

ΚΕΦΑΛΑΙΟ 6

Χωρική Παρεμβολή (*Spatial Interpolation*)

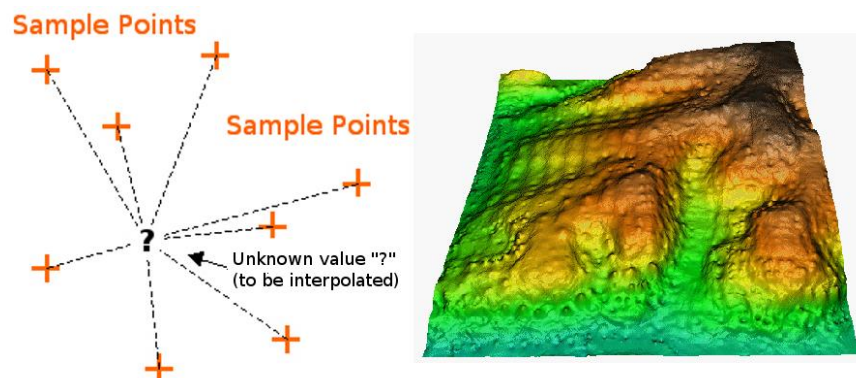
Στόχοι κεφαλαίου

1. Η γνώση για τις έννοιες της Χωρικής Παρεμβολής
2. Η διαμόρφωση προσδιοριστικών μοντέλων χωρικής παρεμβολής (deterministic models)
3. Η διαμόρφωση γεωστατιστικών (στοχαστικών) μοντέλων χωρικής παρεμβολής (geostatistic models)

6.1. Βασικοί ορισμοί

6.1.1. Χωρική παρεμβολή

Χωρική παρεμβολή είναι η εκτίμηση των άγνωστων τιμών μιας μεταβλητής σε ένα οποιαδήποτε σημείο σε μια περιοχή, με βάση γνωστές μετρήσεις.



Γιατί χρησιμοποιείται:

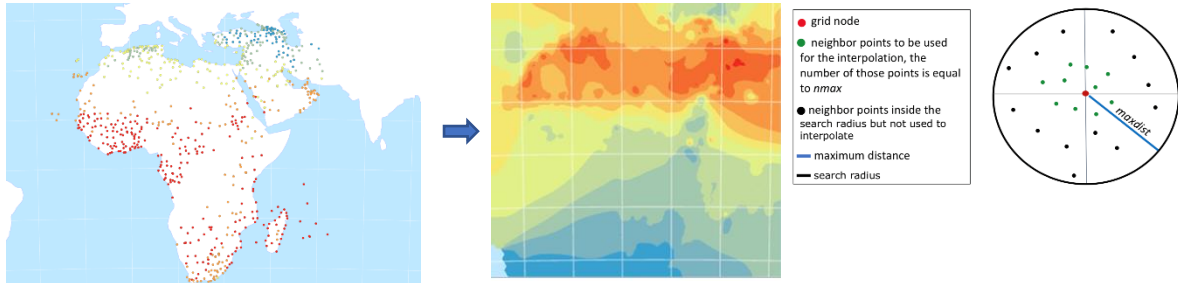
Η εκτίμηση ή πρόβλεψη της χωρικής κατανομής της μεταβλητής σε σημεία που δεν υπάρχουν μετρήσεις (παρατηρήσεις) στην περιοχή μελέτης και η μετατροπή των δεδομένων από σύνολα σημείων σε συνεχείς επιφάνειες

Μερικοί ορισμοί

Καθολική (Global) και Τοπική (Local) χωρική παρεμβολή

Η Καθολική παρεμβολή χρησιμοποιεί όλα τα διαθέσιμα γνωστά σημεία στην περιοχή μελέτης για να εκτιμήσει την άγνωστη τιμή ενός σημείου που βρίσκεται εντός της περιοχής μελέτης. Η τοπική παρεμβολή από την άλλη πλευρά, χρησιμοποιεί ένα δείγμα γνωστών σημείων για να εκτιμήσει την άγνωστη τιμή.

Η καθολική παρεμβολή γενικά παράγει πιο λείες επιφάνειες. Η διαφορά μεταξύ τοπικής και καθολικής παρεμβολής έγκειται στον αριθμό των σημείων που χρησιμοποιούνται για την εκτίμηση της άγνωστης τιμής. Για να εκτελεστεί μια τοπική χωρική παρεμβολή με CDT, πρέπει να προσδιοριστεί ο ελάχιστος ($nmin$) και ο μέγιστος ($nmax$) αριθμός γειτονικών σημείων που θα χρησιμοποιηθούν για την εκτίμηση της άγνωστης τιμής σε έναν κόμβο πλέγματος και τη μέγιστη απόσταση ($maxdist$) εντός της οποίας θα επιλεγούν γειτονικά σημεία.



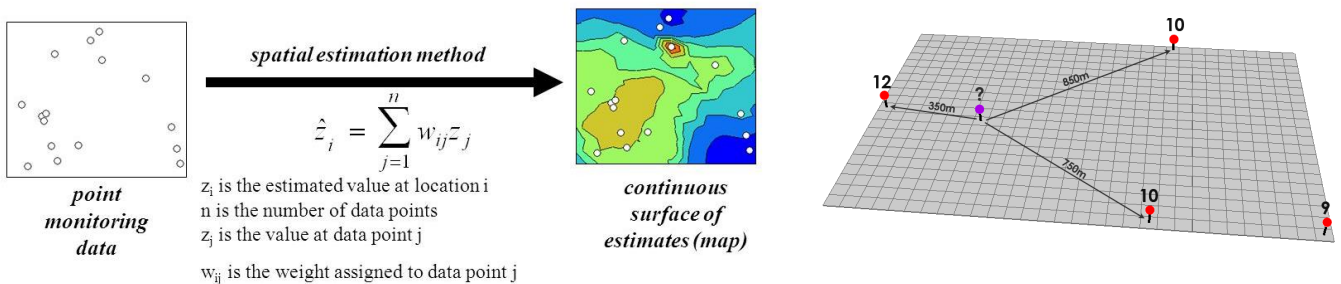
Ντετερμινιστική έναντι στοχαστικής παρεμβολής

Υπάρχουν δύο κύριες ομάδες μεθόδων παρεμβολής: ντετερμινιστική και στοχαστική. Η ντετερμινιστική μέθοδος παρεμβολής βασίζεται σε ακριβώς προκαθορισμένα χωρικά πλαίσια, όπως η έκταση της ομοιότητας ή ο βαθμός εξομάλυνσης. Η μέθοδος στοχαστικής παρεμβολής, από την άλλη πλευρά, λαμβάνει υπόψη τυχαίες συναρτήσεις, συμπεριλαμβανομένης της χωρικής εξάρτησης μεταξύ σημείων, ποσοτικοποιεί τη χωρική αυτοσυσχέτιση μεταξύ των μετρούμενων σημείων και υπολογίζει τη χωρική διαμόρφωση των σημείων του δείγματος γύρω από τη θέση πρόβλεψης.

Η μέθοδος ντετερμινιστικής παρεμβολής δεν παρέχει αξιολόγηση σφαλμάτων με προβλεπόμενες τιμές, ενώ μια μέθοδος στοχαστικής παρεμβολής προσφέρει αξιολόγηση σφαλμάτων προβλέψεων με εκτιμώμενες διακυμάνσεις.

Προσδιορισμός Βαρών

Στη χωρική παρεμβολή, η εκτιμώμενη τιμή z_i , είναι το σταθμισμένο άθροισμα των n γνωστών μετρήσεων. Ο όρος w_{ij} είναι το βάρος που αποδίδεται στη μέτρηση, κατά την εκτίμηση της άγνωστης τιμής.



Μέθοδοι χωρικής παρεμβολής

Καθολική (Global) παρεμβολή

Υπάρχουν κυρίως δύο μέθοδοι καθολικής παρεμβολής:

- **Με μοντέλα ταξινόμησης (*classifications models*)** τα οποία χωρίζουν την περιοχή σε τομείς ή υπο-περιοχές (οι παρατηρήσεις έχουν παρθεί από στατιστικά σταθερό πληθυσμό και συνεπώς η μέση τιμή και η διακύμανση είναι ανεξάρτητες από τη θέση μέσα στους τομείς). *Η ταξινόμηση σε ομογενή πολύγωνα προϋποθέτει ότι η διακύμανση μέσα στους τομείς είναι μικρότερη από αυτή μεταξύ των τομέων. Δηλαδή, οι σημαντικές αλλαγές λαμβάνουν χώρα στα όρια των τομέων.*
- **Με μεθόδους παλινδρόμησης :**
 - **α) Αναλύσεις επιφάνειες τάσης (*Trend surface analysis*)** και
 - **β) Μοντέλα παλινδρόμησης (*Regression model*).** Μπορούμε να προβλέψουμε την χωρική μεταβολή μιας ιδιότητας, εάν γνωρίζουμε την χωρική μεταβολή μιας ή περισσότερων άλλων ιδιοτήτων οι οποίες σχετίζονται με την πρώτη.

Τοπική (Local) παρεμβολή

- **Ντετερμινιστές μέθοδοι**
 - **Μέθοδος πλησιέστερου γείτονα (πολύγωνα *Thiessen, Dirichlet* ή *Voronoi*)**
 - **Τριγωνισμός Delaunay (*Triangulated Irregular Network*)**
 - **Μέθοδος Τοπικού Μέσου Όρου Μετρήσεων**
 - **Μέθοδος Αντίστροφων Αποστάσεων (*Inverse Distance Weighting*)**
- **Γεωστατιστικές (στοχαστικές) μέθοδοι**
 - **Ordinary Kriging (OK)**
 - **Universal Kriging (UK)**

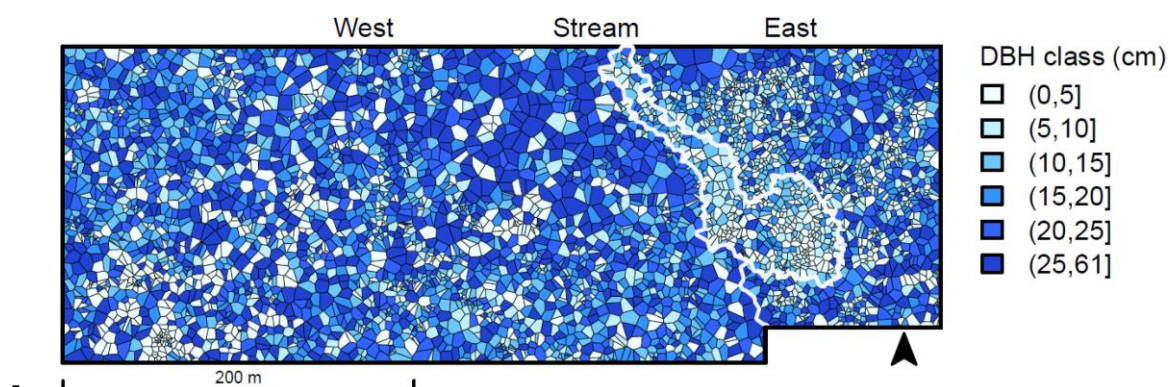
6.2. Μέθοδος πλησιέστερου γείτονα (πολύγωνα Thiessen, Dirichlet ή Voronoi)

Με τη Nearest Neighbours μέθοδο (γνωστή και ως πολύγωνα Thiessen, Dirichlet ή Voronoi) η μελετώμενη περιοχή χωρίζεται σε πολύγωνα έτσι ώστε σε κάθε ένα από αυτά να περιέχεται και ένα σημείο, για το οποίο υπάρχουν δεδομένα.

Αν τα δεδομένα παρουσιάζουν συμμετρική εξάπλωση, τότε τα πολύγωνα Thiessen είναι όμοια μεταξύ τους. Σε αντίθετη περίπτωση, προκύπτει ένα ακανόνιστο δίκτυο πολυγώνων. Είναι μια εύκολη και γρήγορη μέθοδος συσχέτισμού σημειακών δεδομένων στο χώρο, η οποία βρίσκει εφαρμογές στα GIS, στη γεωγραφική ανάλυση και στη μετεωρολογία.

- Ενδείκνυται να χρησιμοποιείται όταν απαιτείται χάρτης ποιοτικών δεδομένων - όπως είναι οι χρήσεις γης - και δεν ενδιαφέρει η παράξενη και απότομη μορφή των ορίων.
- Αντίθετα, αν τα σημεία δειγματοληψίας είναι λιγοστά, η εφαρμογή της μεθόδου αντενδείκνυται, ιδιαίτερα για ιδιότητες που μεταβάλλονται βαθμιαία (π.χ. βροχόπτωση και θερμοκρασία). Αυτό συμβαίνει γιατί :
 - ο (α) η μορφή του τελικού παραγόμενου χάρτη εξαρτάται από
 - ο την κατανομή των παρατηρήσεων
 - ο (β) η μεταβολή των τιμών της ιδιότητας γίνεται στα όρια των πολυγώνων ενώ εντός των πολυγώνων υπάρχει ομοιομορφία, η οποία όμως δεν υφίσταται στην πραγματικότητα.

Η καθολική και τοπική μορφή της μεθόδου «πλησιέστερου γείτονα» ταυτίζονται και παράγουν το ίδιο αποτέλεσμα.

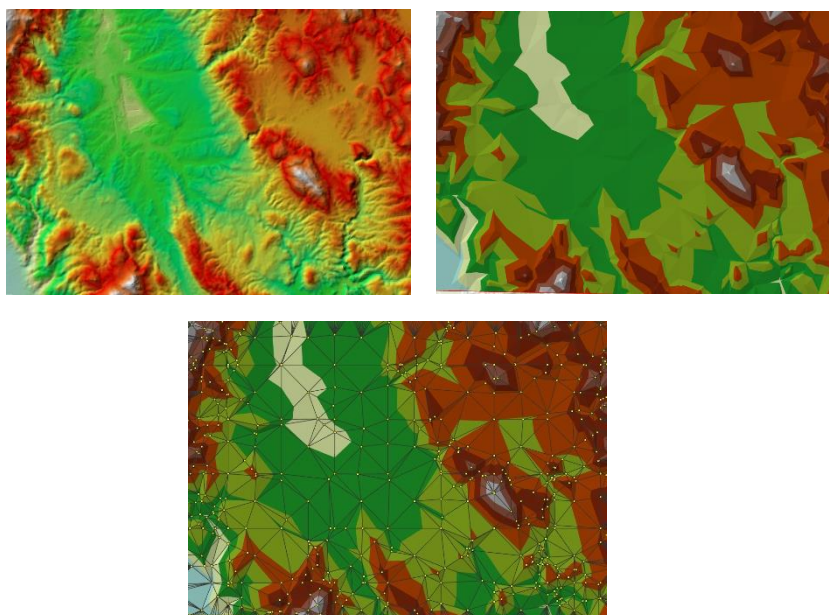
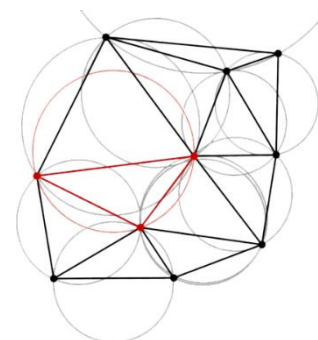


Voronoi tessellation for the living trees in the study area, colored by tree diameter class (Putz et al. 2022)

6.3. Μέθοδος Τριγωνισμός Delaunay (Triangulated Irregular Network)

Σύμφωνα με τη μέθοδο αυτή τρία σημεία παρατήρησης αποτελούν ένα τρίγωνο Delaunay, εάν τα αντίστοιχα πολύγωνα επιρροής τους έχουν μια κοινή κορυφή.

Το τρίγωνο αυτό ορίζει ένα Triangulated Irregular Network, που χρησιμοποιείται ευρέως στην αποτύπωση του αναγλύφου, στη γεωλογία και γεωμορφολογία. Ο Delaunay ορίζει ότι τρία σημεία σχηματίζουν τρίγωνα Delaunay εάν και μόνο αν ο κύκλος που περνάει από τα τρία σημεία δεν περιέχουν άλλα σημεία (κορυφές)



Τα μοντέλα TIN είναι λιγότερο ευρέως διαθέσιμα από τα μοντέλα επιφανειών ράστερ και τείνουν να είναι πιο ακριβά στην κατασκευή και την επεξεργασία τους. Το Τριγωνικό Ακανόνιστο Δίκτυο (TIN) είναι ένας τρόπος αποθήκευσης συνεχών επιφανειών. Ίσως η πιο κοινή εφαρμογή του μοντέλου TIN έγκειται στην τοπογραφική μοντελοποίηση, η οποία σχετίζεται με υψομετρικές αλλαγές σε μια περιοχή. Εδώ, τα μοντέλα TIN χρησιμοποιούνται για την εκτίμηση τοπογραφικά σχετικών πληροφοριών. Ωστόσο, είναι σημαντικό να σημειωθεί ότι η μοντελοποίηση οποιασδήποτε φυσικής επιφάνειας με ένα μαθηματικό μοντέλο είναι δύσκολη. Από αυτή την άποψη, τα TIN είναι ένας απλοποιημένος και βελτιστοποιημένος τρόπος αναπαράστασης ενός εδάφους.

Τα TIN χρησιμοποιούνται συνήθως για μοντελοποίηση μικρότερων περιοχών υψηλής ακρίβειας, όπως σε εφαρμογές μηχανικής, όπου είναι χρήσιμα επειδή επιτρέπουν υπολογισμούς επιπεδομετρικής περιοχής, επιφάνειας και όγκου. Τα TIN είναι επίσης ένα σημαντικό στοιχείο στην ψηφιακή ανάλυση εδάφους (DTA). Τα DEM που βασίζονται σε TIN χρησιμοποιούνται για την αναπαράσταση του υπεδάφους σε μοντέλα διεργασιών υδρολογίας λεκάνης απορροής. Έχει υποστηριχθεί ότι τα DEM που βασίζονται σε TIN ξεπερνούν τα DEM που βασίζονται σε ράστερ στην υδρολογική μοντελοποίηση υψηλής ακρίβειας μικρών λεκανών όταν εξετάζεται η πυκνότητα του σημείου δειγματοληψίας (Ivanov et al., 2004b).

6.4. Μέθοδος Τοπικού Πολυωνύμου (local polynomial interpolation)

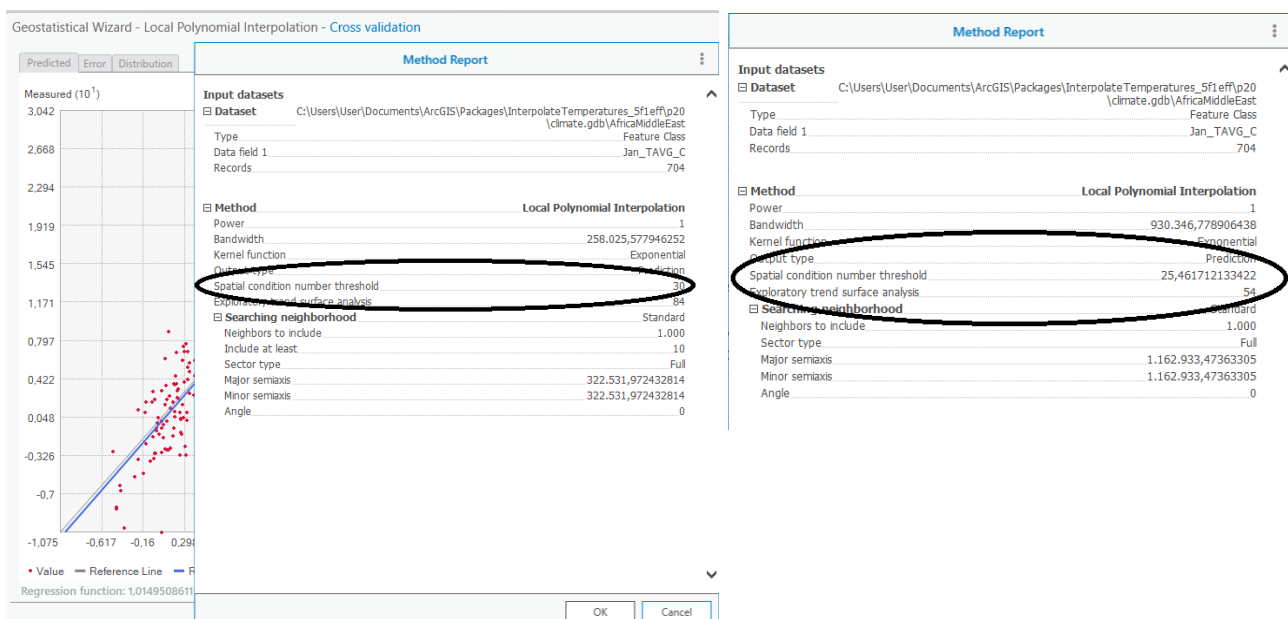
Η τοπική πολυωνυμική παρεμβολή ταιριάζει σε πολλά πολυώνυμα, το καθένα μέσα σε καθορισμένες επικαλυπτόμενες γειτονιές. Η γειτονιά αναζήτησης μπορεί να οριστεί χρησιμοποιώντας το μέγεθος και το σχήμα, τον αριθμό των γειτόνων και τη διαμόρφωση τομέα.

Η τοπική πολυωνυμική παρεμβολή θα είναι πιο ακριβής όταν τα δεδομένα έχουν τις ακόλουθες ιδιότητες:

- Τα δείγματα ελήφθησαν σε πλέγμα (δηλαδή, τα δείγματα έχουν ίση απόσταση).
- Οι τιμές δεδομένων, εντός της γειτονιάς αναζήτησης, κατανέμονται κανονικά.

Στην πράξη, τα περισσότερα σύνολα δεδομένων δεν θα έχουν αυτές τις ιδιότητες. Σε αυτές τις περιπτώσεις, οι προβλεπόμενες τιμές θα επηρεαστούν. Για να αποφασιστεί εάν τα αποτελέσματα σε ορισμένες περιοχές είναι αξιόπιστα ή όχι, ελέγχεται με βάση το βαθμό του πολυώνυμου την κρίσιμη χωρική τιμή. Τιμές που είναι κάτω από τα όρια αριθμών κρίσιμης χωρικής κατάστασης υποδεικνύουν σε ποιες τοποθεσίες είναι αξιόπιστες οι λύσεις. Τιμές κοντά ή ίσες με τις κρίσιμες τιμές είναι αμφισβητήσιμες (θα πρέπει να εξετάζονται με προσοχή) και τιμές πάνω από τα κρίσιμα όρια δεν είναι αξιόπιστες.

Order of polynomial	Critical spatial condition number threshold value
1	10
2	100
3	1000
Greater than 3	Not recommended for most situations



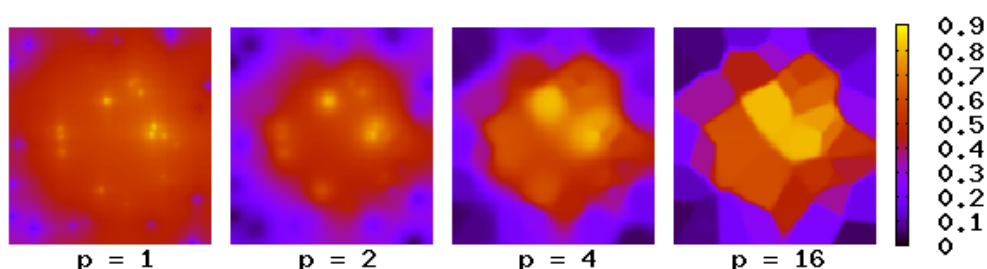
6.5. Μέθοδος Αντίστροφων Αποστάσεων (*Inverse Distance Weighting*)

Η μέθοδος των Αντίστροφων Αποστάσεων δίνει πολύ αξιόπιστα αποτελέσματα και διαθέτει μεγάλη ευελιξία. Η στάθμιση αντίστροφης απόστασης (*IDW*) είναι μια μέθοδος παρεμβολής που εκτιμά τις τιμές κελιών υπολογίζοντας τον μέσο όρο των τιμών των σημείων δεδομένων δείγματος στη γειτονιά του κάθε κελιού επεξεργασίας. Όσο πιο κοντά βρίσκεται ένα σημείο στο κέντρο της θέσης που εκτιμάται, τόσο μεγαλύτερη επιρροή ή βάρος έχει στη διαδικασία υπολογισμού του μέσου όρου.

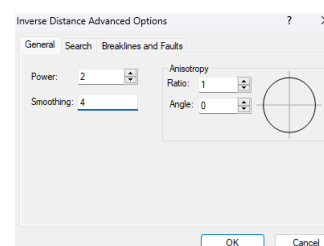
Αυτή η μέθοδος υποθέτει ότι η μεταβλητή που χαρτογραφείται μειώνεται σε επιρροή με την απόσταση από τη θέση του δείγματος. Για παράδειγμα, όταν παρεμβάλλεται μια επιφάνεια αγοραστικής δύναμης καταναλωτή για μια ανάλυση λιανικής πώλησης, η αγοραστική δύναμη μιας πιο απομακρυσμένης τοποθεσίας θα έχει μικρότερη επιρροή, επειδή οι άνθρωποι είναι πιο πιθανό να ψωνίσουν πιο κοντά στην περιοχή διαμονής τους.

Δύναμη (*power*)

Με την *IDW* μπορεί να ελεγχθεί η σημασία των γνωστών σημείων στις παρεμβαλλόμενες τιμές με βάση την απόστασή τους από το σημείο εξόδου. Ορίζοντας υψηλότερη δύναμη (*power*), δίνεται μεγαλύτερη έμφαση στα πλησιέστερα σημεία και η επιφάνεια που προκύπτει θα έχει περισσότερες λεπτομέρειες (θα είναι λιγότερο λεία). Ο καθορισμός χαμηλότερης ισχύος θα δώσει μεγαλύτερη επιρροή στα σημεία που βρίσκονται πιο μακριά, με αποτέλεσμα μια πιο λεία επιφάνεια. Η τιμή $p=2$ χρησιμοποιείται πιο συχνά με το *IDW*.



Υπάρχει και μια παραλλαγή της *IDW*, η **μέθοδος του Shepard** (Shepard, 1968), με δύο διαφορετικές συναρτήσεις στάθμισης που χρησιμοποιούν δύο ξεχωριστές γειτονιές. Π.χ. η προεπιλεγμένη συνάρτηση στάθμισης για τη μέθοδο του Shepard είναι ένας εκθέτης 2 στην εσωτερική γειτονιά και ένας εκθέτης 4 στην εξωτερική γειτονιά (αλλά διαθέσιμο στο λογισμικό *Surfer*)



Ακτίνα αναζήτησης

Τα χαρακτηριστικά της παρεμβαλλόμενης επιφάνειας μπορούν να ελεγχθούν με την εφαρμογή μιας σταθερής ή μεταβλητής ακτίνας αναζήτησης, η οποία περιορίζει τον αριθμό των σημείων εισόδου που μπορούν να χρησιμοποιηθούν για τον υπολογισμό κάθε παρεμβαλλόμενης κυψέλης. Ο περιορισμός του αριθμού των γειτόνων για τον υπολογισμό του κάθε κελιού θα βελτιώσει τις ταχύτητες επεξεργασίας, ενώ και σημεία μακριά από τη θέση του κελιού όπου γίνεται η πρόβλεψη ενδέχεται να μην έχουν χωρική αντιστοιχία (μικρή χωρική αυτοσυσχέτιση).

6.6. Ποιότητα Χωρικής Παρεμβολής

6.6.1. Μετρήσεις σφάλματος

Στη Χωρική Παρεμβολή, η γενική μορφή του **Σφάλματος Εκτίμησης** είναι η διαφορά μεταξύ της εκτιμώμενης τιμής και γνωστής τιμής στο σημείο επικύρωσης (όπου η αντίστοιχη αρχική μέτρηση είναι γνωστή).

Μια συνηθισμένη διαδικασία για τον υπολογισμό του είναι η τεχνική **“Leave one out – Cross validation”**, όπου υπολογίζεται το σφάλμα για ένα – ένα όλα τα αρχικά σημεία (με την αφαίρεση και εκτίμηση κάθε φορά από ένα σημείο και επαναφορά μετά στο σύνολο).

Η διασταυρούμενη επικύρωση είναι μια μέθοδος επαναδειγματοληψίας που χρησιμοποιεί πρώτα όλα τα σημεία εισόδου για να εκτιμήσει τις παραμέτρους ενός μοντέλου παρεμβολής (όπως το ημιβαριόγραμμα του kriging ή την τιμή ισχύος της αντίστροφης στάθμισης απόστασης). Στη συνέχεια αφαιρεί ένα μόνο σημείο εισόδου και χρησιμοποιεί τα υπόλοιπα σημεία για να προβλέψει την τιμή στη θέση του “κρυφού” σημείου και η προβλεπόμενη τιμή συγκρίνεται με τη μετρούμενη τιμή. Στη συνέχεια, το κρυφό σημείο προστίθεται ξανά στο σύνολο δεδομένων και ένα διαφορετικό σημείο κρύβεται και προβλέπεται. Αυτή η διαδικασία επαναλαμβάνεται σε όλα τα σημεία εισόδου.

Τι υπολογίζεται:

- **Μετρούμενη τιμή**—Η μετρούμενη τιμή του κρυφού σημείου.
- **Προβλεπόμενη τιμή** —Η προβλεπόμενη τιμή από τη διασταυρούμενη επικύρωση στη θέση του κρυφού σημείου.
- **Σφάλμα**—Η διαφορά μεταξύ των προβλεπόμενων και των μετρούμενων τιμών (προβλεπόμενη μείον μετρημένη). Ένα θετικό σφάλμα σημαίνει ότι η πρόβλεψη ήταν μεγαλύτερη από τη μετρούμενη τιμή και ένα αρνητικό σφάλμα σημαίνει ότι η πρόβλεψη ήταν μικρότερη από τη μετρούμενη τιμή.
- **Τυπικό σφάλμα**—Το τυπικό σφάλμα της προβλεπόμενης τιμής. Εάν τα σφάλματα κατανέμονται κανονικά, περίπου τα δύο τρίτα των μετρούμενων τιμών θα εμπίπτουν σε ένα τυπικό σφάλμα της προβλεπόμενης τιμής και περίπου το 95 τοις εκατό θα εμπίπτουν σε δύο τυπικά σφάλματα.
- **Τυποποιημένο σφάλμα**—Το σφάλμα διαιρούμενο με το τυπικό σφάλμα. Για να χρησιμοποιηθούν οι τύποι εξόδου ποσοστιαίων ή πιθανοτήτων, οι τυποποιημένες τιμές σφάλματος θα πρέπει να ακολουθούν μια τυπική κανονική κατανομή (μέση τιμή ίση με μηδέν και τυπική απόκλιση ίση με ένα).

Σε επίπεδο συνολικής αξιολόγησης υπολογίζονται και τα:

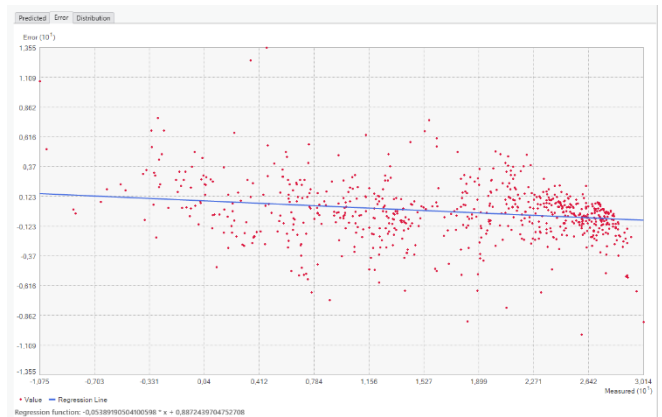
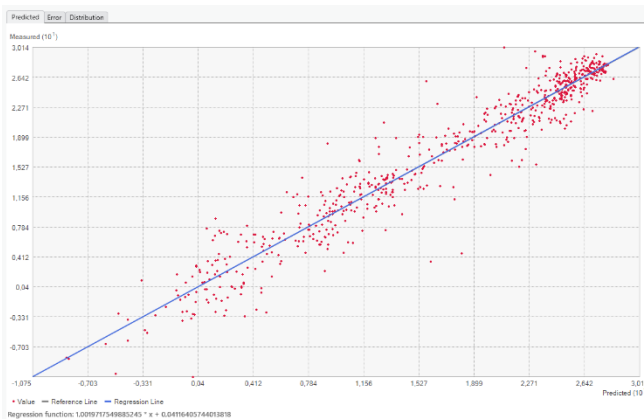
- **Μέσο τετραγωνικό σφάλμα (Mean Squared error – MSE)** – Ο μέσος όρος των τετραγωνισμένων τιμών του σφάλματος, που υπολογίζεται με το άθροισμα του τετραγωνισμένου σφάλματος και της διακύμανσης σφαλμάτων
- **Ρίζα Μέσου Τετραγωνικού Σφάλματος (Root Mean Square Error – RMSE)**

Τα MSE και RMSE εκφράζουν το μέσο σφάλμα εκτίμησης σε μονάδες της μεταβλητής στο εύρος $(0, + \infty)$ και είναι αδιάφορες στην κατεύθυνση (πρόσημο) των σφαλμάτων. Χαμηλότερες τιμές φανερώνουν καλύτερα αποτελέσματα. Η διαφορά του RMSE σε σχέση με το MSE είναι

1. το RMSE δίνει ένα σχετικά υψηλό βάρος σε μεγάλα σφάλματα (εξαιτίας ότι τα σφάλματα τετραγωνίζονται πριν το μέσο όρο).
2. το RMSE δεν αυξάνεται απαραίτητα με τη διακύμανση (εύρος) των σφαλμάτων. Ωστόσο αυξάνεται με τη διακύμανση της κατανομής συχνότητας (πλήθος) των μεγεθών σφάλματος.
3. το RMSE έχει την τάση να είναι όλο και μεγαλύτερο από το MSE καθώς αυξάνεται το πλήθος των αρχικών σημείων (προβληματικό όταν συγκρίνονται αποτελέσματα με διαφορετικό πλήθος σημείων)

6.6.2. Διασποροδιάγραμμα

Η συνολική αξιολόγηση μπορεί να γίνει και οπτικά με ένα διασποροδιάγραμμα τιμών, όπου στον οριζόντιο άξονα μπαίνουν οι εκτιμήσεις και στον κάθετο άξονα οι πραγματικές (μετρημένες) τιμές. Σημεία που βρίσκονται κοντά στη διχοτόμο γραμμή των 45°, παρουσιάζουν μικρή διαφορά μεταξύ των πραγματικών και εκτιμώμενων τιμών. Επίσης μπορεί να αξιολογηθεί το διασποροδιάγραμμα της μεταβολής του σφάλματος, καθώς και το διάγραμμα συχνοτήτων του σφάλματος.



Με βάση το διασποροδιάγραμμα, μπορεί να υπολογιστεί και ο συντελεστής συσχέτισης R^2 του γραμμικού μοντέλου. Μεγαλύτερες τιμές (κοντά στο +1) υποδηλώνουν καλύτερη συνολική συμφωνία μεταξύ παρατηρούμενων και εκτιμώμενων τιμών, χωρίς αυτό να σημαίνει πάντα και καλύτερο αποτέλεσμα της Χωρικής Παρεμβολής (αν η συμφωνία τιμών περιλαμβάνει συστηματικά μεγάλο σφάλμα). Η χρήση όλων των στατιστικών μετρικών είναι απαραίτητη για τη συνολική εκτίμηση της ακρίβειας της Χωρικής Παρεμβολής.

