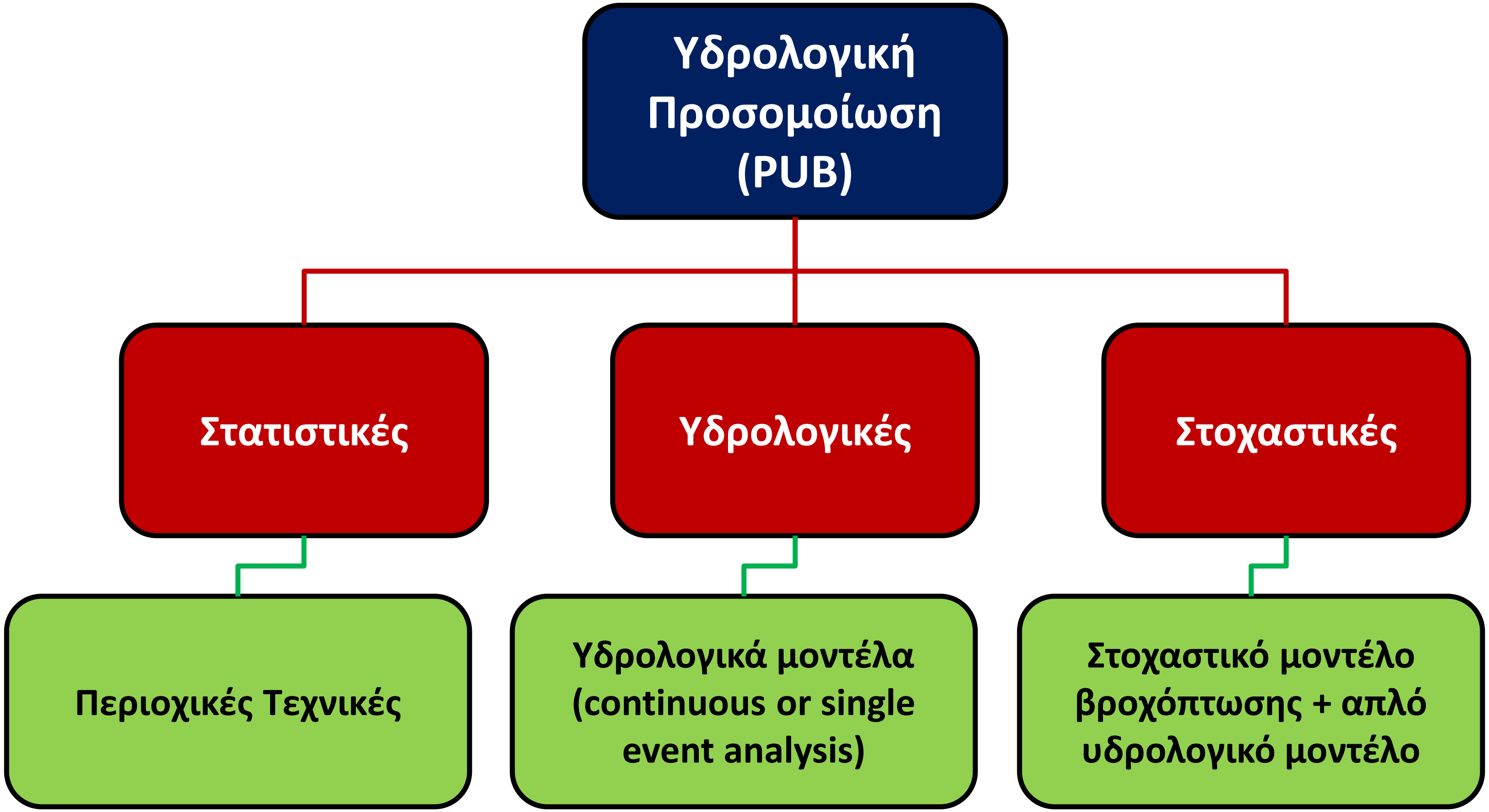
Τμήμα Πολιτικών Μηχανικών
Πανεπιστήμιο Θεσσαλίας

ΥΔΡΟΛΟΓΙΚΗ ΠΡΟΣΟΜΟΙΩΣΗ ΚΑΙ ΠΡΟΓΝΩΣΗ

Ενότητα 4: Γεωστατιστική

Καθ. Αθανάσιος Λουκάς
Εργαστήριο Υδρολογίας και Ανάλυσης Υδατικών Συστημάτων
Τμήμα Πολιτικών Μηχανικών
Πολυτεχνική Σχολή

ΜΕΘΟΔΟΙ ΥΔΡΟΛΟΓΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ



ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΥΔΡΟΛΟΓΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ

Περιοχικές Τεχνικές

- **Μέθοδοι Παλινδρόμησης**
  - **Ολική Παλινδρόμηση**
  - **Περιοχική Παλινδρόμηση**
  - **Περιοχή Επιρροής** (Burn, 1990)
- **Μέθοδοι Δείκτη Πλημμύρας (Dalrymple, 1960)**
  - **Υπολογισμός Δείκτη μ_i** (π.χ. παλινδρόμηση)
  - **Δείκτης $\mu_i * q(F)$**
- **Γεωστατιστικές Μέθοδοι**
  - **Χωρικές Συσχετίσεις** (Ευκλίδεια απόσταση)
  - **Συνδυαστικές Τεχνικές** (π.χ. Χωρικές Συσχετίσεις Υπολοίπων + παλινδρόμηση)

Burn, D. H.: An appraisal of the “region of influence” approach to flood frequency analysis, Hydrolog. Sci. J., 35, 149–165, 1990.
Dalrymple, T.: Flood frequency analysis, Water Supply Paper 1543A, US Geological Survey, Reston, Virginia, 1960.

ΔΙΑΡΘΡΩΣΗ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ: ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ - ΓΕΩΣΤΑΤΙΣΤΙΚΗ

- Εισαγωγή
- Μέθοδοι Χωρικής Παρεμβολής
- Γεωστατιστική
- Μέθοδος Βέλτιστης Παρεμβολής Kriging

ΜΕΘΟΔΟΙ ΧΩΡΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ

[Burrough and McDonnell, 1998; Li and Heap, 2008]

- **Μέθοδοι τοπικών εκτιμήσεων - Local methods**
  - Μέθοδος δικτύου τριγώνων – *Triangulated Irregular Network* (*TIN*)
  - Πολύγωνα Thiessen – *Thiessen Polygons*
  - Πολυωνυμικές Συναρτήσεις τύπου *Splines* – *Radial Basis Functions*
  - Παρεμβολή με απόδοση βαρών σε σημειακά δεδομένα στον πλησιέστερο γείτονα – *Inverse Distance Weighting* (*IDW*)
  - Πολυωνυμικές συναρτήσεις - *Local Polynomial Interpolation*
  - Παλινδρόμηση με γεωγραφική βαρύτητα - Geographically Weighted Regression (Moving window regression)
  - Γεωστατιστικές Μεθόδους (π.χ. μέθοδος βέλτιστης παρεμβολής kriging – Ordinary Kriging, OK)
- **Μέθοδοι γενικευμένων προσεγγίσεων – Global methods**
  - Πολυωνυμικές συναρτήσεις - *Global Polynomial Interpolation*
  - Γραμμική Παλινδρόμηση - Global Linear regression
  - Ανάλυση *Fourier*

ΜΕΘΟΔΟΙ ΧΩΡΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ

- Οι γενικευμένες μέθοδοι σε αντίθεση με τις τοπικές μεθόδους χρησιμοποιούν όλα τα υπάρχοντα στοιχεία από ολόκληρη την περιοχή μελέτης επιτυγχάνοντας εκτίμηση του φαινομένου για το σύνολο της περιοχής ενδιαφέροντος.
- Οι γενικές πολυωνυμικές συναρτήσεις και οι ακολουθίες *Fourier* ενδιαφέρουν μόνο για την αναπαράσταση των γενικών τάσεων και για την ανάλυση περιοδικότητας σε μία επιφάνεια.
- Η γραμμική παλινδρόμηση εφαρμόζεται κυρίως στην αναγωγή των σημειακών μετρήσεων της μεταβλητής (π.χ. θερμοκρασία, βροχόπτωση) σε επιφανειακές τιμές σε επίπεδο υδρολογικής λεκάνης.
- Με τις τοπικές μεθόδους εξετάζονται οι χωρικές διαφοροποιήσεις που δημιουργούνται κοντά στο υπό εκτίμηση σημείο – γειτνίαση, και γι’ αυτό οι τοπικές μέθοδοι εφαρμόζονται συχνότερα από τις γενικευμένες.

ΜΕΘΟΔΟΙ ΧΩΡΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ

- Οι δύο αυτές κατηγορίες μεθόδων (local and global methods) χρησιμοποιούν μαθηματικές συναρτήσεις και λειτουργίες για να παραχθεί η χωρική επιφάνεια.
- Οι γεωστατιστικές μέθοδοι (geostatistical methods) χρησιμοποιούν και μαθηματικές και στατιστικές τεχνικές για οποιαδήποτε εκτίμηση απορρέει από τις σημειακές μετρήσεις. Το σημαντικό πλεονέκτημα των γεωστατιστικών μεθόδων σε σχέση με τις υπόλοιπες μεθόδους είναι ότι ποσοτικοποιούν και ελαχιστοποιούν το σφάλμα εκτίμησης σε άγνωστα σημεία εκτός του δείγματος των σημειακών μετρήσεων.
- Οι μεθοδολογίες χωρικής παρεμβολής που χρησιμοποιούν μόνο μετρήσεις της εξεταζόμενης μεταβλητής (primary variable) καλούνται μέθοδοι μίας μεταβλητής (univariate methods). Όταν χρησιμοποιείται και δευτερογενής πληροφορία άλλων μεταβλητών τότε ονομάζονται μέθοδοι πολλών μεταβλητών (multivariate methods).

ΜΕΘΟΔΟΙ ΧΩΡΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ - ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

- **Εκτίμηση περιοχικής και επιφανειακής βροχόπτωσης**
  - Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, Journal of Hydrology, 228, 113-129.
- **Εκτίμηση περιοχικών υδρολογικών παραμέτρων**
  - Merz, R., and G. Blöschl, 2004. Regionalisation of catchment model parameters, Journal of Hydrology 287, 95-123.
- **Υδρολογικές εφαρμογές**
  - Grayson, R., and G. Bloschl, 2000. Spatial Patterns in Catchment Hydrology: Observations and Modelling, Cambridge University Press
- **Υδρογεωλογικές εφαρμογές**
  - Kitanidis, P.K., 1997. Introduction to Geostatistics: Applications in Hydrogeology, Cambridge University Press, Cambridge.
- **Περιβαλλοντικές μελέτες**
  - Li, J. and A.D. Heap, 2008. A Review of Spatial Interpolation Methods for Environmental Scientists, Geoscience Australia, Record 2008/23, 137 p.
- **Μετεωρολογικές - κλιματολογικές εφαρμογές**
  - Dobesch, H., P. Dumolard, and I. Dyras (eds.). 2007. Spatial Interpolation for Climate Data: The Use of GIS in Climatology and Meteorology, ISTE-Geographical Information Systems Series.

ΑΝΑΛΥΣΗ ΧΩΡΙΚΩΝ ΠΡΟΤΥΠΩΝ ΣΗΜΕΙΩΝ

- Αν τα στοιχεία αποτελούνται από ένα σύνολο σημείων σε καθένα από τα οποία αντιστοιχεί ένα χωρικό γεγονός.

- Η ανάλυση εστιάζεται στην εξέταση των πιθανών χωρικών προτύπων που δημιουργούν οι θέσεις αυτών των γεγονότων.

- Ως αποτέλεσμα, το ενδιαφέρον αφορά στα χωρικά πρότυπα των θέσεων για τις οποίες υπάρχουν παρατηρήσεις.

ΑΝΑΛΥΣΗ ΧΩΡΙΚΩΝ ΠΡΟΤΥΠΩΝ ΣΗΜΕΙΩΝ

- Αν τα στοιχεία αποτελούνται από ένα σύνολο σημείων στα οποία μετράται η τιμή ενός συνεχούς χωρικού φαινομένου.

- Η ανάλυση εστιάζεται στην ανάλυση του προτύπου των τιμών στα σημεία που ανήκουν στην περιοχή μελέτης.

- Ως αποτέλεσμα, το ενδιαφέρον αφορά στα χωρικά πρότυπα των ίδιων των χαρακτηριστικών.

ΔΙΑΔΙΚΑΣΙΑ ΤΗΣ ΧΩΡΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ

Αν…

***$Z(s), s \in A$*:** Χωρική στοχαστική διαδικασία η οποία μεταβάλλεται συνεχώς στην περιοχή **A.**

και …

Z_i, i=1,…,n: Σύνολο παρατηρήσεων για ένα χωρικά συνεχές χαρακτηριστικό, οι οποίες έχουν καταγραφεί σε συγκεκριμένες θέσεις **s** στην περιοχή μελέτης **A**.

Ο στόχος της ανάλυσης είναι να εξαχθούν συμπεράσματα για τη χωρική διαφοροποίηση του χαρακτηριστικού σε ολόκληρη την περιοχή μελέτης **A**, με βάση τις τιμές στα σταθερά σημεία-θέσεις του δείγματος.

ΔΙΑΔΙΚΑΣΙΑ ΤΗΣ ΧΩΡΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ

- Αν δηλαδή το ενδιαφέρον εστιάζεται στη μοντελοποίηση του προτύπου της μεταβλητότητας των τιμών του υπό εξέταση χαρακτηριστικού.

- Απαιτείται η χρήση μοντέλων προκειμένου να επιτύχουμε καλές εκτιμήσεις για την τιμή που παίρνει το χαρακτηριστικό σε σημεία τα οποία δεν ανήκουν στο αρχικό δείγμα.

- Δηλαδή, → διαδικασία της **χωρικής παρεμβολής** (***spatial interpolation***).

ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΧΩΡΙΚΩΝ ΣΤΟΙΧΕΙΩΝ

Η συμπεριφορά των χωρικών φαινομένων είναι συχνά αποτέλεσμα του συνδυασμού των επιπτώσεων δύο διαδικασιών.

- **Πρώτης τάξης (*first order*)*:*** Οι επιπτώσεις της πρώτης τάξης σχετίζονται με τη μεταβλητότητα στη μέση τιμή της υπό εξέταση χωρικής διαδικασίας και εκπροσωπούν γενικευμένες ή μεγάλης κλίμακας τάσεις-διαφοροποιήσεις.

- **Δεύτερης τάξης (*second order*):** Οι επιπτώσεις της δεύτερης τάξης είναι αποτέλεσμα της δομής της χωρικής συσχέτισης ή της χωρικής εξάρτησης στη διαδικασία. Με άλλα λόγια, είναι η τάση για αποκλίσεις στις τιμές από τη μέση τιμή σε γειτονικές θέσεις και εκπροσωπούν τοπικές ή μικρής κλίμακας επιπτώσεις.

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ

ΜΕΘΟΔΟΙ ΤΟΠΙΚΩΝ ΕΚΤΙΜΗΣΕΩΝ

- Χωρικός Κινητός Μέσος
- Ψηφιδοποίηση
  - TIN
  - Πολύγωνα Thiessen (Thiessen Polygons)

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ

- Επιφάνεια Τάσης
- Μοντέλα Ταξινόμησης

ΓΕΩΣΤΑΤΙΣΤΙΚΟΙ ΜΕΘΟΔΟΙ

- Kriging

ΕΠΙΠΤΩΣΕΙΣ ΔΕΥΤΕΡΗΣ ΤΑΞΗΣ

- Ως **στάσιμη** *(stationary)* ή **ομοιογενής** διαδικασία ορίζεται η χωρική διαδικασία της οποίας οι στατιστικοί δείκτες είναι ανεξάρτητοι από την απόλυτη θέση στην περιοχή μελέτης **Α**. Δηλαδή:
  - **$\in$(Z(s))**, **var(Z(s))** σταθερές στην περιοχή μελέτης **Α.**
  - **C(si,sj)** εξαρτάται από τη σχετική θέση, από την απόσταση και την κατεύθυνση μεταξύ τους και όχι από την απόλυτη θέση τους στην **Α**.
  - Μια στάσιμη διαδικασία ορίζεται ως **ισοτροπική**, εάν η συνδιασπορά εξαρτάται μόνο από την απόσταση μεταξύ των σημείων s_i και s_j και όχι από την κατεύθυνση κατά την οποία διαχωρίζονται.
- **Μη στάσιμη** ή **ετερογενής** θεωρείται η διαδικασία αν η μέση τιμή, η διασπορά ή η συνδιασπορά διαφοροποιούνται στην περιοχή μελέτης.

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ

ΥΠΟΘΕΣΕΙΣ

- Η επιφάνεια που εκφράζει το υπό εξέταση χαρακτηριστικό είναι συνεχής.
- Υπάρχει χωρική εξάρτηση των τιμών του υπό εξέταση χαρακτηριστικού.

ΚΑΤΗΓΟΡΙΕΣ

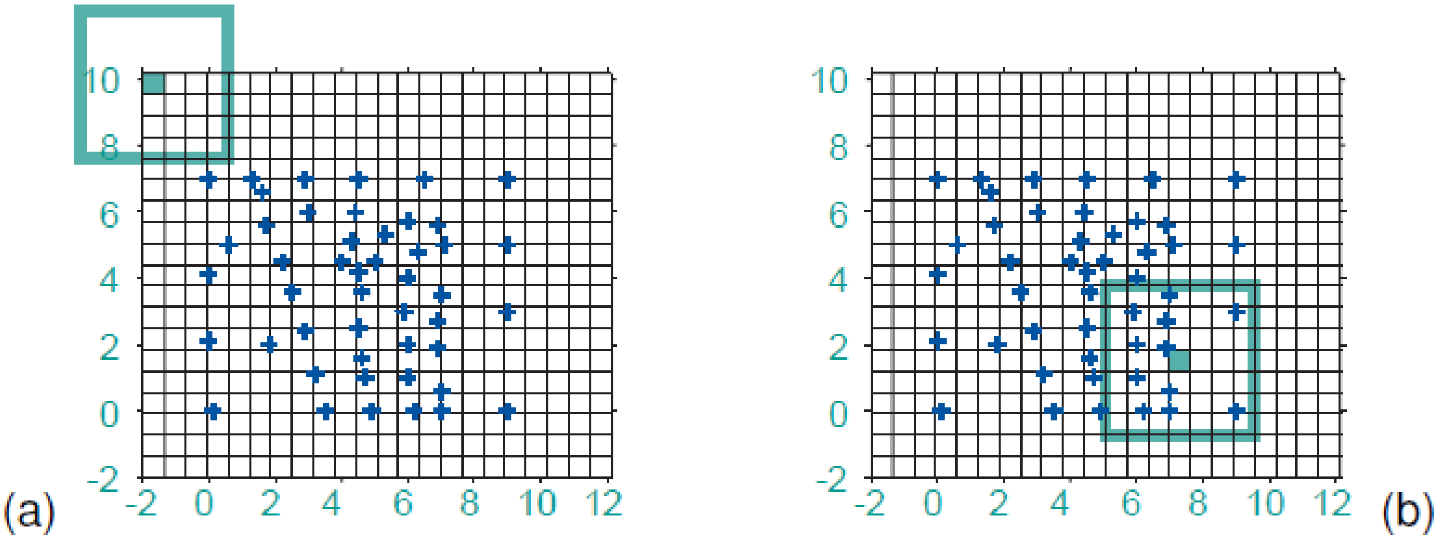
- Μέθοδοι τοπικών εκτιμήσεων *(local estimation)*
- Μέθοδοι γενικευμένων προσεγγίσεων *(global approximation)* και
- Γεωστατιστικές μέθοδοι χωρικής συσχέτισης *(kriging)*

ΜΕΘΟΔΟΙ ΤΟΠΙΚΩΝ ΕΚΤΙΜΗΣΕΩΝ

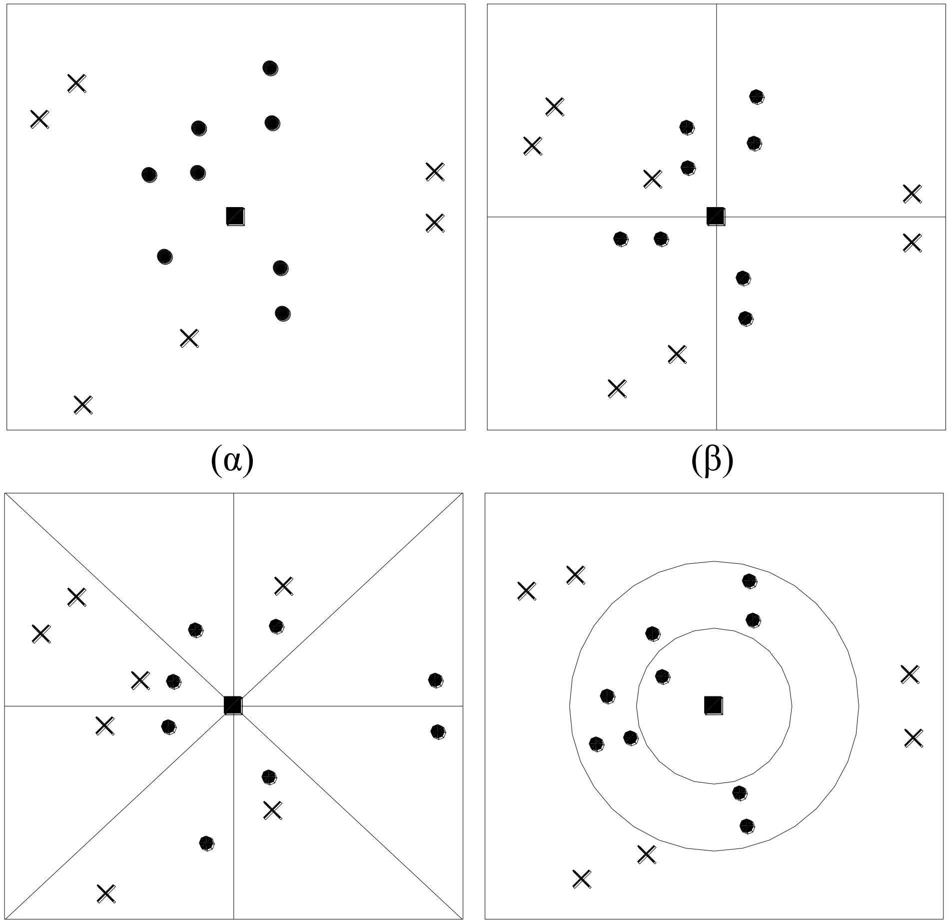
Οι μέθοδοι αυτές αναφέρονται στην εκτίμηση της τιμής ενός χαρακτηριστικού σε μια συγκεκριμένη θέση, με βάση στοιχεία τα οποία προέρχονται από σημεία που βρίσκονται στην άμεση γειτονική περιοχή του. Η διαδικασία απαιτεί τα εξής βήματα:

- Τον ορισμό της τοπικής περιοχής εκτίμησης ή την αποδεκτή «γειτονία» γύρω από το υπό εκτίμηση σημείο.
- Την εύρεση του αριθμού των σημείων που οι τιμές τους θα πρέπει να ληφθούν υπόψη για την εκτίμηση.
- Την επιλογή των σημείων αυτών από το σύνολο των σημείων της περιοχής μελέτης.
- Την επιλογή της μαθηματικής συνάρτησης που αντιπροσωπεύει τη διαφοροποίηση της τιμής του χαρακτηριστικού, δηλαδή τη διαδικασία εκτίμησης.

ΜΕΘΟΔΟΙ ΤΟΠΙΚΩΝ ΕΚΤΙΜΗΣΕΩΝ



ΕΠΙΛΟΓΗ ΣΗΜΕΙΩΝ



(α) (β)

(γ) (δ)

ΧΩΡΙΚΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ

- Η μέθοδος εκτίμησης αυτή βασίζεται στην εύρεση του μέσου όρου των τιμών που παρατηρούνται γύρω από το σημείο που έχει επιλεγεί.

- Στατιστικά θεωρείται ότι:

  - Υπάρχει μια σειρά από παρατηρήσεις **Z_i i=1,…,n** για ένα χωρικά συνεχές χαρακτηριστικό.
  - Οι μετρήσεις **Z_i** είναι παρατηρήσεις μιας χωρικής στοχαστικής διαδικασίας **$\{Z_i , s \in A\}$**, οι οποίες έχουν καταγραφεί σε αντίστοιχες χωρικές θέσεις **s_i** στην περιοχή μελέτης **A**.
  - Υποτίθεται ότι η εκτίμηση της απλής *(χωρίς βάρη)* μέσης τιμής, είναι μια ισοτροπική διαδικασία *(δεν παρατηρούνται «τάσεις» προς καμία κατεύθυνση)*.

ΧΩΡΙΚΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ

Η υπόθεση της ισοτροπίας δεν είναι πάντοτε σωστή και δεν επιτρέπει τη χωρική διαφοροποίηση στην κατανομή των σημείων που έχουν επιλεγεί.

Για την αντιμετώπιση του προβλήματος αυτού χρησιμοποιείται ο μέσος όρος των τιμών των παρατηρήσεων με βάρη. Πιο συγκεκριμένα:

$$\hat{\mu}(s) = \sum_{i=1}^{n} w_i(s) Z_i$$

όπου:

$$\sum_{i=1}^{n} w_i(s) = 1$$

Βασικό μέλημα της προσέγγισης αυτής είναι ο προσδιορισμός της συνάρτησης της απόστασης **$w_i(s)$**.

ΧΩΡΙΚΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ

Οι Bailey and Gatrell έχουν προτείνει τις εξής μορφές:

$$w_i(s) = h_i^{-a}$$

και

$$w_i(s) = e^{-ah_i}$$

Όπου:

h_i = η απόσταση μεταξύ του υπό εκτίμηση σημείου **s** και του σημείου **s_i**.

α = είναι μια παράμετρος που παίρνει τιμές, ώστε η εκτιμηθείσα επιφάνεια να είναι όσο το δυνατό πιο ομαλή.

Συνήθως, η **$w_i(s)$** ορίζεται να παίρνει τιμή μηδέν πέρα από μια συγκεκριμένη απόσταση.

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ

- Προδιαγραφή της πιο πιθανής εξασθένισης με την απόσταση
  - γραμμική: $w_{ij} = -b\ d_{ij}$
  - ύψωση σε αρνητική δύναμη: $w_{ij} = d_{ij}^{-b}$
  - αρνητική εκθετική: $w_{ij} = e^{-bdij}$
- Ισοτροπική και κανονική – ισχύει αυτό σε όλα τα γεωγραφικά φαινόμενα?
  - Αναγωγικές ή επαγωγικές προσεγγίσεις

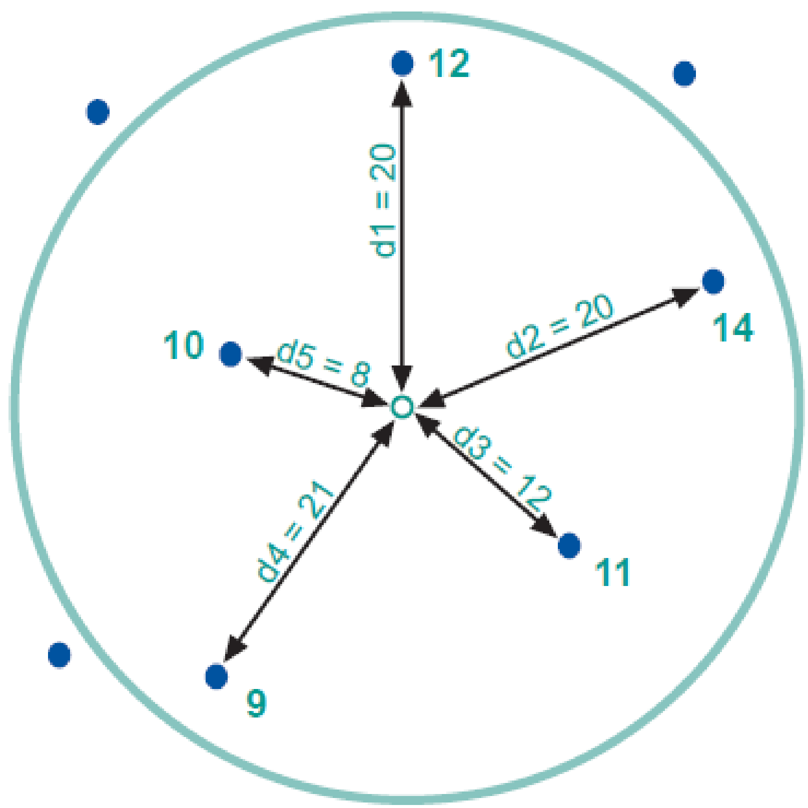
ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ

- Οι τιμές ενός συνεχούς πεδίου έχουν μετρηθεί σε ένα αριθμό σημείων δειγματοληψίας
- Προκύπτει η ανάγκη μέτρησης όλου του συνεχούς πεδίου
  - Εκτίμηση τιμών σε σημεία που δεν έγινε μέτρηση
  - Δημιουργία ενός χάρτη ισοϋψών καμπυλών με σχεδίαση ισοϋψών ανάμεσα στις θέσεις που έχουν μετρηθεί
- Οι μέθοδοι χωρικής παρεμβολής έχουν σχεδιαστεί για να επιλύουν αυτό το πρόβλημα

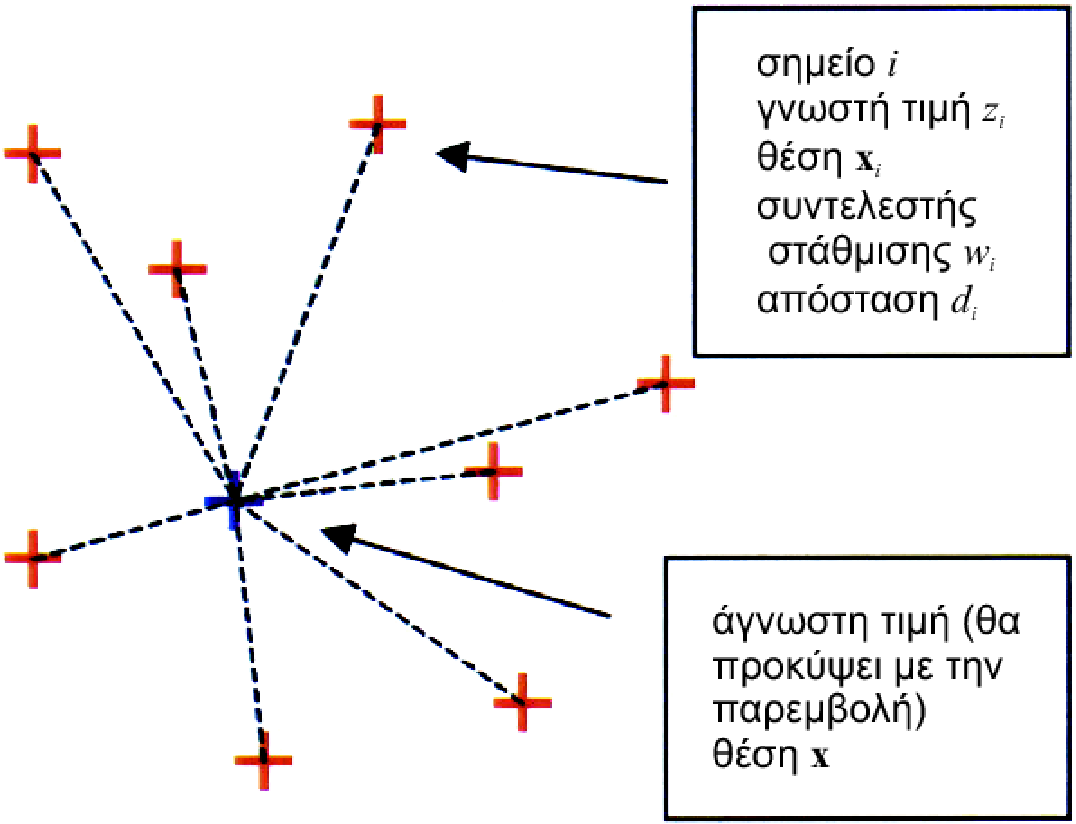
Στάθμιση αντίστροφης απόστασης

- **Inverse distance weighting (IDW)**
- Η εκτίμηση της άγνωστης τιμής του συνεχούς πεδίου σε ένα σημείο υπολογίζεται παίρνοντας μια μέση τιμή πάνω στις γνωστές τιμές
  - Κάθε γνωστή τιμή συμμετέχει με ένα βάρος σύμφωνα με την απόσταση από το σημείο, με τα πιο κοντινά σημεία να έχουν μεγαλύτερο βάρος
  - Αποτελεί μια υλοποίηση του Νόμου του Tobler

Στάθμιση αντίστροφης απόστασης IDW



Στάθμιση αντίστροφης απόστασης IDW



$$z(\mathbf{x}) = \sum_i w_i z_i \Big/ \sum_i w_i$$

Η εκτίμηση είναι μια σταθμισμένη μέση τιμή

$$w_i = 1 \big/ d_i^2$$

Το βάρος μειώνεται με την απόσταση

Μέθοδος σταθμισμένων αντίστροφων αποστάσεων (ΣΑΑ)

Η παρεμβολή γίνεται με βάση τη σχέση:

$$h = \frac{d_1^{-k}}{\sum_{n=1}^{N} d_n^{-k}} h_1 + \frac{d_2^{-k}}{\sum_{n=1}^{N} d_n^{-k}} h_2 + \ldots\ldots + \frac{d_N^{-k}}{\sum_{n=1}^{N} d_n^{-k}} h_N$$

όπου :

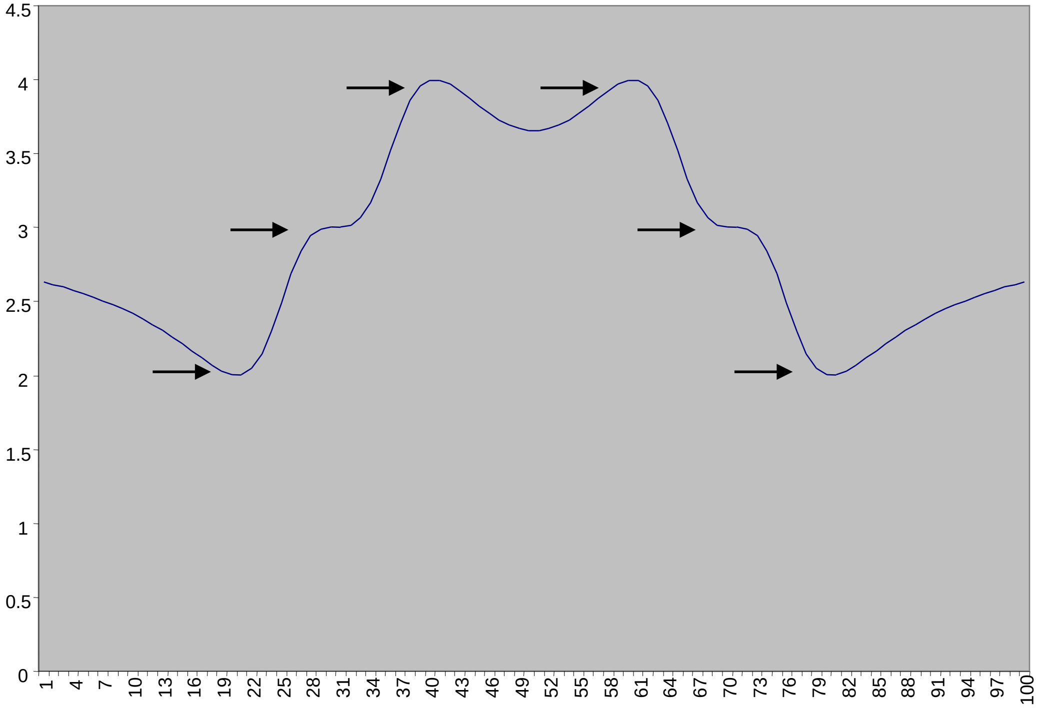
h	η τιμή της μεταβλητής στη ζητούμενη θέση
N	ο αριθμός των σημείων που συμμετέχουν
$h_1, h_2, h_3, \ldots, h_N$	οι σημειακές μετρήσεις στα σημεία 1, 2, 3, …, N
$d_1, d_2, d_3, \ldots, d_N$	οι αποστάσεις του κυττάρου από τα σημεία 1, 2, 3, …, N
k	ο συντελεστής επιρροής της απόστασης

Η τιμή του εκθέτη k συνήθως λαμβάνεται 1 ή 2 [Dingman, 1994].

Ζητήματα σχετικά με τη μέθοδο IDW

- Το εύρος των υπολογιζόμενων τιμών δεν μπορεί να ξεπερνά το εύρος των τιμών που έχουν μετρηθεί
  - Είναι σημαντικό στα σημεία δειγματοληψίας να συμπεριλαμβάνονται τα ακραία σημεία του συνεχούς πεδίου
  - Κάτι τέτοιο μπορεί να είναι πολύ δύσκολο

Πιθανά ανεπιθύμητο χαρακτηριστικό της παρεμβολής IDW



Αυτό το σύνολο των έξι σημείων δείχνει καθαρά ένα προφίλ λόφου. Αλλά σε περιοχές με λίγα ή καθόλου δεδομένα η παρεμβολή θα κινηθεί προς τη συνολική μέση τιμή. Η μπλε γραμμή δείχνει το προφίλ παρεμβολής με τη μέθοδο IDW

Μέθοδος Inverse Distance Weighting - IDW

- Η εξομάλυνση μπορεί να εξισορροπηθεί με τη χρήση ενός μέσου όρου με βάρη, κατά την οποία το βάρος που αποδίδεται σε σημείο είναι αντιστρόφως ανάλογο της απόστασής του από το σημείο παρεμβολής.
- Η αντίστροφη αναλογία μπορεί να είναι γραμμική ή εκθετική (δύναμη του δύο) ώστε να παρέχει μία αντίστροφη τετραγωνική σχέση.
- Ουσιαστικά η μέθοδος *IDW* ακολουθεί το συμπέρασμα ότι κάθε εκτιμώμενο σημείο παρεμβολής έχει μία τοπική επιρροή που μικραίνει με την αύξηση της απόστασης από τα μετρημένα σημεία του δείγματος.
- Η διαδικασία της εφαρμογής της μεθόδου απαιτεί τα εξής βήματα:
  1. τον ορισμό της τοπικής περιοχής εκτίμησης – περιοχής πλησιέστερου γείτονα
  2. την εύρεση του αριθμού των σημείων που οι τιμές τους θα πρέπει να ληφθούν υπόψη για την εκτίμηση
  3. την επιλογή των σημείων αυτών από το σύνολο των σημείων της περιοχής μελέτης και
  4. την επιλογή της μαθηματικής συνάρτησης που αντιπροσωπεύει τη διαφοροποίηση της τιμής του χαρακτηριστικού δηλαδή την διαδικασία εκτίμησης.

Μέθοδος Inverse Distance Weighting - IDW

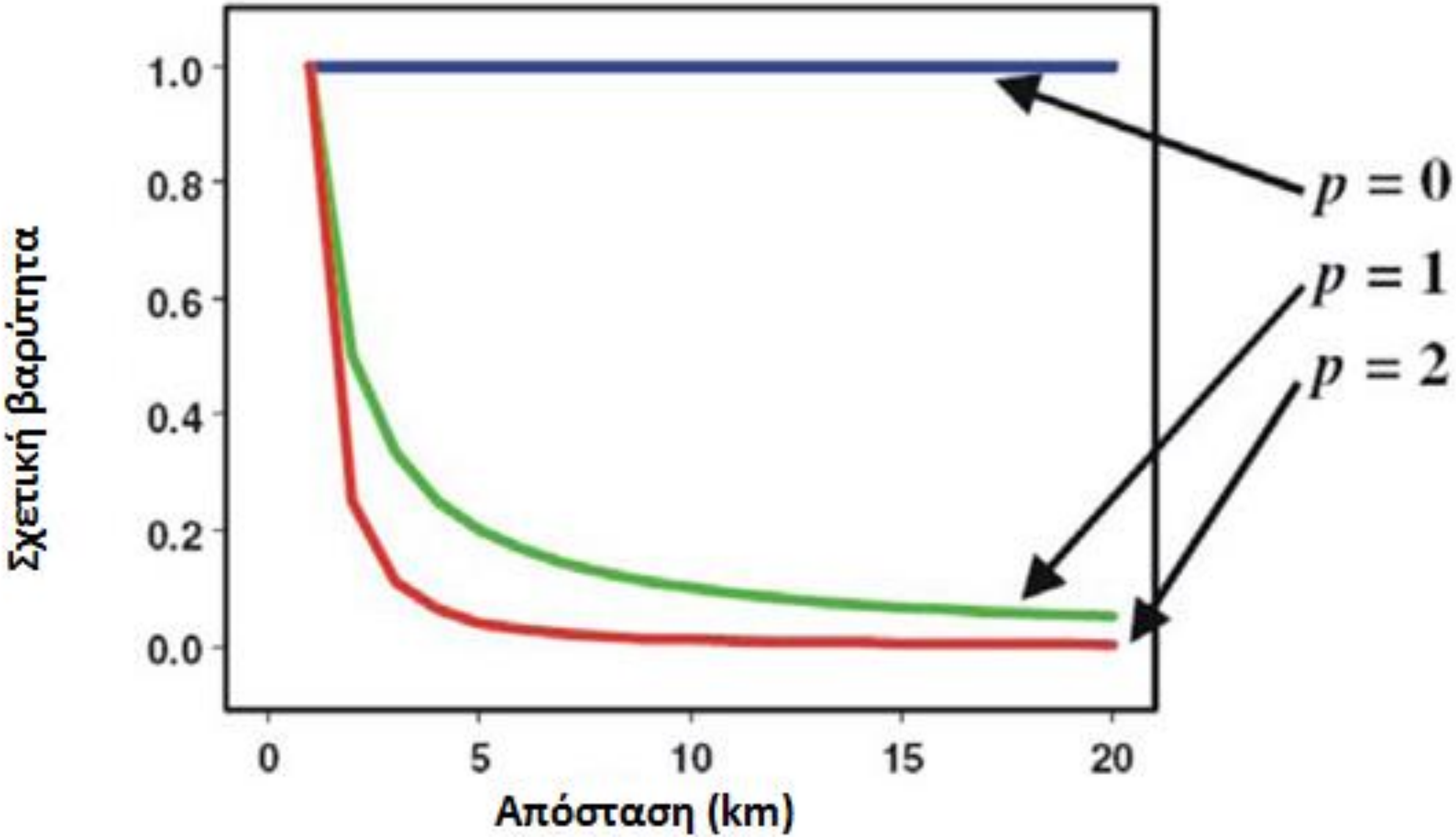
$$Z(u)=\frac{1}{\sum_{i=1}^{n}\lambda_{ui}}\sum_{i=1}^{n}\lambda_{ui}Z(x_i) \qquad \lambda_{ui}=\frac{1}{h_{ui}{}^{p}}$$

όπου:

- n είναι ο αριθμός των μετρημένων σημείων του δείγματος που χρησιμοποιούνται ως πλησιέστεροι γείτονες για την παρεμβολή της άγνωστης θέσης,
- λ_{ui} το βάρος που πρέπει να αποδοθεί σε κάθε σημείο του δείγματος στο σημείο x_i,
- p είναι ο εκθέτης (δύναμη) της απόστασης h_{ui}. Συνήθως η τιμή του εκθέτη είναι ίση με δύο (2) ώστε το βάρος να είναι αντιστρόφως ανάλογο προς την απόσταση από το σημείο του δείγματος.
- $Z(u)$ είναι η υπολογισμένη τιμή της παρεμβολής, $Z(x_i)$η τιμή των γνωστών δεδομένων στη θέση x_i, h_{ui} η απόσταση μεταξύ της άγνωστης (unsampled) θέσης u και της γνωστής θέσης x_i.

Πρέπει να σημειωθεί ότι: $\sum_{i=1}^{n}\lambda_{ui}=1$

Η επίδραση της δύναμης στη μέθοδο ΣΑΑ (Inverse Distance Weighting, IDW) Διάγραμμα βάρους – απόστασης



ΨΗΦΙΔΟΠΟΙΗΣΗ: TIN

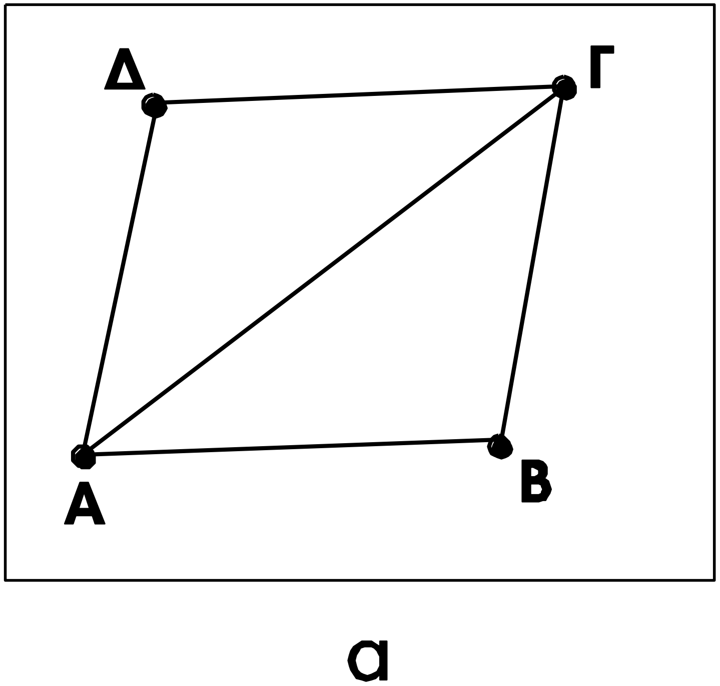
Η διαδικασία της δημιουργίας του TIN αφορά:

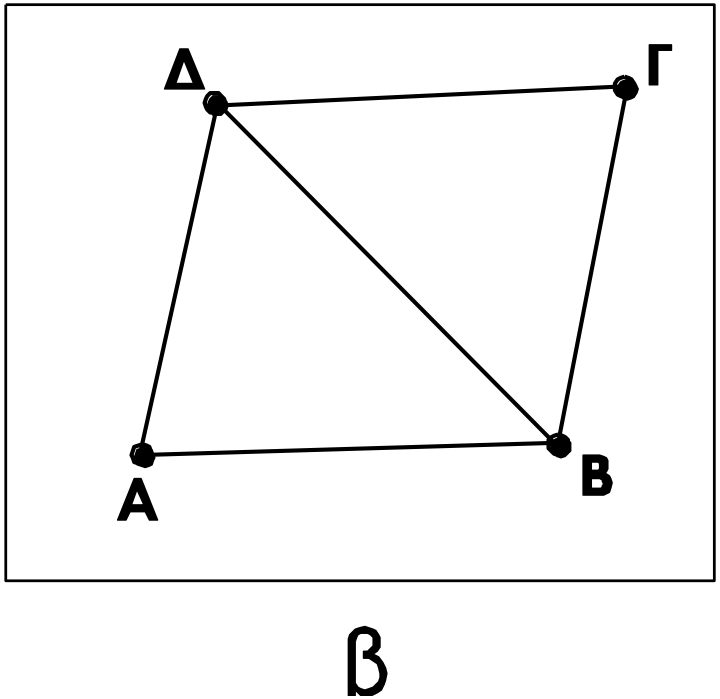
- Στη δημιουργία του πλέγματος των τριγώνων (όλα τα σημεία ενώνονται μεταξύ τους, μετατρεπόμενα σε ένα σύνολο πλευρών τριγώνων).
- Στον καθορισμό της συνάρτησης της χωρικής διαφοροποίησης των τιμών (η τιμή του χαρακτηριστικού μεταξύ των δύο κορυφών της πλευράς μεταβάλλεται με έναν καθορισμένο και σταθερό τρόπο).

ΨΗΦΙΔΟΠΟΙΗΣΗ: TIN

Η διαδικασία «**τριγωνοποίησης**» μπορεί να επιτευχθεί με διαφορετικούς τρόπους ανάλογα με το κριτήριο σύνδεσης των σημείων.

Από τις μεθόδους αυτές η πλέον γνωστή είναι η **μέθοδος Delaunay**, γνωστή και ως κριτήριο μέγιστη-ελάχιστη γωνία.

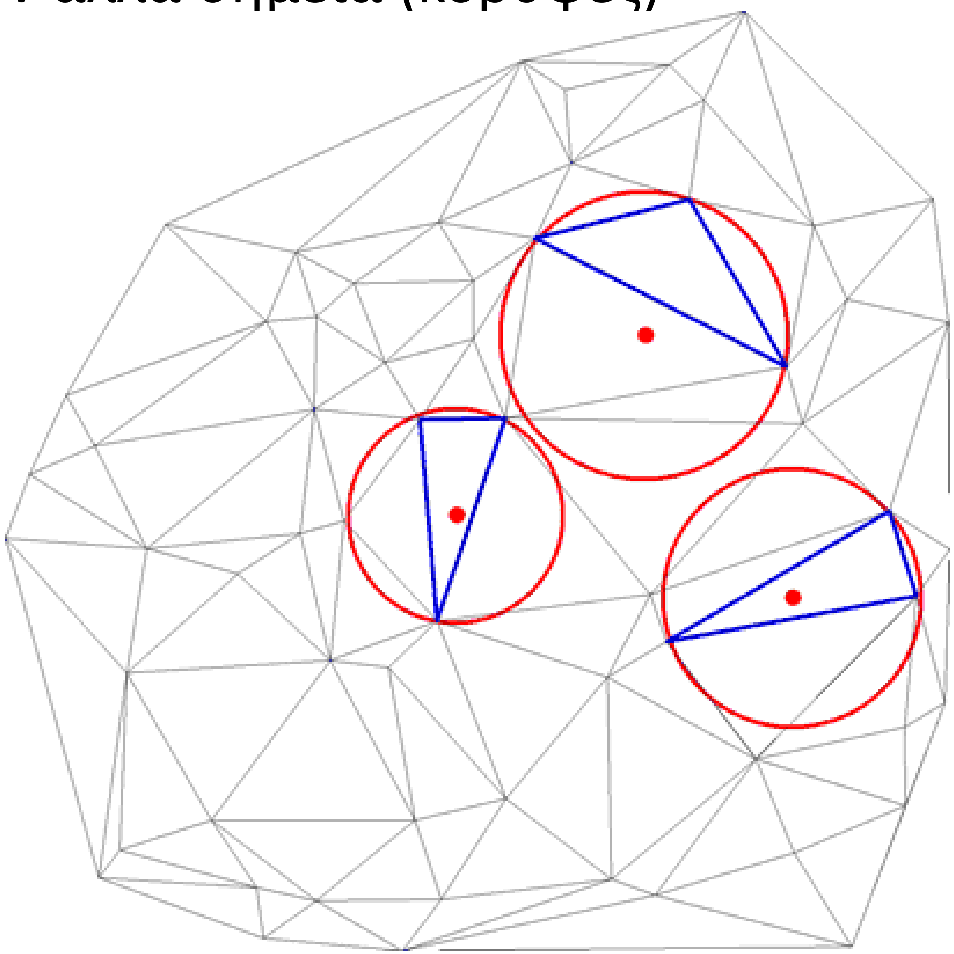




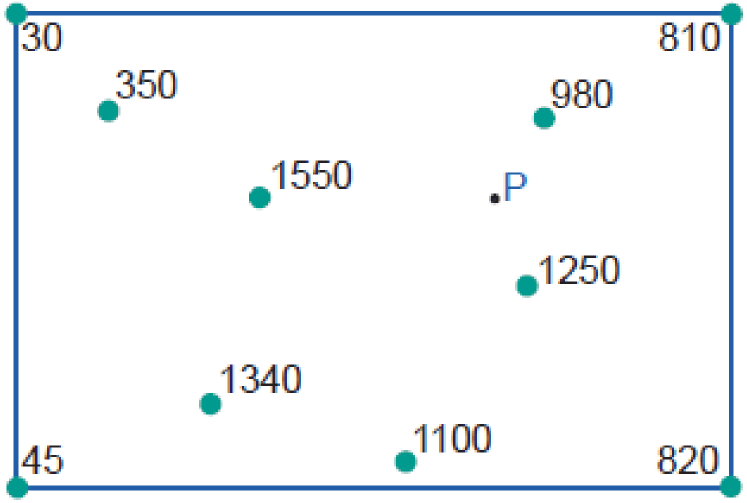
Ο **τριγωνισμός Delaunay** μεγιστοποιεί την ελάχιστη γωνία από όλες τις γωνίες του τριγώνου ώστε να αποφεύγονται τρίγωνα με μικρό ύψος (skinny triangles)

ΨΗΦΙΔΟΠΟΙΗΣΗ: ΤΙΝ

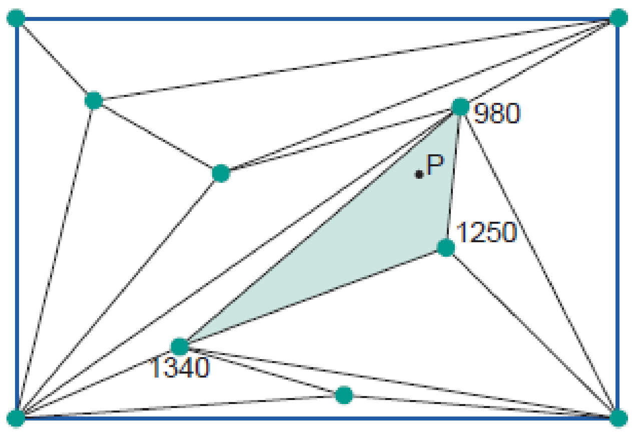
Ο **Delaunay** ορίζει ότι τρία σημεία σχηματίζουν τρίγωνα **Delaunay** εάν και μόνο αν ο κύκλος που περνάει από τα τρία σημεία δεν περιέχουν άλλα σημεία (κορυφές)

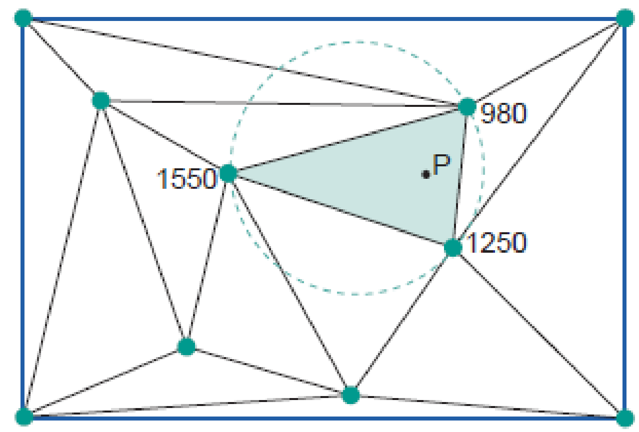


ΨΗΦΙΔΟΠΟΙΗΣΗ: TIN



Two triangulations based on the input locations of top Figure (a) one with many ‘stretched’ triangles; (b) the triangles are more equilateral; this is a Delaunay triangulation.

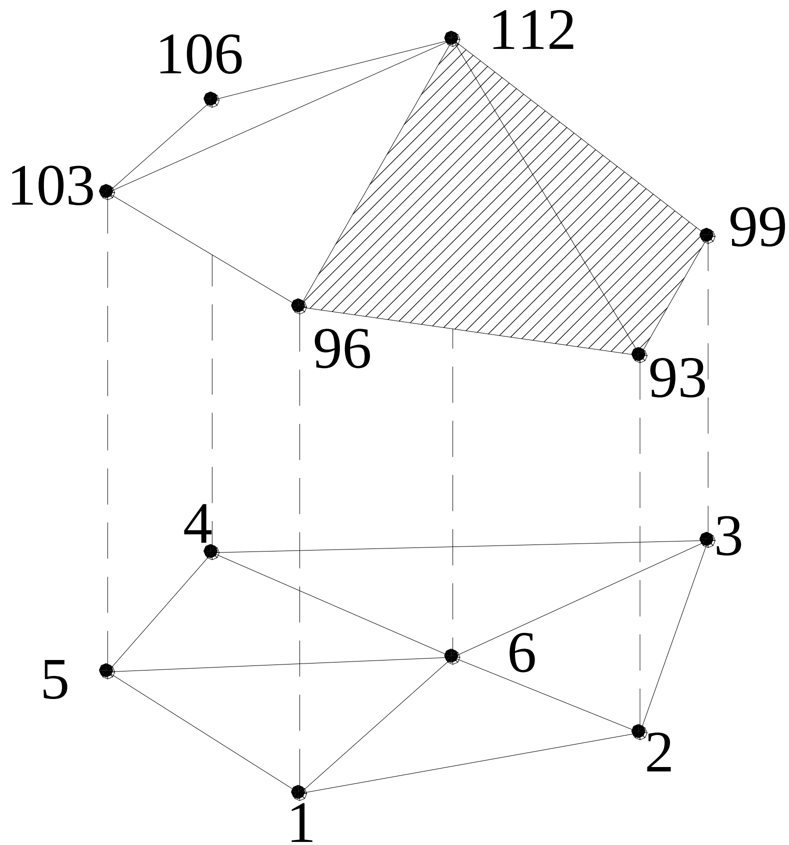




ΨΗΦΙΔΟΠΟΙΗΣΗ: TIN

Με την ολοκλήρωση της «τριγωνοποίησης», ο ορισμός της χωρικής συνάρτησης, μπορεί να αρχίσει και μπορεί να πάρει πολλές μορφές.

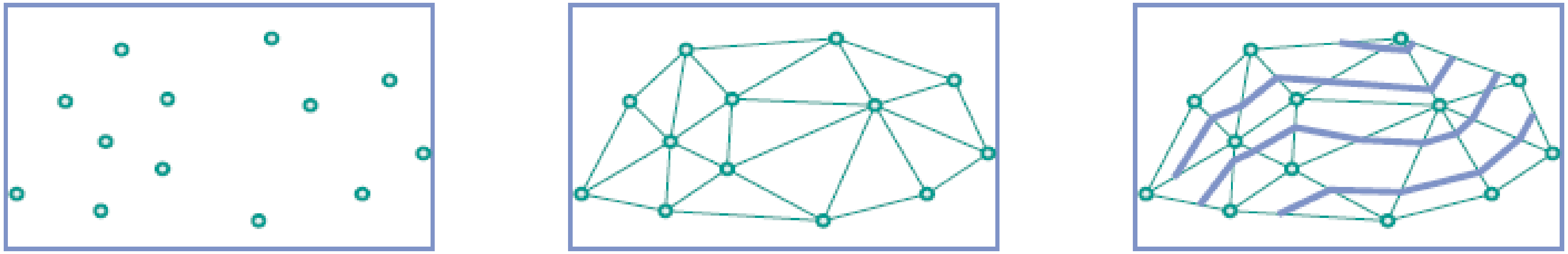
Η πιο απλή: γραμμική παρεμβολή **$Z_i = \alpha X_i + \beta Y_i + \gamma$**



ΨΗΦΙΔΟΠΟΙΗΣΗ: TIN

Triangulation ως μέθοδος **χωρικής παρεμβολής**

(a) known point measurements;
(b) constructed triangulation on known points;
(c) isolines constructed from the triangulation.

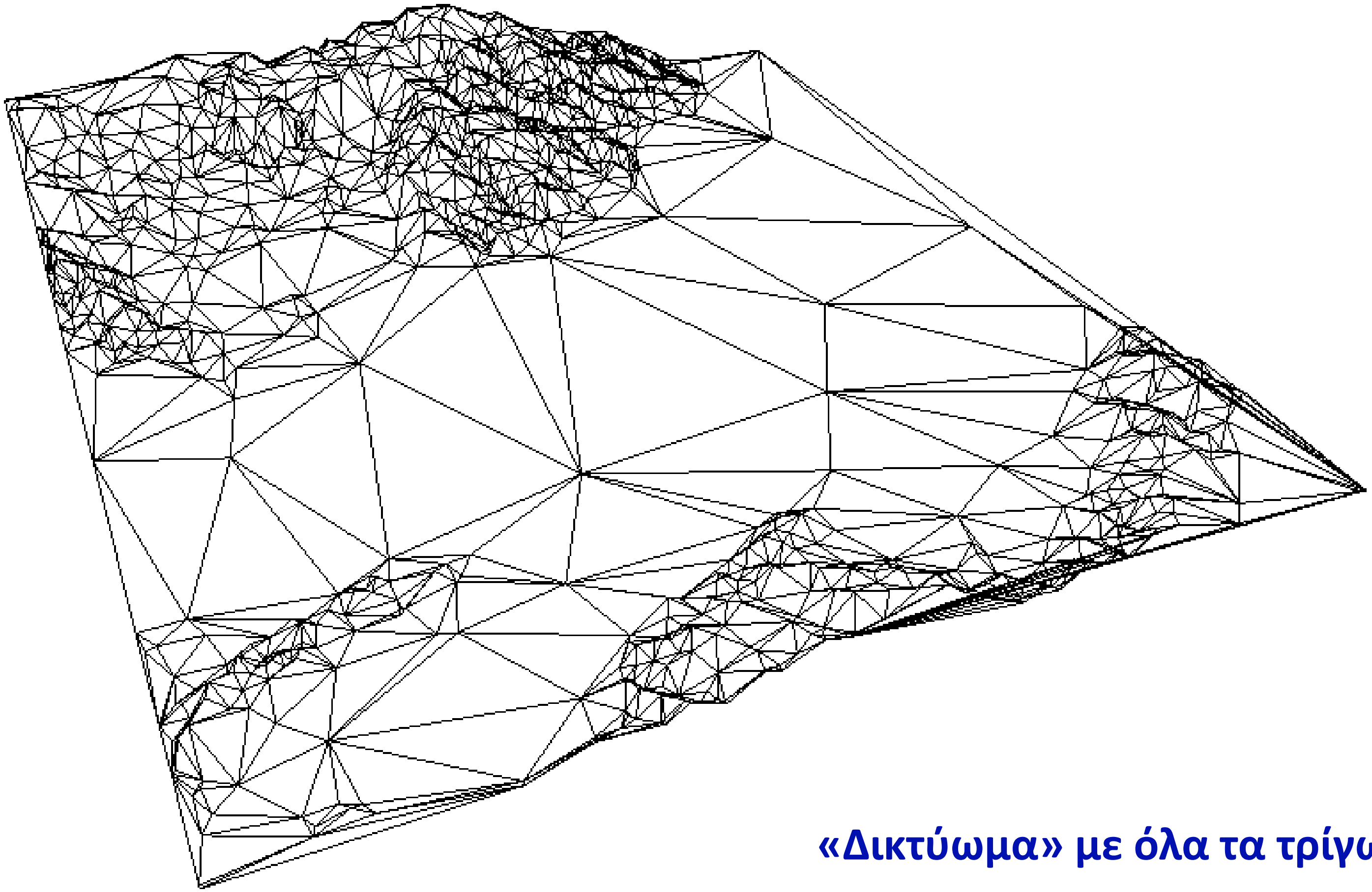


(a) (b) (c)

Δημοφιλής μέθοδος χωρικής παρεμβολής για TINs:

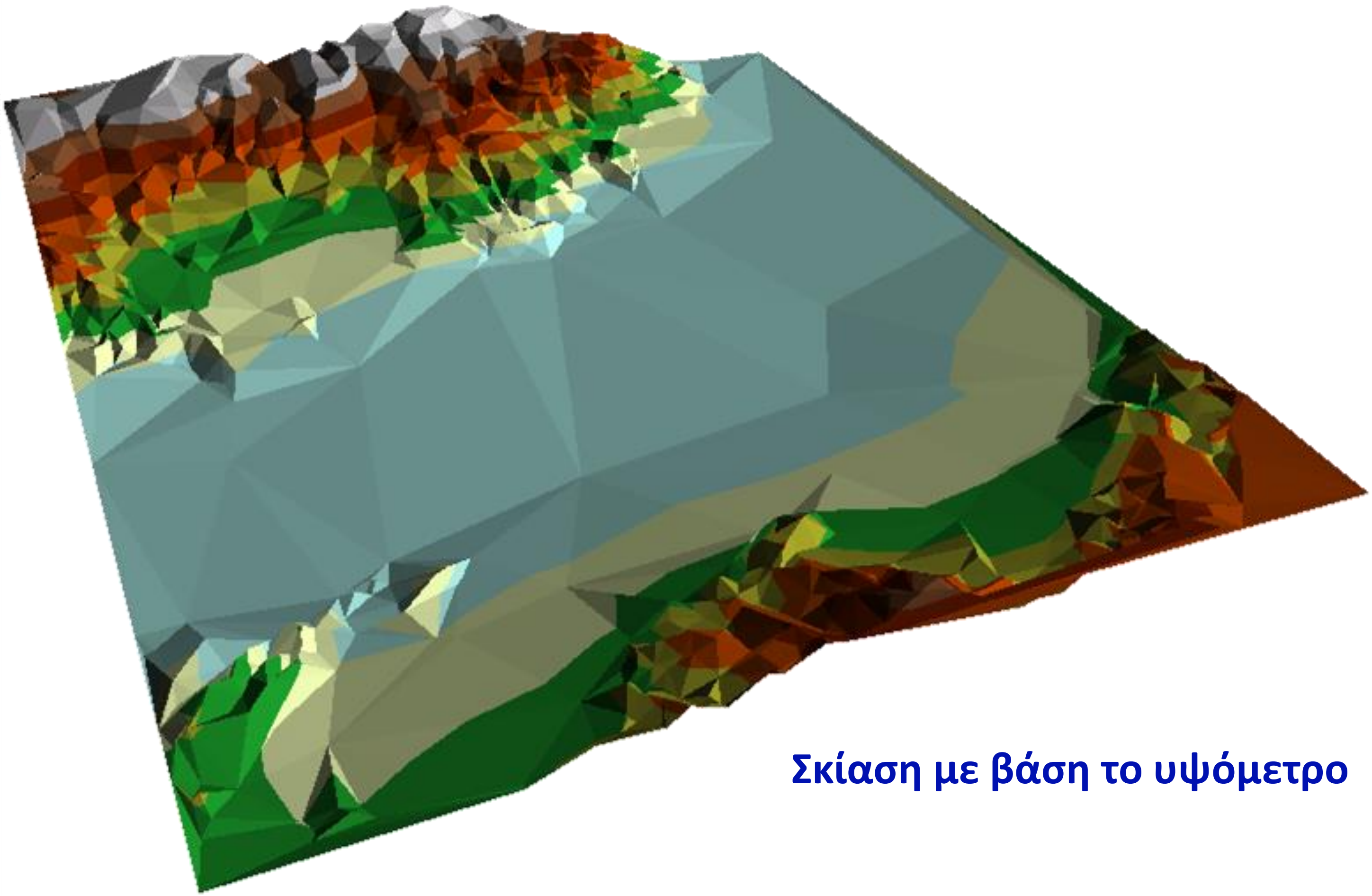
- **Διμεταβλητή 5ου βαθμού πολυωνυμική εξίσωση των συντεταγμένων X και Y**

Η επιφάνεια της Death Valley, California, σε μορφή TIN



«Δικτύωμα» με όλα τα τρίγωνα

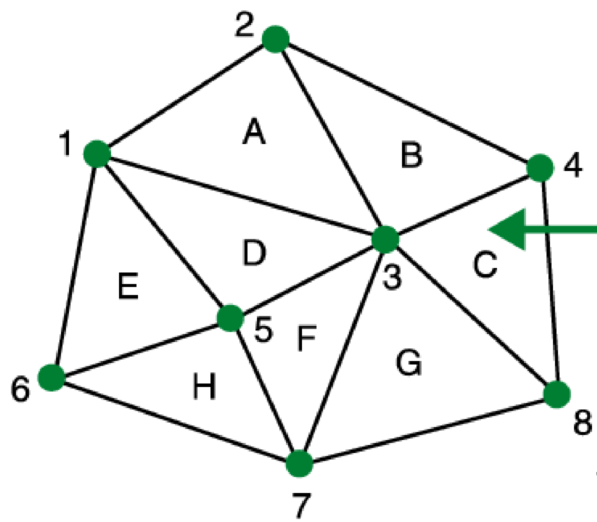
Η επιφάνεια της Death Valley, California, σε μορφή TIN



Σκίαση με βάση το υψόμετρο

Triangulated Irregular Network – TIN

Ένα TIN είναι μια δομή τοπολογικών δεδομένων
η οποία διαχειρίζεται πληροφορίες για τους κόμβους
που απαρτίζουν κάθε τρίγωνο και τα γειτονικά του



Τρίγωνο	Κατάλογος κόμβων	Γείτονες
A	1, 2, 3	-, B, D
B	2, 4, 3	-, C, A
C	4, 8, 3	-, G, B
D	1, 3, 5	A, F, E
E	1, 5, 6	D, H, -
F	3, 7, 5	G, H, D
G	3, 8, 7	C, -, F
H	5, 7, 6	F, -, E

Τα τρίγωνα έχουν πάντα τρεις κόμβους και συνήθως έχουν τρία γειτονικά τρίγωνα.
Τα τρίγωνα στην περιφέρεια του TIN μπορούν να έχουν ένα ή δύο γειτονικά τρίγωνα.

ΨΗΦΙΔΟΠΟΙΗΣΗ:
ΠΟΛΥΓΩΝΑ THIESSEN

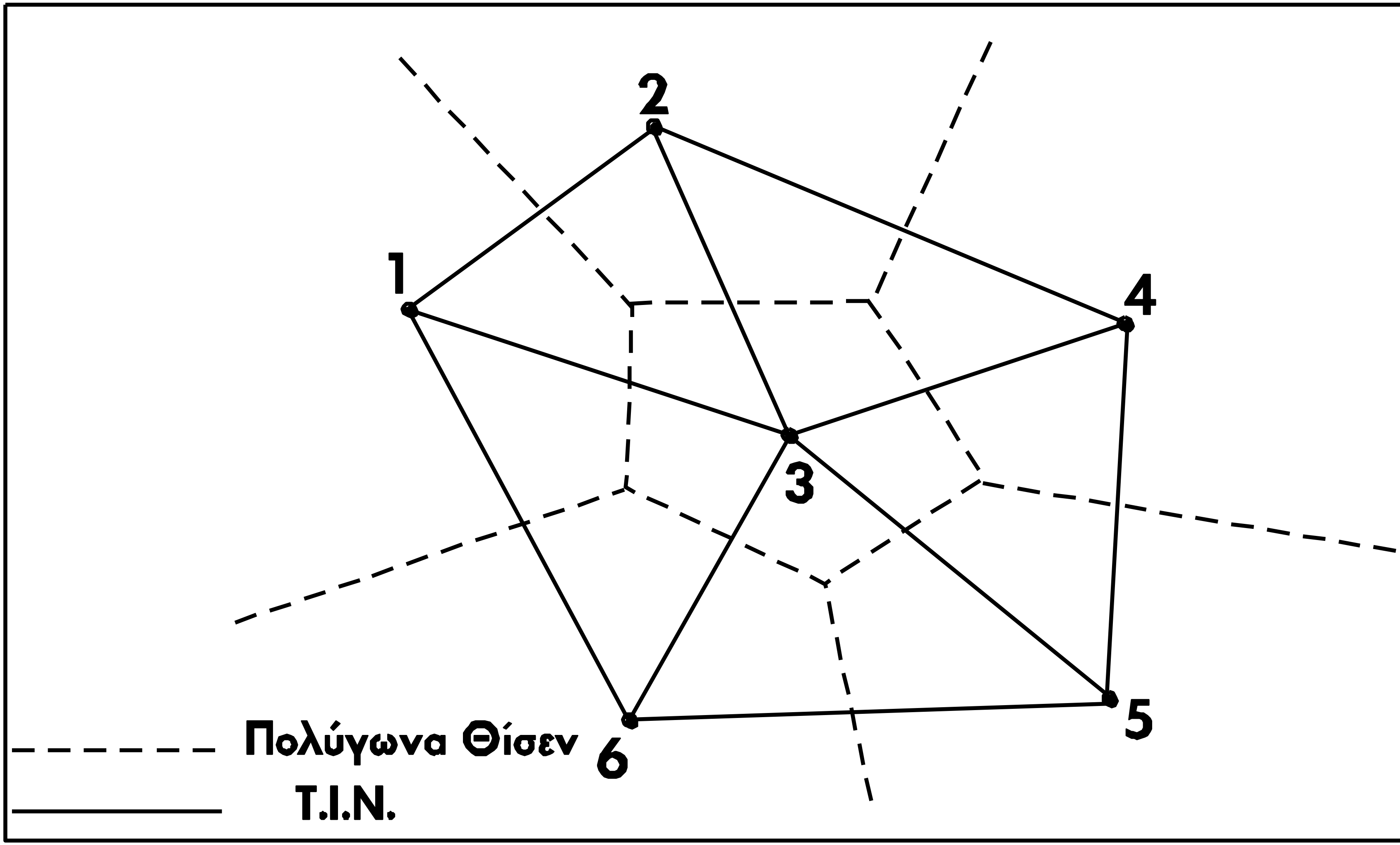
- Αν υπάρχουν **n** γνωστά σημεία σε μια περιοχή μελέτης **A,** σε κάθε σημείο s_i κατανέμεται ένα τμήμα της **A**, έτσι ώστε κάθε σημείο του τμήματος αυτού να είναι πλησιέστερο στο s_i περισσότερο από κάθε άλλο σημείο s_j.

- Οι γραμμές μεταξύ δύο σημείων αποτελούν το γεωμετρικό τόπο των σημείων που ισαπέχουν από τα σημεία αυτά.

- Κάθε πλευρά ενός πολυγώνου **Thiessen** είναι ένα γραμμικό τμήμα που τέμνει κάθετα τη γραμμή που ενώνει το σημείο του πολυγώνου με το γειτονικό του, που ενώνει δηλαδή την πλευρά ενός τριγώνου του **TIN**.

ΨΗΦΙΔΟΠΟΙΗΣΗ:

ΠΟΛΥΓΩΝΑ THIESSEN



ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ

- Οι **γενικευμένες μέθοδοι** εκτίμησης χρησιμοποιούν όλα τα υπάρχοντα στοιχεία, επιτυγχάνοντας εκτιμήσεις για το σύνολο της περιοχής που ενδιαφέρει.

- Οι **γενικευμένοι εκτιμητές** χρησιμοποιούνται συνήθως εμμέσως για χωρικές παρεμβολές (αποτελούν εργαλεία εξέτασης και απομάκρυνσης των γενικευμένων χωρικών διαφοροποιήσεων).

- Οι **διαδικασίες εκτίμησης** των γενικευμένων μεθόδων υπολογίζονται σχετικά απλά και βασίζονται κυρίως σε συνηθισμένες στατιστικές μεθόδους ανάλυσης διασποράς.

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

- Όταν η διαφοροποίηση σε ένα χαρακτηριστικό είναι συνεχής στο χώρο, υπάρχει ανάγκη για τη δημιουργία ενός μοντέλου, ώστε να μπορεί:
  - Να εξηγηθεί αυτή η διαφοροποίηση.
  - Να εκτιμηθεί η τιμή του χαρακτηριστικού σε θέσεις εκτός αυτών για τις οποίες υπάρχουν στοιχεία.

- Η μέθοδος της **ανάλυσης επιφάνειας τάσης** επιτυγχάνει το διαχωρισμό των παρατηρήσεων μιας χωρικά κατανεμημένης μεταβλητής σε:
  - Ένα τμήμα που σχετίζεται με τις γενικευμένες τάσεις που υπάρχουν.
  - Ένα τμήμα που είναι το αποτέλεσμα των τοπικών επιδράσεων.

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

- Επομένως, κάθε τιμή **Z(s)** ενός φαινομένου μπορεί να διαχωριστεί σε δύο μέρη, που το καθένα είναι αποτέλεσμα μιας ξεχωριστής, σε διαφορετική κλίμακα, χωρικής διαδικασίας.
  - Το πρώτο τμήμα είναι αποτέλεσμα μιας διαδικασίας **μεγάλης κλίμακας**, λειτουργεί δηλαδή σε μια μεγάλη περιοχή και δημιουργεί την επιφάνεια τάσης.
  - Το δεύτερο τμήμα συνδυάζει τις τυχαίες μεταβολές και τα σφάλματα μέτρησης και είναι αποτέλεσμα μιας χωρικής διαδικασίας που επιδρά σε μια σημαντικά **μικρότερη περιοχή** από την περιοχή μελέτης.

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

Για τη χωρική διαφοροποίηση σε ένα χαρακτηριστικό ισχύει η σχέση:

παρατηρούμενη τιμή σε ένα σημείο	**=**	**τιμή επιφάνειας τάσης στο σημείο αυτό**	**+**	**υπόλοιπο στο σημείο αυτό**

Η βασική εξίσωση κάθε χωρικού προτύπου να δίνεται από τη σχέση:

$$\mathbf{Z_i = f(X_i, Y_i) + \varepsilon_i}$$

όπου:

Z_i = η παρατηρούμενη τιμή της μεταβλητής Z στο σημείο s_i (X_i, Y_i)

$f(X_i, Y_i)$ = η τιμή της επιφάνειας τάσης στο σημείο s_i (X_i, Y_i)

ε_i = το υπόλοιπο

Η μορφή της συνάρτησης **f** μπορεί να ποικίλλει από απλή γραμμική μέχρι και πολύπλοκου τετάρτου ή μεγαλύτερου βαθμού.

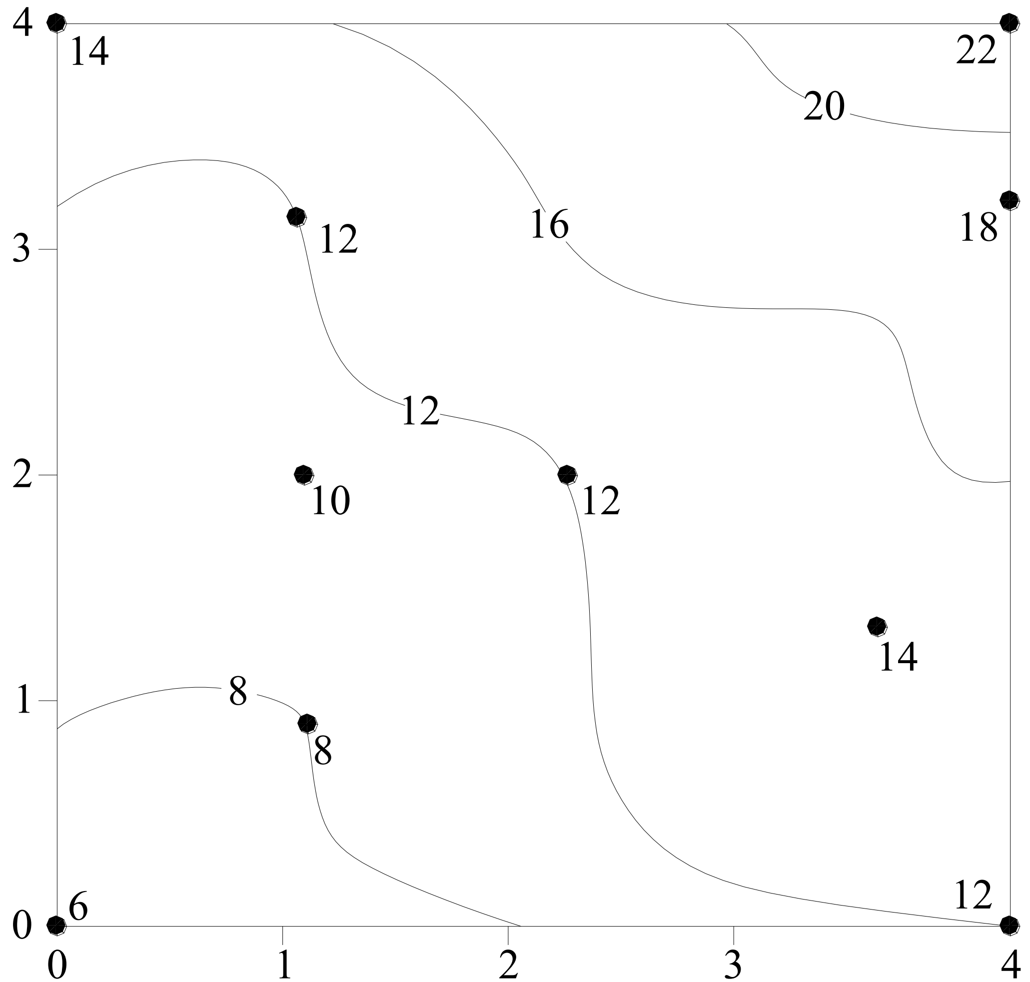
ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

ΠΑΡΑΔΕΙΓΜΑ

Αρ. Σημείου	X	Y	Z	$\hat{Z}$ (εκτιμηθείσα)	ε
1	0	0	6	4.94593	+1.05407
2	1	1	8	8.92296	-0.92296
3	2	1	11	11.02943	-0.022944
4	3	1	12	13.13591	-1.13591
5	4	0	14	13.37182	+0.62818
6	2	2	12	12.90000	-0.90000
7	1	3	14	12.66409	+1.33591
8	0	4	12	12.42818	-0.42818
9	3	4	18	18.74760	-0.74760
10	4	4	22	20.85407	+1.14593

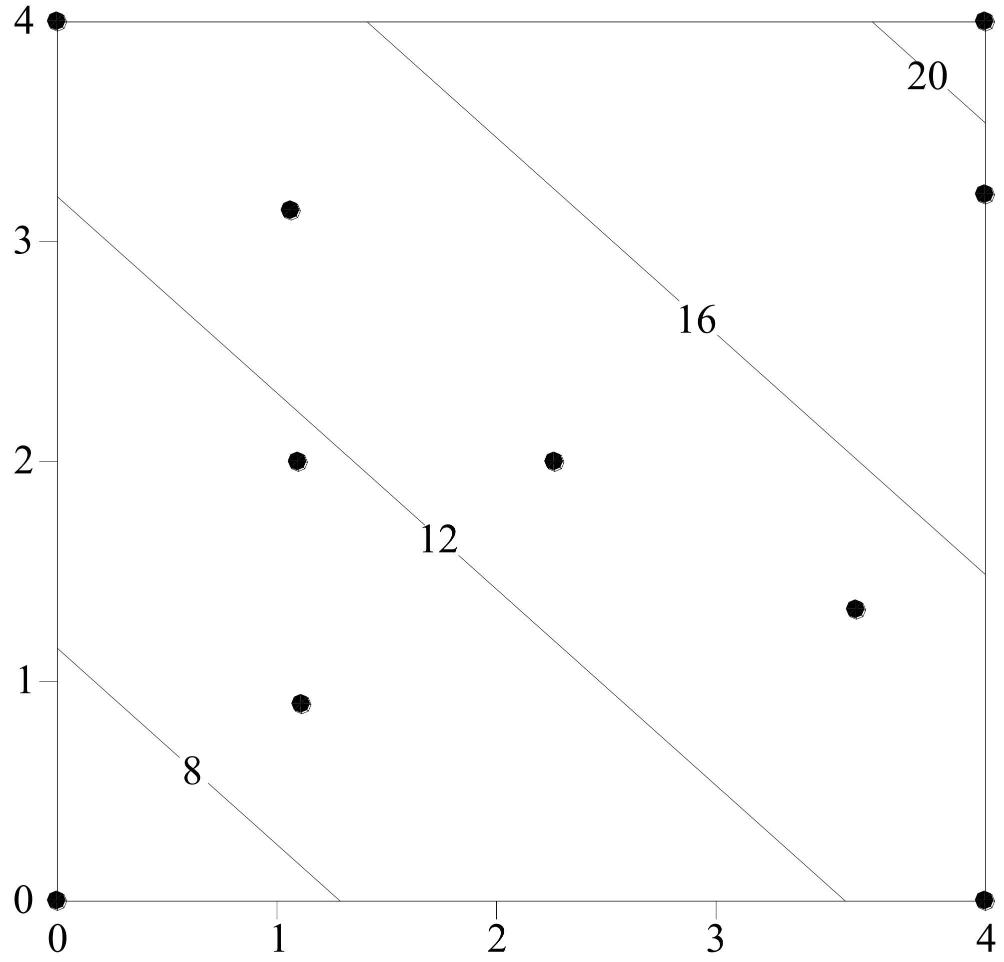
ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

Παράδειγμα: Αρχικός Χάρτης



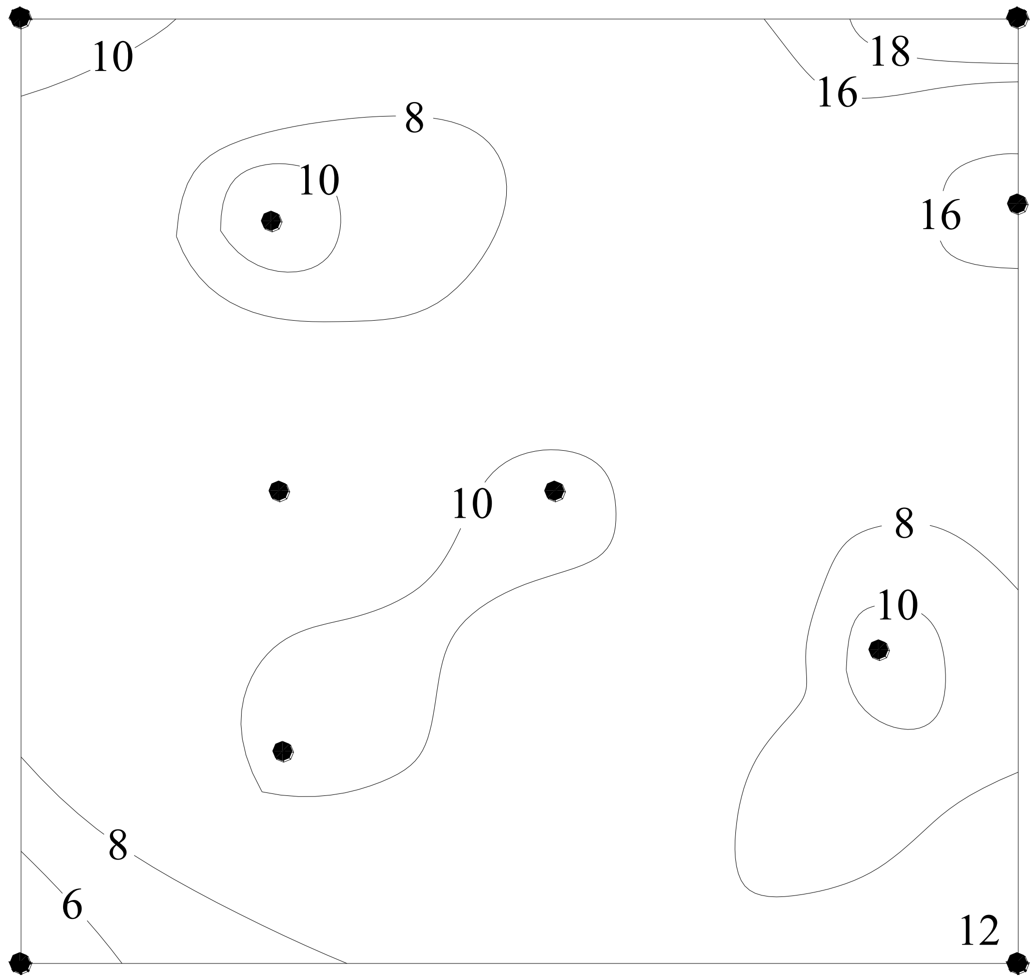
ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

Παράδειγμα: Γραμμική Επιφάνειας Τάσης



ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

Παράδειγμα: Υπόλοιπα

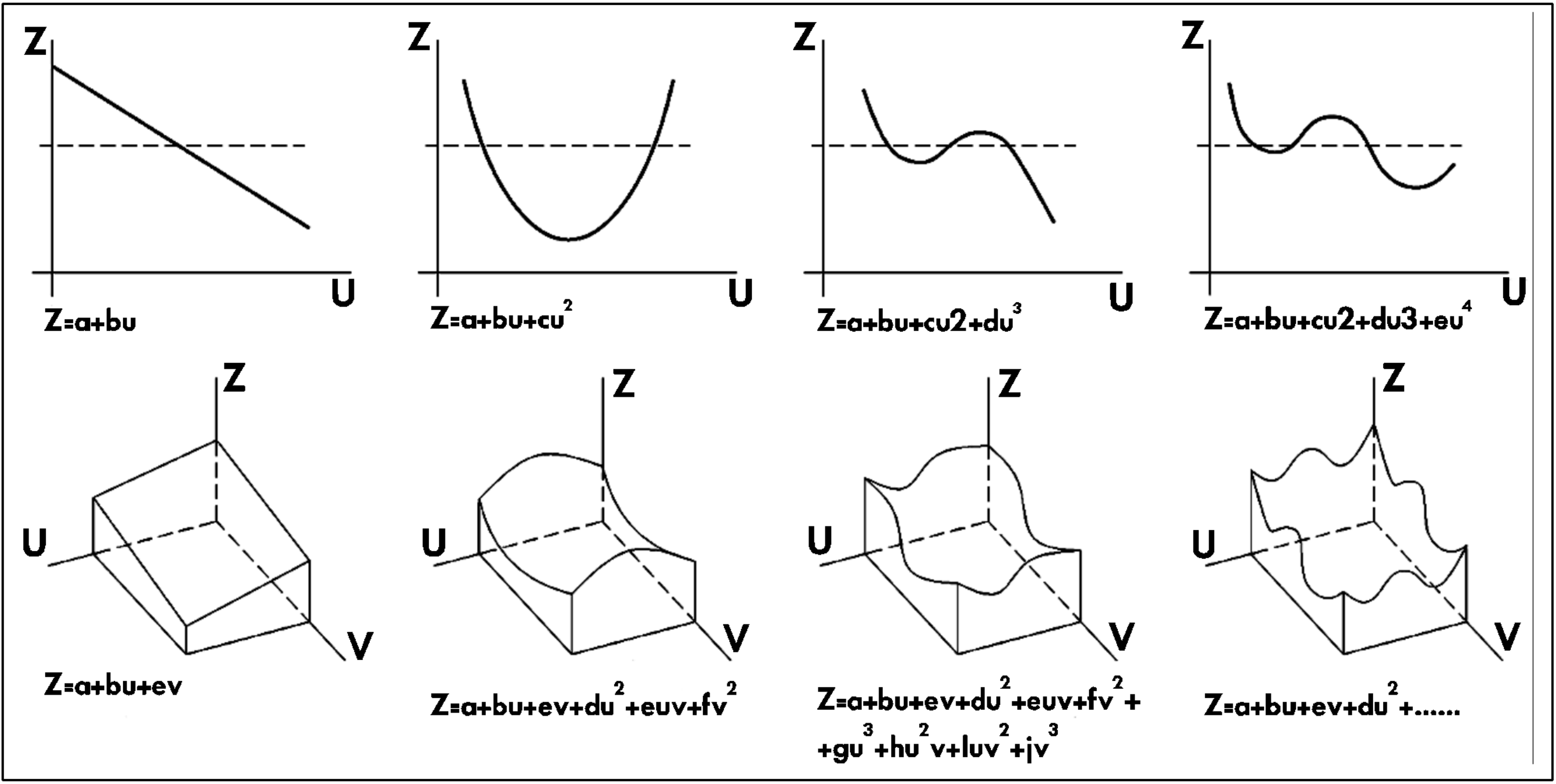


ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ - ΜΕΙΟΝΕΚΤΗΜΑΤΑ

- Στην παλινδρόμηση θεωρείται ότι οι τιμές των σφαλμάτων ε_i είναι ανεξάρτητες.
- Υποτίθεται ότι στη χωρική στοχαστική διαδικασία που εξετάζεται, παρατηρούνται μόνον πρώτου και όχι δεύτερου βαθμού διαφοροποιήσεις.
- Η διασπορά των ε_i δεν είναι σταθερή σε ολόκληρη την περιοχή μελέτης.

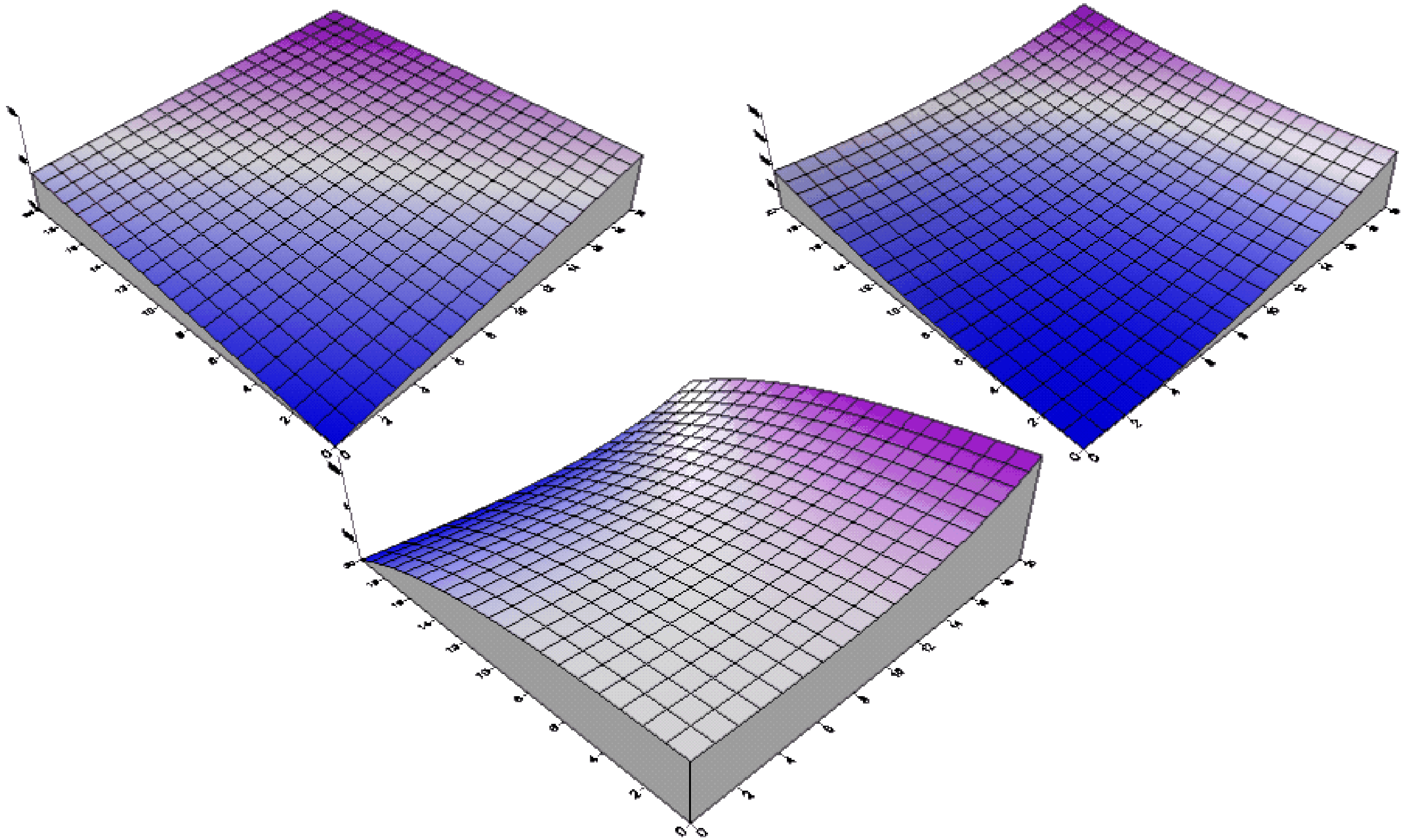
Ο στατιστικός έλεγχος και οι εκτιμήσεις της παλινδρόμησης δεν μπορούν να γίνουν αποδεκτές.

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ



ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

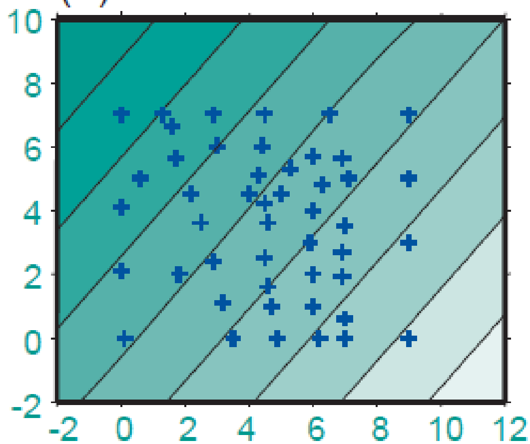
First, second and third order mathematical surfaces



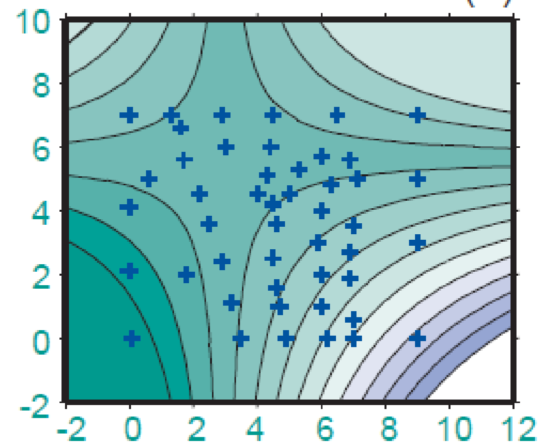
ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

(a) simple tilted plane; (b) bilinear saddle;
(c) quadratic surface; (d) cubic surface

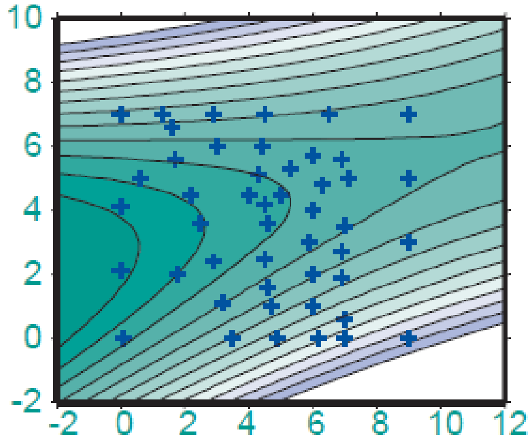
(a) $f(x,y) = c_1 \cdot x + c_2 \cdot y + c_3$



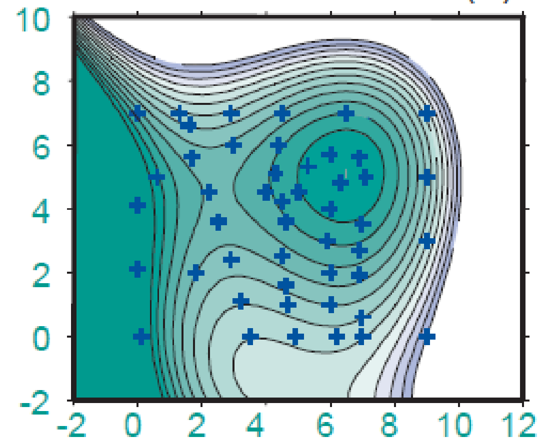
(b) $f(x,y) = c_1 \cdot x + c_2 \cdot y + c_3 \cdot xy + c_4$



(c) $f(x,y) = c_1 \cdot x^2 + c_2 \cdot x + c_3 \cdot y^2 + c_4 \cdot y + c_5 \cdot xy + c_6$



(d)
$$\begin{aligned} f(x,y) &= c_1 \cdot x^3 + c_2 \cdot x^2 + c_3 \cdot x + \\ &\quad c_4 \cdot y^3 + c_5 \cdot y^2 + c_6 \cdot y + \\ &\quad c_7 \cdot x^2 y + c_8 \cdot xy^2 + c_9 \cdot xy + c_{10} \end{aligned}$$



ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΕΠΙΦΑΝΕΙΑ ΤΑΣΗΣ

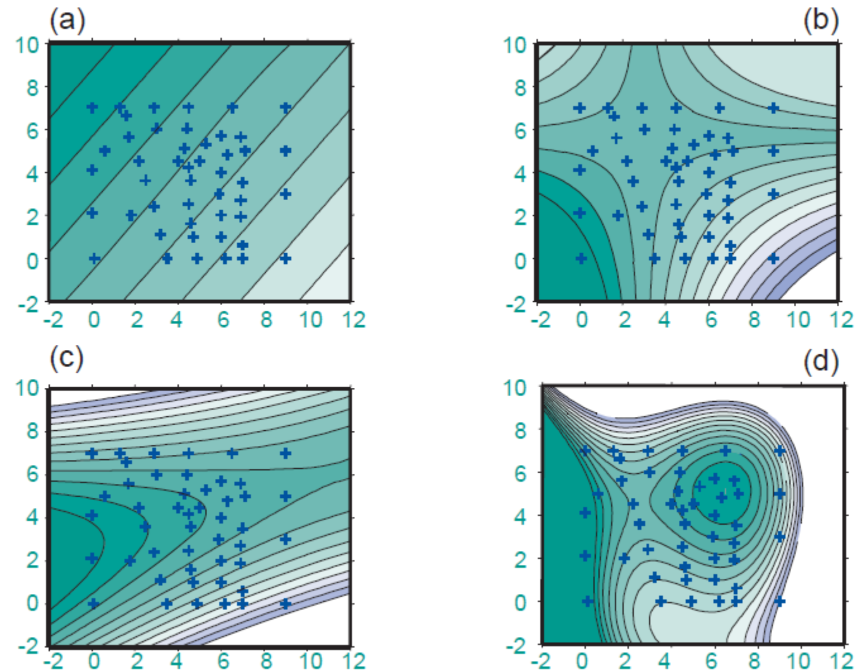
(a) simple tilted plane; (b) bilinear saddle;
(c) quadratic surface; (d) cubic surface

$$f(x,y) = c_1 \cdot x + c_2 \cdot y + c_3$$

$$f(x,y) = c_1 \cdot x + c_2 \cdot y + c_3 \cdot xy + c_4$$

$$f(x,y) = c_1 \cdot x^2 + c_2 \cdot x + c_3 \cdot y^2 + c_4 \cdot y + c_5 \cdot xy + c_6$$

$$\begin{aligned} f(x,y) \;=\; & c_1 \cdot x^3 + c_2 \cdot x^2 + c_3 \cdot x + \\ & c_4 \cdot y^3 + c_5 \cdot y^2 + c_6 \cdot y + \\ & c_7 \cdot x^2 y + c_8 \cdot x y^2 + c_9 \cdot xy + c_{10} \end{aligned}$$



Μοντέλα

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
(a)	-1.83934	1.61645	70.8782							
(b)	-5.61587	-2.95355	0.993638	89.0418						
(c)	0.000921084	-5.02674	-1.34779	7.23557	0.813177	76.9177				
(d)	-0.473086	6.88096	31.5966	-0.233619	1.48351	-2.52571	-0.115743	-0.052568	2.16927	96.8207

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΜΟΝΤΕΛΑ ΤΑΞΙΝΟΜΗΣΗΣ

- Οι τάσεις θεωρούνται ότι είναι το αποτέλεσμα της ύπαρξης ενός υπόβαθρου αποτελούμενου από μια ομάδα περιοχών οι οποίες διακρίνονται μεταξύ τους από τις μέσες τιμές του υπό εξέταση χαρακτηριστικού.
- Υποθέτει ότι η παρατηρούμενη δομή μιας συγκεκριμένης χωρικής διαφοροποίησης καθορίζεται από εξωγενώς καθοριζόμενες χωρικές ενότητες.
- Στη χρησιμοποιούμενη μέθοδο της ανάλυσης διασποράς, θεωρείται ότι εντός κάθε χωρικής ενότητας η διασπορά των τιμών είναι μικρότερη από τη διασπορά των τιμών μεταξύ των ενοτήτων.
- Επομένως, σημαντικές αλλαγές στις τιμές της παρατηρούμενης μεταβλητής παρατηρούνται στα όρια μεταξύ των χωρικών μονάδων.

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΜΟΝΤΕΛΑ ΤΑΞΙΝΟΜΗΣΗΣ

Το στατιστικό μοντέλο ταξινόμησης έχει ως εξής:

$$Z(s) = \mu + v_\kappa + \varepsilon$$

όπου:

Z(s) = η τιμή του χαρακτηριστικού στη θέση **s**.

μ = η συνολική, για ολόκληρη την περιοχή μελέτης, μέση τιμή.

v_κ = η απόκλιση από τη **μ**, της μέσης τιμής κάθε χωρικής μονάδας **κ**, και

ε = το υπόλοιπο, σφάλμα, που είναι γνωστό και ως «θόρυβος».

Το μοντέλο θεωρεί ότι για κάθε ομάδα κ, οι τιμές του χαρακτηριστικού έχουν κανονική κατανομή που η μέση τιμή της είναι ίση με $\mu + v_\kappa$.

ΓΕΝΙΚΕΥΜΕΝΕΣ ΜΕΘΟΔΟΙ: ΜΟΝΤΕΛΑ ΤΑΞΙΝΟΜΗΣΗΣ

Ο δείκτης που έχει σημασία είναι η σχετική διασπορά

$$\sigma_\varepsilon^2 / \sigma_\sigma^2$$

Όπου: σ_σ^2 = η συνολική διασπορά που είναι ίση με $\sigma_\varepsilon^2 + \sigma_\mu^2$

σ_μ^2 = η διασπορά μεταξύ κατηγοριών, που είναι η ίδια για όλες τις κατηγορίες.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ

Οι γενικευμένες μέθοδοι χωρικής παρεμβολής παρουσιάζουν μια σειρά από μειονεκτήματα, όπως:

- Δεν μπορούν να δώσουν άμεσες εκτιμήσεις για την ποιότητα των προβλέψεων.
- Δεν γνωρίζουμε ότι οι επιλεγείσες τιμές για τις διάφορες παραμέτρους είναι πραγματικά οι βέλτιστες.
- Δεν υπάρχει αντικειμενικός τρόπος επιλογής του αριθμού των σημείων που είναι αναγκαία για την εφαρμογή των μεθόδων τοπικών εκτιμήσεων.
- Δεν είναι γνωστά τα λάθη *(αβεβαιότητες)* που σχετίζονται με τις εκτιμηθείσες τιμές της χωρικής παρεμβολής.
- Οι μέθοδοι αυτές αναφέρονται αποκλειστικά μόνο σε χωρικές διαφοροποιήσεις πρώτου βαθμού.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Από τη στατιστική→ **δείκτης συνδιασποράς** (εκφράζει την κατεύθυνση της σχέσης μεταξύ δύο μεταβλητών X και Y)

- θετικός δείκτης: ⬆ τιμές X σχετίζονται με ⬆ τιμές Y
- αρνητικός δείκτης: ⬇ τιμές X σχετίζονται με ⬆ τιμές Y και

Συντελεστής συσχέτισης (εκφράζει την κατεύθυνση και το μέγεθος της σχέσης μεταξύ δύο μεταβλητών X και Y)

$$\rho_{xy} = \frac{\operatorname{cov}(X, Y)}{\sigma_x\ \sigma_y} = \frac{E(\mathrm{X} - \mu_\chi)(\mathrm{Y} - \mu_y)}{\sigma_x\ \sigma_y}$$

Παράμετροι κατανομής

- Η από κοινού **συνάρτηση πυκνότητας δύο ΤΜ** δίνεται από τη σχέση:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

- Πέρα από τις μέσες τιμές και τις διασπορές των δύο ΤΜ χωριστά, ορίζονται η **συνδιασπορά** Cov(X,Y) και ο **συντελεστής συσχέτισης**:

$$C(X,Y) = E\left\{\left(X - \overline{X}\right)\left(Y - \overline{Y}\right)\right\} \qquad r_{xy} = \frac{C(X,Y)}{\sigma_x \sigma_y}$$

Τι είναι η συνδιασπορά;

- Όπως φαίνεται, η **συνδιασπορά** είναι και αυτή μία ροπή τάξης 2.
- Εάν παρατηρήσουμε την σχέση ορισμού θα διαπιστώσουμε ότι η C(X,Y) είναι ένα άθροισμα γινομένων το οποίο μεγιστοποιείται σε απόλυτη τιμή όταν στις περιπτώσεις που η ΤΜ Χ παίρνει τιμές κοντά στη μέση τιμή της, ταυτόχρονα συμβαίνει το ίδιο και για την Υ και αντιθέτως.
- Αυτό όμως σημαίνει ότι οι δύο ΤΜ είναι συσχετισμένες, δηλαδή όταν έχουμε πληροφόρηση για την μία, πχ ότι βρίσκεται κοντά στην μέση τιμή της, τότε μπορούμε να συμπεράνουμε με μεγάλη ασφάλεια ότι και η δεύτερη θα κινείται κοντά στην δική της μέση τιμή, άρα έχουμε πληροφόρηση και για αυτήν. Αντιθέτως, όταν η συνδιασπορά κυμαίνεται γύρω από το μηδέν, οι ΤΜ είναι ασυσχέτιστες.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Επειδή πολλές μεταβλητές εμφανίζουν «χωρική εξάρτηση», υπάρχει ενδιαφέρον για τον τρόπο με τον οποίο σε διαφορετικά σημεία στο χώρο οι αποκλίσεις των παρατηρούμενων τιμών από τη μέση τιμή **συμμεταβάλλονται** ή **συσχετίζονται**.

Για παράδειγμα, η απόκλιση από το μέσο ύψος βροχής σε μια θέση είναι πιθανότερο να είναι παρόμοια με αυτήν που παρατηρείται σε μια κοντινή θέση απ' ότι σε μια άλλη 10 χλμ μακριά.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Για μια χωρική στοχαστική διαδικασία **{Z(s), s∈A}**,

όπου **∈(Z(s)) =μ(s)** και **var(Z(s)) = σ²(s)**,

η συνδιασπορά αυτής της διαδικασίας σε δύο σημεία s_i και s_j ορίζεται ως:

$$\mathbf{C}(\mathbf{s_i},\mathbf{s_j}) = \mathbf{E}\left[\left(\mathbf{Z}(\mathbf{s_i}) - \boldsymbol{\mu}(\mathbf{s_i})\right)\left(\mathbf{Z}(\mathbf{s_j}) - \boldsymbol{\mu}(\mathbf{s_j})\right)\right]$$

ενώ η αντίστοιχη συσχέτιση ορίζεται ως:

$$\boldsymbol{\rho}(\mathbf{s_i},\mathbf{s_j}) = \frac{\mathbf{C}(\mathbf{s_i},\mathbf{s_j})}{\boldsymbol{\sigma}(\mathbf{s_i})\boldsymbol{\sigma}(\mathbf{s_j})}$$

βέβαια η $\mathbf{C}(\mathbf{s_i},\mathbf{s_j})$ **= σ²(s)**

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Η παραπάνω διαδικασία θεωρείται στάσιμη αν

$\mu(s) = \mu$ και $\sigma^2 (s) = \sigma^2$ και επιπλέον:

$C(s_i, s_j) = C(s_i - s_j) = C(h)$ όπου βέβαια: $C(h=0) = \sigma^2$

- Η $C(h)$ αναφέρεται ως συνάρτηση **συνδιασποράς** ή **συνβαριόγραμμα** της διαδικασίας.
- Η $\rho(h)$ η αντίστοιχη συσχέτιση αναφέρεται ως **συσχετόγραμμα**.

Η παραπάνω εξίσωση δείχνει ότι η $C(s_i, s_j)$ εξαρτάται μόνο από τη διανυσματική διαφορά h μεταξύ s_i και s_j και όχι από την απόλυτη θέση τους, δηλαδή η διαδικασία είναι ισοτροπική.

Στην περίπτωση αυτή: $C(s_i, s_j) = C(h)$ και $\rho(s_i, s_j) = \rho(h)$

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Αν υποτεθεί ότι η μέση τιμή και η διασπορά των διαφορών μεταξύ των τιμών σε δύο σημεία, που η θέση τους απέχει μια συγκεκριμένη απόσταση και διεύθυνση, είναι σταθερές (εσωτερική στασιμότητα),τότε:

$$\in(Z(s+h) - Z(s)) = 0$$

$$VAR(Z(s+h) - Z(s)) = 2\gamma h$$

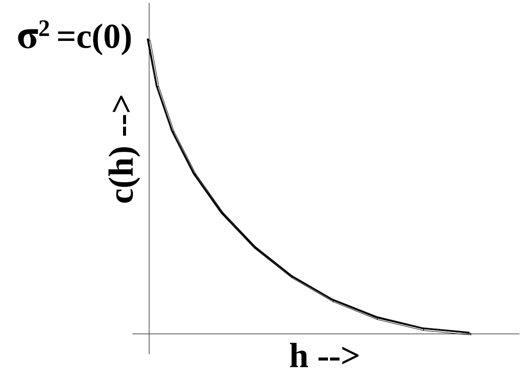
Η **γ(h)** αναφέρεται ως **βαριόγραμμα** (ή **ημιμεταβλητογράφημα**)

Σε μια στάσιμη χωρική διαδικασία το συνβαριόγραμμα, το συσχετόγραμμα και το βαριόγραμμα σχετίζονται άμεσα ως εξής:

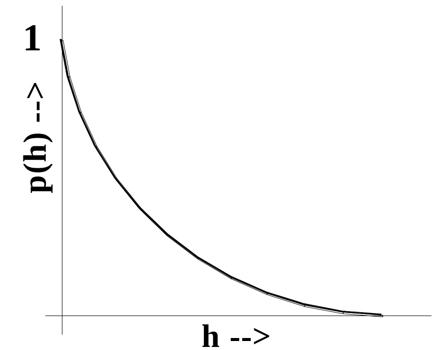
$$\mathbf{p}(\mathbf{h}) = \frac{\mathbf{C}(\mathbf{h})}{\boldsymbol{\sigma}^2}$$

$$\boldsymbol{\gamma}(\mathbf{h}) = \boldsymbol{\sigma}^2 - \mathbf{C}(\mathbf{h})$$

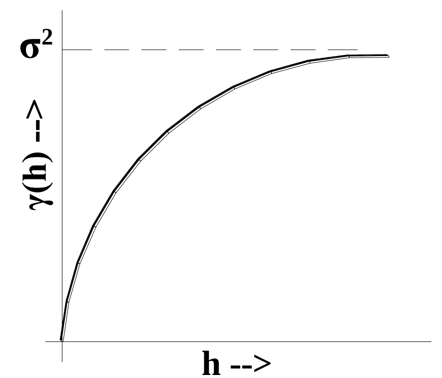
ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ:
ΒΑΡΙΟΓΡΑΜΜΑ



Συνβαριόγραμμα



Συσχετόγραμμα



Βαριόγραμμα

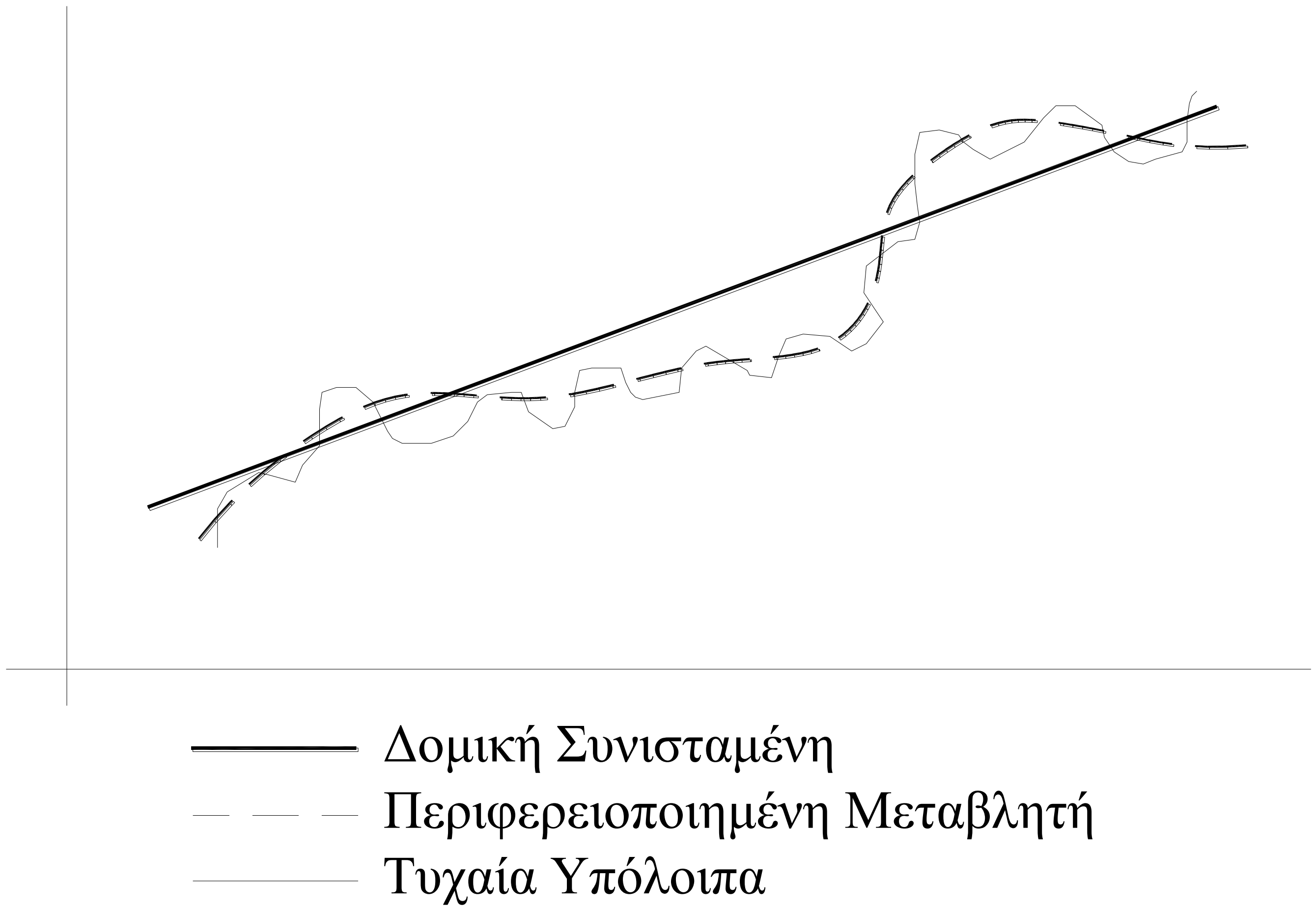
ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Η θεωρία των περιφερειοποιημένων (ή περιοχικών, regionalized variable theory) μεταβλητών υποθέτει ότι η χωρική διαφοροποίηση μιας οποιασδήποτε μεταβλητής μπορεί να θεωρηθεί ως το άθροισμα των εξής τριών βασικών συνισταμένων:

- Της **δομικής συνισταμένης**, η οποία έχει μια σταθερή μέση τιμή ή «τάση».
- Μιας τυχαίας αλλά χωρικά συσχετισμένης συνισταμένης, γνωστής και ως **περιφερειοποιημένη ή περιοχική μεταβλητή**, και
- Ενός μη χωρικά συσχετισμένου τυχαίου θορύβου ή ενός **υπόλοιπου τυχαίου σφάλματος**.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Χωρική Διαφοροποίηση Μεταβλητής



ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Η τιμή της μεταβλητής **Z** σε ένα σημείο s_i δίνεται από τη σχέση:

$$Z(s_i) = m(s_i) + U(s_i) + \varepsilon$$

όπου:

$m(s_i)$ = μια αιτιοκρατική συνάρτηση που περιγράφει τη δομική συνισταμένη της **Z** στο σημείο s_i.

$U(s_i)$ = ένας στοχαστικός όρος που εκφράζει τα υπόλοιπα από την **$m(s_i)$** που μεταβάλλονται τοπικά αλλά είναι χωρικά συσχετισμένα, δηλαδή την περιφερειοποιημένη μεταβλητή, και

ε = τα χωρικά ανεξάρτητα σφάλματα που ακολουθούν μια κανονική κατανομή με μέση τιμή μηδέν και διασπορά σ^2.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

- Αν υποτεθεί ότι η προηγούμενη εξίσωση εκφράζει μια εσωτερική στάσιμη χωρική διαδικασία.
- Αν οι δομικές επιπτώσεις *(π.χ. η μέση τιμή στην περιοχή μελέτης)* μπορούν να εξηγηθούν (η εναπομείνασα διασπορά είναι ομοιογενής ως προς τη διαφοροποίησή της).

Οι διαφορές μεταξύ σημείων του δείγματος είναι απλώς μια συνάρτηση της απόστασης μεταξύ τους.

Η εξίσωση μπορεί να πάρει τη μορφή:

$$\mathbf{Z(s_i) = m(s_i) + \gamma(h) + \varepsilon}$$

για την κατανόηση της οποίας αρκεί η εκτίμηση της **γ(h)**.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Το βαριόγραμμα μπορεί να εκτιμηθεί από τα στοιχεία ενός δείγματος ως εξής:

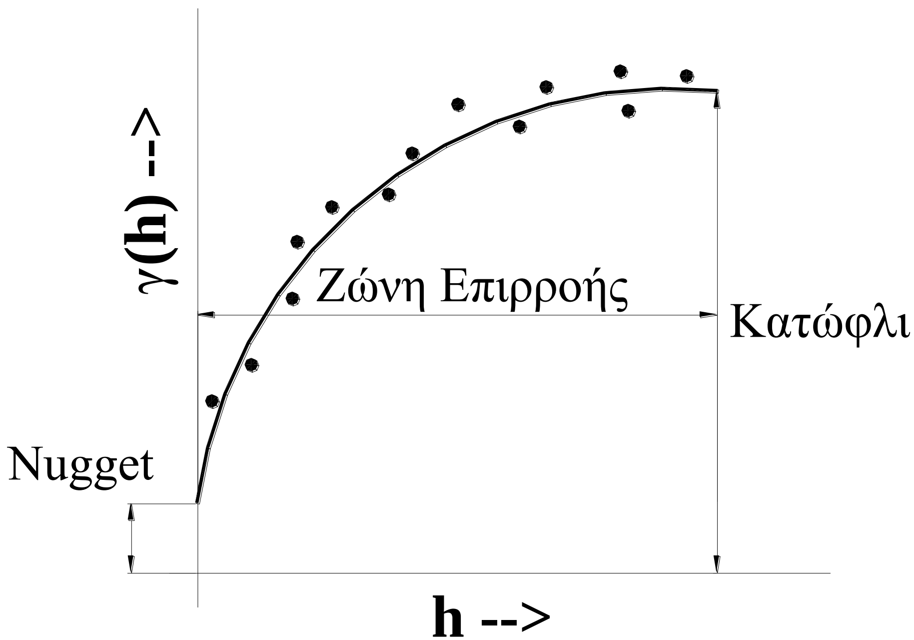
$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{s_i - s_j = h}^{n} \left(Z(s_i) - Z(s_i + h) \right)^2$$

όπου: $\sum\limits_{s_i - s_j = h}^{n} =$ το άθροισμα για όλα τα ζευγάρια των παρατηρούμενων σημείων που απέχουν μεταξύ τους κατά το διάνυσμα **h**.

n(h) = ο αριθμός αυτών των ζευγαριών.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Καθώς το **h** μεταβάλλεται, μια σειρά από τιμές $\hat{\gamma}$ εκτιμώνται, δημιουργώντας το **δειγματικό** ή **πειραματικό** βαριόγραμμα.



ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

Εξετάζοντας το σχήμα παρατηρούνται τα εξής:

- Καθώς οι τιμές του **h** αυξάνουν, η τιμή της **γ(h**) αυξάνεται ασυμπτωτικά προς ένα ανώτερο όριο που ονομάζεται ***κατώφλι (sill)***.
- Η καμπύλη της **γ(h)** αυξάνεται μέχρι να φτάσει στο κατώφλι, οπότε λαμβάνει τη μέγιστη τιμή της σε μια απόσταση που ορίζει τη ***ζώνη επιρροής (range)***, η οποία με τη σειρά της καθορίζει το χώρο εντός του οποίου οι μεταξύ των σημείων διαφοροποιήσεις είναι χωρικά εξαρτημένες.

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

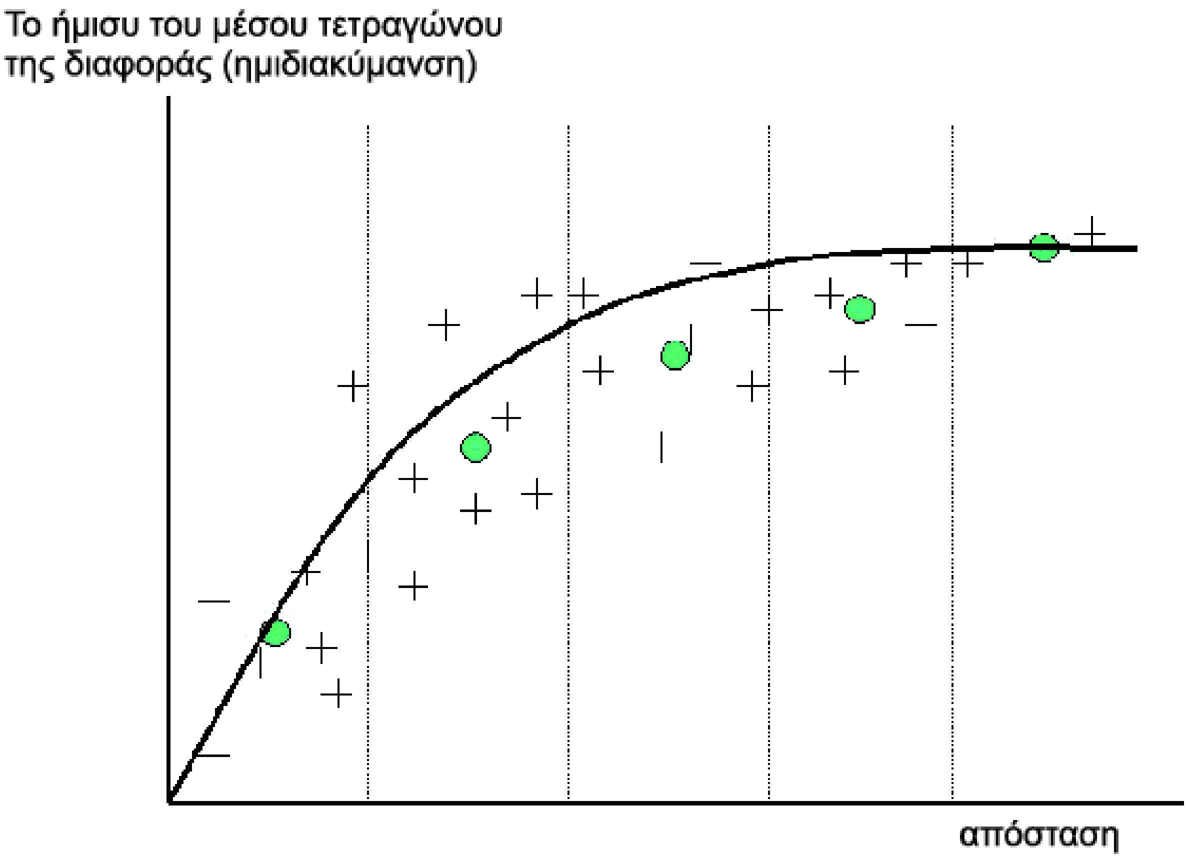
Εξετάζοντας το σχήμα παρατηρούνται τα εξής:

- Η εκτιμηθείσα καμπύλη $\hat{\gamma}(\mathbf{h})$ δεν περνά από την αρχή των αξόνων αλλά τέμνει τον κάθετο άξονα σε ένα σημείο με θετική τιμή. Επομένως, θετική τιμή $\hat{\gamma}(\mathbf{h})$ **h→0** είναι μια εκτίμηση του **ε**, δηλαδή του μη χωρικού θορύβου ή υπόλοιπου, που είναι γνωστό ως **nugget**.
- Στην περίπτωση που παρατηρείται ένα **βαριόγραμμα** όπου οι τιμές των διασπορών είναι ευρέως διασκορπισμένες, αυτό δηλώνει καθαρά ότι η εκτίμηση της $\hat{\gamma}(\mathbf{h})$ έγινε με τη χρήση ενός μικρού δείγματος.

Μέθοδος Kriging

- Είναι μια τεχνική χωρικής παρεμβολής με γερά θεμέλια στη θεωρία της γεω-στατιστικής
- Ένα *ημιμεταβλητογράφημα* (*semivariogram*) αντανακλά το Νόμο του Tobler
  - Οι διαφορές σε μια μικρή γειτονιά αναμένεται να είναι μικρές
  - Οι διαφορές αυξάνουν όσο μεγαλώνει η απόσταση

Μέθοδος Kriging



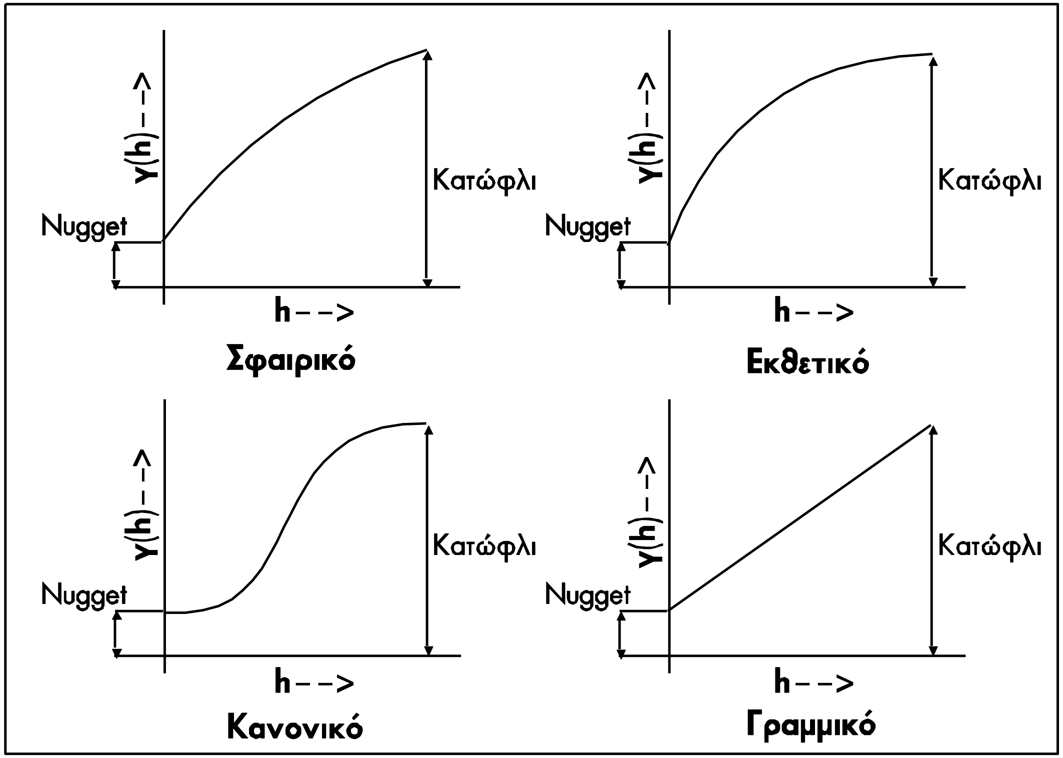
Ένα ημιβαριόγραμμα. Κάθε σταυρός αντιπροσωπεύει ένα ζεύγος σημείων. Οι συμπαγείς κύκλοι προκύπτουν από τον υπολογισμό του μέσου όρου μέσα στις περιοχές ή τα *δοχεία* (*bins*) του άξονα της απόστασης. Η συμπαγής γραμμή είναι η βέλτιστη προσέγγιση αυτών των πέντε σημείων, με τη χρήση μιας τυπικής μαθηματικής συνάρτησης.

Τα στάδια της μεθόδου Kriging

- Ανάλυση των τιμών που έχουν μετρηθεί για να παραχθεί ένα ημιμεταβλητογράφημα
- Υπολογισμός των τιμών στα άγνωστα σημεία ως σταθμισμένες μέσες τιμές
  - Τα βάρη προκύπτουν από το ημιμεταβλητογράφημα
  - Η επιφάνεια παρεμβολής επαναλαμβάνει στατιστικές ιδιότητες του βαριογράμματος ή ημιμεταβλητογραφήματος

ΓΕΩΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ: ΒΑΡΙΟΓΡΑΜΜΑ

- Η μορφή του **βαριογράμματος** είναι ιδιαίτερα κατατοπιστική:
  - Για το είδος της χωρικής διαφοροποίησης που υπάρχει στην περιοχή μελέτης
  - Για να βοηθήσει σημαντικά στη διαδικασία χωρικής παρεμβολής.
- Τέσσερα είναι τα βασικά μοντέλα (μορφές) του **βαριογράμματος**:



ΒΑΡΙΟΓΡΑΜΜΑ:
ΣΦΑΙΡΙΚΟ ΜΟΝΤΕΛΟ

$$\gamma(h) = \begin{cases} \alpha + \left(\sigma^2 - \alpha\right)\left(\dfrac{3h}{2r} - \dfrac{h^3}{2r^3}\right) & \text{για } 0 < h < r \\ 0 & \text{για } h = 0 \\ \sigma^2 & \text{για } h \geq r \end{cases}$$

όπου **r =** η ζώνη επιρροής *(range)*

σ^2 = το κατώφλι-διασπορά

h = η απόσταση μεταξύ των σημείων

α = το τμήμα της διασποράς που αφορά το nugget

ΒΑΡΙΟΓΡΑΜΜΑ

Εκθετικό Μοντέλο

$$\gamma(h) = \begin{cases} \alpha + \left(\sigma^2 - \alpha\right)\left(1 - e^{-3h/r}\right) & \text{για } h > 0 \\ 0 & \text{για } h = 0 \end{cases}$$

Κανονικό ή Μοντέλο Γκάους

$$\gamma(h) = \begin{cases} \alpha + \left(\sigma^2 - \alpha\right)\left(1 - e^{-3h^2/r^2}\right) & \text{για } h > 0 \\ 0 & \text{για } h = 0 \end{cases}$$

Γραμμικό Μοντέλο

$\gamma(h) = \sigma^2 + \beta h$

όπου: **β =** η κλίση της γραμμής του βαριογράμματος

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:
ΣΥΝΗΘΗΣ KRIGING – ORDINARY KRIGING

Η τιμή μιας συνεχούς χωρικής μεταβλητής δίνεται από το άθροισμα δύο συνισταμένων, μιας πρώτου και μιας δεύτερου βαθμού.

Η συνισταμένη πρώτου βαθμού είναι ο γενικευμένος μέσος όρος **μ(s)**, ο οποίος είναι σταθερός και δεν χρειάζεται να εκτιμηθεί από τις παρατηρηθείσες τιμές

Αν η τιμή αυτή αφαιρεθεί από τις παρατηρούμενες τιμές σε κάθε σημείο του δείγματος, τα υπόλοιπα δίνουν τις τιμές **U(s)** με γνωστή διακύμανση **σ^2** και συνδιασπορά **C( )**, που εκφράζονται από το βαριόγραμμα.

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:
ΣΥΝΗΘΗΣ KRIGING – ORDINARY KRIGING

Το πρόβλημα που επιλύεται με τη μέθοδο της συνήθους Kriging είναι η εκτίμηση:

$\hat{V}(s) =$ η τιμή της τυχαίας μεταβλητής **V(s)** στη θέση **s** (με βάση τις παρατηρηθείσες τιμές **$V(s_i)$** σε **n** θέσεις **s_i** του δείγματος).

Z(s) = που εκτιμάται αυτόματα (αρκεί στην τιμή **V(s)** για το σημείο **s** να προστεθεί η γνωστή μέση τιμή ή τάση **μ(s)**).

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:
ΣΥΝΗΘΗΣ KRIGING – ORDINARY KRIGING

Θεωρείται ότι οι εκτιμήσεις αποτελούν ένα γραμμικό, με βάρη συνδυασμό των παρατηρούμενων τιμών των περιφεριοποιημένων μεταβλητών, που παίρνει τη μορφή:

$$\hat{V}(s_o) = \sum_{i=1}^{n} \lambda_i V(s_i)$$

όπου $\hat{V}(s_o) =$ Η εκτιμηθείσα τιμή στη θέση s_o.

$V(s_i)$ = Η παρατηρηθείσα τιμή στο σημείο s_i.

λ_i = Τα βάρη που αντιστοιχούν σε κάθε σημείο του δείγματος s_i, δηλαδή εξαρτώνται από τη θέση τους σε σχέση με την υπό εκτίμηση θέση s_o

Τα βάρη λ_i επιλέγονται έτσι ώστε η εκτίμηση $\hat{V}(s_o)$ να τηρεί τον όρο της μη προκατάληψης ότι η εκτιμηθείσα διακύμανση είναι μικρότερη από κάθε άλλο γραμμικό συνδυασμό των παρατηρούμενων τιμών.

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:

ΣΥΝΗΘΗΣ KRIGING

Λόγω της απαίτησης για μη προκατάληψη και επειδή η αναμενόμενη τιμή είναι γραμμικός τελεστής, ισχύουν:

$$E\left[V(s_0) - \hat{V}(s_0)\right] = 0 \;\Rightarrow\; \mu_{s_0} - \sum_{i=1}^{n} \lambda_i \cdot \mu_{s_0} = 0$$

Επειδή η μέση τιμή είναι σταθερή, η απαίτηση για μη προκατάληψη σημαίνει ότι:

$$\sum_{i=1}^{n} \lambda_i = 1.0$$

Η απαίτηση για βέλτιστη πρόβλεψη οδηγεί στην ελαχιστοποίηση της ποσότητας **Q** ως εξής:

$$Q = E\left[\left(V(s_0) - \hat{V}(s_0)\right)^2\right] = E\left[\left(V(s_0) - \sum_{i=1}^{n} \lambda_i\, V(s_i)\right)^2\right]$$

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:
ΣΥΝΗΘΗΣ KRIGING

Μετά από μετασχηματισμούς η εξίσωση γίνεται:

$$Q = E\left[\left(V(s_0)-\mu\right)^2\right] + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j E\left[\left(V(s_i)-\mu\right)\left(V(s_j)-\mu\right)\right] - $$
$$-2\sum_{i=1}^{n}\lambda_i\, E\left[\left(V(s_i)-\mu\right)\left(V(s_0)-\mu\right)\right]$$

Από τους τρεις όρους:

- Ο πρώτος είναι η διασπορά της τυχαίας συνάρτησης, δηλαδή C_0.
- Ο μεσαίος όρος είναι η συνδιασπορά των μεταβλητών στις θέσεις s_i, s_j.
- Ο τρίτος όρος είναι η συνδιασπορά της μεταβλητής που θέλουμε να προβλεφθεί και της μεταβλητής στη θέση s_i, δηλαδή $C_i = C(s_i, s_0)$.

Επομένως:

$$Q = C_0 + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\, C_{ij} - 2\sum_{i=1}^{n}\lambda_i\, C_i$$

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:

ΣΥΝΗΘΗΣ KRIGING

Με δεδομένο ότι $\sum_{i=1}^{n} \lambda_i = 1.0$ και με τη χρήση πολλαπλασιαστών **Lagrange**, δημιουργείται ένα σύστημα εξισώσεων ως εξής:

$$\begin{cases} \dfrac{\partial}{\partial \lambda_\kappa}\left[Q - 2\varphi\left(1 - \sum_{i=1}^{n} \lambda_i\right)\right] = 0 \quad \text{για } \kappa = 1, 2, 3, \ldots, n \\ \sum_{i=1}^{n} \lambda_i = 1.0 \end{cases}$$

Αν υλοποιηθεί η μερική παράγωγος για κάθε **κ**, τότε:

$$\begin{cases} 2\sum_{i=1}^{n} \lambda_i C_{\kappa i} - 2C_\kappa + 2\varphi = 0 \quad \text{για } \kappa = 1, 2, 3, \ldots, n \\ \sum_{i=1}^{n} \lambda_i = 1.0 \end{cases}$$

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:

ΣΥΝΗΘΗΣ KRIGING

Το σύστημα των **n+1** εξισώσεων σε αναλυτική μορφή είναι:

$$\begin{cases} \lambda_1 C_0 + \lambda_2 C_{1,2} + \ldots + \lambda_n C_{1,n} + \varphi = C_1 \\ \lambda_1 C_{1,2} + \lambda_2 C_0 + \ldots + \lambda_n C_{2,n} + \varphi = C_2 \\ \vdots \\ \lambda_1 C_{1,n} + \lambda_2 C_{2,n} + \ldots + \lambda_n C_0 + \varphi = C_n \\ \lambda_1 + \lambda_2 + \ldots + \lambda_n = 1.0 \end{cases}$$

που μπορεί βέβαια να γραφεί σε μορφή σύνθετων πινάκων ως εξής:

$$\begin{bmatrix} C & 1 \\ 1^T & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ \varphi \end{bmatrix} = \begin{bmatrix} C \\ 1 \end{bmatrix}$$

όπου: **1 =** διάνυσμα μήκους **n** που περιλαμβάνει μονάδες,

C = πίνακας με συνδιασπορές ανάμεσα στα σημεία με γνωστές τιμές, και

λ = το διάνυσμα των αγνώστων γραμμικών συντελεστών.

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING:
ΣΥΝΗΘΗΣ KRIGING

Η λύση του συστήματος είναι:

$$\begin{bmatrix} \lambda \\ \varphi \end{bmatrix} = \begin{bmatrix} C & 1 \\ 1^T & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} C \\ 1 \end{bmatrix}$$

Για απλοποίηση η εξίσωση μπορεί να γραφεί και ως εξής:

$$\begin{bmatrix} \lambda \\ \varphi \end{bmatrix} = A^{-1} \cdot \beta$$

Η εξίσωση αυτή δίνει και τη λύση για τη συνήθη Kriging.

Ordinary Kriging → “**Best Linear Unbiased Estimator**” ή **BLUE** μέθοδος

ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ ΣΦΑΛΜΑΤΟΣ KRIGING

Η μεθόδος Kriging μαζί με την πρόβλεψη δίνει και την εκτίμηση του σφάλματος. Το σφάλμα **σ_κ**, γνωστό και ως σφάλμα **Kriging**, είναι:

$$\sigma_\kappa = \sqrt{C_0 - \sum_{\kappa=1}^{n} \lambda_\kappa C_\kappa - \varphi}$$

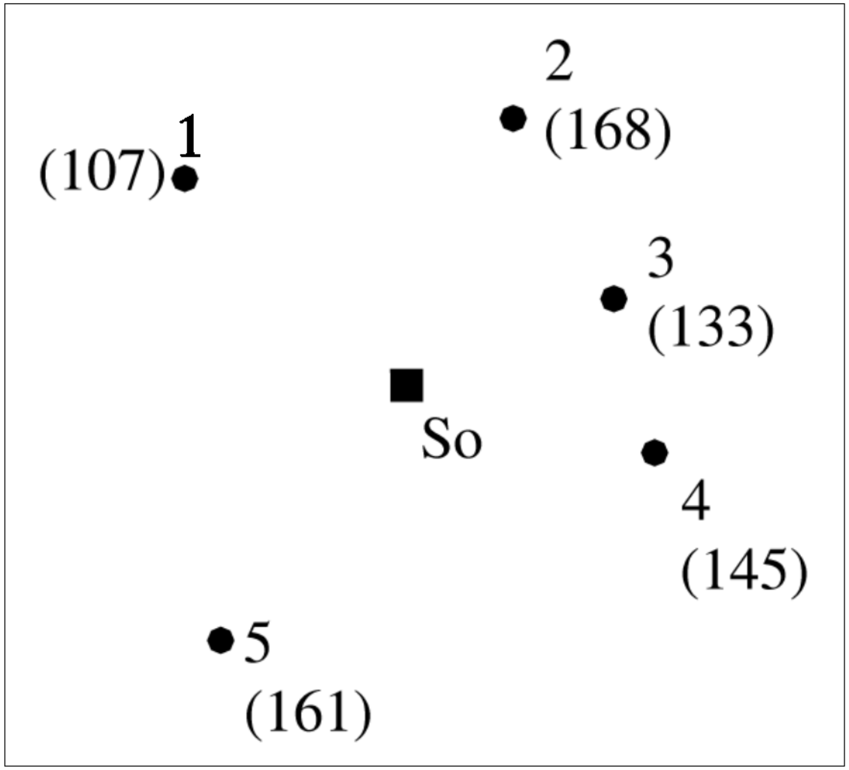
Μπορεί να εκτιμηθεί και το διάστημα τιμών με συγκεκριμένη πιθανότητα. Για παράδειγμα:

$$\mathbf{Pr} = \left\{ \mathbf{V}(\mathbf{s}_0) \in \left(\sum_{\kappa=1}^{n} \lambda_i\, \mathbf{z}\;(\mathbf{s}_i) - 1{,}96\sqrt{C_0 - \sum_{\iota=1}^{n} \lambda_\kappa C_\kappa - \varphi},\; \sum_{\kappa=1}^{n} \lambda_i\, \mathbf{z}\;(\mathbf{s}_i) + 1{,}96\sqrt{C_0 - \sum_{\iota=1}^{n} \lambda_\kappa C_\kappa - \varphi} \right) \right\} = \mathbf{0{,}95}$$

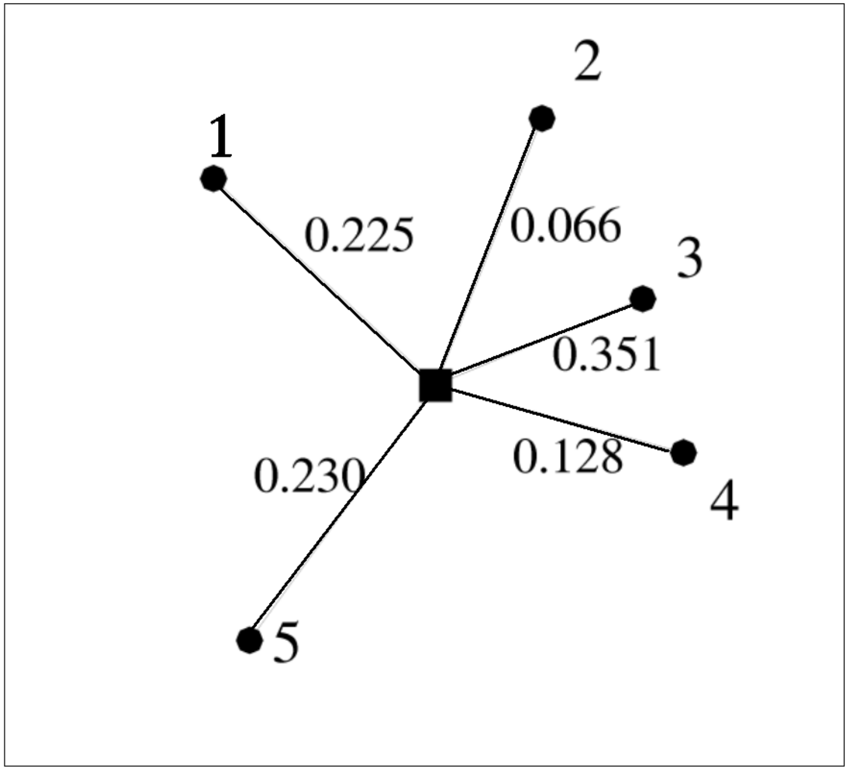
ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ ΣΦΑΛΜΑΤΟΣ KRIGING

- Ένας άλλος προσδιορισμός του σφάλματος είναι η μέθοδος **Cross Validation**.
    - Σε αυτήν μια τιμή αφαιρείται από το δείγμα και η αναμενόμενη τιμή αφαιρείται από την πραγματική.
    - Η διαδικασία αυτή επαναλαμβάνεται για όλες τις τιμές και έτσι λαμβάνονται **n** υπόλοιπα τα οποία θεωρητικά έχουν κανονική κατανομή.

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING: ΠΑΡΑΔΕΙΓΜΑ



α



β

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING: ΠΑΡΑΔΕΙΓΜΑ

Υποτίθεται ότι η διαδικασία που δίνει τις τιμές στα σημεία δεν παρουσιάζει καμιά τάση στην περιοχή μελέτης και ότι απλώς έχει μια μέση τιμή ίση με 145,37. Επομένως, είναι δυνατόν να αφαιρεθεί η μέση αυτή τιμή από τις αρχικές δειγματικές τιμές **$Z(s_i)$**, ώστε να υπολογιστούν τα υπόλοιπα **$V(s_i)$** ως εξής:

Θέσεις	$Z(s_i)$		Θέσεις	$V(s_i)$
1	107		1	-38,37
2	168		2	22,63
3	133	→	3	-12,37
4	145		4	-0,37
5	161		5	15,63

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING: ΠΑΡΑΔΕΙΓΜΑ

Υποτίθεται, ότι η δομή των συνδιασπορών είναι γνωστή, **ισοτροπική**, και εκφράζεται από ένα εκθετικό βαριόγραμμα, το οποίο παρουσιάζει **κατώφλι** ίσο με 20 μονάδες και **περιοχή επιρροής** ίση με 100 μονάδες, ενώ δεν υπάρχουν επιπτώσεις **nugget**.

Επομένως, το συνβαριόγραμμα έχει τη μορφή:

$$C(h) = 20e^{-3b/100}$$

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING: ΠΑΡΑΔΕΙΓΜΑ

Αν μετρηθούν οι αποστάσεις μεταξύ όλων των σημείων του δείγματος και μεταξύ του σημείου $\mathbf{s_0}$ και των υπόλοιπων σημείων και χρησιμοποιηθούν στην εξίσωση, μπορούν να υπολογιστούν οι όροι **A** και β που οδηγούν στον υπολογισμό των βαρών $\boldsymbol{\lambda(s_i)}$.

Υπολογισμός A

Θέσεις	1	2	3	4	5	6
1	20,00	4,571	3,970	2,828	3,228	1.000
2	4,571	20,00	5,790	3,739	1,086	1.000
3	3,970	5,970	20,00	12,45	2,300	1.000
4	2,828	3,739	12,45	20,00	2,479	1.000
5	3,228	1,086	2,300	2,479	20,00	1.000
6	1.000	1,000	1,000	1,000	1,000	0,000

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING: ΠΑΡΑΔΕΙΓΜΑ

Υπολογισμός β

Θέσεις	s
1	6,895
2	5,032
3	10,15
4	8,83
5	4,079
6	1,000

Υπολογισμός $\lambda(s_i)$

Θέσεις	λ(s)
1	0,225
2	0,066
3	0,351
4	0,128
5	0,230
6	0,123

$\xrightarrow[\varphi]{}$

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING: ΠΑΡΑΔΕΙΓΜΑ

Η εκτίμηση της $\hat{V}(s)$ έχει ως εξής:

$$\hat{V}(s) = \mathbf{0{,}225 \times (-38{,}37) + 0{,}066 \times 22{,}63 + 0{,}351 \times (-12{,}37) + 0{,}128 \times (-0{,}37) + 0{,}230 \times 15{,}63 \approx -7{,}937}$$

με αποτέλεσμα η εκτιμηθείσα τιμή στο σημείο s_0, αφού προστεθεί η γνωστή μέση τιμή, να είναι:

$$\hat{Z}(s) = \mathbf{145{,}370 - 7{,}937 = 137{,}433}$$

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ KRIGING: ΠΑΡΑΔΕΙΓΜΑ

Επιπλέον, η **διασπορά Kriging** υπολογίζεται ως εξής:

$$\sigma_\varepsilon^2 = \mathbf{20 - (0{,}225x6{,}895 + 0{,}066x5{,}032 + 0{,}351x10{,}15 + 0{,}128x8{,}83 + 0{,}230x4{,}079) - 0{,}123 =}$$

$$\mathbf{= 20 - 7{,}504 - 0{,}13 = 12{,}373}$$

οπότε το **σφάλμα Kriging** είναι ίσο με: $\sqrt{12{,}373}$ **= 3,517,** που οδηγεί σε ένα διάστημα για επίπεδο εμπιστοσύνης 95% ίσο με

$$\mathbf{145{,}37 \pm 1{,}96x3{,}517 = 145{,}37 \pm 6{,}88}$$

ΜΕΘΟΔΟΙ KRIGING

- ***Ordinary – Simple Kriging:*** Η πλέον διαδεδομένη, έχει τις παρακάτω παραδοχές: α) η μεταβλητή ακολουθεί κανονική κατανομή, β) η εκτίμηση είναι αμερόληπτη, γ) μονιμότητα δευτέρου βαθμού, δ1) ο τοπικός μέσος όρος είναι γνωστός (**simple**), ή δ2) ο τοπικός μέσος είναι άγνωστος (**ordinary**).
- ***Neighbourhood Kriging:*** Αν και η τοπική μέση τιμή και διασπορά είναι σταθερές σε όλη την περιοχή, στις περισσότερες εφαρμογές τα δεδομένα περιέχουν τοπικές διακυμάνσεις. Για το λόγο αυτό στην εκτίμηση της άγνωστης τιμής συμμετέχουν τα κοντινότερα σημεία ή αυτά που περιλαμβάνονται στη γύρω περιοχή.
- ***Block Kriging:*** Η διαφορά της προσέγγισης αυτής από τις προηγούμενες είναι ότι η εκτίμηση της παρεμβολής δεν αναφέρεται σε ένα σημείο αλλά στο μέσο όρο των τιμών που μετρώνται σε μια συγκεκριμένη περιοχή.
- ***Μη Γραμμική Kriging:*** Στην περίπτωση αυτή η διαδικασία θεωρείται ότι είναι λογαριθμοκανονική, δηλαδή ο λογάριθμός της έχει κανονική κατανομή.

ΑΛΛΕΣ ΜΟΡΦΕΣ KRIGING

- ***Universal Kriging:*** Στην περίπτωση αυτή η εκτίμηση της παρεμβολής γίνεται κάτω από την παραδοχή ότι η μέση τιμή δεν είναι σταθερή, όπως στη συνήθη Kriging, αλλά μεταβάλλεται ως συνάρτηση της θέσης.
- ***Disjunctive Kriging:*** Υπολογίζει για κάθε εκτίμηση και την πιθανότητα η αληθινή τιμή να υπερβαίνει ένα συγκεκριμένο κατώφλι.
- ***Co Kriging:*** Η μέθοδος αυτή είναι μια πολυδιάστατη εκδοχή της συνήθους Kriging και αφορά την περίπτωση όπου στο ίδιο σημείο της περιοχής μελέτης μετρώνται περισσότερες από μια μεταβλητές.
- ***Multivariate Kriging:*** Η μέθοδος του Kriging πολυμεταβλητών αναφέρεται, όπως είναι ευνόητο, στην εφαρμογή της γεωστατιστικής σε μεθόδους ανάλυσης πολυμεταβλητών.
- ***Space time Kriging:*** Σχετίζεται με την εισαγωγή της χρονικής διάστασης των δεδομένων.

Βέλτιστη παρεμβολή Ordinary Kriging

- Η **μέση τιμή είναι άγνωστη**:
- Τελικά, οι (n+1) άγνωστοι λ_i, μ, υπολογίζονται από το (n+1) γραμμικό σύστημα, το οποίο καλείται «**Ordinary Kriging**» σύστημα και είναι:

$$\begin{cases} \sum_{j=1}^{n} \lambda_j C(x_i - x_j) + \mu = C(x_i - x_0), i = 1,\ldots n \\ \sum_{j=1}^{n} \lambda_j = 1 \end{cases}$$

- το οποίο μπορεί να γραφτεί και σε **όρους ημιμεταβλητογραφήματος** μιας και αυτό είναι που χρησιμοποιείται στη γεωστατική:

$$\begin{cases} \sum_{j=1}^{n} \lambda_j \gamma(x_i - x_j) + \mu = \gamma(x_i - x_0), i = 1,\ldots n \\ \sum_{j=1}^{n} \lambda_j = 1 \end{cases}$$

Βέλτιστη παρεμβολή Simple Kriging

- Η **τάση** (**μέση τιμή**) είναι **γνωστή** και **σταθερή** σε όλο το πεδίο μελέτης:
- Η **Ordinary Kriging** OK υπολογίζει την άγνωστη τιμή του υπό εξέταση χαρακτηριστικού σε μία άγνωστη θέση u ως γραμμικούς συνδυασμούς των γύρω παρατηρημένων τιμών:

$$Z(u) = \sum_{i=1}^{n} \lambda_{ui} Z(x_i)$$

- Η μέθοδος **Simple Kriging** SK υπολογίζει την άγνωστη τιμή του υπό εξέταση χαρακτηριστικού ως :

$$Z(u) - m(u) = \sum_{i=1}^{n} \lambda_{ui} \left[Z(x_i) - m(x_i) \right]$$

Μέθοδος Kriging σε συνδυασμό με παλινδρόμηση (Regression Kriging)

- **Χωρικό ομοίωμα: $Z(x) = m(x) + e(x)$**

  όπου $m(x) = \sum_{k=0}^{K} \beta_k y_k(x)$

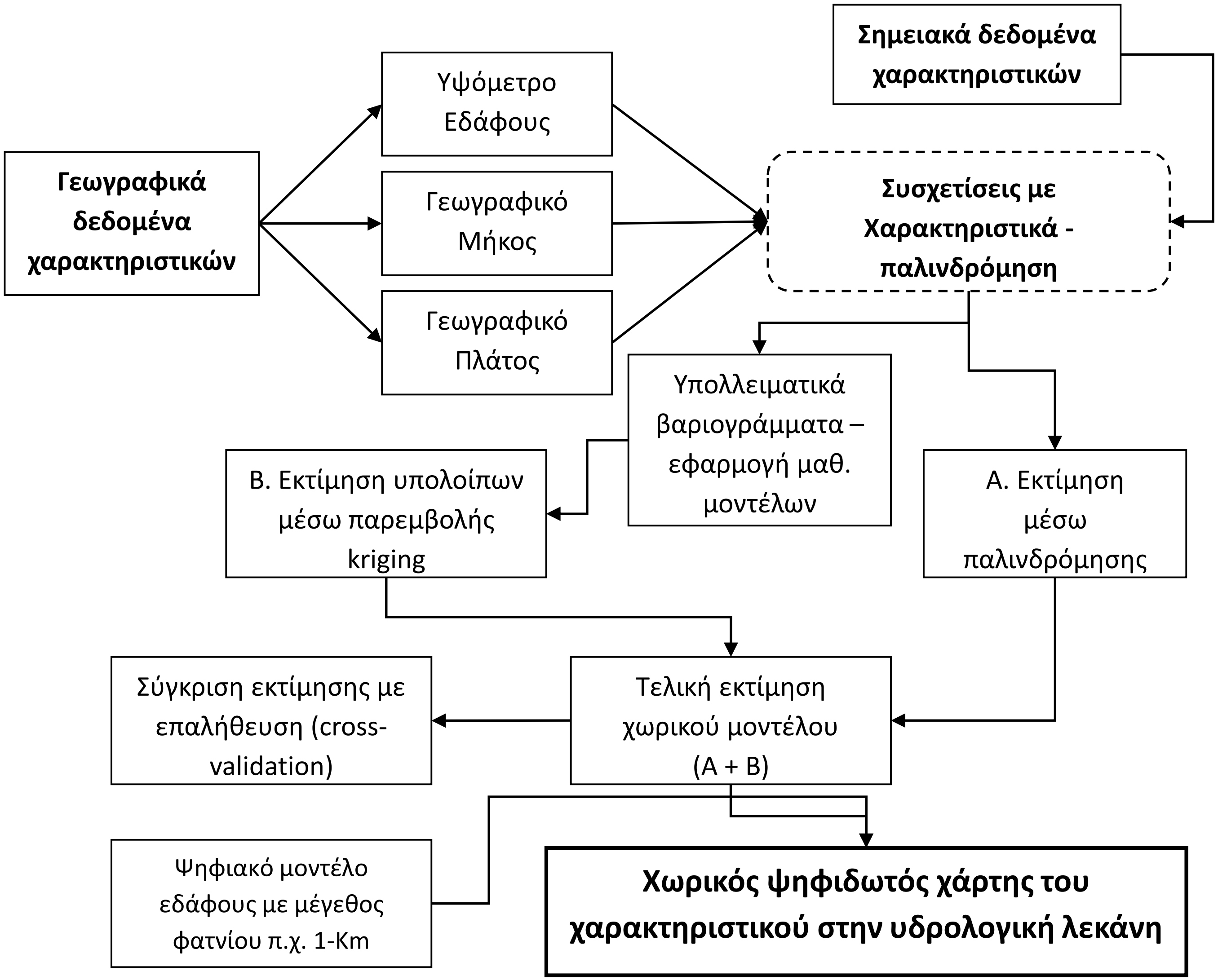
- Η βασική σχέση της **τάσης $m(x)$** εκφράζεται ως **συνάρτηση** των **γνωστών εξωτερικών (external) μεταβλητών $y_1(x)$, $y_2(x)$,…, $y_K(x)$** και των **αγνώστων συντελεστών β_κ**.

- Ο **μέσος όρος** της δεν είναι πλέον σταθερός όπως στην κανονική μέθοδο Kriging αλλά **συνάρτηση των εξωτερικών μεταβλητών** πχ. των συντεταγμένων των χωρικών σημείων $s(x, y)$, ή χωρική τάση

Μέθοδος Kriging σε συνδυασμό με παλινδρόμηση (Regression Kriging)

- Γενικότερα, η έκφραση αυτή του μέσου όρου *m(x)* μπορεί να είναι συνάρτηση όχι μόνο των συντεταγμένων των χωρικών σημείων *s*(*x*, *y*) (περίπτωση της **Universal Kriging**), αλλά και οποιασδήποτε **εξωγενούς μεταβλητής** όπως αυτές που εκφράζουν την επίδραση των παραγόντων στην εξεταζόμενη μεταβλητή (μορφολογικά, υδρολογικά δεδομένα, γεωλογικά κ.α).
- Η **διερεύνηση των συσχετίσεων** μεταξύ των εξωγενών ή επικουρικών μεταβλητών γίνεται συνήθως με **γραμμική παλινδρόμηση**.
  - Η υπόθεση που υιοθετείται στην περίπτωση αυτή είναι ότι η αιτιοκρατικής φύσεως διακύμανση ***m(x)*** της υπό εξέταση μεταβλητής λαμβάνεται από το μοντέλο παλινδρόμησης, ενώ η υπολειμματική διακύμανση ***e*(*x*)** εμφανίζει χωρική συσχέτιση.
  - Εφόσον η εξωγενής μεταβλητή είναι διαθέσιμη σε πολύ περισσότερα σημεία στο χώρο από την υπό εξέταση μεταβλητή, μπορεί να χρησιμοποιηθεί για την πρόβλεψη του ***m(x)*** από το μοντέλο παλινδρόμησης. Παράλληλα, η υπολειμματική διακύμανση ***e*(*x*)** μπορεί να εκτιμηθεί από το σύστημα εξισώσεων της μεθόδου ΟΚ. Τελικά το άθροισμα των δύο εκτιμήσεων $Z(x) = m(x) + e(x)$ μας δίνει της εκτίμηση της υπό εξέταση μεταβλητής

Διάγραμμα ροής συνδυασμένης μεθόδου Kriging με παλινδρόμηση (Regression Kriging)



Βιβλιογραφία

Burrough, P.A., and R.A., McDonnell, 1998. Principles of Geographical Information Systems, Oxford University Press, Oxford.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Li, J. and A.D. Heap, 2008. A Review of Spatial Interpolation Methods for Environmental Scientists, Geoscience Australia, Record 2008/23, 137 p.

Κουτσόπουλος, Κ.Χ., 2005. «Γεωγραφικά Συστήματα Πληροφοριών και Ανάλυση Χώρου», 2η Έκδοση, Α. Παπασωτηρίου & Σία ΟΕ.

Μιμίκου, Μ.Α., 2006. «Τεχνολογία Υδατικών Πόρων», 3η Έκδοση, Α. Παπασωτηρίου & Σία ΟΕ.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Θεσσαλίας**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.