Uncertainty and Parameter Space Analysis in Visualization

-

Session 4: Structural Uncertainty

Analyzing the effect of uncertainty on the appearance of structures in scalar fields

Rüdiger Westermann and Tobias Pfaffelmoser
Computer Graphics & Visualization

Specific literature

[1] C. Johnson and A. Sanderson, "A next step: Visualizing errors and uncertainty," Computer Graphics and Applications, IEEE, vol. 23, no. 5, pp. 6–10, 2003.

[2] A. Pang, C. Wittenbrink, and S. Lodha, "Approaches to uncertainty visualization," The Visual Computer, vol. 13, no. 8, pp. 370–390,1997.

[3] T. Pfaffelmoser, M. Reitinger, and R. Westermann, "Visualizing the positional and geometrical variability of isosurfaces in uncertain scalar fields," in Computer Graphics Forum, vol. 30, no. 3. Wiley Online Library, 2011, pp. 951–960.

[4] T. Pfaffelmoser and R. Westermann, "Visualization of global correlation structures in uncertain 2d scalar fields," in Computer Graphics Forum, vol. 31, no. 3. Wiley Online Library, 2012, pp. 1025–1034.

[5] T. Pfaffelmoser and R. Westermann, "Correlation Visualization for Structural Uncertainty Analysis", in Journal of Uncertainty Quantification, 2012

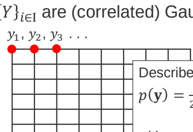
Session 4: Structural uncertainty

- Part 1: Structural uncertainty [3] [5]
  - Stochastic modelling of uncertainty
  - Variance as an uncertainty indicator
  - Limitations of variance as an uncertainty indicator
  - Definition of structural uncertainty
  - Stochastic dependence and correlation
- Part 2: Local and global correlation visualization [4] [5]
  - Requirements and challenges
  - Local correlation tensor
  - Glyph based correlation visualization
  - Global correlation clustering
  - Concluding remarks

Part 1: Structural uncertainty

Uncertain scalar fields

- Uncertain scalar field $Y: M \to \mathbb{R}$ on compact domain $M \subset \mathbb{R}^n$
- Discretely sampled on nodes $\{x\}_{i \in \mathrm{I}}$ with values $\{Y\}_{i \in \mathrm{I}}$
- We assume a stochastic uncertainty model: $\{Y\}_{i \in \mathrm{I}}$ are (correlated) Gaussian distributed random variables



Described by a multivariate Gaussian distribution
$p(\mathbf{y}) = \frac{1}{2\pi\sqrt{\det\Sigma}} \exp(-0.5(\mathbf{y}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{y}-\boldsymbol{\mu}))$

with mean $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$
and covariance matrix Σ

- $Y(x_i) = \mu_i$ is the most likely configuration

Multivariate Gaussian distribution

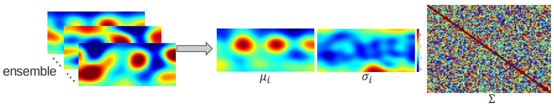
- Uncertainty parameters in the multivariate case
  - The degree of uncertainty is modeled by the Covariance Matrix Σ
    $p(\mathbf{y}) = \frac{1}{2\pi\sqrt{\det\Sigma}} \exp(-0.5(\mathbf{y}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{y}-\boldsymbol{\mu}))$
  - For 3 grid points:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

$$\sigma_i^2 = var(Y_i) = E(Y_i - \mu_i)^2$$
$$\sigma_{ij} = Cov(Y_i, Y_j) = E((Y_i - \mu_i)(Y_j - \mu_j))$$

Multivariate Gaussian distribution

- Compute μ_i, σ_{ij} from a given ensemble of data sets
  - Ensemble members are treated as realizations of a corresponding multivariate random variable



- Specify μ_i, σ_{ij} and map independent normally distributed random number vectors to a corresponding realization

http://en.wikipedia.org/wiki/Multivariate_normal_distribution#Drawing_values_from_the_distribution

Gaussian covariance matrix

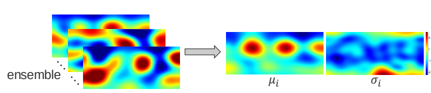
- Variance information is often used as primary indicator of the uncertainty (*square of the standard deviation* σ)

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

  - Measure of the amount of variation of the values of a random variable
  - Often visualized directly via confidence regions, uncertainty glyphs, or specific color or opacity mappings

Variance as uncertainty indicator

- Example:
  an ensemble of 2D seismic tomography wave velocities
  - Relative velocities are color coded from blue (negative) to red (positive))



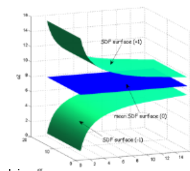
  - Two prominent circular features are observed in the mean values
  - Mapping the standard deviations to colors shows low, and almost constant uncertainty in both regions

Variance as uncertainty indicator

- Uncertainty mapping in 3D via Stochastic Distance Functions

$$\Psi_\theta(x_i) := \frac{\mu(x_i) - \theta}{\max(\sigma(x_i), \sigma_{min})}$$

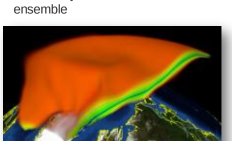
Mahalanobis distance:
in numbers of standard deviation
to the level-θ surface in the mean field
(http://en.wikipedia.org/wiki/Mahalanobis_distance)

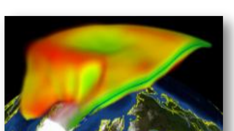


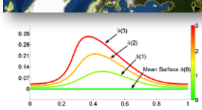
- SDF surfaces are level-sets in the Mahalanobis distance field, enclosing the volume which contains the uncertain surface with a certain probability
  - For Gaussian distributions, the less standard deviations an observation is from the mean, the higher the probability of this observation

Variance as uncertainty indicator

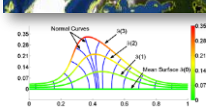
- Uncertainty visualization for an iso-surface in the mean values of a 3D temperature ensemble







Color coding of stochastic distance

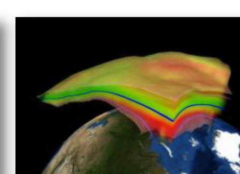


Color coding of Euclidian distance along normal curves [3]

Variance as uncertainty indicator

- Uncertainty visualization for an iso-surface in the mean values of a 3D temperature ensemble

Iso-surface in the mean values of a 3D temperature field



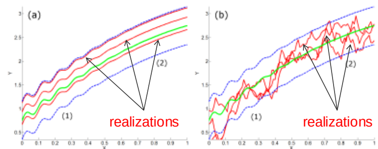
Positional variability of the surface due to uncertainty

Limitations of variance as uncertainty indicator

- The standard deviation describes the local uncertainty but does not allow inferring on possible variations at different positions relative to each other
  - e.g., if mean values and standard deviations at two adjacent points are identical, but the realizations of the random variables at both points behave independently, it cannot be predicted whether there is a positive, zero, or negative derivative of the data between the two points
  - In general, the effect of uncertainty on structures which depend on random values at multiple points cannot be predicted from the mean values and standard deviations alone
    - The effect is to a large extent arbitrary if the realizations of the random values are stochastically independent, while a structure's shape can be assumed stable if the realizations are stochastically dependent

Limitations of variance as uncertainty indicator

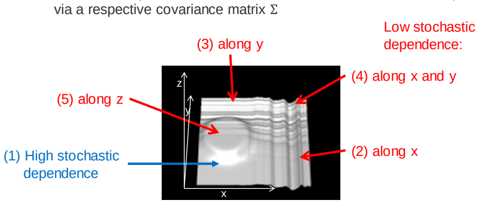
- Variations of an uncertain curve due to uncertainty
  - Curve points $Y_i = Y(x_i)$ are modelled via a multivariate Gaussian random variable with smoothly varying mean values (green curve) and a constant standard deviation (blue curves show corresponding confidence interval)



In a) and b), high and low stochastic dependence between the values at adjacent x_i was modelled

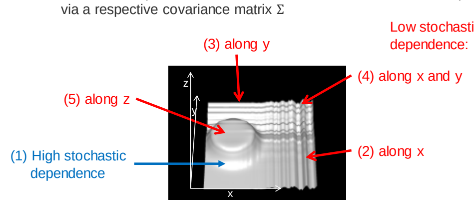
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



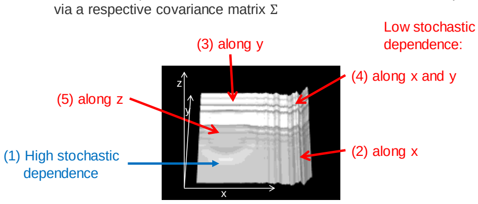
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



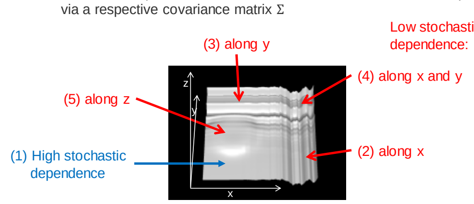
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



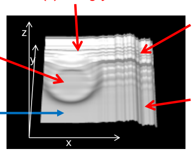
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



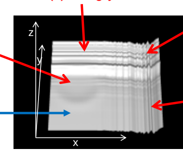
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



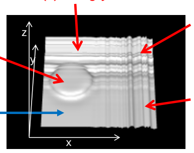
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



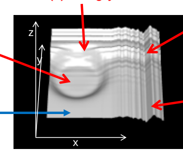
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



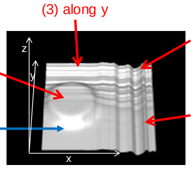
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



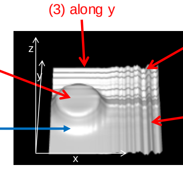
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



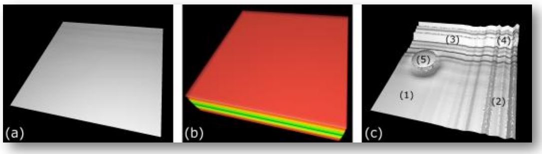
Limitations of variance as uncertainty indicator

- Variations of an iso-surface in an uncertain 3D scalar field
  - Stochastic dependence between scalar values was modelled explicitly via a respective covariance matrix Σ



Limitations of variance as uncertainty indicator

- Random values in the 3D scalar field shown before were generated using a multivariate Gaussian distribution with constant standard deviation and linearly increasing mean (along z)



(a) Iso-surface in the mean values
(b) Confidence volume containing all points that belong to the surface with a certain probability
(c) Occurrence of the surface in (a) for one possible realization of the random values

Limitations of variance as uncertainty indicator

- Conclusion: just by looking at the standard deviation it is impossible to infer on the stability (or confidence) of structures in the mean values
  - A structure is very likely to change its shape due to uncertainty, if the stochastic dependence between the data values making up the region is low
  - A structure is rather unlikely to change (it is stochastically stable), if the stochastic dependence between the data values making up the structure is high, even if the standard deviations are high
- Thus, one important goal in uncertainty visualization is to convey the stochastic dependence and conclude on the stability of structures in the data

Structural uncertainty

- We call the effect of uncertainty in the data values on structures depending on the values at multiple points a structural uncertainty
  - It is associated with the occurrence of particular structures in the data which are affected by the degree of dependence between the values at two or more data points
- Indicators for structural uncertainty are given by the sub-diagonals of Σ, which contain covariance information on relative uncertainties between random variables

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

Structural uncertainty

- For Gaussian distributed random variables, the stochastic dependence of random values at two points is given by the correlation:

$$\rho_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}, \qquad -1 \le \rho_{ij} \le 1$$

Positive correlation: $\rho_{ij} > 0$

Stochastic independence: $\rho_{ij} = 0$

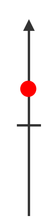
Negative/Inverse correlation: $\rho_{ij} < 0$

Stochastic dependence

- Variability of a scalar value around its mean position:

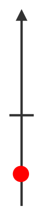
Stochastic dependence

- Variability of a scalar value around its mean position:



Stochastic dependence

- Variability of a scalar value around its mean position:



Stochastic dependence

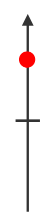
- Variability of a scalar value around its mean position:

Stochastic dependence

- Variability of a scalar value around its mean position:

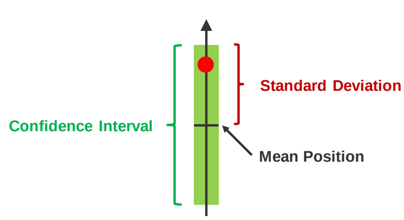
Stochastic dependence

- Variability of a scalar value around its mean position:



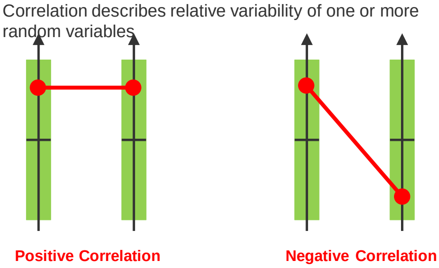
Stochastic dependence

- Variability of a scalar value around its mean position:



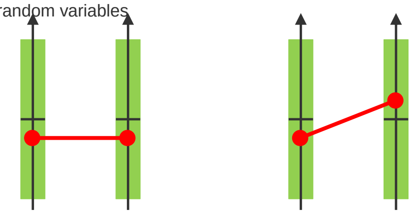
Stochastic dependence

- Correlation describes relative variability of one or more random variables



Stochastic dependence

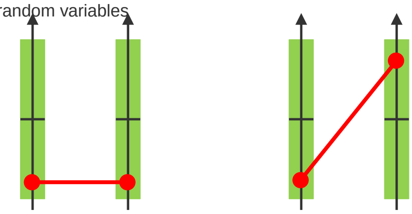
- Correlation describes relative variability of one or more random variables



Positive Correlation **Negative Correlation**

Stochastic dependence

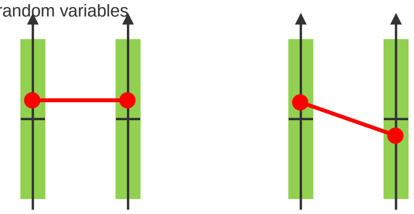
- Correlation describes relative variability of one or more random variables



Positive Correlation **Negative Correlation**

Stochastic dependence

- Correlation describes relative variability of one or more random variables



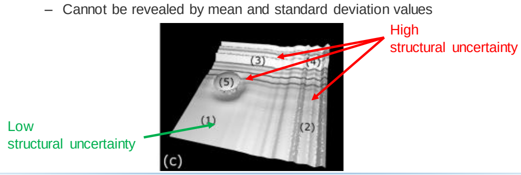
Positive Correlation **Negative Correlation**

Part 2: Local and global correlation visualization

2.1 Local Correlation Visualization

Structural uncertainty

- Remember the example in Part 1: locally varying structural variations of an iso-surface in different realizations of an uncertain 3D scalar field
  - Caused by different correlation structures
  - Resulting in stochastically stable and unstable surface parts
  - Cannot be revealed by mean and standard deviation values

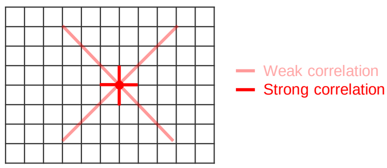


Structural uncertainty

- Problems addressed in this part of the tutorial:
  1. How can one visualize correlations to analyze the structural uncertainty of particular features in 2D and 3D scalar fields, given a set of realizations or a (Gaussian based) stochastic uncertainty model
- Two different approaches will be discussed, both aim at conveying correlation structures in the data:
  1. Local correlation analysis using **characteristic correlation tensors**
  2. Global correlation analysis showing longe-range dependencies such as inverse correlations via **correlation clusters**

Local correlation analysis

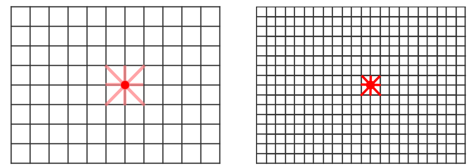
- Assumption of a **distance dependent** correlation model: higher correlations between the data values at points with shorter Euclidean distance



Domain of a 2D scalar data set

Local correlation analysis

- Assumption of a **distance dependent** correlation model implies that correlation strength depends on grid resolution

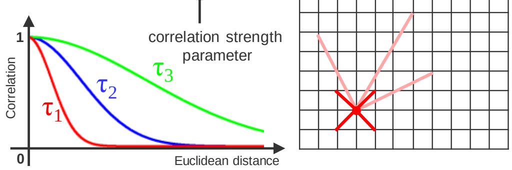


Low correlation to 1-ring neighborhood **High correlation to 1-ring neighborhood**

Distance dependent correlation model

- The Gaussian correlation function relates correlation to spatial distance

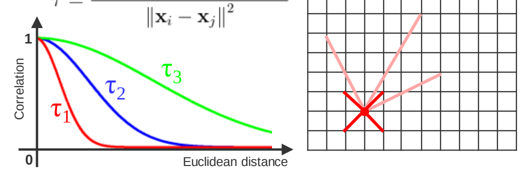
$$\rho(Y(\mathbf{x}_i), Y(\mathbf{x}_j)) = \exp(-\tau \, \|\mathbf{x}_i - \mathbf{x}_j\|^2)$$



Distance dependent correlation model

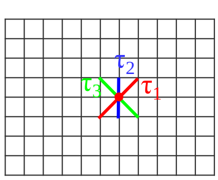
- Correlation strength parameter τ relates correlation to distance and is independent of grid resolution for local correlation analysis

$$\tau = \frac{-\log\left(\rho(Y(\mathbf{x}_i), Y(\mathbf{x}_j))\right)}{\|\mathbf{x}_i - \mathbf{x}_j\|^2}$$



Correlation anisotropy

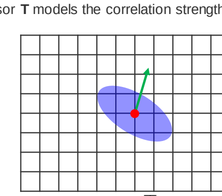
- Problem: Correlation values – and thus correlation strength parameters – depend on direction, yielding a **directional correlation distribution** at each grid point



A single correlation value can not represent the stochastic dependences!

Correlation strength tensor

- Idea: The (anisotropic) distribution of the correlation strength parameter is modelled by a rank-2 tensor **T**
  - The correlation strength tensor **T** models the correlation strength to 8 (2D) or 26 (3D) neighbors
  - Given **T**, correlation strength $\boldsymbol{\tau}$ at position $\boldsymbol{x}$ into direction $\boldsymbol{r}$ can be computed via tensor-vector multiplication



$$\tau(\mathbf{x}, \mathbf{r}) = \mathbf{r}^\top \mathbf{T}(\mathbf{x})\mathbf{r}$$

Correlation strength tensor

- Given the tensor **T**, the distance and direction dependent correlation becomes

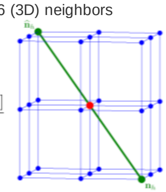
$$\rho(Y(\mathbf{x}_i), Y(\mathbf{x}_j)) = \exp\left(-\frac{(\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{T}(\mathbf{x}_i - \mathbf{x}_j)}{\|\mathbf{x}_i - \mathbf{x}_j\|}\right)$$

$$\mathbf{T} \quad 0.5(\mathbf{T}(\mathbf{x}_i) + \mathbf{T}(\mathbf{x}_j))$$

- At every grid point a correlation tensor is computed
  - i.e., correlation data is transformed into a distance dependent tensor model
  - Reduces memory requirments because it avoids storing (global) correlation values explicitly

Correlation strength tensor computation

- **T** is symmetric and has 3 (2D) or 6 (3D) free values to be determined
- **T** models the correlation strength to 8 (2D) or 26 (3D) neighbors
- **T** is obtained for every grid point by solving an over-determined linear system using a least squares approach



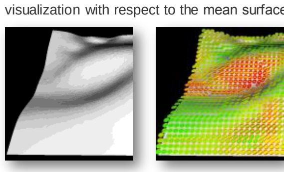
$$\mathbf{r}^\top \mathbf{T} \mathbf{r} = \frac{-\log\left[0.5(|\rho(Y(\mathbf{x}_i), Y(\mathbf{n}_h))| + |\rho(Y(\mathbf{x}_i), Y(\hat{\mathbf{n}}_h))|)\right]}{\|\mathbf{n}_h - \mathbf{x}_i\|}$$

$$\mathbf{r} = \frac{\mathbf{n}_h - \hat{\mathbf{n}}_h}{\|\mathbf{n}_h - \hat{\mathbf{n}}_h\|} \quad , \quad \mathbf{n}_h \in \mathbb{N}(\mathbf{x}_i), \quad h \in \{1, 2, ..., 13\}$$

- **T** is independent of the grid resolution, because correlation is set in relation to Euclidean distances

Correlation visualization on iso-surfaces

- **Goal**: Visualizing the structural uncertainty of specific features in the mean data, i.e. an iso-surface
  - Requires correlation visualization with respect to the mean surface's geometric shape



- **Idea**: Distinguish between correlations in the surface's local tangent planes and normal direction, and visualize these correlations via glyphs

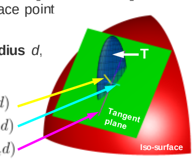
Correlation visualization on iso-surfaces

- Approach: Correlation values are extracted for the two principal correlation directions in the surface's tangent plane (highest and lowest strength) and normal direction [4]
  - Basis transformation is necessary to obtain correlation strength parameters in principal tangential directions (τ_{tan_max}) and (τ_{tan_min}), and normal direction (τ_{normal}) at every surface point
  - By interactively specifying an **Euclidean radius** *d*, correlation values along the three principal directions are computed as

$$\rho_{normal} = \exp\left(-\tau_{normal} d\right)$$

$$\rho_{tan\_min} = \exp\left(-\tau_{tan\_min} d\right)$$

$$\rho_{tan\_max} = \exp\left(-\tau_{tan\_max} d\right)$$



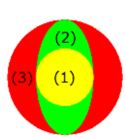
and used to model the appearance of circular correlation glyphs

Correlation glyphs

- Correlation glyphs encode the correlation values in the **first and second pricipal tangent** and **surface normal direction**

Glyph coloring wrt correlation in surface …

(1) … normal direction
(2) … first principal tangent direction
(3) … second principal tangent direction



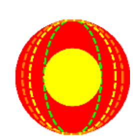
0 → Correlation → 1

Correlation glyphs

- **Correlation ratio** between first and second tangent direction is encoded in the shape of the elliptic zone (2)

Glyph coloring wrt correlation in surface …

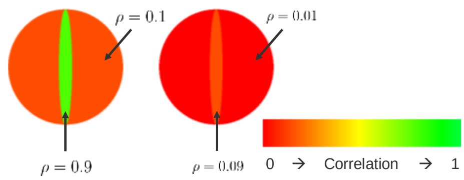
(1) … normal direction
(2) … first principal tangent direction
(3) … second principal tangent direction



0 → Correlation → 1

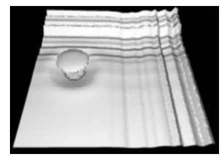
Correlation glyphs

- **Absolute** and **relative** correlation anisotropy in the surface tangent plane are encoded via color differences

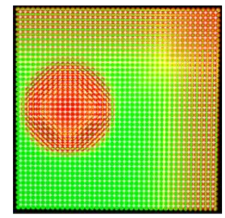


Structural uncertainty analysis

- Surface realization and glyph based correlation visualization on mean surface

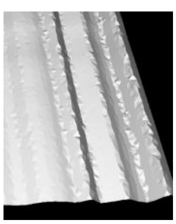


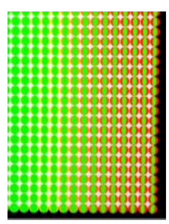
Surface realization



Correlation glyphs

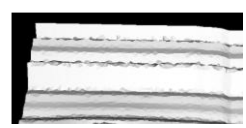
Structural uncertainty analysis

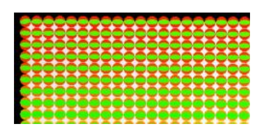




High and low correlation in vertical and horizontal direction

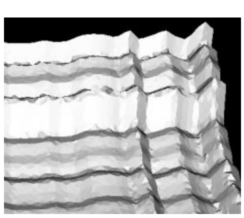
Structural uncertainty analysis

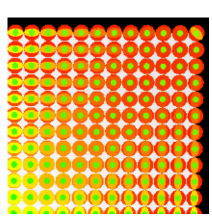




High and low correlation in horizontal and vertical direction

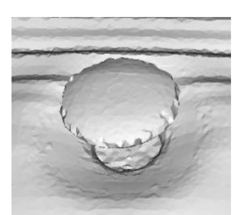
Structural uncertainty analysis

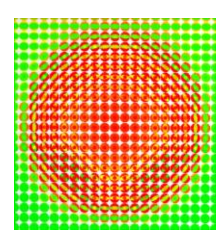




Low correlation in both vertical and horizontal direction

Structural uncertainty analysis

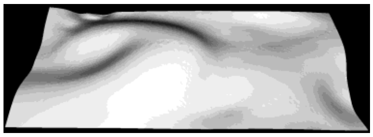




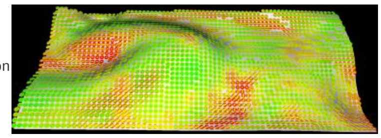
Low correlation especially in normal direction

Structural uncertainty analysis

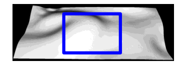
Mean surface
in a simulated
temperature ensemble

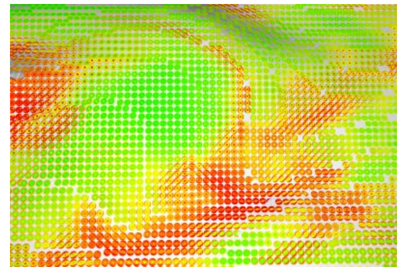


Glyph based
correlation visualization

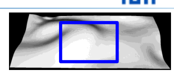


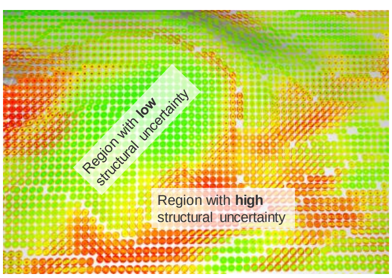
Structural uncertainty analysis





Structural uncertainty analysis





Local correlation analysis summary

- Local distance dependent correlation tensor model allows **memory reduction to *O(n)*** for *n* grid points
- Visualization of **local correlation anisotropy** possible
- Simultaneous visualization of correlation **ratios** as well as **absolute** correlation values
- Interactive analysis of **anisotropy** in surface tangent plane and between tangent plane and surface normal direction

Local correlation analysis summary

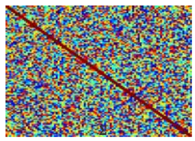
- No integration of **absolute uncertainty information** (e.g. standard deviations)
- **Global dependences** are ignored
- **Inverse correlation** cannot be visualized as correlation model only accounts for correlation strengths (magnitudes)

Part 2: Local and global correlation visualization

2.2 Global correlation visualization

Global correlation analysis

- To analyze **long-range spatial dependencies** between the data values in uncertain scalar fields, in principle the **full covariance matrix** needs to be visualized
- This is problematic due to the following reasons
  - For $\boldsymbol{n}$ spatial grid points it requires $\boldsymbol{O(n^2)}$ **memory**
  - Entries in $\boldsymbol{\Sigma}$ are not in a **spatial contex**, making it difficult to infer on spatial dependencies
  - **Uncertainty information** has to be visually separated from correlation information

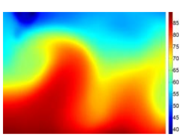


Global correlation analysis

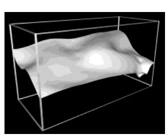
- Requirements due to the aforementioned problems
  - The correlation information first has to be condensed
    - Requires to define and seek for the **most prominent short- and long-range dependencies**
  - Correlation structures have to be embedded into standard visualizations providing **spatial relationships**
- Approach
  - Definition of a measure for the **degree of dependency** of a random variable to its local and global spatial surrounding
  - **Spatial clustering** of random variables based on this measure
  - **Color coding** of data points wrt. cluster IDs

Example data set

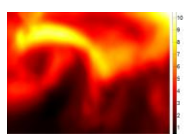
- 2D temperature ensemble simulated by the European Centre for Medium-Range Weather Forecasts



Mean values



Mean values as heightfield



Standard deviations

Correlation strength indicator

- Assumption: the more „correlation partners“ a particular point in the domain has to which the correlation is larger than a given level, the more important this point is
  - The sets of partners spread anisotropically across the domain
  - The expansion in different directions is directly related to the correlation distribution in the respective region
- For a given level and point x_i, the number of partners indicates the degree of dependency between the random variable $Y(x_i)$ to its local and global spatial surrounding
  - It counts the most prominent partners of x_i, independent of their position in the domain
  - The number of partners is used as the correlation strength indicator

Correlation strength indicator

- Definitions

Correlation neighborhood:

$$\eta_{\rho^+}(x_i) := \left\{x_j \in D \mid \rho\left(Y(x_i), Y(x_j)\right) \geq \rho^+\right\}$$

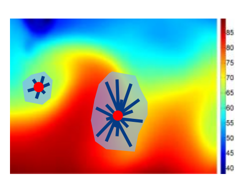
D: $data\ domain$

ρ_1: $correlation\ level$

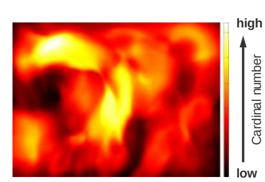
Cardinal number: $|\eta_{\rho^+}(x_i)|$

Correlation strength indicator

- Example showing relation between correlation neighborhoods and cardinal numbers



Correlation neighborhoods of 2 grid points



Cardinal numbers

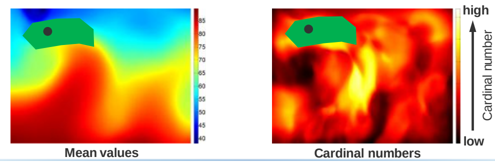
Spatial correlation clustering

- **Basic idea**: use correlation neighborhoods as clusters
- **Problem**: correlation neighborhoods of different points can overlap
  - Results in ambiguities in the assignment of points to clusters
- **Solution**: process correlation neighborhoods sequentially and exclude those neighborhoods which overlap any previously processed neighborhood

- The strategy avoids overlapping clusters but, depending on selection order, will exclude large clusters which overlap small cluster
- Thus, neighborhoods are selected in descending order of cardinality such that the largest clusters are always selected first

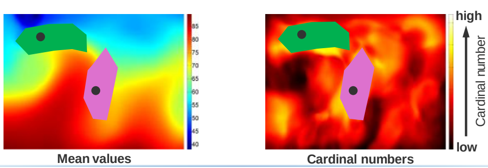
Correlation clustering algorithm

- **Step 1:** Select correlation level ρ^+
- **Step 2:** Compute cardinal numbers for all points in the 2D domain
- **Step 3:** Select domain point with largest cardinal number as first cluster center
- **Step 4:** Create cluster (correlation neighborhood)



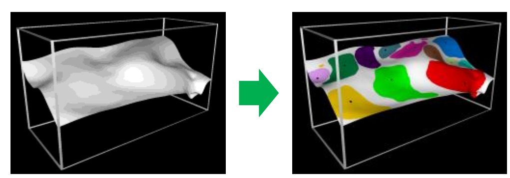
Correlation clustering algorithm continued

- **Step 5:** In the set of points not yet assigned to a cluster, search for point *p* with highest cardinal number whose correlation neighborhood *c* does not overlap any existing cluster
- **Step 6:** Create cluster using *p* and *c* and repeat with Step 5
- **…** generate custers for different levels ρ^+ **(multilevel clustering)**

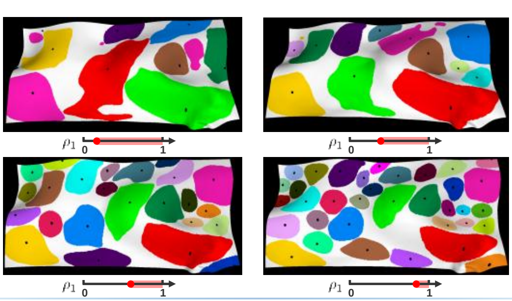


Correlation clustering results

- Visualization of correlation clusters and cluster centers on the mean height surface for a particular correlation level ρ^+

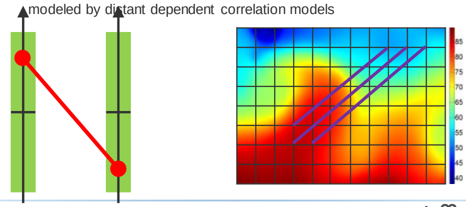


Correlation clustering results for varying levels



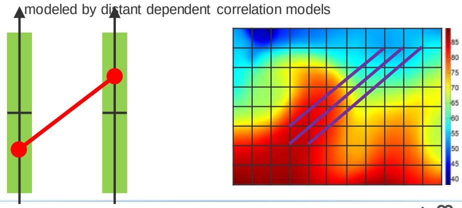
Inverse dependencies

- Inverse stochastic dependencies typically exist between **spatially separated regions**; ie. pairs which are inversely correlated to each other
  - **Inverse correlation structures** are global features and cannot be modeled by distant dependent correlation models



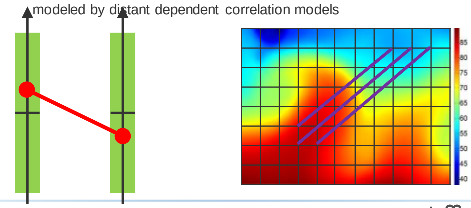
Inverse dependencies

- Inverse stochastic dependencies typically exist between **spatially separated regions**; ie. pairs which are inversely correlated to each other
  - **Inverse correlation structures** are global features and cannot be modeled by distant dependent correlation models



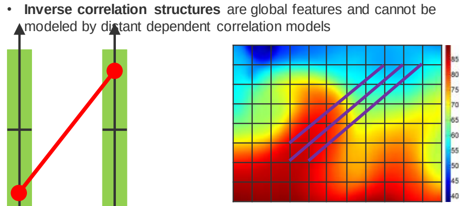
Inverse dependencies

- Inverse stochastic dependencies typically exist between **spatially separated regions**; ie. pairs which are inversely correlated to each other
  - **Inverse correlation structures** are global features and cannot be modeled by distant dependent correlation models



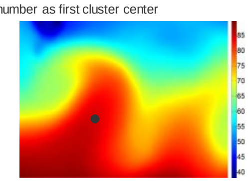
Inverse dependencies

- Inverse stochastic dependencies typically exist between **spatially separated regions**; ie. pairs which are inversely correlated to each other
  - **Inverse correlation structures** are global features and cannot be modeled by distant dependent correlation models



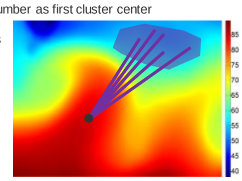
Inverse correlation clustering algorithm

- **Step 1:** Select negative correlation level ρ^-
- **Step 2**: Compute correlation neighborhoods $\left(\rho\left(Y(x_i),Y(x_j)\right) \le \rho^-\right)$ and corresponding cardinal numbers for all points
- **Step 3**: Select point with largest cardinal number as first cluster center



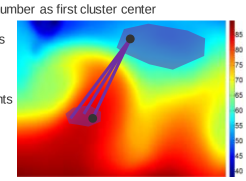
Inverse correlation clustering algorithm

- **Step 1:** Select negative correlation level ρ^-
- **Step 2**: Compute correlation neighborhoods $\left(\rho\left(Y(x_i),Y(x_j)\right) \le \rho^-\right)$ and corresponding cardinal numbers for all points
- **Step 3**: Select point with largest cardinal number as first cluster center
- **Step 4:** Assign negatively correlated points to first cluster



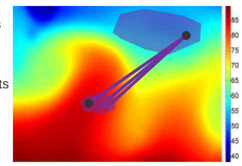
Inverse correlation clustering algorithm

- **Step 1:** Select negative correlation level ρ^-
- **Step 2**: Compute correlation neighborhoods $\left(\rho\left(Y(x_i),Y(x_j)\right) \le \rho^-\right)$ and corresponding cardinal numbers for all points
- **Step 3**: Select point with largest cardinal number as first cluster center
- **Step 4:** Assign negatively correlated points to first cluster
- **Step 5:** Select point in cluster with largest number of inversely correlated points



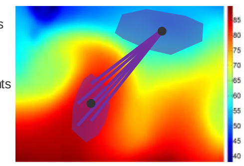
Inverse correlation clustering algorithm

- **Step 1:** Select negative correlation level ρ^-
- **Step 2**: Compute correlation neighborhoods $(\rho\left(Y(x_i),Y(x_j)\right) \leq \rho^-)$ and corresponding cardinal numbers for all points
- **Step 3**: Select point with largest cardinal number as first cluster center
- **Step 4:** Assign negatively correlated points to first cluster
- **Step 5:** Select point in cluster with largest number of inversely correlated points



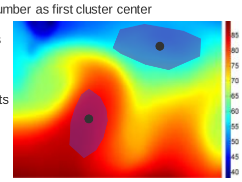
Inverse correlation clustering algorithm

- **Step 1:** Select negative correlation level ρ^-
- **Step 2**: Compute correlation neighborhoods $(\rho\left(Y(x_i),Y(x_j)\right) \leq \rho^-)$ and corresponding cardinal numbers for all points
- **Step 3**: Select point with largest cardinal number as first cluster center
- **Step 4:** Assign negatively correlated points to first cluster
- **Step 5:** Select point in cluster with largest number of inversely correlated points



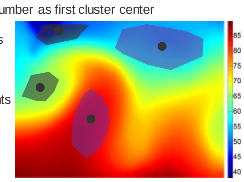
Inverse correlation clustering algorithm

- **Step 1:** Select negative correlation level ρ^-
- **Step 2**: Compute correlation neighborhoods $(\rho\left(Y(x_i),Y(x_j)\right) \leq \rho^-)$ and corresponding cardinal numbers for all points
- **Step 3**: Select point with largest cardinal number as first cluster center
- **Step 4:** Assign negatively correlated points to first cluster
- **Step 5:** Select point in cluster with largest number of inversely correlated points
- **Step 6:** Assign both clusters to one pair



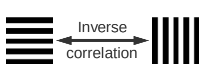
Inverse correlation clustering algorithm

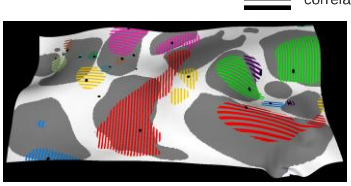
- **Step 1:** Select negative correlation level ρ^-
- **Step 2**: Compute correlation neighborhoods $(\rho\left(Y(x_i),Y(x_j)\right) \leq \rho^-)$ and corresponding cardinal numbers for all points
- **Step 3**: Select point with largest cardinal number as first cluster center
- **Step 4:** Assign negatively correlated points to first cluster
- **Step 5:** Select point in cluster with largest number of inversely correlated points
- **Step 6:** Assign both clusters to one pair
- **Step 7:** Proceed with Step 3, taking into account that clusters do not overlap



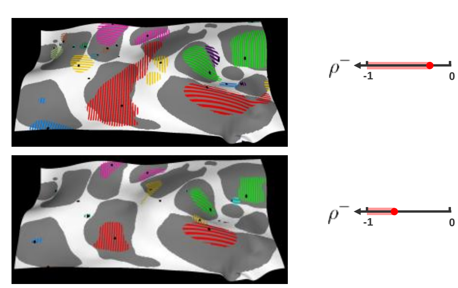
Inverse dependencies

- Each pair of inversely correlated spatial regions gets assigned one distinct color
- Clusters in each pair are distinguished by differently oriented stripe patterns



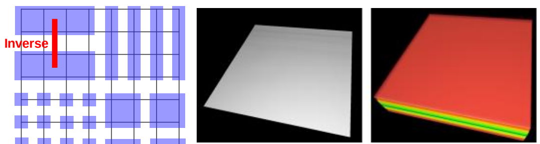


Inverse dependencies for varying negative levels ρ^-



Global correlation analysis results

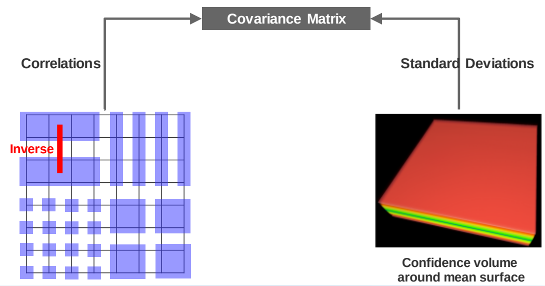
- Example: An uncertain 2D scalar field with **constant** means and standard deviations. Specific orrelation structures have been enforced:
  - **Strong positive correlation** within the sets of points colored blue.
  - **Zero correlation** between these sets
  - **Inverse correlation** between the two sets in the upper left region



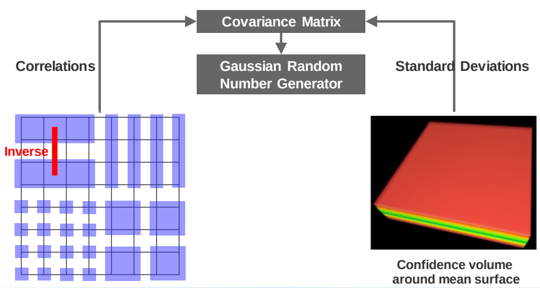
Mean surface

Confidence volume around mean surface

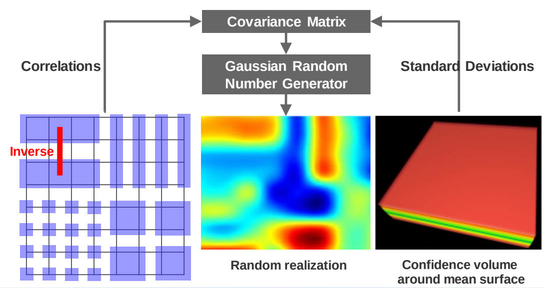
Global correlation analysis results



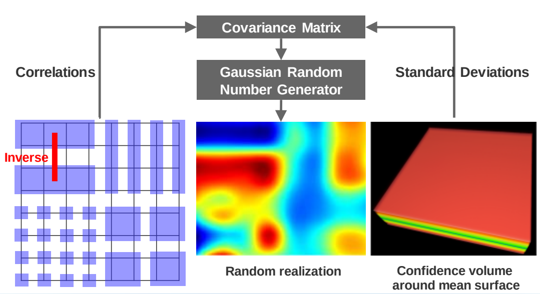
Global correlation analysis results



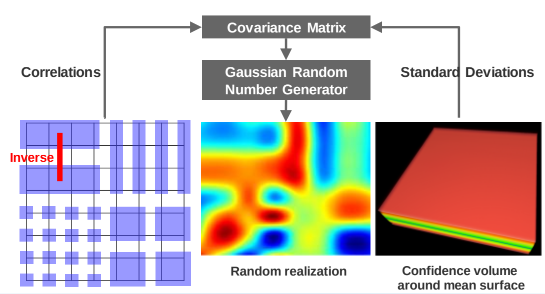
Global correlation analysis results



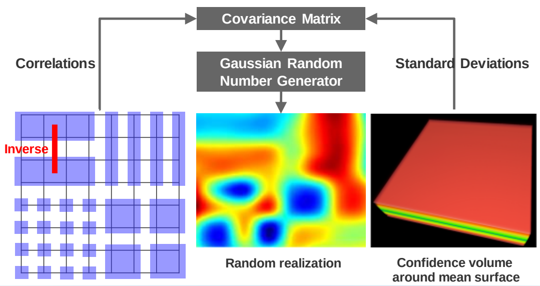
Global correlation analysis results



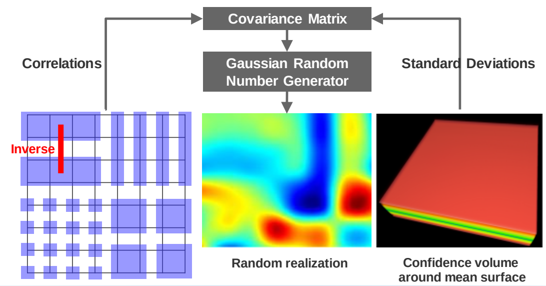
Global correlation analysis results



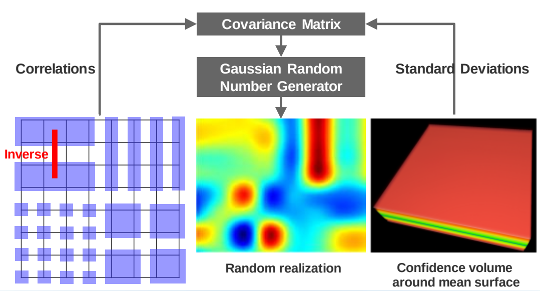
Global correlation analysis results



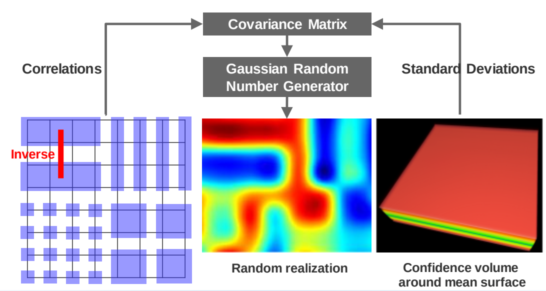
Global correlation analysis results



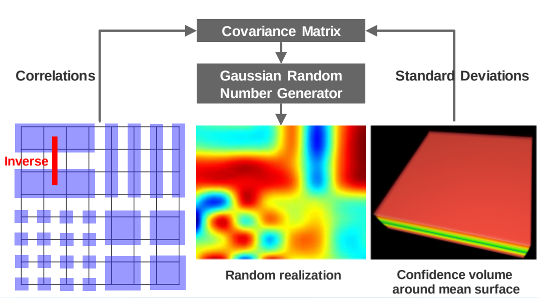
Global correlation analysis results



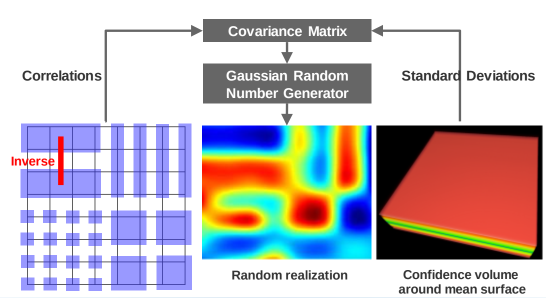
Global correlation analysis results



Global correlation analysis results

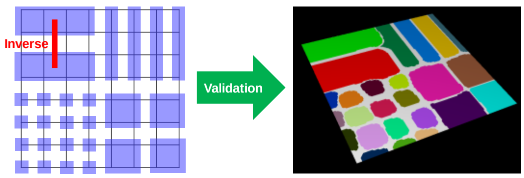


Global correlation analysis results



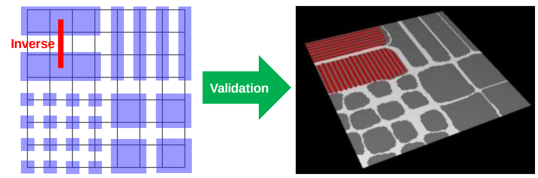
Global correlation analysis results

- Correlation clustering algorithm identifies **positively correlated regions** correctly



Global correlation analysis results

- Correlation clustering algorithm identifies **inversely correlated regions** correctly



Global correlation analysis summary

- Correlation clustering allows analyzing **global correlation structures** such as inverse correlations, requiring an amount of memory that is linear in the number of data points
- The clusters‘ distributions reveal **anisotropic correlation effects**
- **Uncertainty information** can be integrated easily, for instance, by extruding clusters depending on standard deviation [5]
- Extension to 3D is possible, but special projection or restriction schemes are required when used to analyze correlation structures on iso-surfaces

Future work and challenges in correlation visualization

- Approaches for visualizing correlation structures in 3D
- Correlation analysis/visualization for other data types (e.g. vector fields)
- Integration of correlation information in existing uncertainty visualization approaches (e.g. positional uncertainty of features)
- Quantification of the effect of correlation on the occurence of differential quantities and higher order features, like critical points
- Modeling and interpretation of stochastic dependencies in non-Gaussian distributed random fields