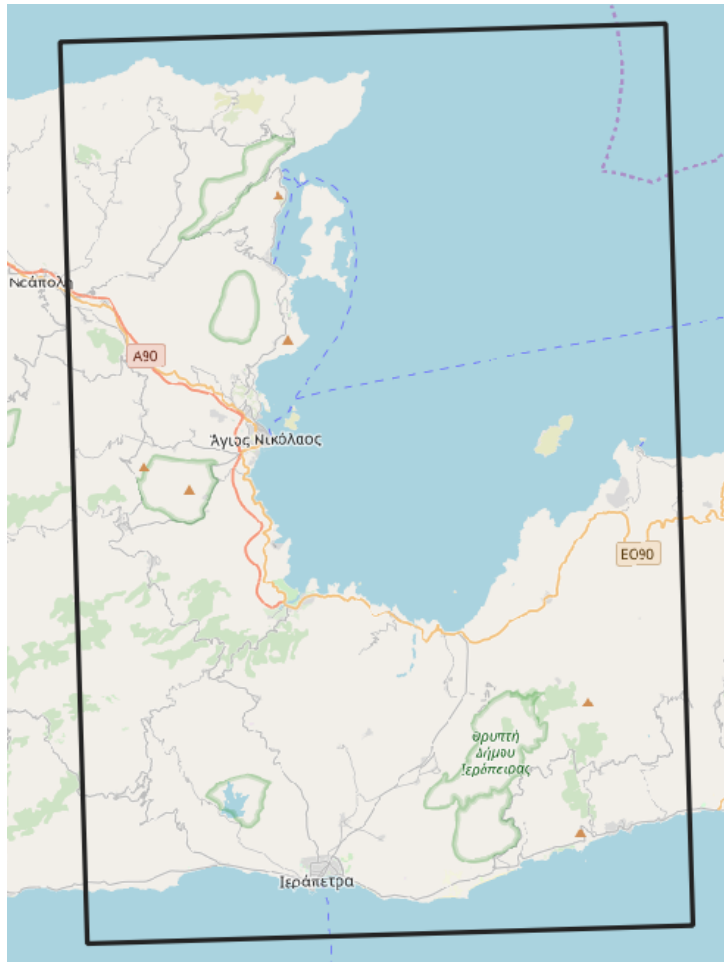
4.2.1 Στόχοι της εφαρμογής

Η παρούσα εφαρμογή παρατίθεται στην διπλωματική εργασία, με στόχο να αναδείξει τα χωρικά δεδομένα, τις επιστήμες που τα αξιοποιούν και την χρησιμότητα των εφαρμογών που προκύπτουν. Για την ολοκλήρωση της εφαρμογής ακολουθήθηκαν ορισμένες διεργασίες όπως αυτές προκύπτουν από τροποποίηση της μεθοδολογίας που αναπτύχθηκε στο πλαίσιο της διπλωματικής εργασίας 'Χαρτογράφηση Κάλυψης Γης και Καλλιεργειών με Τεχνικές Μηχανικής Μάθησης από Διαχρονικά Δεδομένα Sentinel-2', Ο.Γούναρη (2018) και το Εργαστήριο Τηλεπισκόπησης του Εθνικού Μετσόβιου Πολυτεχνείου.

4.2.2 Περιγραφή εφαρμογής και δεδομένων

Η εφαρμογή αναδεικνύει την αύξηση της αστικής δόμησης στην περιοχή γύρω από τον Άγιο Νικόλαο (βόρεια και νότια προάστια) και την Ιεράπετρα Κρήτης (ανατολικά προάστια), συνολικού εμβαδού 1142.860 τετραγωνικά χιλιόμετρα. Οι περιοχές αυτές τα τελευταία χρόνια έχουν δεχθεί μεγάλη αύξηση τουριστικών υποδομών και ξενοδοχειακών μονάδων. Παρόλο που δεν είναι εύκολο να αναδειχθεί με την συγκεκριμένη εφαρμογή η χρήση γης, πραγματοποιήθηκε μια προσπάθεια να αναδειχθεί η μεταβολή της κάλυψης γης που έχει υποστεί το τοπίο. Χρησιμοποιήθηκαν πολυφασματικά δεδομένα που αντιστοιχούν σε 4 λήψεις για κάθε έτος, επί 8 έτη. Το αρχικό μέγεθος του αρχείου μιας λήψης ανέρχεται στα 518.6 MB για τον Landsat 5 και στα 1.8 GB για τον Landsat 8, άρα το μέγεθος των λήψεων που πραγματοποιήθηκαν ανέρχεται περίπου στα 27 GB. Συγκεκριμένα οι χρονολογίες έχουν επιλεγεί ανά πέντε χρόνια και καλύπτουν έναν χρονικό ορίζοντα 33 χρόνων, ως εξής: 1985, 1990, 1995, 2000, 2005, 2010, 2015, 2018.

Οι τέσσερις λήψεις κάθε έτους, επιχειρήθηκε να καλύπτουν όλο το εύρος του έτους, έτσι ώστε να λαμβάνεται υπόψη η περιοδικότητα των φαινομένων που συμβαίνουν στην γήινη επιφάνεια. Διαδικασίες που θα ακολουθούσαν σε μια άρτια επιστημονική μεθοδολογία τηλεπισκόπησης, όπως για παράδειγμα ατμοσφαιρικές διορθώσεις (π.χ. αλγόριθμος ATCOR), εντοπισμός και απαλοιφή σύννεφων και σκιών (π.χ. αλγόριθμος FMASK), δεν πραγματοποιήθηκαν για λόγους απλοποίησης της διαδικασίας. Συμπερασματικά, τα αποτελέσματα που προκύπτουν προέρχονται από μια απλοποιημένη μεθοδολογία, με την λιγότερη δυνατή προεπεξεργασία των εικόνων. Οι εικόνες επιλέχθηκαν από τον ιστότοπο της USGS έτσι ώστε να έχουν την λιγότερη δυνατή συννεφοκάλυψη, να μην αποκρύπτονται μέρη της περιοχής μελέτης και κυρίως να επηρεάζεται όσο το δυνατόν λιγότερο ο αλγόριθμος της ταξινόμησης.



Εικόνα 25 - Περιοχή μελέτης

Παράλληλα, αποκτήθηκαν τα δεδομένα του προγράμματος CORINE για τα έτη 2006 και 2012, τα οποία χρησιμοποιήθηκαν ως βοηθητική πληροφορία, στο στάδιο συλλογής δεδομένων για την εκπαίδευση (ground truth) του αλγορίθμου ταξινόμησης ο οποίος χρησιμοποιήθηκε κατά την ταξινόμηση των εικόνων. Η συλλογή δεδομένων εκπαίδευσης του αλγορίθμου πραγματοποιήθηκε με παρατήρηση της περιοχής και επιλογή ορισμένων εικονοστοιχείων που αντιστοιχούν στις ακόλουθες κατηγορίες, μέσω ψηφιοποίησης στο πρόγραμμα QGIS. Σαν βοηθητικό υπόβαθρο αξιοποιήθηκε το CORINE.

- "Δομημένες Περιοχές"
- "Άμμος / Βράχια"
- "Μόνιμα Αρδευόμενες Καλλιέργειες"
- "Όπωροφόρα / Ελαιώνες"
- "Σύνθετες Καλλιέργειες"
- "Κωνοφόρα"
- "Χαμηλή βλάστηση"
- "Στάσιμα Ύδατα"
- "Θάλασσα"

Τα δεδομένα τα οποία συλλέχθηκαν, χωρίστηκαν σε σύνολο δεδομένων εκπαίδευσης, και σύνολο δεδομένων ελέγχου σε ποσοστό 80% και 20% αντίστοιχα. Και στα δύο σύνολα περιέχονται όλες οι περιπτώσεις των διαφορετικών κατηγοριών.

4.2.3 Προ-επεξεργασία εικόνων

Οι εικόνες που λήφθηκαν από το site της USGS, αρχικά κόπηκαν στις συντεταγμένες $minx=25.6239$, $maxx=25.9216$, $miny=34.9867$, $maxy=35.355$ του συστήματος αναφοράς ΕΓΣΑ'87, ώστε να διατηρηθεί μόνο η περιοχή μελέτης. Στη συνέχεια, για κάθε δορυφόρο, βρέθηκαν οι ελάχιστες και οι μέγιστες τιμές για κάθε κανάλι από όλες τις χρονολογίες των λήψεων και για κάθε μια από τις τέσσερις λήψεις της εκάστοτε χρονολογίας, και κανονικοποιήθηκαν οι τιμές των εικόνων στο διάστημα 0-1. Τα κανάλια που χρησιμοποιήθηκαν είναι τα B1, B2, B3, B4, B5, B7 για τον Landsat_5 και τα αντίστοιχα B2, B3, B4, B5, B6, B7 για τον Landsat_8. Ακολουθεί ο υπολογισμός των φασματικών δεικτών, οι οποίοι υπολογίστηκαν για κάθε εικονοστοιχείο της εικόνας έτσι ώστε να χρησιμοποιηθούν σαν συμπληρωματική πληροφορία κατά την διαδικασία της ταξινόμησης. Οι φασματικοί δείκτες που χρησιμοποιήθηκαν είναι οι εξής:

$$\begin{aligned}ndvi &= (nir-red)/(nir+red) \\msavi &= (2 * nir + 1 - np.sqrt((2 * nir + 1)**2 - 8*(nir-red))) / 2 \\ndbi &= (swir-nir)/(swir+nir) \\ndwi &= (nir-swir)/(nir+swir)\end{aligned}$$

Ουσιαστικά, οι δείκτες είναι εικόνες οι οποίες προκύπτουν μετά από μαθηματικές πράξεις μεταξύ των φασματικών καναλιών της ίδιας εικόνας ή διαφορετικών εικόνων. Οι λόγοι καναλιών βασίζονται ιδιαίτερα στις φασματικές ιδιότητες των υλικών του εδάφους όπως π.χ. απορρόφηση / ανάκλαση στα διαφορετικά μήκη κύματος. Η απορρόφηση εξαρτάται από τη μοριακή δομή της επιφάνειας που προσπίπτει η ηλεκτρομαγνητική ακτινοβολία. Ως αποτέλεσμα προκύπτει ότι οι λόγοι των φασματικών καναλιών δίνουν πληροφορίες για την χημική σύσταση του κάθε στόχου. (Περάκης, 2015) Παρακάτω, αναλύεται ο δείκτης NDVI και ενδεικτικά οι υπόλοιποι.

NDVI (Normalized Difference Vegetation Index) (Rouse et al. 1974)

Αποτελεί το δημοφιλέστερο δείκτης βλάστησης, είναι απλός αλλά αποτελεσματικός στο να ποσοτικοποιεί τη βλάστηση. Χρησιμοποιεί το κόκκινο και το εγγύς-υπέρυθρο μήκος κύματος το οποίο ανακλάται από την επιφάνεια της γης. Κανονικοποιεί την σκέδαση των φυλλωμάτων στο εγγύς-υπέρυθρο και την απορρόφηση της χλωροφύλλης στο κόκκινο μήκος κύματος. Περιγράφει την πυκνότητα της βλάστησης αλλά και την κατάσταση υγείας του φυτού, γεγονός από το οποίο μπορούμε να αντλήσουμε δεδομένα για τον φαινολογικό κύκλο του υπό μελέτη στοιχείου ή την απόδοση καλλιεργειών κ.α. Στα πλεονεκτήματα του είναι η ελαχιστοποίηση των τοπογραφικών επιδράσεων. Στα μειονεκτήματα του δείκτη είναι ότι παρουσιάζει δείγματα κορεσμού (saturation) σε πολύ υψηλές συγκεντρώσεις βλάστησης (Περάκης, 2015). Ακόμη, όπως επισημαίνουν οι Banpari et al (σύμφωνα με τους Jackson et al), ο δείκτης παρουσιάζει ευαισθησία στις ατμοσφαιρικές επιδράσεις, ενώ η δυναμική του εξασθενεί όταν η φυτοκάλυψη είναι αραιή (συγκέντρωση μικρότερη του 50%), εξαιτίας της

ανάκλαστικότητας του εδάφους. Οι τιμές του δείκτη NDVI, ανάλογα με το διάστημα που κυμαίνονται, υποδεικνύουν:

NDVI = -1 έως 0	-> υδάτινες μάζες.
NDVI = -0.1 έως 0.1	-> γυμνοί βράχοι, άμμος ή χιόνι.
NDVI = 0.2 έως 0.5	-> θάμνοι, γρασίδι, ώριμες καλλιέργειες.
NDVI = 0.6 έως 1.0	-> πυκνή βλάστηση ή τροπικό δάσος.

MSAVI (Modified soil adjusted vegetation index) (Qi et al. 1994)

Ο δείκτης SAVI δημιουργήθηκε ώστε να βελτιώσει την ευαισθησία του NDVI όταν υπάρχει στο φόντο της βλάστησης το υπέδαφος το οποίο μπερδεύει την φασματική της υπογραφή. Ο MSAVI αποτελεί παραλλαγή του SAVI, και διευκολύνει τη διάκριση βλάστησης -εδάφους και την απομάκρυνση του «θορύβου» από την ανάκλαση του εδάφους που δεν κατορθώνουν άλλοι δείκτες βλάστησης όπως ο NDVI.

NDWI (Normalized Difference Water Index) (McFeeters 1996)

Αποτελεί τον καταλληλότερο δείκτη για την απεικόνιση των υδάτινων μαζών. Το νερό έχει δυνατή απορρόφηση και χαμηλή ακτινοβολία στο εύρος μεταξύ του ορατού και του υπέρυθρου μήκους κύματος. Ο δείκτης χρησιμοποιεί το πράσινο και το εγγύς υπέρυθρο, μπορεί να εξάγει χρήσιμες πληροφορίες σχετικά με τις υδάτινες μάζες στις περισσότερες των περιπτώσεων. Είναι ευαίσθητος σε οικιστική γη, προσφέρει την ανάδειξη των υδάτινων σωμάτων και μπορεί να χρησιμοποιηθεί για την εκτίμηση των επιπτώσεων πλημμύρας.

NDBI (Normalized Difference Built-up Index) (Zha, Gao, Ni 2003)

Ο NDBI χρησιμοποιείται στον εντοπισμό αστικών περιοχών με υψηλότερη ανάκλαση στην περιοχή των μικροκυματικών υπέρυθρων (SWIR), σε σύγκριση με την περιοχή της εγγύς υπέρυθρης ακτινοβολίας (NIR). Οι εφαρμογές περιλαμβάνουν προβλέψεις απορροής ποταμών και χωροταξικό σχεδιασμό.

Στη συνέχεια, για κάθε χρονολογία κατασκευάζεται ο κύβος, ο οποίος αποτελεί έναν τρισδιάστατο πίνακα τιμών διαστάσεων 922x1374x40 και μέγεθος αρχείου 2002.7 MB. Ενδεικτικά, η δομή του κύβου για το έτος 2018 είναι:

Κύβος για την χρονολογία 2018 (922x1374x40)

- Κύβος την λήψη 14-04-2018 (922x1374x10)
 - ➔ 6 κανάλια για την συγκεκριμένη λήψη (922x1374x6)
 - ➔ 4 δείκτες για την συγκεκριμένη λήψη (922x1374x4)
- Κύβος την λήψη 03-07-2018 (922x1374x10)
 - ➔ 6 κανάλια για την συγκεκριμένη λήψη (922x1374x6)
 - ➔ 4 δείκτες για την συγκεκριμένη λήψη (922x1374x4)

- Κύβος την λήψη 05-09-2018 (922x1374x10)
 - ➔ 6 κανάλια για την συγκεκριμένη λήψη (922x1374x6)
 - ➔ 4 δείκτες για την συγκεκριμένη λήψη (922x1374x4)
- Κύβος την λήψη 08-11-2018 (922x1374x10)
 - ➔ 6 κανάλια για την συγκεκριμένη λήψη (922x1374x6)
 - ➔ 4 δείκτες για την συγκεκριμένη λήψη (922x1374x4)

Ακολουθεί η εκπαίδευση του αλγορίθμου Random Forest. Στο σημείο αυτό είναι απαραίτητο να υπάρχει το σύνολο των δεδομένων που θα υποδείξουν στον αλγόριθμο τις σωστές κατηγορίες, γνωστά και ως ground truth. Η ποιότητα της πληροφορίας που περιέχεται στο ground truth είναι καθοριστικής σημασίας για την απόδοση οποιουδήποτε αλγορίθμου. Όπως αναφέρθηκε, η εκπαίδευση πραγματοποιήθηκε με το 80% του ground truth, και το 20% κρατήθηκε 'κρυφό' από τον αλγόριθμο, ώστε να ακολουθήσει η εξέταση της απόδοσής του με αυτό.

4.2.4 Random Forest Classifier

Ο αλγόριθμος ταξινόμησης που χρησιμοποιήθηκε είναι τα Τυχαία Δάση (Random Forest). Η λογική στην οποία βασίζεται ο ταξινομητής είναι πως συνδυάζοντας διαφορετικούς 'αδύναμους' ταξινομητές, μπορεί να επιτευχθεί εν τέλη μια καλή ακρίβεια ταξινόμησης (bagging). Με τη βοήθεια της βιβλιοθήκης sklearn, ο ταξινομητής συνδυάζει πολλά δένδρα απόφασης, για αυτό και ονομάζεται 'Τυχαίο Δάσος'. Κάθε δένδρο απόφασης προκύπτει από ένα τυχαίο υποσύνολο (bootstrap sample) του train_set το οποίο αποτελείται από τον ίδιο αριθμό δειγμάτων εκπαίδευσης με το αρχικό αριθμό, αλλά με λιγότερα χαρακτηριστικά για κάθε δείγμα εκπαίδευσης. Έτσι, κάθε δέντρο απόφασης έχει εκπαιδευτεί σε τυχαίο υποσύνολο των δεδομένων εκπαίδευσης και όχι σε όλα τα δείγματα των δεδομένων εκπαίδευσης.

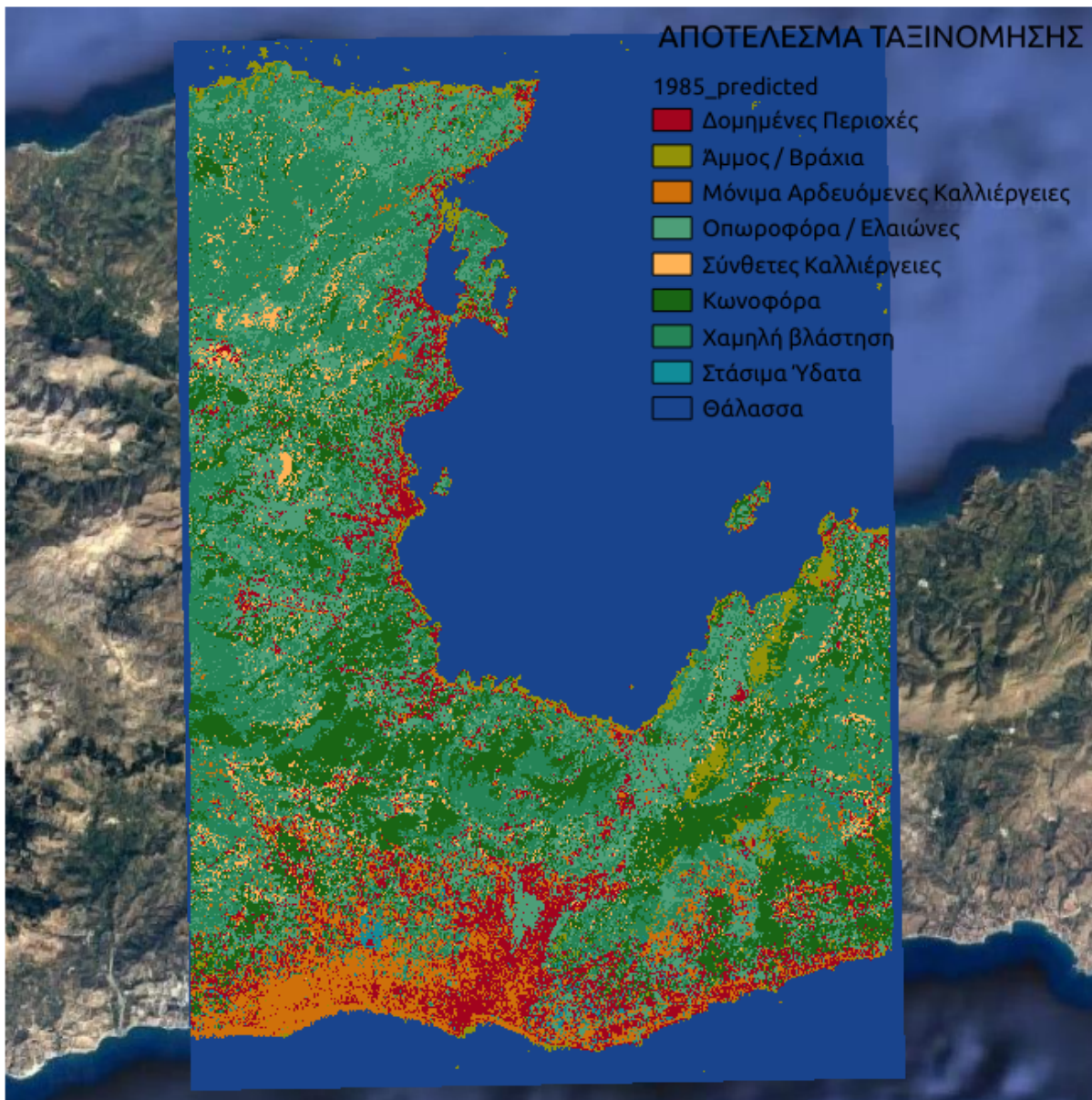
Εξ' αιτίας αυτής της τυχειότητας, αυξάνεται το bias του δάσους, σε σύγκριση με μόνο ένα δέντρο απόφασης. Το γεγονός αυτό αντισταθμίζεται, διότι η τελική απόφαση του δάσους για την κατάταξη ενός καινούργιου εικονοστοιχείου προκύπτει από τον μέσο όρο των αποφάσεων των δέντρων, και άρα μειώνεται η διακύμανση του μοντέλου, αυξάνοντας την απόδοσή του σε σχέση με ένα μόνο δέντρο απόφασης. Προφανώς, το εικονοστοιχείο κατατάσσεται στην κατηγορία που θα υποδείξουν τα περισσότερα δέντρα απόφασης. Πιο συγκεκριμένα όμως, η βιβλιοθήκη sklearn, δεν αξιοποιεί την μία κατηγορία που θα υποδείξει κάθε δέντρο απόφασης, αλλά λαμβάνει υπόψιν την πιθανότητα με βάση την οποία το δέντρο απόφασης ταξινομεί το εικονοστοιχείο σε κάποια κατηγορία. Τέλος, οι παράμετροι που χρησιμοποιήθηκαν είναι οι by default της βιβλιοθήκης sklearn (π.χ. δέντρα απόφασης=10).

Λόγω της αρχιτεκτονικής των τυχαίων δασών, ο ταξινομητής είναι ταχύτατος ακόμα και όταν αυξάνουν τα δεδομένα εκπαίδευσης, ενώ αξίζει να σημειωθεί ότι επιλέχθηκε ο συγκεκριμένος αλγόριθμος διότι σταθερά σε όλα τα πειράματα της διπλωματικής εργασίας της Ο.Γούναρη (2018) σημειώνει ελαφρώς μεγαλύτερα ποσοστά επιτυχίας, σε συνδυασμό με τον απαιτούμενο χρόνο.

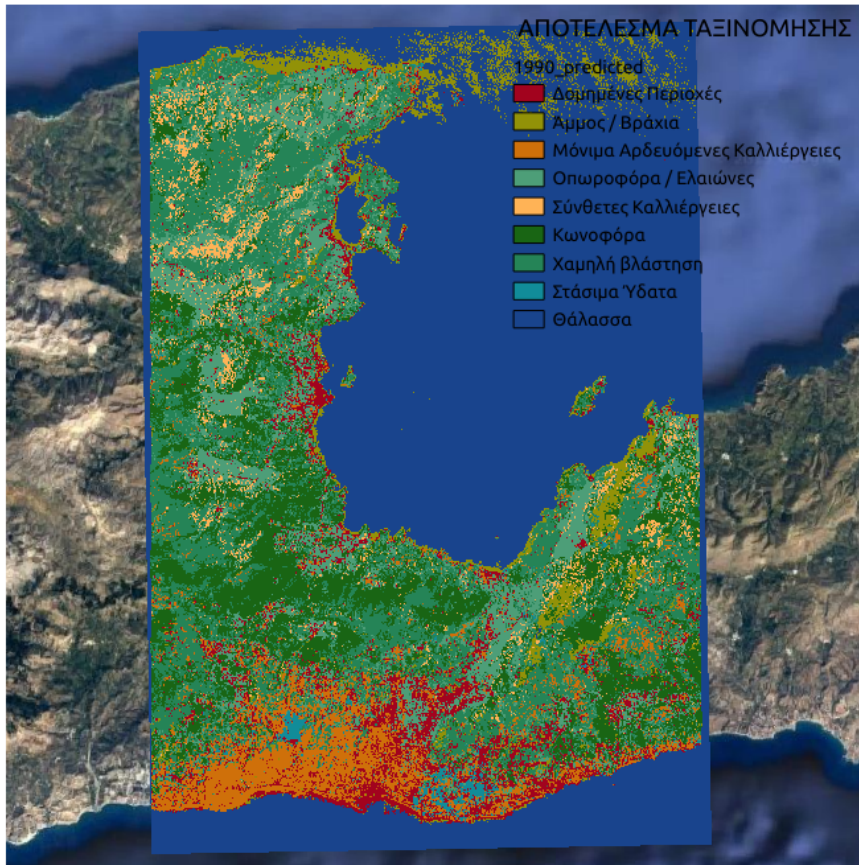
Η προεπεξεργασία, η ταξινόμηση, η διαχείριση των εικόνων και η αξιολόγηση της απόδοσης του αλγορίθμου πραγματοποιήθηκε σε προγραμματιστική γλώσσα Python 3.6.8. Οι βασικότερες βιβλιοθήκες που χρησιμοποιηθήκαν είναι οι os, rasterio, numpy, sklearn, pandas και matplotlib.

4.2.5 Αποτελέσματα

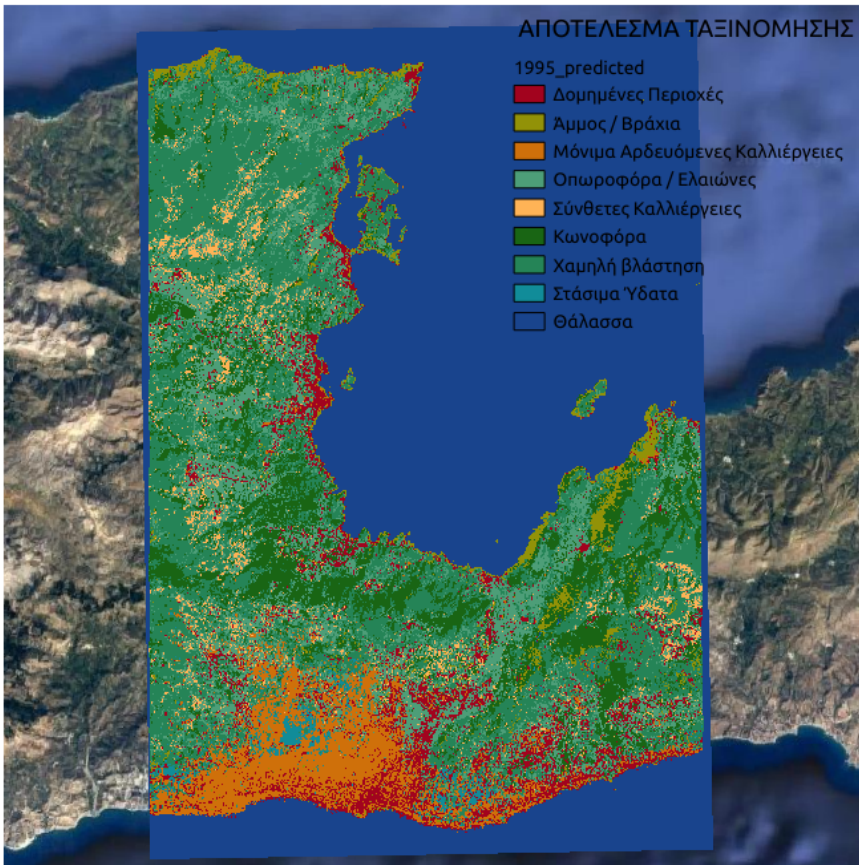
Παρατίθενται οι εικόνες, όπως προκύπτουν από τον αλγόριθμο. Προστέθηκε υπόμνημα με τα χρώματα των αντίστοιχων κατηγοριών για την ευκολότερη ανάγνωση του θεματικού χάρτη.



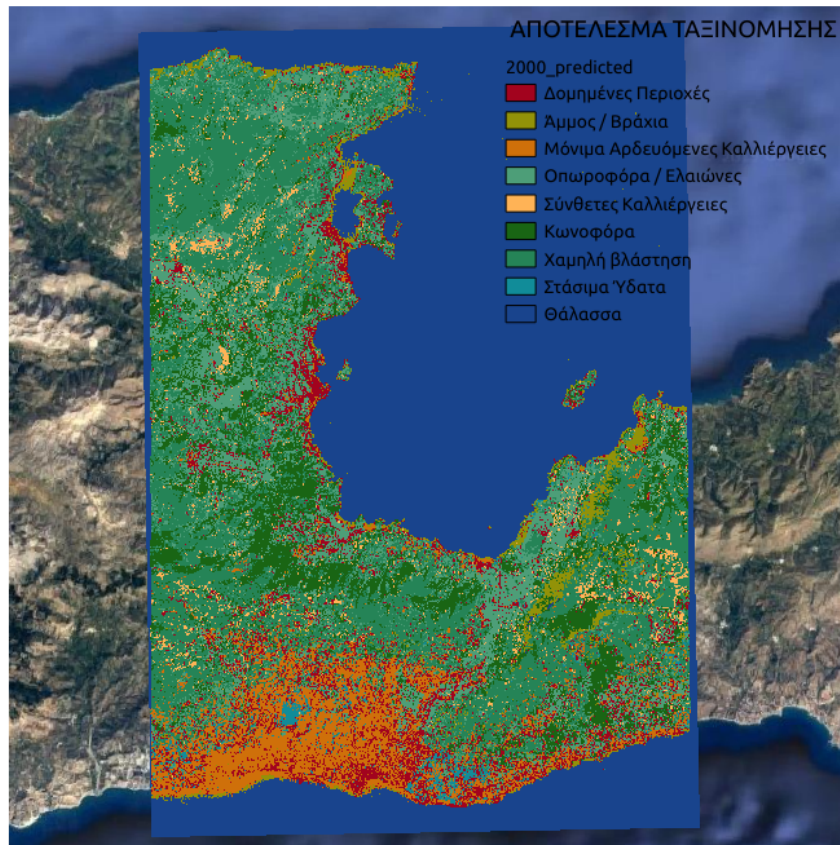
Εικόνα 26 - Αποτελέσματα Ταξινόμησης 1985



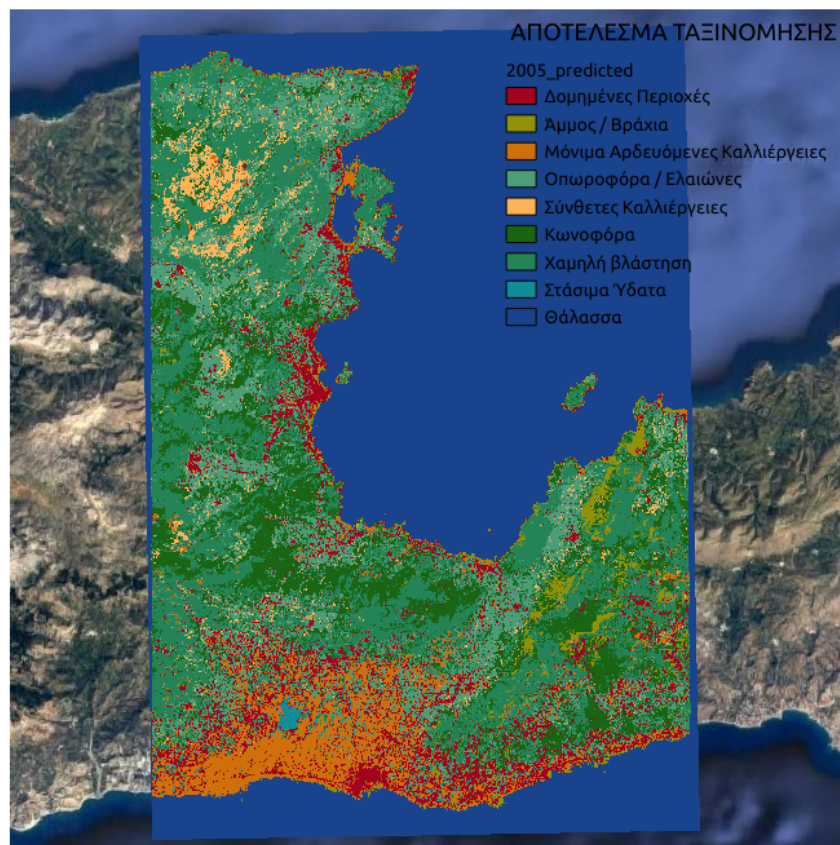
Εικόνα 27 - Αποτελέσματα Ταξινόμησης 1990



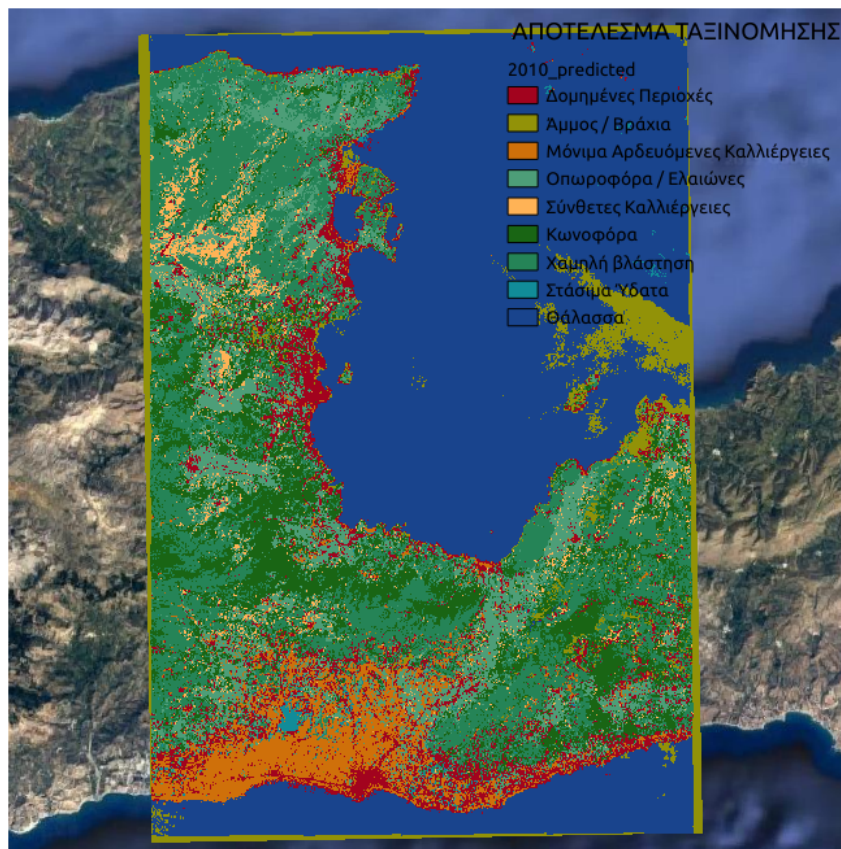
Εικόνα 28 - Αποτελέσματα Ταξινόμησης 1995



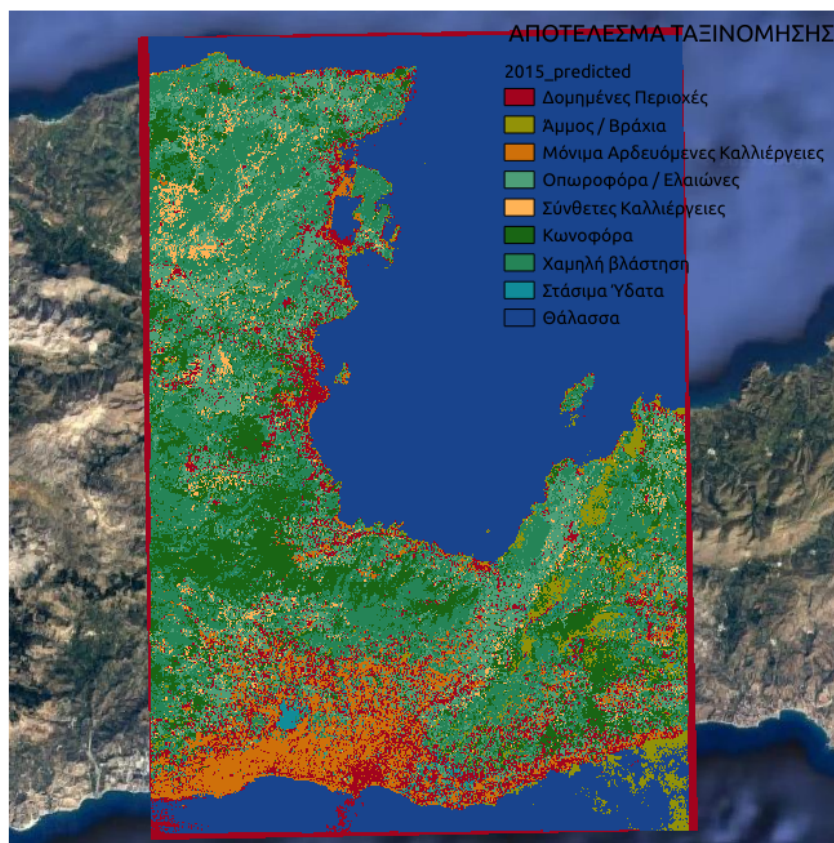
Εικόνα 29 - Αποτελέσματα Ταξινόμησης 2000



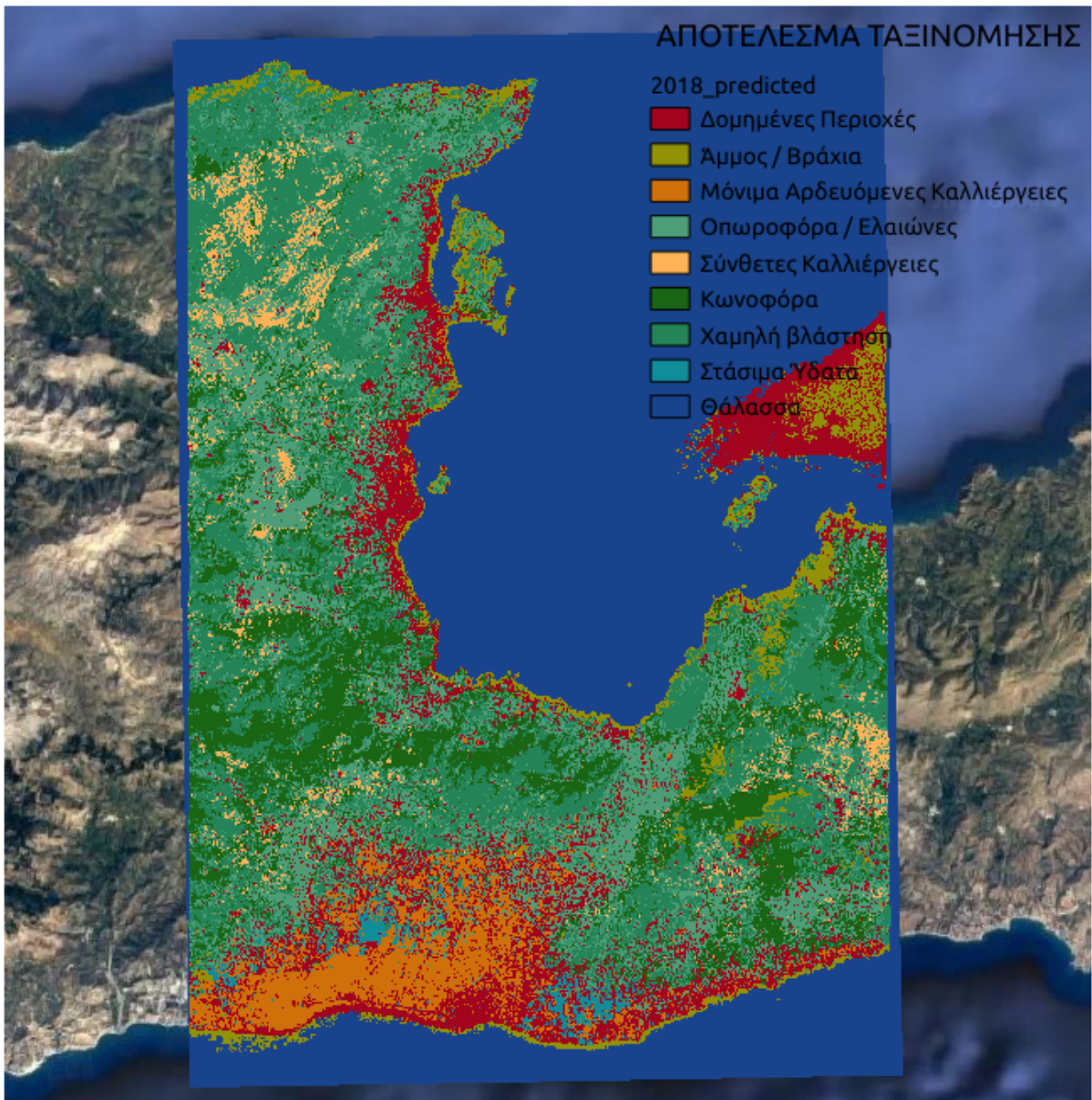
Εικόνα 30 - Αποτελέσματα Ταξινόμησης 2005



Εικόνα 31 - Αποτελέσματα Ταξινόμησης 2010



Εικόνα 32 - Αποτελέσματα Ταξινόμησης 2015



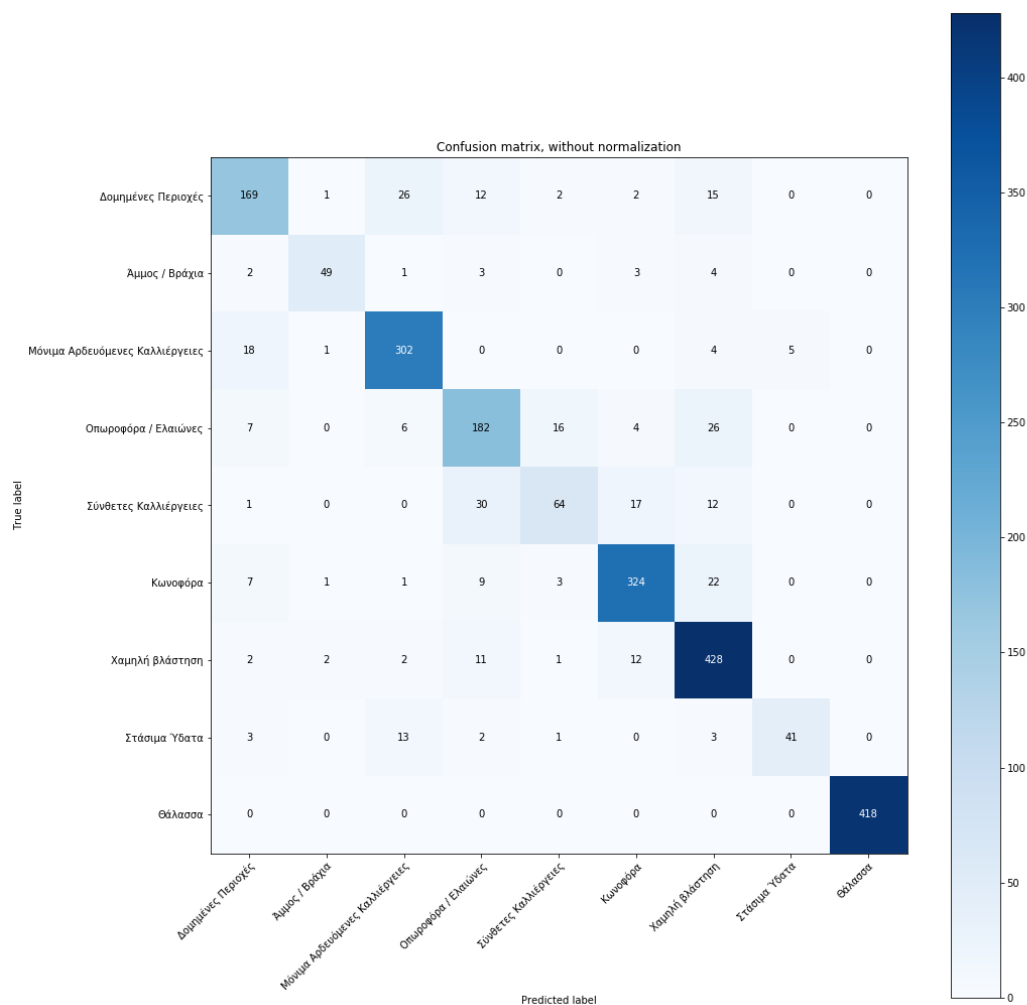
Εικόνα 33 - Αποτελέσματα Ταξινόμησης 2018

4.2.6 Αξιολόγηση αποτελεσμάτων

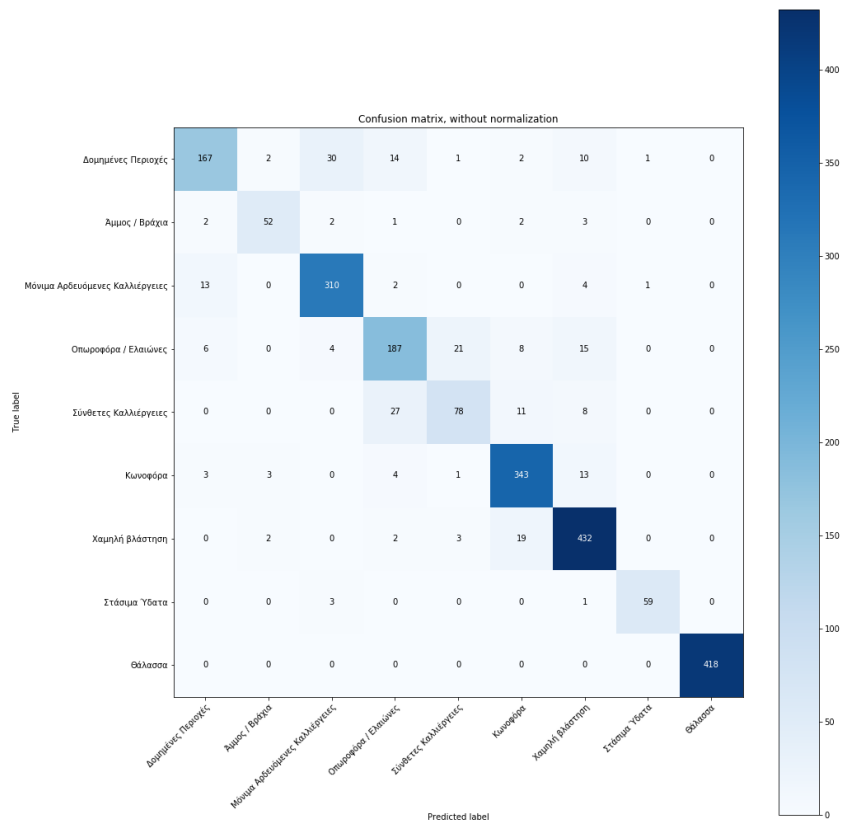
Για κάθε χρονολογία πραγματοποιήθηκε νέα εκπαίδευση του αλγορίθμου με βάση τα ίδια δεδομένα και παράχθηκε ξεχωριστό μοντέλο για κάθε έτος. Οι ακρίβειες που προκύψαν φαίνονται παρακάτω, ενώ στη συνέχεια παρατίθενται οι Πίνακες Σύγκυσης (Confusion Matrices), ώστε να γίνει κατανοητό ποιες κατηγορίες συγχέονται περισσότερο με ποιες άλλες.

Μοντέλο	Ακρίβεια
1985	86.33 %
1990	89.34 %
1995	85.28 %
2000	85.67 %
2005	88.90 %
2010	90.00 %
2015	86.72 %
2018	88.55 %

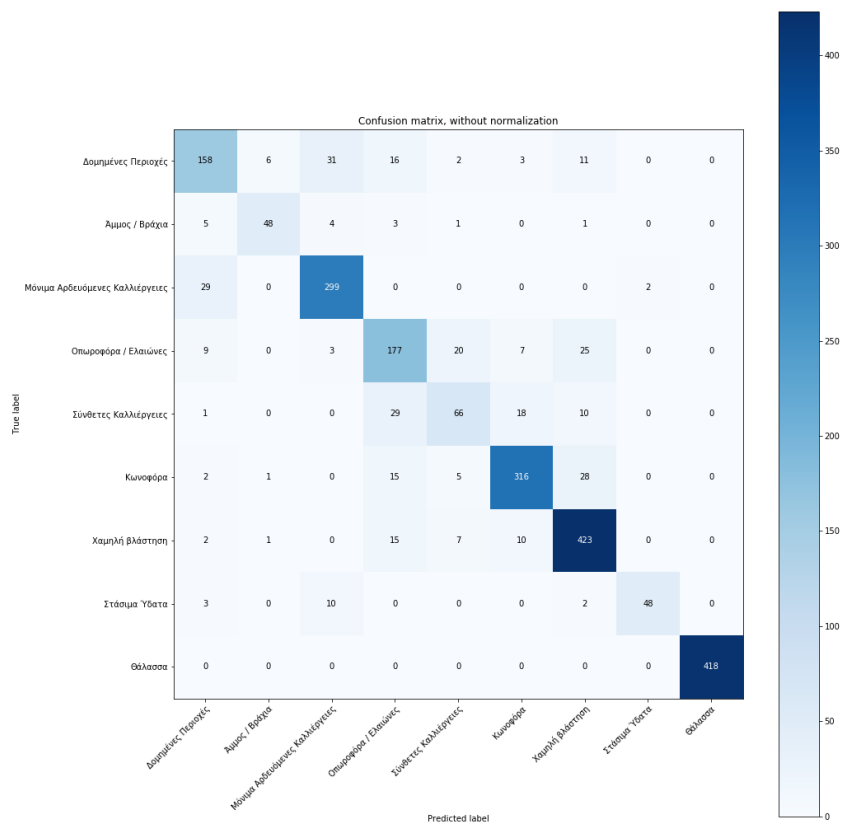
Πίνακας 5 - Ακρίβεια Ταξινόμησης



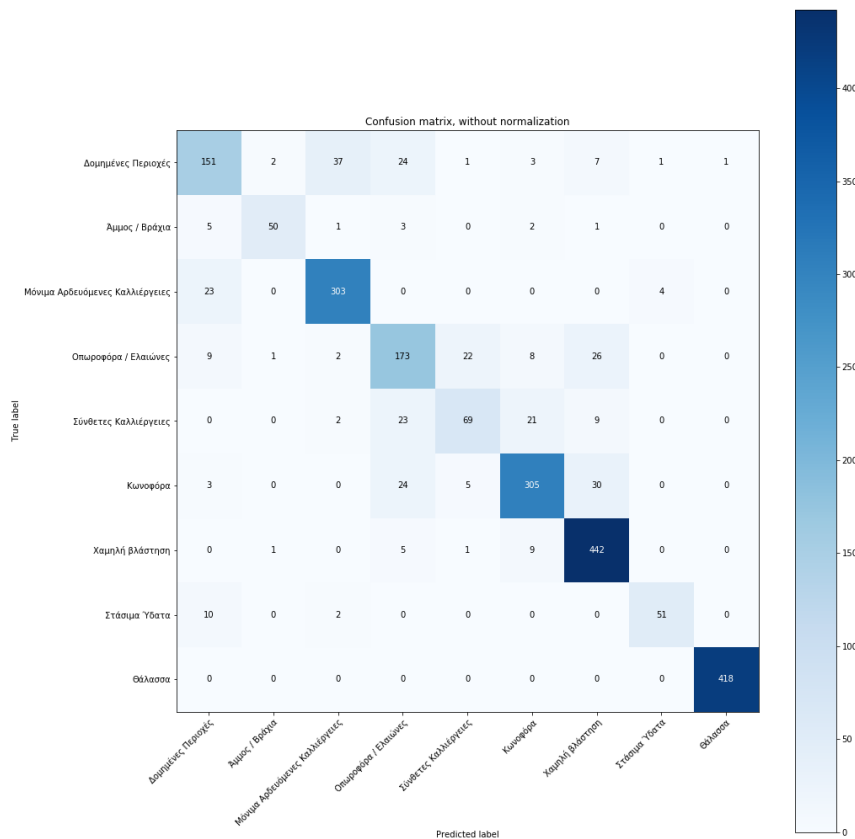
Πίνακας 6 - Πίνακας Σύγκυσης 1985



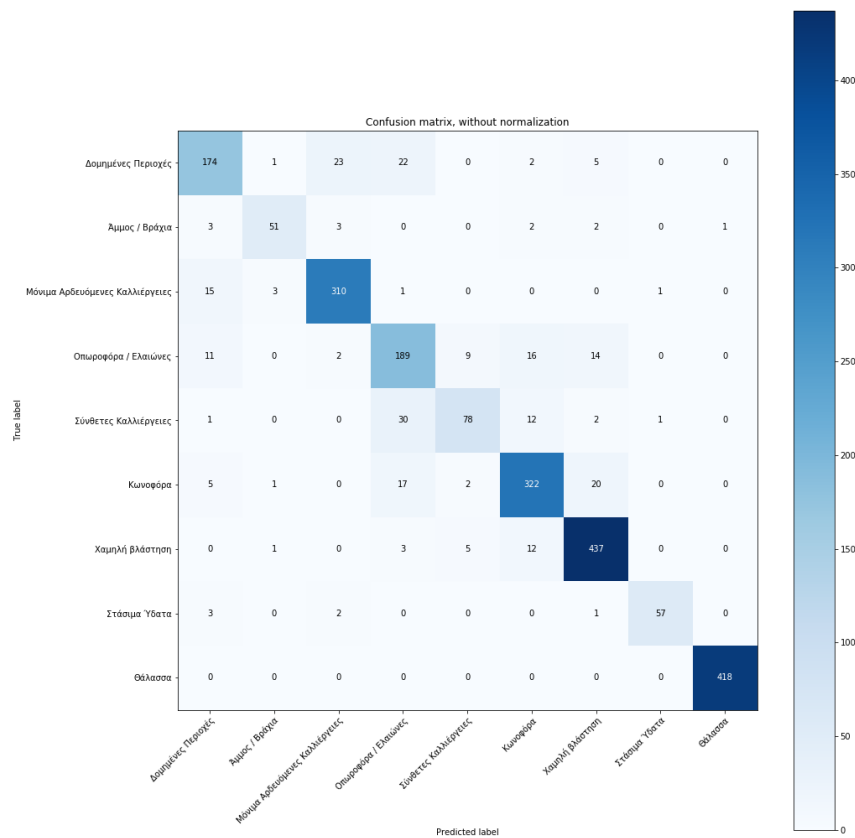
Πίνακας 7 - Πίνακας Σύγχυσης 1990



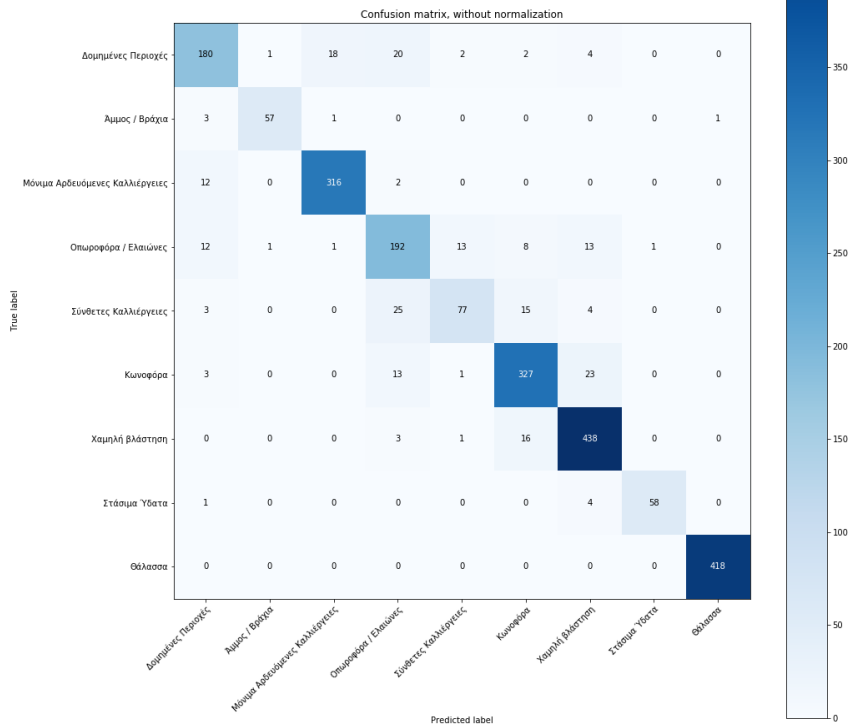
Πίνακας 8 - Πίνακας Σύγχυσης 1995



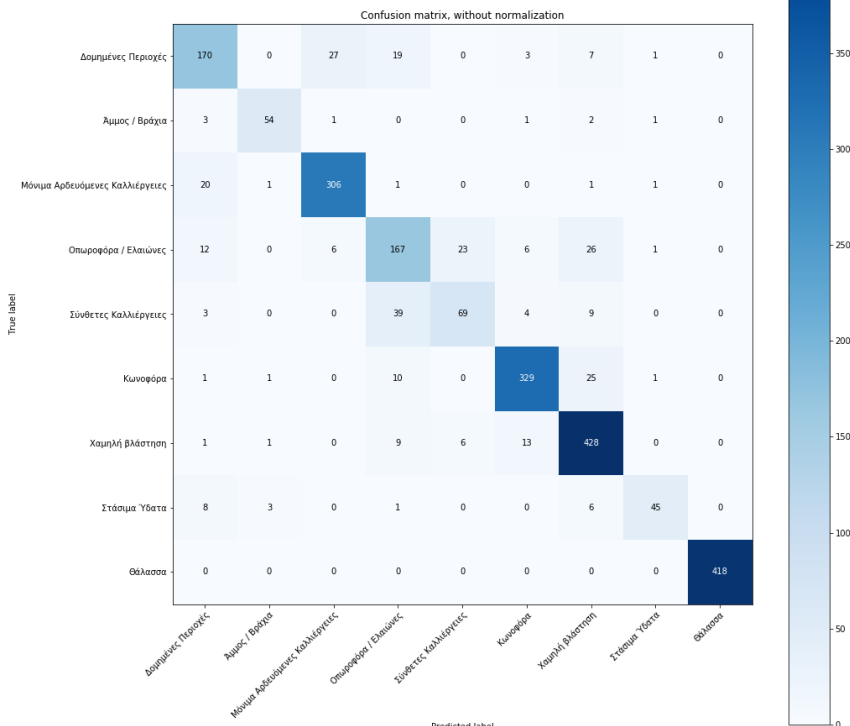
Πίνακας 9 - Πίνακας Σύγχυσης 2000



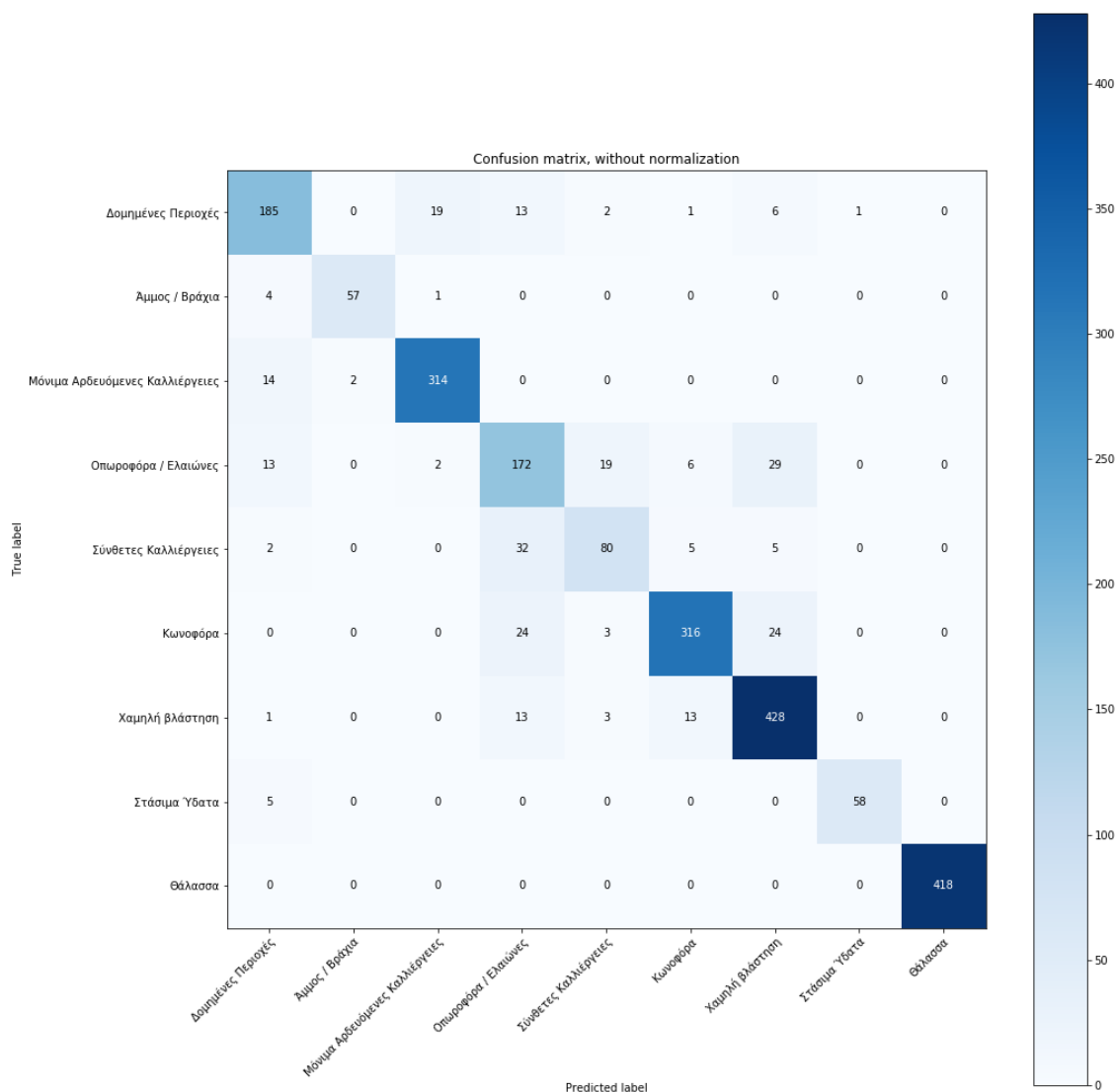
Πίνακας 10 - Πίνακας Σύγχυσης 2005



Πίνακας 11 - Πίνακας Σύγχυσης 2010



Πίνακας 12 - Πίνακας Σύγχυσης 2015



Πίνακας 13 - Πίνακας Σύγχυσης 2018

Με παρατήρηση των Πινάκων Σύγχυσης για κάθε έτος, προκύπτουν τα εξής:

1985

"Δομημένες Περιοχές" με "Μόνιμα Αρδευόμενες Καλλιέργειες"
 "Οπωροφόρα / Ελαιώνες" με "Χαμηλή βλάστηση" και "Σύνθετες Καλλιέργειες"
 "Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες" και "Κωνοφόρα"
 "Στάσιμα Ύδατα" με "Μόνιμα Αρδευόμενες Καλλιέργειες"

1990

"Δομημένες Περιοχές" με "Μόνιμα Αρδευόμενες Καλλιέργειες".
 "Οπωροφόρα / Ελαιώνες" με "Χαμηλή βλάστηση" και "Σύνθετες Καλλιέργειες".
 "Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες" και "Κωνοφόρα".

1995

"Δομημένες Περιοχές" με "Μόνιμα Αρδευόμενες Καλλιέργειες".
"Μόνιμα Αρδευόμενες Καλλιέργειες" με "Δομημένες Περιοχές".
"Οπωροφόρα / Ελαιώνες" με "Χαμηλή βλάστηση" και "Σύνθετες Καλλιέργειες".
"Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες" και "Κωνοφόρα".
"Στάσιμα Ύδατα" με "Μόνιμα Αρδευόμενες Καλλιέργειες".

2000

"Δομημένες Περιοχές" με "Μόνιμα Αρδευόμενες Καλλιέργειες" και "Οπωροφόρα/Ελαιώνες".
"Μόνιμα Αρδευόμενες Καλλιέργειες" με "Δομημένες Περιοχές".
"Οπωροφόρα / Ελαιώνες" με "Χαμηλή βλάστηση" και "Σύνθετες Καλλιέργειες".
"Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες" και "Κωνοφόρα".
"Στάσιμα Ύδατα" με "Δομημένες Περιοχές".

2005

"Δομημένες Περιοχές" με "Μόνιμα Αρδευόμενες Καλλιέργειες" και "Οπωροφόρα / Ελαιώνες".
"Οπωροφόρα / Ελαιώνες" με "Κωνοφόρα".
"Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες".

2010

"Δομημένες Περιοχές" με "Οπωροφόρα / Ελαιώνες" και "Μόνιμα Αρδευόμενες Καλλιέργειες".
"Οπωροφόρα / Ελαιώνες" με "Χαμηλή βλάστηση" και "Σύνθετες Καλλιέργειες".
"Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες" και "Κωνοφόρα".

2015

"Δομημένες Περιοχές" με "Μόνιμα Αρδευόμενες Καλλιέργειες" και "Οπωροφόρα / Ελαιώνες".
"Οπωροφόρα / Ελαιώνες" με "Χαμηλή βλάστηση", "Σύνθετες Καλλιέργειες" και "Δομημένες Περιοχές".
"Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες".
"Στάσιμα Ύδατα" με "Δομημένες Περιοχές" και "Χαμηλή βλάστηση".

2018

"Δομημένες Περιοχές" με "Μόνιμα Αρδευόμενες Καλλιέργειες".
"Οπωροφόρα / Ελαιώνες" με "Χαμηλή βλάστηση" και "Σύνθετες Καλλιέργειες".
"Σύνθετες Καλλιέργειες" με "Οπωροφόρα / Ελαιώνες".
"Κωνοφόρα" με "Οπωροφόρα / Ελαιώνες" και "Χαμηλή βλάστηση".
"Στάσιμα Ύδατα" με "Δομημένες Περιοχές".

Συμπερασματικά, παρά την απλοποιημένη μεθοδολογία που ακολουθήθηκε για την ταξινόμηση των οκτώ χρονολογιών, τα μοντέλα φαίνεται να έχουν την ικανότητα να σημειώσουν σχετικά υψηλά ποσοστά επιτυχίας. Αρχικά, φαίνεται ότι μπορούν να ξεχωρίσουν τις τρεις βασικές κατηγορίες, δηλαδή έδαφος, νερό και βλάστηση. Είναι

αναμενόμενες ορισμένες συμπεριφορές, όπως για παράδειγμα το ότι συγχέονται κατηγορίες όπως "Χαμηλή βλάστηση", "Σύνθετες Καλλιέργειες" και "Οπωροφόρα / Ελαιώνες", λόγω ομοιότητας του προτύπου που ακολουθούν. Σε αυτές τις κατηγορίες συμβάλει σημαντικά η παρεμβολή του εδάφους μεταξύ της βλάστησης. Επιπροσθέτως, είναι αναμενόμενο από την συγκεκριμένη εφαρμογή το γεγονός ότι συγχέονται οι περιοχές με δέντρα όπως για παράδειγμα "Κωνοφόρα" με "Οπωροφόρα / Ελαιώνες".

Το γεγονός που προκαλεί εντύπωση, είναι η σύγχυση του μοντέλου μεταξύ της κατηγορίας "Δομημένες Περιοχές" με άλλες κατηγορίες που περιέχουν ύδατα, για παράδειγμα "Μόνιμα Αρδευόμενες Καλλιέργειες". Οι φασματικές αποκρίσεις των κατηγοριών 'νερό' και 'έδαφος' ή ακόμα και 'πέτρα-τσιμέντο', έχουν πολύ μεγάλη διαφοροποίηση μεταξύ τους και σπάνια συμβαίνει να συγχέονται. Στη παρούσα εφαρμογή το φαινόμενο αυτό παρατηρείται διότι ήδη από τα δεδομένα εκπαίδευσης υπάρχουν εικονοστοιχεία που περιέχουν ύδατα και υποδείχθηκαν στον αλγόριθμο σαν "Δομημένες Περιοχές". Τα ύδατα αυτά προκύπτουν από τις πισίνες που βρίσκονται μεταξύ των κτιρίων των ξενοδοχείων και των μικρότερων ενοικιαζόμενων βιλών.

4.3 Εφαρμογή Χωροθέτησης

Το αντικείμενο της συγκεκριμένης εφαρμογής είναι υποθετικό και αφορά τη χωροθέτηση μιας εταιρείας παροχής πρόσθετων υπηρεσιών σε ενοικιαζόμενα μέσω Airbnb στην περιοχή της Ιεράπετρας. Στόχος είναι η χρήση διαθέσιμων δεδομένων και τεχνολογιών ανάλυσης πάνω σε διανυσματικά δεδομένα για την επίλυση ενός προβλήματος χωροθέτησης – που αποτελεί ένα σύνθετο πρόβλημα χωρικής ανάλυσης.

4.3.1 Δεδομένα

Τα δεδομένα που χρησιμοποιήθηκαν για τη συγκεκριμένη εφαρμογή προήλθαν από 2 πηγές. Για την ανάλυση του οδικού δικτύου χρησιμοποιήθηκαν δεδομένα από το OpenStreetMap. Τα δεδομένα αυτά είναι ελεύθερο να τα κατεβάσει κανείς ή να τα τροποποιήσει από το API του OpenStreetMap ή από την ιστοσελίδα του OpenStreetMap ή από υπηρεσίες τρίτων όπως το Geofabrik (εταιρεία που παρέχει υπηρεσίες πάνω στα δεδομένα του OpenStreetMap).

Τα δεδομένα για τα σημεία ενδιαφέροντος (δωμάτια Airbnb) είναι δυνατό να συγκεντρωθούν με τρεις τρόπους:

- Από το API του Airbnb
- Μέσω λογισμικού εξόρυξης δεδομένων που να εντοπίζει τα σημεία πάνω στο χάρτη στην ιστοσελίδα του Airbnb
- Με αναλογική ψηφιοποίηση

Τη στιγμή συγγραφής του παρόντος η πρόσβαση στο API του Airbnb ήταν κλειστή και η ανάπτυξη λογισμικού εξόρυξης δεδομένων (που δεν θα ξαναχρησιμοποιηθεί) για μερικές δεκάδες σημεία δε θεωρείται δόκιμη οπότε τα δεδομένα του αυτά ψηφιοποιήθηκαν με το χέρι.

4.3.2 Τεχνολογίες

QGIS

Πρόκειται για ΣΓΠ ανοιχτού λογισμικού.

Ο Gary Sherman ξεκίνησε την ανάπτυξη του QGIS στις αρχές του 2002 και έγινε έργο του Open Society Geospatial Foundation το 2007. Η έκδοση 1.0 κυκλοφόρησε τον Ιανουάριο του 2009. Το QGIS συντηρείται και αναπτύσσεται από προγραμματιστές - εθελοντές οι οποίοι εκδίδουν τακτικά ενημερώσεις και εκδόσεις που διορθώνουν σφάλματα. Από το 2012 έχει μεταφραστεί σε 48 γλώσσες και η εφαρμογή χρησιμοποιείται διεθνώς σε ακαδημαϊκό και επαγγελματικό περιβάλλον.

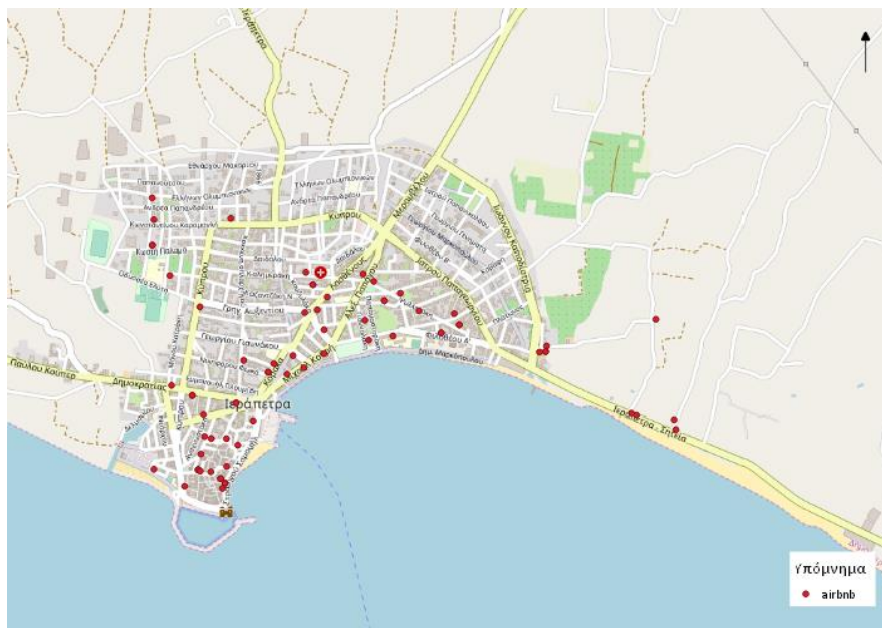
Openrouteservice

Το εργαλείο που χρησιμοποιήθηκε για το κομμάτι της χωρικής ανάλυσης που έχει να κάνει με την εύρεση των συντομότερων αποστάσεων και την παραγωγή ισόχρονων καμπυλών είναι το OPENROUTESERVICE. Πρόκειται για λογισμικό ανοιχτού κώδικα που έχει αναπτυχθεί σε C++ και αποτελεί ένα εργαλείο δρομολόγησης υψηλής απόδοσης σε οδικά δίκτυα.

Συνδυάζει τους εξελιγμένους αλγορίθμους δρομολόγησης με τα ανοικτά και δωρεάν δεδομένα οδικού δικτύου του έργου OpenStreetMap. Ο υπολογισμός συντομότερης διαδρομής σε δίκτυο ηπειρωτικών διαστάσεων μπορεί να διαρκέσει μερικά δευτερόλεπτα. Το OPENROUTESERVICE χρησιμοποιεί μια εφαρμογή των ιεραρχιών συστολής (τεχνική επιτάχυνσης υπολογισμού της συντομότερης διαδρομής) και είναι σε θέση να υπολογίσει και να εξάγει μια συντομότερη διαδρομή μεταξύ οποιασδήποτε προέλευσης και προορισμού μέσα σε λίγα χιλιοστά του δευτερολέπτου, οπότε ο καθαρός υπολογισμός της διαδρομής απαιτεί πολύ λιγότερο χρόνο. Οι περισσότερες προσπάθειες δαπανώνται για την σχολιασμό της διαδρομής και τη μετάδοση της γεωμετρίας μέσω του δικτύου.

4.3.3 Μεθοδολογία

Η αρχική κατάσταση παρουσιάζεται παρακάτω με τα ψηφιοποιημένα δεδομένα να απεικονίζονται πάνω στο χάρτη από το OpenStreetMaps.



Εικόνα 34 - Εφαρμογή Χωροθέτησης - Αρχική κατάσταση

Το επόμενο βήμα έχει να κάνει με τον εντοπισμό των περιοχών συγκέντρωσης των σημείων ενδιαφέροντος. Αυτό γίνεται εφικτό με τη δημιουργία ενός χάρτη έντασης. Αυτό αφορά την δημιουργία των ζωνών επιρροής του καθενός σημείου ενδιαφέροντος, το άθροισμα αυτών των περιοχών και την απεικόνισή τους σε μια συνεχή χρωματική παλέτα:



Εικόνα 35 - Εφαρμογή Χωροθέτησης - Χάρτης έντασης

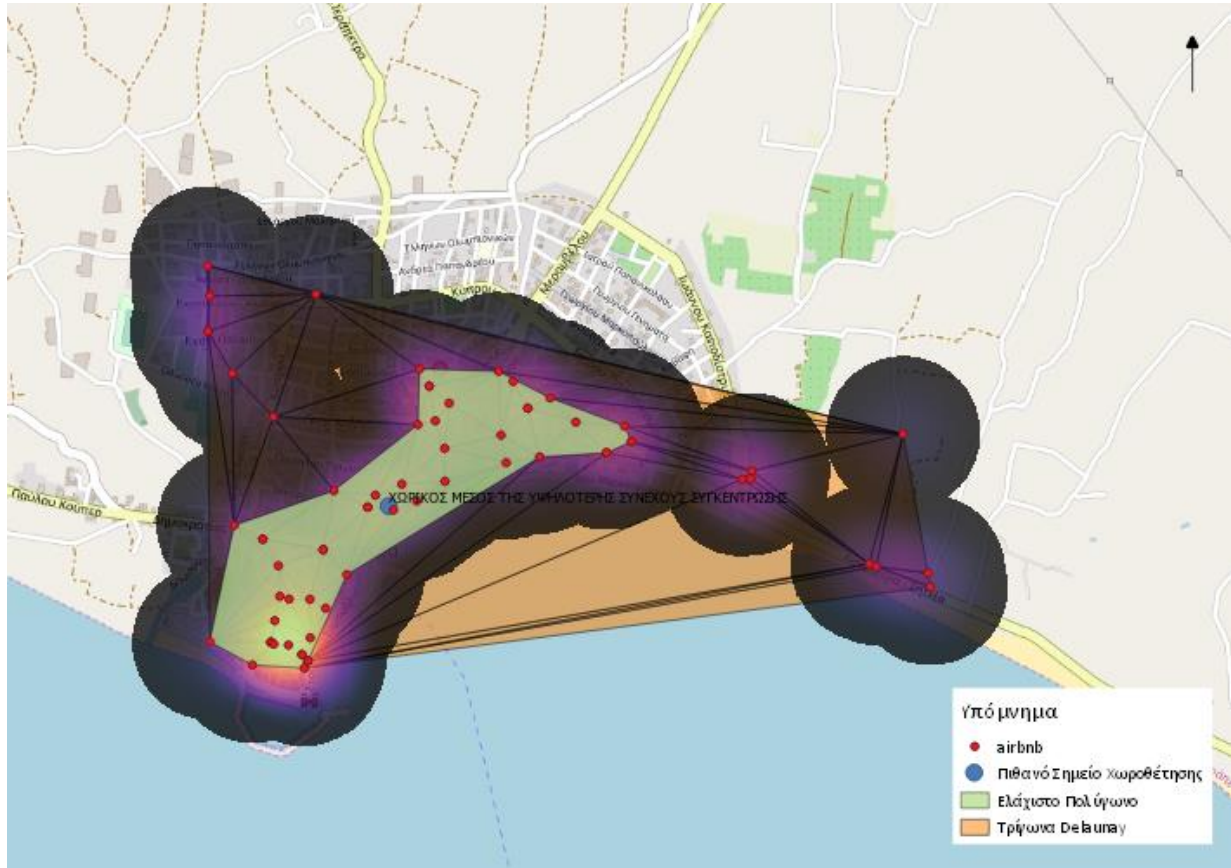
Με τα σημεία συγκέντρωσης να φαίνονται ξεκάθαρα χρειάζεται να αποφασίσουμε τα πιθανά σημεία χωροθέτησης τα οποία θα εξετάσουμε με βάση το οδικό δίκτυο. Θα πρέπει σίγουρα να δοκιμαστούν ο χωρικός μέσος όλων των σημείων ενδιαφέροντος καθώς και το σημείο υψηλότερης συγκέντρωσης. Θα ελεγχθεί επίσης και σημείο το οποίο να αποτελεί το χωρικό μέσο όλων εκείνων των σημείων που περιγράφουν την υψηλότερη συγκέντρωση σημείων ενδιαφέροντος ανά εμβαδό.

Για τον υπολογισμό του γενικού χωρικού μέσου θα χρησιμοποιηθεί το κεντροειδές του ελάχιστου πολυγώνου που περιέχει τα σημεία ενδιαφέροντος:



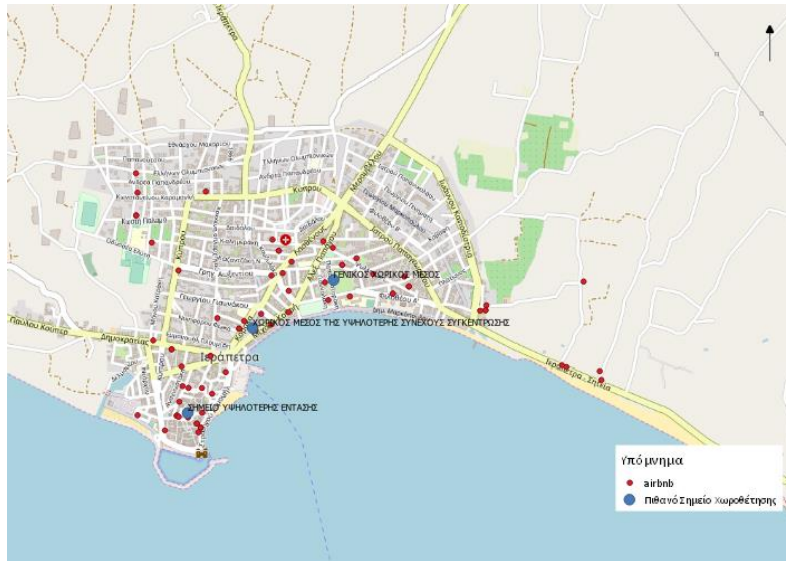
Εικόνα 36 - Εφαρμογή Χωροθέτησης - Υπολογισμός Χωρικού Μέσου

Για τον υπολογισμό του χωρικού μέσου του πολυγώνου με την υψηλότερη συγκέντρωση σημείων ενδιαφέροντος θα χρησιμοποιηθεί το κεντροειδές του πολυγώνου που προκύπτει από τα τρίγωνα Delaunay στις περιοχές της υψηλότερης έντασης (η οποία παρατηρείται ότι περιέχει περίπου το 75% των σημείων ενδιαφέροντος):



Εικόνα 37 - Εφαρμογή Χωροθέτησης - Υπολογισμός Χωρικού Μέσου της περιοχής υψηλής έντασης

Άρα τα τρία σημεία που είναι υποψήφια είναι τα παρακάτω:



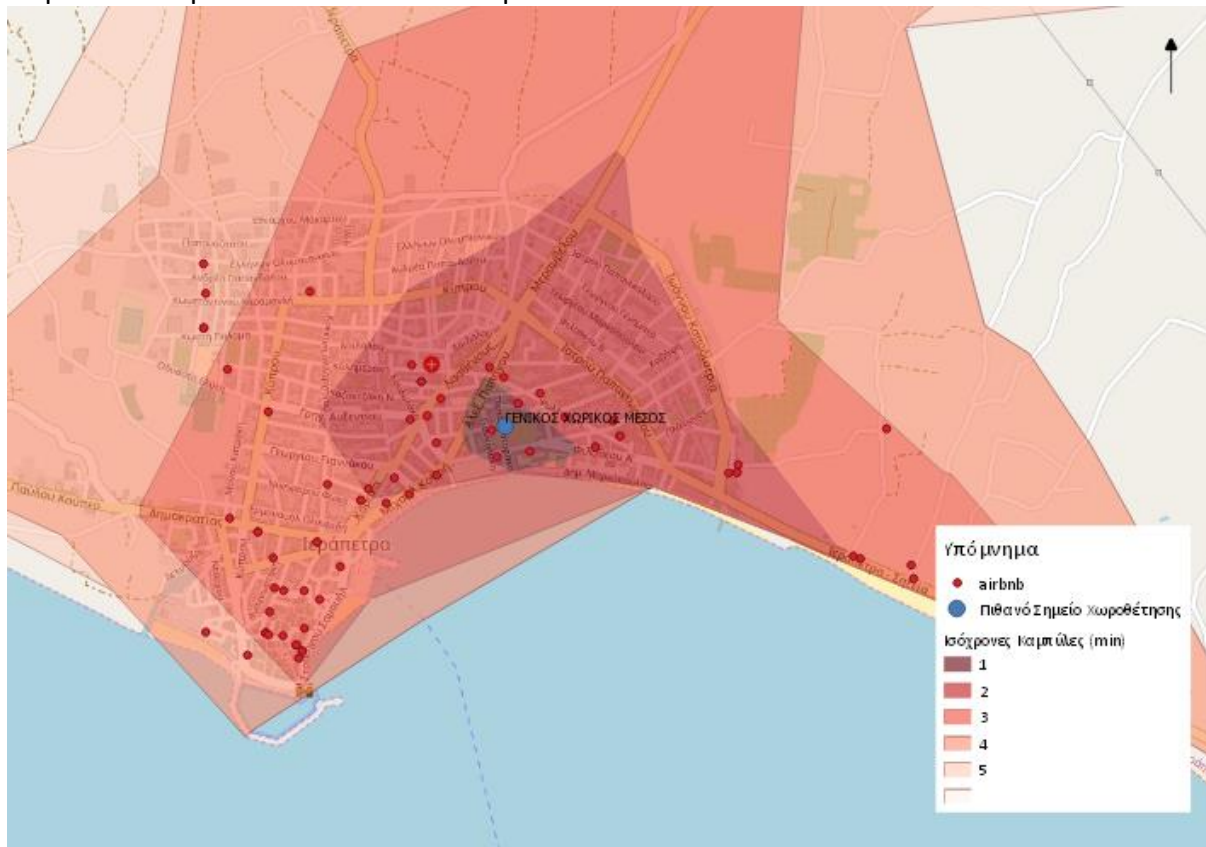
Εικόνα 38 - Εφαρμογή Χωροθέτησης - Πιθανά Σημεία Χωροθέτησης

Έπειτα για αυτά τα τρία σημεία παράχθηκαν οι ισόχρονες καμπύλες για την οδήγηση με αμάξι ώστε να επιλεγεί η ιδανικότερη από τις τρεις τοποθεσία. Με τον όρο ισόχρονες καμπύλες νοούνται οι επιφάνειες εκείνες στις οποίες ο χρόνος μετάβασης από το σημείο αφετηρίας είναι ο ίδιος. Ο τρόπος που υλοποιείται είναι η επίλυση πολλαπλών προβλημάτων δρομολόγησης ακτινωτά από το σημείο αφετηρίας, η αξιολόγησή

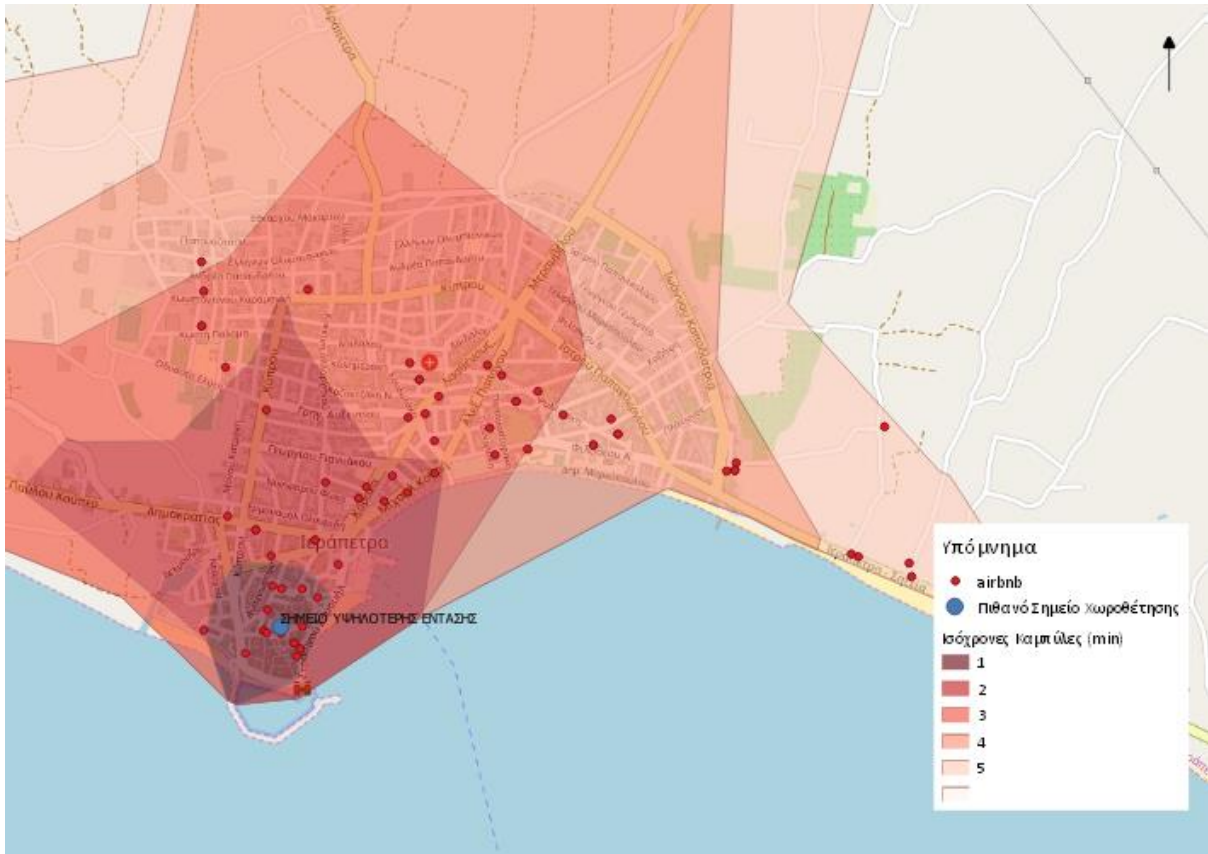
τους χρονικά και τέλος η παραγωγή των ελάχιστων πολυγώνων ανά χρόνο που αποτελούν τις ισόχρονες καμπύλες.

Αυτή η τεχνική προκρίνεται για προβλήματα χωροθέτησης στα οποία ενδιαφέρει η μετακίνηση από και προς το (προς χωροθέτηση) σημείο καθώς παράγει ζώνες επιρροής με βάρος το χρόνο ή την απόσταση που όταν εμπλέκονται δίκτυα μεταφοράς δεν είναι ευθεία.

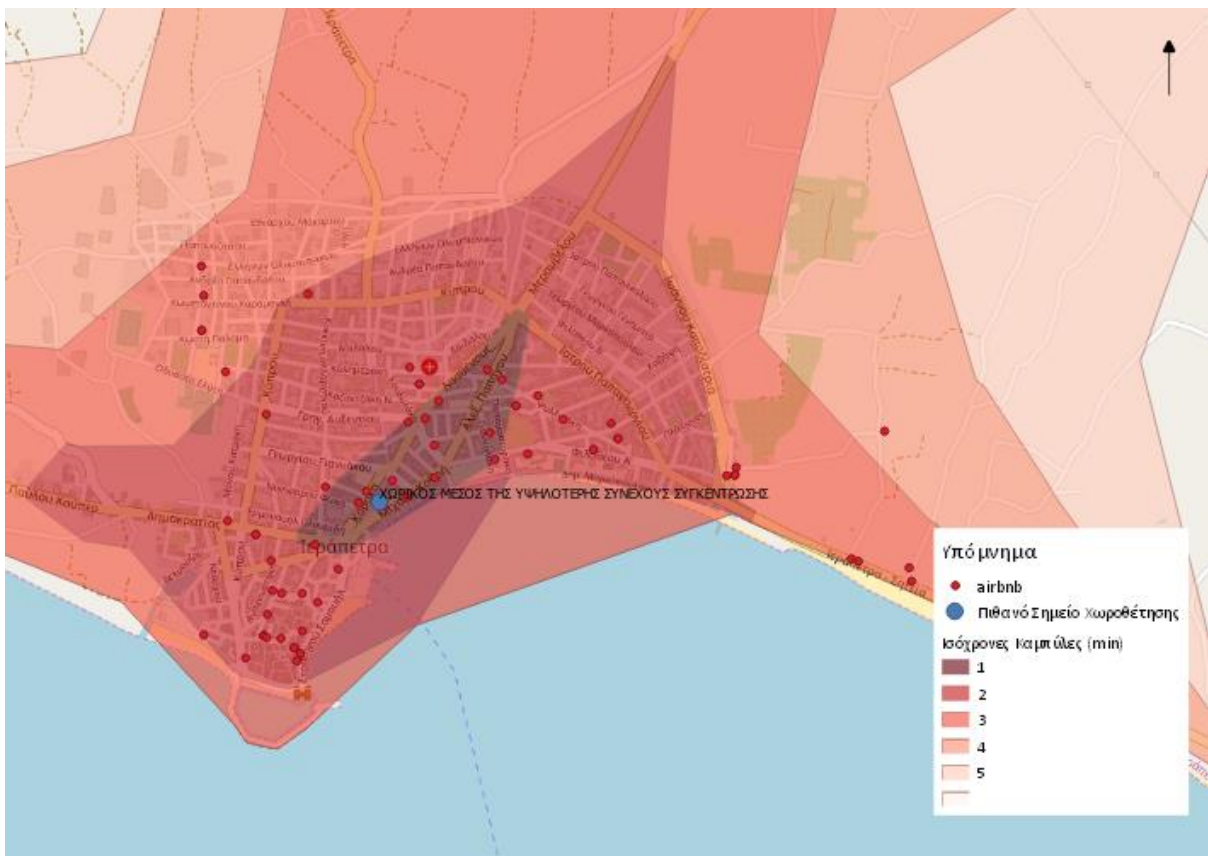
Παρακάτω παρατίθενται τα αποτελέσματα:



Εικόνα 39 - Εφαρμογή Χωροθέτησης - Ισόχρονες Καμπύλες - Περίπτωση 1



Εικόνα 40 - Εφαρμογή Χωροθέτησης - Ισόχρονες Καμπύλες - Περίπτωση 2



Εικόνα 41 - Εφαρμογή Χωροθέτησης - Ισόχρονες Καμπύλες - Περίπτωση 3

4.3.4 Συμπεράσματα

Από την κατανομή των σημείων ενδιαφέροντος στις ισόχρονες καμπύλες της κάθε περίπτωσης που εξετάστηκε προκύπτουν οι παρακάτω πίνακες:

Καμπύλες	1min	2min	3min	4min	5min
Περίπτωσης					
1. Γενικός Χωρικός Μέσος	4	28	60	71	71
2. Σημείο Υψηλότερης Έντασης	18	35	54	66	71
3.Χωρικός Μέσος 75%	17	56	70	71	71

Πίνακας 14 - Κατανομή σημείων ενδιαφέροντος στις ισόχρονες καμπύλες κάθε περίπτωσης σε απόλυτους αριθμούς.

Καμπύλες	1min	2min	3min	4min	5min
Περίπτωσης					
1. Γενικός Χωρικός Μέσος	5.63	39.43	84.51	100	100
2. Σημείο Υψηλότερης Έντασης	25.35	42.30	76.06	92,95	100
3.Χωρικός Μέσος 75%	23.94	78.87	98,59	100	100

Πίνακας 15 - Κατανομή σημείων ενδιαφέροντος στις ισόχρονες καμπύλες κάθε περίπτωσης σε ποσοστά.

Στη περίπτωση του Γενικού Χωρικού Μέσου παρατηρείται ότι τα σημεία ενδιαφέροντος κοντά σε αυτόν είναι λίγα, που σημαίνει ότι δεν ακολουθούν κανονική κατανομή. Είναι εμφανές από τους χάρτες ότι συγκεντρώνονται στη δυτική μεριά της περιοχής μελέτης. Γι' αυτό και τα αποτελέσματα στις κατηγορίες 1min & 2min είναι τα χειρότερα. Είναι επίσης απολύτως λογικό (αφού πρόκειται για το χωρικό μέσο) όσο προχωρά στις μεγαλύτερες κατηγορίες να βελτιώνεται σε σχέση με άλλες περιπτώσεις.

Στην περίπτωση του Σημείου Υψηλότερης Έντασης παρατηρείται το καλύτερο αποτέλεσμα στην μικρότερη κατηγορία - αυτή του 1min - το οποίο και είναι εντελώς αναμενόμενο. Σε όλες τις υπόλοιπες κατηγορίες υστερεί έναντι των άλλων περιπτώσεων καθώς η υψηλότερη συγκέντρωση απέχει πολύ από το Γενικό Χωρικό Μέσο και τείνει να είναι χωρικά ακραία τιμή.

Η περίπτωση Χωρικός Μέσος 75% πετυχαίνει τα καλύτερα αποτελέσματα το οποίο είναι εντελώς αναμενόμενο. Η περίπτωση αυτή αποτελεί χωρικά μια μέση κατάσταση μεταξύ του Σημείου Υψηλότερης Έντασης και του Γενικού Χωρικού Μέσου και, δεδομένου του ότι το οδικό δίκτυο δεν έχει κάτι ιδιαίτερο που να τροποποιεί φοβερά τις ισόχρονες καμπύλες, αποτελεί τη σωστότερη λύση για το πρόβλημα. Εντύπωση τέλος προκαλεί το ότι ακόμη και στην μόνη κατηγορία που υστερεί έναντι του Σημείου Υψηλότερης Έντασης (1min) αυτό είναι μόνο κατά 1,46%.

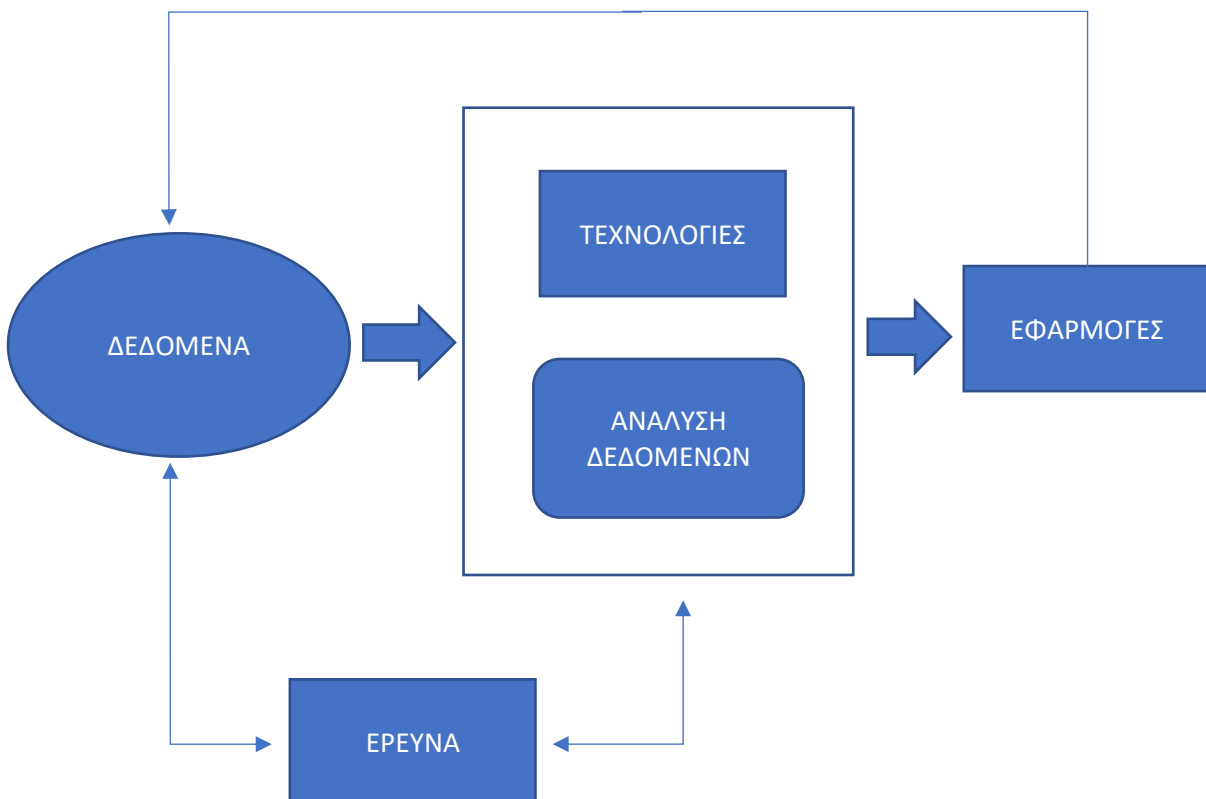
ΚΕΦΑΛΑΙΟ 5

ΣΥΜΠΕΡΑΣΜΑΤΑ

Η παρούσα εργασία ασχολήθηκε με τα χωρικά δεδομένα, τα εργαλεία χωρικής ανάλυσης, τις τεχνολογίες, τις τεχνικές και τη θεωρία ανάλυσης χωρικών δεδομένων, ξύνοντας ίσως μόνο την επιφάνειά τους. Όσο όμως περισσότερο καταπιάνεται κανείς με αυτά τόσο πιο προφανής γίνεται η σχέση μεταξύ τους και οι δυναμικές που εμφανίζονται.

Η αφετηρία της διαδικασίας δεν μπορεί παρά να είναι τα δεδομένα – άλλωστε αυτά είναι η ίδια η ζωή αποτυπωμένη σε χαρτί, φιλμ, megabytes. Η ύπαρξή τους είναι που γεννά την έρευνα, τις τεχνολογίες, την ανάγκη για ειδικούς. Και η κάθε κρίσιμη ποσοτική τους ανάπτυξη είναι που οδηγεί σε ένα ποιοτικό τεχνολογικό άλμα σε σχέση με την αποθήκευση, μεταφορά, ανάλυση και αξιοποίησή τους. Ένα τέτοιο κατώφλι έχουμε διασχίσει αυτή την περίοδο.

Η σχέση όλων των εμπλεκόμενων μερών είναι ανατροφοδοτική (είτε το καθένα ενισχύει τα υπόλοιπα είτε τα επιβραδύνει) και σε όλες αυτές τις σχέσεις συνεκτικό ιστό αποτελεί ο επιστήμονας δεδομένων.



Εικόνα 42 - Οι σχέσεις δεδομένων, ανάλυσης και τεχνολογίας

Με αυτή την αναγνώριση οδηγούμαστε αβίαστα στο συμπέρασμα ότι οι τεχνικές, τεχνολογικές και όποιες άλλες διαδικασίες δυσχεραίνουν αυτόν το αυτοτροφοδοτούμενο κύκλο οφείλουν να εγκαταλειφθούν ή/και να αντικατασταθούν με άλλες που θα λειτουργούν ακριβώς στην αντίθετη κατεύθυνση.

Το τεχνολογικό πρόβλημα είναι δεδομένο, η πληθώρα εργαλείων συλλογής, εργαλείων ανάλυσης, μορφότυπων, βάσεων και εξυπηρετητών δεδομένων δημιουργούν ένα μωσαϊκό το οποίο είναι δύσκολο να διατρέξει κανείς.

Αυτό ενσωματώνει από τη μία μεριά τον τρόπο που αναπτύχθηκαν αυτά τα εργαλεία (που χαρακτηρίζεται από τον κατακερματισμό και την εξειδίκευση των εργαλείων) και από την άλλη μια σειρά από πρακτικές που έχουν ως στόχο να εμποδίσουν τη λειτουργία αυτού του κύκλου.

Τέτοιες πρακτικές είναι:

- η μη διάθεση (δημόσιων) δεδομένων είτε από ιδιωτικές εταιρείες είτε από κρατικούς φορείς
- η χρήση τεχνικών και μεθόδων για την απόκρυψη εργαλείων ή και δεδομένων
- η χρήση ιδιότροπων/σπάνιων μορφότυπων, βάσεων και εξυπηρετητών χωρίς να υπάρχει τεχνικού τύπου αιτιολόγηση
- η μη δημοσίευση υλικού που να εξηγεί τα εργαλεία και τα δεδομένα

Κάτι τέτοιο είναι αναμενόμενο αν αναλογιστεί κανείς 1) ότι τη δεκαετία που διανύουμε θεωρείται ότι ο τομέας των δεδομένων ξεπέρασε σαν κομμάτι της παγκόσμιας οικονομίας αυτόν του πετρελαίου και 2) ότι πολλές φορές τα δεδομένα αφορούν τομείς όπως π.χ. ο στρατός. Άρα όποιος έχει ένα ειδικό ρόλο στη ροή αυτού του κύκλου έχει ένα σαφές συγκριτικό πλεονέκτημα και το αποκτά σε βάρος όλων των υπολοίπων οντοτήτων και πιθανών εφαρμογών, των δεδομένων και των εργαλείων.

Αντίβαρο στη σημερινή κυρίαρχη πρακτική μπορεί να αποτελέσει ένα πλέγμα πολιτικών και τεχνολογιών που ενσωματώνουν αυτές τις πολιτικές. Πολιτικές που θα ρυθμίζουν την πρόσβαση στα δεδομένα και καλές πρακτικές που θα ενισχύουν τα αποτελέσματα για την επιστήμη και την κοινωνία.

1. Ανοιχτά δημόσια δεδομένα

Τα ανοιχτά δημόσια δεδομένα είναι μια διαδικασία που μπορεί να οδηγήσει στον εκδημοκρατισμό πολλών όψεων την ανθρώπινης δραστηριότητας και να οριοθετήσει την ισχύ (όποιου) που έρχεται ως αποτέλεσμα της συσσώρευσης και του ελέγχου δεδομένων. Πυρήνας αυτής της διαδικασίας είναι η δημοσίευση των χωρικών δεδομένων του δημοσίου (Πολεοδομίας, Κτηματολογίου, Δασαρχείου κτλ), διαδικασία που έχει ξεκινήσει τα προηγούμενα χρόνια αλλά μένει ακόμα σε εμβρυακό στάδιο (π.χ. [epoleodomia](#)).

Δευτερεύον αλλά όχι ασήμαντο ρόλο σε αυτή τη διαδικασία έρχονται να παίξουν παραδείγματα εθελοντικής γεωγραφικής πληροφορίας όπως το OpenStreetMap ([βλ. Παράρτημα](#)).

2. Ανοιχτές τεχνολογίες

Η ύπαρξη ανοιχτών δεδομένων τροφοδοτεί την δημιουργία ανοιχτών τεχνολογιών που να τα αξιοποιούν. Όσο δε καλύτερη η ποιότητα των δεδομένων τόσο καλύτερα και τα αποτελέσματα των τεχνολογιών. Οι ανοιχτές τεχνολογίες οδηγούν στην επιτάχυνση της έρευνας και της ανάπτυξης εφαρμογών καθώς δεν απαιτείται η εκ νέου δημιουργία λογισμικού παρά μόνο για να επιλυθούν επιπλέον προβλήματα ανάλυσης.

Για παράδειγμα το λογισμικό της Γούναρης Ολύνας και το Openrouteservice, που χρησιμοποιήθηκαν στις υποθέσεις εργασίας του παρόντος, αναπτύχθηκαν πάνω σε ανοιχτές βιβλιοθήκες χωρικής ανάλυσης και αποτελούν εργαλεία που μπορούν να ενσωματωθούν σε οποιαδήποτε τελική εφαρμογή ή σε άλλο τελικό προϊόν/εργαλείο ανάλυσης.

Βασικός τόπος οργάνωσης, σχεδιασμού και υλοποίησης ανοιχτών τεχνολογιών χωρικής ανάλυσης είναι το [Open Source Geospatial Foundation](#).

3. Καλές πρακτικές οργάνωσης και διάθεσης των δεδομένων

Η οργάνωση και ταξινόμηση των υφιστάμενων δεδομένων με τρόπο που να άρει τον κατακερματισμό των δεδομένων και των τρόπων αποθήκευσής τους είναι απαραίτητη για να “καθαρίσει” την υφιστάμενη χωρική πληροφορία (που υπάρχει σε πολλά σημεία διπλότυπη, μη ενημερωμένη κτλ.) καθώς και να την καταστήσει άμεσα αξιοποιήσιμη (ασχέτως του πως υπάρχει ως τώρα). Βασικό ρόλο σε αυτή τη διαδικασία είναι η αξιοποίηση μεταδεδομένων πάνω στα υπάρχοντα δεδομένα τα οποία θα εισάγουν ένα επίπεδο αφαίρεσης αποσυνδέοντας την πληροφορία από τη μορφή στην οποία είναι παγιδευμένη και καθιστώντας την από εδώ και στο εξής άμεσα αξιοποιήσιμη.

Ακόμα και όταν τα δεδομένα είναι οργανωμένα σωστά και ανοιχτά πρέπει η υποδομή που τα διαθέτει να είναι προσπελάσιμη. Αφαιρώντας από τη συζήτηση το κομμάτι των μηχανημάτων (Hardware, Bandwidth κτλ.) θα πρέπει η πληροφορία να διατίθεται μέσω εργαλείων που να είναι διαδεδομένα π.χ. SOAP & REST υπηρεσίες ώστε να αποσυνδέονται από τις εσωτερικές εργασίες και δομές αυτού που τα παρέχει και να είναι καλά τεκμηριωμένες (documentation).

Χαρακτηριστικό παράδειγμα τέτοιας προσπάθειας αποτελεί η οδηγία INSPIRE ([βλ. Παράρτημα](#)) και το [Open Geospatial Consortium](#).

4. BLOCKCHAIN

Κομμάτι της έρευνας γύρω από τα χωρικά δεδομένα και τις τεχνολογίες οφείλει να αποτελέσει (ως υπόδειγμα τουλάχιστον) η τεχνολογία Blockchain. Ενσωματώνει τεχνικά μια σειρά από τα ζητούμενα όπως η δημόσια φύση των δεδομένων, ο συγχρονισμός τους, η κατανομημένη φύση του συστήματος.

Η τεχνολογία αυτή εμφανίστηκε με το κρυπτονόμισμα Bitcoin και επέτρεψε ασφαλείς και αμετάβλητες συναλλαγές μεταξύ των μελών (peer-to-peer) χωρίς μεσάζοντες για να εγγυηθούν την συναλλαγή (trusted authorities). Ουσιαστικά αποτελεί μια κατανεμημένη βάση δεδομένων όπου αποθηκεύεται κρυπτογραφημένη πληροφορία με τρόπο που να δημιουργεί μια διαδοχή δεδομένων. Εγγυητές μιας συναλλαγής είναι οι χρήστες, οι οποίοι αποτελούν τους κόμβους της κατανεμημένης βάσης, και άρα οι συναλλαγές είναι δημόσια προσβάσιμες. Συγκεκριμένα, το Bitcoin αποτελεί παράδειγμα δημόσιου κατάλογου Blockchain. Υπάρχει δυνατότητα και ιδιωτικού καταλόγου.

Η τεχνολογία Blockchain σήμερα βρίσκεται σε πρώιμο στάδιο. Δεν αφορά μόνο τις χρηματικές συναλλαγές, αλλά αναμένεται να τροποποιήσει την λειτουργία κάθε υπηρεσίας που διατηρεί δεδομένα (π.χ. ληξιαρχείο, κτηματολόγιο, ασφαλιστικά ταμεία, εφορία)

5. Εκπαίδευση

Τέλος, βασικό κομμάτι είναι η συζήτηση, η έρευνα μέσα στο πανεπιστήμιο και η εκπαίδευση των νέων επιστημόνων γύρω από ζητήματα ηθικής και δικαιώματος χρήσης δεδομένων και λογισμικών (licencing). Πρόκειται για κάτι που λείπει από τη ζωή του ΕΜΠ με εξαίρεση ίσως την κοινότητα ανοιχτού λογισμικού του ΕΜΠ και είναι δεδομένο ότι θα απασχολήσει τον σύγχρονο επιστήμονα δεδομένων και μηχανικό.

Η επιστημονική κοινότητα οφείλει να βγάλει αυτή τη συζήτηση και στην κοινωνία. Πρέπει να εκπαιδευτεί κόσμος γύρω από το γιατί μια εταιρεία που του παρέχει υπηρεσίες δρομολόγησης “δωρεάν” μπορεί να καταγράφει την κίνησή του μέσα στο μήνα και να μπορεί να την διαθέτει σε τρίτους. Ή το πως η παρουσία του στα social media και οι αναζητήσεις που κάνει στον ιστό καταλήγουν μέσω της ανάλυσης δεδομένων με χωρο-χρονική υπόσταση σε εξατομικευμένες διαφημίσεις, ειδήσεις, διασκέδαση με ότι αυτό μπορεί να συνεπάγεται.

Αντί επιλόγου

Το πρόβλημα της παραγωγής, της αποθήκευσης, της ανάλυσης, της διαχείρισης της χωρικής πληροφορίας και των εφαρμογών της, της δημοκρατίας της ασφάλειας και της ιδιοκτησίας των δεδομένων δεν είναι τεχνικό αλλά κοινωνικό. Είναι επιλογή μιας κοινωνίας το τι κάνει με τα δεδομένα που παράγει και αυτή η συζήτηση πρέπει να γίνει. Η απουσία αυτής της συζήτησης ενισχύει την υφιστάμενη κατάσταση. Οι ανοιχτές τεχνολογίες και τα ανοιχτά δεδομένα έχουν την δυνατότητα να ενσωματώσουν τεχνικά πιο δημοκρατικές πρακτικές με μεγαλύτερο έλεγχο και να προάγουν ένα πλαίσιο πιο δυναμικής οριζόντιας ανάπτυξης.

Κατάλογος Εικόνων

Εικόνα 1 - Παραγωγή χωρικών δεδομένων από τις τηλεπικοινωνίες (www.cellmapper.net)	8
Εικόνα 2 - Παραγωγή χωρικών δεδομένων από τη ναυσιπλοΐα (www.marinetraffic.com)	8
Εικόνα 3 - Παραγωγή χωρικών δεδομένων από τις τηλεπικοινωνίες (www.cellmapper.net)	8
Εικόνα 4 - Παραγωγή χωρικών δεδομένων από τη ναυσιπλοΐα (www.marinetraffic.com)	9
Εικόνα 5 - Διαδρομή πτήσης των δορυφόρων LANDSAT (https://earthnow.usgs.gov/observer/)	9
Εικόνα 6 - Παραγωγή χωρικών δεδομένων από τη ναυσιπλοΐα (www.marinetraffic.com).....	9
Εικόνα 7 - Παραγωγή χωρικών δεδομένων από την αεροπλοΐα (www.flightradar24.com)	9
Εικόνα 8 - Διαδρομή πτήσης των δορυφόρων LANDSAT (https://earthnow.usgs.gov/observer/)	9
Εικόνα 9 - Λήψη από βίντεο της συνεχής απεικόνισης του Landsat 7 πάνω από το Μεξικό (https://earthnow.usgs.gov/observer/)	9
Εικόνα 10 - Ιστορικό διαδρομών σε παραλίες και καταστήματα κατανάλωσης στα Χανιά, από χρήση της Google.....	10
Εικόνα 11 - Επεξεργασία στο OpenStreetMap του πολυγώνου που απεικονίζει το 3ο Γυμνάσιο Ιεράπετρας.....	10
Εικόνα 12 - Παράδειγμα απεικόνισης γεωγραφικών δεδομένων σύμφωνα με την κανονικοποιημένη δομή.....	17
Εικόνα 13 - Η διανυσματική εκδοχή της Εικόνας 8	18
Εικόνα 14 - Παράδειγμα τοπολογικής δομής.....	19
Εικόνα 15 - Δημιουργία τριγωνισμού Delaunay από σημειακά δεδομένα με πληροφορία υψομέτρου κα σχηματισμός ΔΑΤ από υψομετρικά δεδομένα.....	20
Εικόνα 16 - Τα πολύγωνα Thiessen που προκύπτουν από το δίκτυο σημείων της Εικόνας 11.....	21
Εικόνα 17 - Ιστογράμματα με τις πιο αγαπητές, μισητές και "ποθητές" βάσεις, βάσει του ερωτηματολογίου του Stack Overflow για το2019	28
Εικόνα 18- Η Χωρική Ανάλυση και Χωρική στατιστική Beyond Mapping III - innovativegis.....	31
Εικόνα 19 - Αρχικός χάρτης από τον John Snow που δείχνει τις συστάδες περιπτώσεων χολέρας στην επιδημία του Λονδίνου του 1854, όπως σχεδιάστηκε και λιθογραφήθηκε από τον Charles Cheffins.	32
Εικόνα 20 - Συστατικά στοιχεία ενός ΣΓΠ	44
Εικόνα 21 - Ιστόγραμμα Δημοφιλίας λογισμικών ΓΣΠ από το G2 Crowd (ιστότοπος αξιολόγησης λογισμικών από το κοινό).....	46
Εικόνα 22 - Διάγραμμα διασποράς δημοφιλίας GIS λογισμικών από το G2 Crowd (ιστότοπος αξιολόγησης λογισμικών από το κοινό)	47
Εικόνα 23 - Διαστημικό Πρόγραμμα LANDSAT (https://landsat.gsfc.nasa.gov)	48
Εικόνα 24 - Φασματική απόκριση πέντε βασικών επιφανειακών γήινων στοιχείων (www.researchgate.net).....	51
Εικόνα 25 - Περιοχή μελέτης	55
Εικόνα 26 - Αποτελέσματα Ταξινόμησης 1985.....	59
Εικόνα 27 - Αποτελέσματα Ταξινόμησης 1990.....	60
Εικόνα 28 - Αποτελέσματα Ταξινόμησης 1995.....	60
Εικόνα 29 - Αποτελέσματα Ταξινόμησης 2000.....	61
Εικόνα 30 - Αποτελέσματα Ταξινόμησης 2005.....	61
Εικόνα 31 - Αποτελέσματα Ταξινόμησης 2010.....	62
Εικόνα 32 - Αποτελέσματα Ταξινόμησης 2015.....	62
Εικόνα 33 - Αποτελέσματα Ταξινόμησης 2018.....	63
Εικόνα 34 - Εφαρμογή Χωροθέτησης - Αρχική κατάσταση	72
Εικόνα 35 - Εφαρμογή Χωροθέτησης - Χάρτης έντασης	73
Εικόνα 36 - Εφαρμογή Χωροθέτησης - Υπολογισμός Χωρικού Μέσου	73

Εικόνα 37 - Εφαρμογή Χωροθέτησης - Υπολογισμός Χωρικού Μέσου της περιοχής υψηλής έντασης	74
Εικόνα 38 - Εφαρμογή Χωροθέτησης - Πιθανά Σημεία Χωροθέτησης	75
Εικόνα 39 - Εφαρμογή Χωροθέτησης - Ισόχρονες Καμπύλες - Περίπτωση 1.....	75
Εικόνα 40 - Εφαρμογή Χωροθέτησης - Ισόχρονες Καμπύλες - Περίπτωση 2.....	76
Εικόνα 41 - Εφαρμογή Χωροθέτησης - Ισόχρονες Καμπύλες - Περίπτωση 3.....	76
Εικόνα 42 - Οι σχέσεις δεδομένων, ανάλυσης και τεχνολογίας	78

Κατάλογος Πινάκων

Πίνακας 1 - Κατηγοριοποίηση τεχνικών συλλογής σε σχέση με τα χαρακτηριστικά των συλλεγμένων γεωγραφικών δεδομένων.	11
Πίνακας 2 - Παραδείγματα εξωτερικών πηγών γεωγραφικών δεδομένων στις ΗΠΑ και στην Ευρώπη	13
Πίνακας 3 - Κανάλια δορυφόρου Landsat 5 (www.usgs.gov)	49
Πίνακας 4 - Κανάλια δορυφόρου Landsat 8 (www.usgs.gov)	51
Πίνακας 5 - Ακρίβεια Ταξινόμησης	64
Πίνακας 6 - Πίνακας Σύγκυσης 1985	64
Πίνακας 7 - Πίνακας Σύγκυσης 1990	65
Πίνακας 8 - Πίνακας Σύγκυσης 1995	65
Πίνακας 9 - Πίνακας Σύγκυσης 2000	66
Πίνακας 10 - Πίνακας Σύγκυσης 2005	66
Πίνακας 11 - Πίνακας Σύγκυσης 2010	67
Πίνακας 12 - Πίνακας Σύγκυσης 2015	67
Πίνακας 13 - Πίνακας Σύγκυσης 2018	68
Πίνακας 14 - Κατανομή σημείων ενδιαφέροντος στις ισόχρονες καμπύλες κάθε περίπτωσης σε απόλυτους αριθμούς.	77
Πίνακας 15 - Κατανομή σημείων ενδιαφέροντος στις ισόχρονες καμπύλες κάθε περίπτωσης σε ποσοστά.	77

Βιβλιογραφία

- Anselin, L. (n.d.). SPATIAL DATA ANALYSIS WITH GIS: AN INTRODUCTION TO APPLICATION IN THE SOCIALSCIENCES. Santa Barbara, California, USA: University of California.
- Bounie, D., & Gille, L. (2012). International Production and Dissemination of Information: Results, Methodological Issues, and Statistical Perspectives. *SSRN Electronic Journal*.
- Domo. (2018). *Data Never Sleeps*.
- Fischer, M. M. (2015). International Encyclopedia of the Social & Behavioral Sciences (Second Edition). Elsevier.
- G2. (2019). Ανάκτηση από g2.com
- GISGeography. (2019). Ανάκτηση από gisgeography.com/gis-formats/
- Hilbert, M., & Lopez, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information.
- Lyman, P., & Varian, H. R. (2000). How Much Information. San Fransisco, California, United States of America: Berkeley.
- OpenStreetMap Wiki. (2019). Ανάκτηση από wiki.openstreetmap.org/wiki/Main_Page
- Rider, F. (1944). The Scholar and the Future of the Research Library: A Problem and its Solution.
- Siegel, E. (2013). *Predictive Analytics: The Power to Predict who Will Click Buy Lie Or Die*. John Wiley & Sons.
- Stack Overflow. (2019). Ανάκτηση από insights.stackoverflow.com/survey/2019
- Γούναρη, Ο. (2018, Οκτώβριος). Χαρτογράφηση Κάλυψης Γης και Καλλιεργειών με Τεχνικές. Αθήνα, Ελλάδα: ΕΜΠ.
- ΚΑΒΟΥΡΑΣ, Μ., ΔΑΡΡΑ, Α., ΚΟΚΛΑ, Μ., ΚΟΝΤΑΞΑΚΗ, Σ., ΠΑΝΟΠΟΥΛΟΣ, Γ., & ΤΟΜΑΗ, Ε. (2015). *Επιστήμη Γεωγραφικής Πληροφορίας - Ολοκληρωμένη Προσέγγιση και Ειδικά Θέματα*. Αθήνα: ΣΥΝΔΕΣΜΟΣ ΕΛΛΗΝΙΚΩΝ ΑΚΑΔΗΜΑΪΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ.
- Κάβουρας, Μ., Δάρρα, Α., Κονταξάκη, Σ., & Τομαή, Ε. (2015). *Επιστήμη Γεωγραφικής Πληροφορίας - Αρχές και Τεχνολογίες*. Αθήνα: ΣΥΝΔΕΣΜΟΣ ΕΛΛΗΝΙΚΩΝ ΑΚΑΔΗΜΑΪΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ.
- Περάκης, Κ. Γ. (2015). Η Τηλεπισκόπηση σε 13 Ενότητες- Θεωρία, Μέθοδοι και Εφαρμογές. Ελλάδα: ΣΥΝΔΕΣΜΟΣ ΕΛΛΗΝΙΚΩΝ ΑΚΑΔΗΜΑΪΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ.

Παράρτημα

OpenStreetMap

Το OpenStreetMap αποτελεί ένα από τα πιο χαρακτηριστικά παραδείγματα εθελοντικής γεωγραφικής πληροφορίας. Δημιουργήθηκε το 2004 από τον Steve Cost για την περιοχή του Ηνωμένου Βασιλείου. Από το 2006 υποστηρίζεται από το OpenStreetMap Foundation (OSFM) που είναι ένας διεθνής μη κερδοσκοπικός οργανισμός με έδρα στη Μεγάλη Βρετανία. Παρατηρείται μια συνεχώς αυξανόμενη τάση συμμετοχής του κοινού στην πλατφόρμα και εξέλιξης της πλατφόρμας. Αναλυτικά δεδομένα και γραφήματα μπορεί να ανατρέξει κανείς στο OpenStreetMapStatistics. Στο <http://www.geofabrik.de/gallery/history/#germany> μπορεί να δει κανείς τον εμπλουτισμό των χαρτών για διάφορες χώρες στη μορφή βίντεο. Τα δεδομένα διατίθεται δωρεάν στο κοινό με τη άδεια Open Data Commons Open Database License (ODbL) του OSMF. Βασικότερο χαρακτηριστικό της εθελοντικής συνεισφοράς δεδομένων στο OpenStreetMap είναι το γεγονός ότι τα δεδομένα είναι ελεύθερα και δωρεάν διαθέσιμα και επεξεργάσιμα. Ο ρόλος των χρηστών είναι πολλαπλός μιας και μπορούν: α) να συνεισφέρουν χωρικά δεδομένα π.χ. με απλές συσκευές GPS χαρτογραφούν μια δεδομένη περιοχή, β) να χρησιμοποιούν δεδομένα που έχουν παραχθεί με κάποιο ΕΓΠ τρόπο π.χ. ανατρέχοντας στο OpenStreetMap για την εξεύρεση κάποιου σημείου ενδιαφέροντος και γ) να δημιουργούν επιπρόσθετη αξία στα δεδομένα με τη δημιουργία εφαρμογών επί αυτών. Η δυναμική της εθελοντικής γεωγραφικής πληροφορίας γίνεται άμεσα αισθητή μέσα από το παράδειγμα της Αϊτής στο OpenStreetMap. Μέσα σε πολύ μικρό χρονικό διάστημα περιοχές που δεν απεικονίζονται στο OpenStreetMap, εμπλουτίστηκαν με πληροφορία. Βασική αρχή στην επιτυχημένη συλλογή πληροφοριών από το κοινό είναι η παροχή μιας πλατφόρμας που είναι ευπρόσιτη και εύκολα κατανοητή στο χρήστη και επίσης του επιτρέπει να διατηρεί την προσωπική του ελευθερία και ευελιξία ως προς τον τρόπο διάθεσης και περιγραφής των δεδομένων του. Στο OpenStreetMap η συνθήκη αυτή τηρείται με την παροχή απλών συντακτών περιεχομένου οι πιο συνηθισμένοι εκ των οποίων είναι οι iD, Potlatch και JOSM για την προσθήκη υλικού. Ανάλογα τη συσκευή (προσωπικός υπολογιστής, έξυπνη συσκευή κτλ.) και το λογισμικό διατίθεται πλήθος συντακτών (μια συγκεντρωτική λίστα μπορεί να ανατρέξει κανείς στο <https://wiki.openstreetmap.org/wiki/Editors>). Το χωρικό μοντέλο του OpenStreetMap είναι ιδιαίτερα απλό και επιτρέπει στο χρήστη να ψηφιοποιεί σημειακά, γραμμικά και επιφανειακά στοιχεία. Ακόμα, ο χρήστης μπορεί να διαθέσει τροχιές GPS. Τα στοιχεία αυτά επισημειώνονται με ετικέτες που τους προσθέτουν σημασιολογική πληροφορία. Οι ετικέτες αυτές έχουν τη μορφή κλειδί=τιμή (key=value). Για παράδειγμα highway=residential, name=Syntagma Square. Οι ετικέτες αυτές συγκεντρώνονται σε λίστες, οι οποίες έχουν προκύψει από το κοινό και είναι δυναμικές, προσαρμόζονται δηλαδή στις εκάστοτε ανάγκες χαρτογράφησης. Οι νέες ετικέτες προς ενσωμάτωση καταχωρούνται σε λίστες μέχρι την αποδοχή τους από την κοινότητα. Ομοίως και για τις ετικέτες που έχουν απορριφθεί από την κοινότητα. Στην περιγραφή κάθε ετικέτας περιλαμβάνονται: η ονομασία του κλειδιού, η τιμή του, τι είδους χωρικό στοιχείο είναι, σχόλιο-μικρή περιγραφή, τρόπος που απεικονίζεται στο χάρτη και φωτογραφία. Επιπλέον συμβουλές για τη σωστή απεικόνιση των στοιχείων και την επιλογή των κατάλληλων συμβόλων και ετικετών παρέχονται στο «How to map». Ειδικά για τη σημασία και την ορθή χρήση των ετικετών ο χρήστης μπορεί να ανατρέξει στο Taginfo. Η επικοινωνία μεταξύ των εθελοντών πραγματοποιείται μέσω λιστών ταχυδρομείου, λίστα talk, irc chat κανάλι, forum, wiki. Ζητήματα σημασιολογίας και διαλειτουργικότητας είναι ιδιαίτερα περίπλοκα όταν γίνεται αναφορά σε εθελοντική γεωγραφική πληροφορία. Αυτό προκύπτει κυρίως από το γεγονός ότι η πληροφορία που παρέχεται είναι άναρχα

δομημένη και δεν ακολουθεί κάποια συγκεκριμένα πρότυπα. Επίσης, το γεγονός ότι δεν είναι γνωστό το υπόβαθρο των συντελεστών και συνεπώς ο τρόπος που αντιλαμβάνονται τον κόσμο περιπλέκει ακόμα περισσότερο το θέμα. Συνεπώς, η σημασιολογική ετερογένεια (semantic heterogeneity) (Oukseil & Sheth 1999, Brodeur 2012) και η ευέλικτη δομή (Anand κ.ά. 2010, Cross & Hu 2011) αποτελούν τροχοπέδη στη ενοποίηση των δεδομένων από πολλές συλλογές μιας και σε κάθε συλλογή τα ίδια φαινόμενα μπορεί να αντιμετωπίζονται εννοιολογικά διαφορετικά. Ο Scheider και οι συνεργάτες του (2011) εισήγαγε μια μεθοδολογία για την αγκίστρωση (grounding) της σημασιολογικής πληροφορίας του OpenStreetMap. Με τον όρο σημασιολογική πληροφορία εννοείται το σύνολο των ετικετών που χρησιμοποιούνται για να περιγράψει ο γεωγραφικός χώρος καθώς και ο τρόπος που κατανοούνται και χρησιμοποιούνται από τους χρήστες. Σύμφωνα με τους συγγραφείς, δυσκολίες εμφανίζονται όταν π.χ. χρησιμοποιείται η λέξη Cafe τόσο για να επισημειώσει καταστήματα καφέ στη Ν. Υόρκη, όσο και για να επισημειώσει καφετέριες Kaffeehauser στη Βιέννη. Ομοίως πολλές καφετέριες σερβίρουν και αλκοόλ τις βραδινές ώρες που τις καθιστά συνώνυμες με μπαρ. Η έννοια της ασάφειας της σημασιολογίας της εθελοντικής γεωγραφικής πληροφορίας αποτελεί αντικείμενο μελέτης και των Ballatore και Bertolotto. Στη μελέτη τους (Ballatore & Bertolotto 2011) χαρακτηριστικά αναφέρουν ότι η πλούσια συνεισφορά των χρηστών έχει οδηγήσει σε έναν πολύ μεγάλο όγκο σημασιολογικά αδόμητων εννοιών. Οι έννοιες αυτές σε αντίθεση με οργανωμένες δομές όπως οι οντολογίες και οι ταξινομίες είναι ασύνδετες και όχι σαφώς ορισμένες. Το αποτέλεσμα είναι να εμφανίζονται σημασιολογικά κενά ανάμεσα στις έννοιες των χρηστών και στις σαφώς ορισμένες έννοιες των οντολογιών που δυσχεραίνουν τη σημασιολογική διαλειτουργικότητα.

Ευρωπαϊκή Χωρική Υποδομή - Οδηγία INSPIRE

Η πολιτική της Ευρωπαϊκής Ένωσης (ΕΕ), σύμφωνα με το 6ο Πρόγραμμα Δράσης για το Περιβάλλον, ορίζει ότι θα πρέπει να εφαρμόζονται πολιτικές, οι οποίες περιορίζουν τη διπλή συλλογή των ιδίων δεδομένων και στηρίζουν και προωθούν την εναρμόνιση, ευρεία διάδοση και χρήση των πληροφοριών για το γεωγραφικό χώρο. Τέτοιες πολιτικές αναμένεται να οδηγήσουν σε μεγαλύτερη αποτελεσματικότητα, και τα οφέλη που θα προκύψουν μπορούν να αξιοποιηθούν για τη βελτίωση της διαθεσιμότητας και της ποιότητας των πληροφοριών. Η χωρική πληροφορία μπορεί να διαδραματίσει συγκεκριμένο ρόλο στο πλαίσιο της νέας αυτής προσέγγισης, δεδομένου ότι επιτρέπει το συνδυασμό πληροφοριών από διάφορους επιστημονικούς κλάδους, και για ποικίλες χρήσεις. Μια συνεκτική και ευρέως προσπελάσιμη χωρική περιγραφή όλης της γεωγραφικής έκτασης ενός κράτους-μέλους ή ολόκληρης της ΕΕ θα συμβάλει στη διαμόρφωση του απαραίτητου πλαισίου για το σχεδιασμό περιβαλλοντικής πολιτικής. Για όλους αυτούς τους λόγους, η επιτροπή αποφάσισε να υποβάλει την οδηγία INSPIRE στο Ευρωπαϊκό Κοινοβούλιο και στο Συμβούλιο της Ευρωπαϊκής Ένωσης, ώστε η χωρική πληροφορία να καταστεί άμεσα διαθέσιμη και αξιοποιήσιμη σε επίπεδο τόσο εθνικής όσο και κοινοτικής πολιτικής, και να επιτραπεί η πρόσβαση του κοινού στην πληροφορία αυτή. Η οδηγία διαμορφώνει νομικό πλαίσιο για τη δημιουργία και λειτουργία SDI στην Ευρώπη, με σκοπό τη χάραξη, εφαρμογή, παρακολούθηση και αξιολόγηση των κοινοτικών πολιτικών, σε όλα τα επίπεδα, και την παροχή πληροφοριών του δημοσίου τομέα. Η οδηγία INSPIRE δεν δρομολογεί ένα εκτεταμένο πρόγραμμα συλλογής νέων χωρικών δεδομένων στα κράτη μέλη. Ωστόσο, θεωρείται πως το κάθε κράτος μέλος θα αναπτύξει τη δική του SDI.

Στόχος της Ευρωπαϊκής Χωρικής Υποδομής INSPIRE Οι πρόσφατες εξελίξεις στην ανάπτυξη ψηφιακών δεδομένων έχουν δημιουργήσει άνευ προηγουμένου δυνατότητες για την ελεύθερη πρόσβαση και χρήση πληροφοριών. Ωστόσο τα δεδομένα είναι συχνά χαμηλής ή απροσδιόριστης ποιότητας, εξαρτώμενα από συγκεκριμένα ΣΓΠ και μη προσπελάσιμα από ιδιώτες ή χρήστες άλλων επιπέδων. Οι εργασίες ομογενοποίησης δεδομένων προερχόμενα από πολλαπλές πηγές ώστε να συνταχθούν πληροφορίες ικανές για τη διατύπωση αποφάσεων πολιτικής είναι συνήθως πολύπλοκες, χρονοβόρες και υψηλού κόστους. Συνεπώς υπάρχει η ανάγκη διατύπωσης επακριβών αρχών, οι οποίες θα ελαττώσουν τη διπλή εμφάνιση δεδομένων, θα βελτιώσουν την ποιότητά τους και θα προάγουν τη διαθεσιμότητά τους. Στόχος του INSPIRE είναι η σταδιακή εναρμόνιση των υποδομών χωρικών δεδομένων των κρατών μελών σε μία ενιαία ευρωπαϊκή υποδομή. Έχει σχεδιαστεί για να βελτιστοποιήσει τις δυνατότητες αξιοποίησης των δεδομένων που διατίθενται ήδη, μέσω της τεκμηρίωσης τους, της λειτουργίας υπηρεσιών που αποσκοπούν στη διευκόλυνση της πρόσβασης σε αυτά, στην αύξηση της διαλειτουργικότητά τους, και της αντιμετώπισης των δυσκολιών στις οποίες προσκρούει η χρήση τους. Η οδηγία INSPIRE σκοπεύει να βελτιώσει την τρέχουσα κατάσταση με την ίδρυση μιας ESDI, η οποία θα διασφαλίζει την πρόσβαση και χρήση χωρικών πληροφοριών βασιζόμενη στις παρακάτω αρχές:

- Τα δεδομένα θα συλλέγονται μια μόνο φορά και θα συντηρούνται στο πλέον κατάλληλο επίπεδο και από τον πλέον κατάλληλο φορέα.
- Εναρμόνιση χωρικών πληροφοριών από διάφορες πηγές σε όλη την Ευρώπη διάθεσή στους προβλεπόμενους χρήστες και εφαρμογές.
- Η πληροφορία θα πρέπει να διατίθεται σε πολλαπλά επίπεδα λεπτομέρειας.
- Η γεωγραφική πληροφορία η οποία κρίνεται απαραίτητη για ομαλή διακυβέρνηση θα πρέπει να είναι αδρή και ευρέως διαθέσιμη.
- Εύκολη εύρεση της διαθέσιμης γεωγραφικής πληροφορίας, εάν πληροί τις ανάγκες του κάθε χρήστη και υπό ποιες συνθήκες μπορεί να προσπελαστεί.

- Η γεωγραφική πληροφορία θα πρέπει να οπτικοποιείται με φιλικούς προς το χρήστη τρόπους, μέσω περιβάλλοντος που κάνει κατανοητή τη φύση της.

Σκοπός των ανωτέρω είναι να γίνει δυνατή η πρόσβαση σε ενιαία και υψηλής ποιότητας Γεωγραφική πληροφορία, σε τοπικό, περιφερειακό, εθνικό και διεθνές επίπεδο, για τη διαμόρφωση, υλοποίηση, επιτήρηση και αξιολόγηση Εθνικών και Ευρωπαϊκών πολιτικών. Στόχος λοιπόν της οδηγίας INSPIRE είναι η σταδιακή εναρμόνιση των υποδομών χωρικών δεδομένων των κρατών μελών σε μία ενιαία ευρωπαϊκή υποδομή. Έχει σχεδιαστεί για να βελτιστοποιήσει τις δυνατότητες αξιοποίησης των δεδομένων που διατίθενται ήδη, μέσω της τεκμηρίωσης τους, της λειτουργίας υπηρεσιών που αποσκοπούν στη διευκόλυνση της πρόσβασης σε αυτά, στην αύξηση της διαλειτουργικότητά τους, και της αντιμετώπισης των δυσκολιών στις οποίες προσκρούει η χρήση τους

Για την επίτευξη των στόχων της, η Οδηγία εστιάζει σε πέντε περιοχές-κλειδί: τα μεταδεδομένα, τη διαλειτουργικότητα και εναρμόνιση των χωρικών δεδομένων και των υπηρεσιών τους, τις διαδικασίες διάθεσης των δεδομένων και των υπηρεσιών, τις υπηρεσίες και τεχνολογίες δικτύου, και τις διαδικασίες συντονισμού και παρακολούθησης. Για να διασφαλίσει ότι οι NSDIs του κάθε κράτους-μέλους θα είναι συμβατές και χρησιμοποιήσιμες σε μια πανευρωπαϊκή, διασυνοριακή υποδομή, η οδηγία απαιτεί την υιοθέτηση κοινών Εκτελεστικών Διατάξεων (Implementing Rules – IRs) σε ένα σύνολο πεδίων. Οι κανόνες υλοποίησης αυτοί υιοθετήθηκαν ή θα υιοθετηθούν ως αποφάσεις ή κανονισμοί της επιτροπής, και θα είναι δεσμευτικοί στο σύνολό τους. Την επιτροπή βοηθάει στο έργο αυτό μια επιτροπή αντιπροσώπων των κρατών μελών (γνωστή και ως διαδικασία της κοιτολογίας). Οι διατάξεις υλοποίησής του INSPIRE προάγουν τη διαλειτουργικότητα μεταξύ χωρικών δεδομένων τα οποία προέρχονται από διαφορετικούς φορείς συλλογής, παραγωγής και χρήσης γεωγραφικής πληροφορίας. Ως αποτέλεσμα, στην ευρωπαϊκή υποδομή θα συμπεριληφθούν χωρικές πληροφορίες πολλών θεματικών ειδών. Η οδηγία ταξινομεί τις διαφορετικές θεματικές ενότητες των χωρικών πληροφοριών σε τρία χωριστά παραρτήματα. Το κάθε παράρτημα, πέρα του ότι περιέχει σχετικά ομοειδείς θεματικές κατηγορίες δεδομένων, διακρίνεται και από διαφορετικές προθεσμίες για την εφαρμογή των απαιτήσεων του INSPIRE, καθώς και διαφορετικά επίπεδα αυστηρότητας για την εναρμόνιση. Οι τεχνικές προδιαγραφές του INSPIRE ανά θεματικό επίπεδο ουσιαστικά αποτελούν τυποποιημένες γεωγραφικές οντολογίες για πολλαπλά πεδία εφαρμογών: πρόκειται για ακριβείς, σαφείς και τυποποιημένες προδιαγραφές, για σχετικά μικρό αριθμό οντοτήτων ανά πεδίο εφαρμογής. Οι κανονισμοί της Ευρωπαϊκής Επιτροπής για την υλοποίηση του INSPIRE ουσιαστικά αποτελούνται από δύο κείμενα: τις ίδιες τις εκτελεστικές διατάξεις και τις συνοδευτικές Τεχνικές Οδηγίες (Tech. Guidelines). Οι εκτελεστικές διατάξεις αποδίδουν ότι πρέπει να επιτευχθεί, λαμβάνοντας υπόψη τις απαιτήσεις κόστους-οφέλους, και είναι νομικά δεσμευτικές προς τα κράτη μέλη. Το κείμενο είναι νομικό, και οι όποιες τεχνικές λεπτομέρειες εκφράζονται σε φυσική γλώσσα (κι αν απαιτείται επεξηγούνται). Ουσιαστικά περιλαμβάνουν τις υποχρεωτικές παραμέτρους των προδιαγραφών, και η εφαρμογή τους είναι υποχρεωτική εντός των χρονικών ορίων που ορίζει η οδηγία. Οι τεχνικές οδηγίες αποτελούν τεχνικές προδιαγραφές που αναπτύσσονται από τις DT και TWG. Ως καθαρά τεχνικά κείμενα, είναι μη δεσμευτικά νομικώς. Περιγράφουν το σύνολο των μέτρων που πρέπει να ληφθούν ώστε να εξασφαλισθεί η διαλειτουργικότητα εντός του INSPIRE, περιέχοντας απαιτήσεις και προτάσεις (υποχρεωτικές και προαιρετικές), επεξηγήσεις και παραδείγματα. Η πλήρη εφαρμογή τους δεν είναι υποχρεωτική, και μπορούν να χρησιμοποιηθούν για την περαιτέρω επέκταση και βελτίωση των ΕΥΓΕΠ στα κράτη μέλη. 9.3.3. Χρονοδιάγραμμα και Εξέλιξη Η ΕΕ υιοθέτησε στις 14 Μαρτίου 2007 την οδηγία INSPIRE (2007/2/EC), μετά τη συνεργασία των κρατών μελών και των ενδιαφερομένων φορέων. Η οδηγία θεωρείται εν ισχύ από τις 15 Μαΐου 2007. Τα κράτη μέλη όφειλαν να θέσουν σε ισχύ εθνικούς νόμους, διατάγματα και διοικητικές πράξεις που να συμμορφώνονται και να ενσωματώνουν την οδηγία έως την 15η Μαΐου 2009 (European Parliament and Council 2007). Το χρονοδιάγραμμα της Οδηγίας INSPIRE αποτελείται από

τρεις φάσεις. Την προπαρασκευαστική (2005-2006), τη μεταβατική (2007-2008) και την φάση εφαρμογής (2009-2013). Η προπαρασκευαστική φάση βρίσκεται στα τελικά στάδια (έχει παραταθεί κατά αρκετά έτη λόγω της πολυπλοκότητας των θεματικών επιπέδων), με τη συγκρότηση Ομάδων Εμπειρογνομόνων (Drafting Teams), υπεύθυνων για τη σύνθεση των προσχεδίων των εκτελεστικών διατάξεων της οδηγίας INSPIRE. Έχουν οριστεί πέντε ομάδες εμπειρογνομόνων, που ασχολούνται από τον Οκτώβριο 2005 με τους κανόνες υλοποίησης των εξής θεμάτων:

- καθορισμός προδιαγραφών δεδομένων (data specifications) 39
- μεταδεδομένα (Metadata)
- υπηρεσίες δικτύου δεδομένων (Network Services)
- διανομή και πρόσβαση δεδομένων και υπηρεσιών (Data and Service Sharing)
- παρακολούθηση και συντονισμός του έργου

Κάθε ομάδα εμπειρογνομόνων εκτελεί σειρά συναντήσεων, με σκοπό να εξετάσει το υπάρχον υλικό (μοντέλα και χωρικά δεδομένα) που έχει το κάθε κράτος μέλος της ΕΕ και να παραδώσει σε μορφή προσχεδίου τις εκτελεστικές διατάξεις του τμήματος του Inspire που την αφορά, ώστε το προσχέδιο να μελετηθεί, να αναθεωρηθεί και τελικά να γίνει επίσημη οδηγία – νόμος υποχρεωτικός για τα κράτη μέλη. Ήδη την προκειμένη στιγμή η πλειοψηφία των προσχεδίων από τις πέντε βασικές ομάδες εμπειρογνομόνων ψηφίστηκαν ως κανονισμοί (regulations) της Ε.Ε., ενώ οι τεχνικές προδιαγραφές των σχημάτων εφαρμογών (application schemas) για τα Παραρτήματα II και III ολοκλήρωσαν το καθεστώς διαβούλευσης. Καθώς οι εν ισχύ ομάδες εμπειρογνομόνων έχουν ως υποχρέωση τη σύνθεση γενικών κανόνων υλοποίησης, προέκυψε πως οι γενικές προδιαγραφές δεδομένων και το γενικευμένο εννοιολογικό μοντέλο που θα προτεινόταν έπρεπε να υποστούν περαιτέρω εξειδίκευση, ώστε να ικανοποιούν τις ανάγκες των χρηστών και των εφαρμογών που αφορούν σε συγκεκριμένες θεματικές ενότητες των τριών παραρτημάτων. Ως εκ τούτου, αποφασίστηκε ως συνέχεια της προπαρασκευαστικής φάσης να συντεθούν ειδικευμένες ομάδες εργασίας για καθένα από τα θεματικά επίπεδα των παραρτημάτων. Οι ομάδες αυτές ονομάζονται Thematic Working Groups (TWGs), και η κάθε μία αποτελείται από επιστήμονες ειδικευμένους στο θεματικό επίπεδο που αναφέρονται και μόνο. Στόχος τους είναι να συγκεκριμενοποιήσουν τις γενικές προδιαγραφές για τα χωρικά δεδομένα της δικής τους θεματικής περιοχής. Σε πρώτη φάση σχηματίστηκαν, κατά προτεραιότητα, οι εννέα για τις θεματικές περιοχές του Παραρτήματος I, των οποίων οι εργασίες ολοκληρώθηκαν και 39 Οι βασικοί κανόνες υλοποίησης του INSPIRE, που θα καθορίσουν πως θα συλλέξει και θα καταγράψει τα δεδομένα της η κάθε χώρα. αποτέλεσαν μέρος του Commission Regulation (EU) No. 1089/2010 που προδιαγράφει τις τεχνικές προδιαγραφές για τη διαλειτουργικότητα των χωρικών δεδομένων και υπηρεσιών, και το οποίο τέθηκε εν ισχύ στις 8/12/2010 (European Commission 2010a). Οι θεματικές ομάδες εργασίας λειτουργούν υπό την αρωγή και καθοδήγηση της Data Specifications Drafting Team (DTDS)