Data driven parametrization of random fields in large structures

Christian Bucher
Center of Mechanics and Structural Dynamics
Vienna University of Technology
& DYNARDO in Austria

Introduction

- Spatial variation of structural properties lead to uncertainties in the structural performance
- Slightly different variation of results in each node/element of the structural mesh
- Large number of random variables, can be described as random field
- For statistical analyses it is important to reduce the number of random variables
- For engineering interpretation it is helpful to reduce noise and keep essential features

Random field

- Real-valued function in n-dimensional space

$$H \in \mathbb{R}; \quad \mathbf{x} = [x_1, x_2, \ldots x_n]^T \in \mathcal{D} \subset \mathbb{R}^n$$

- Mean value function

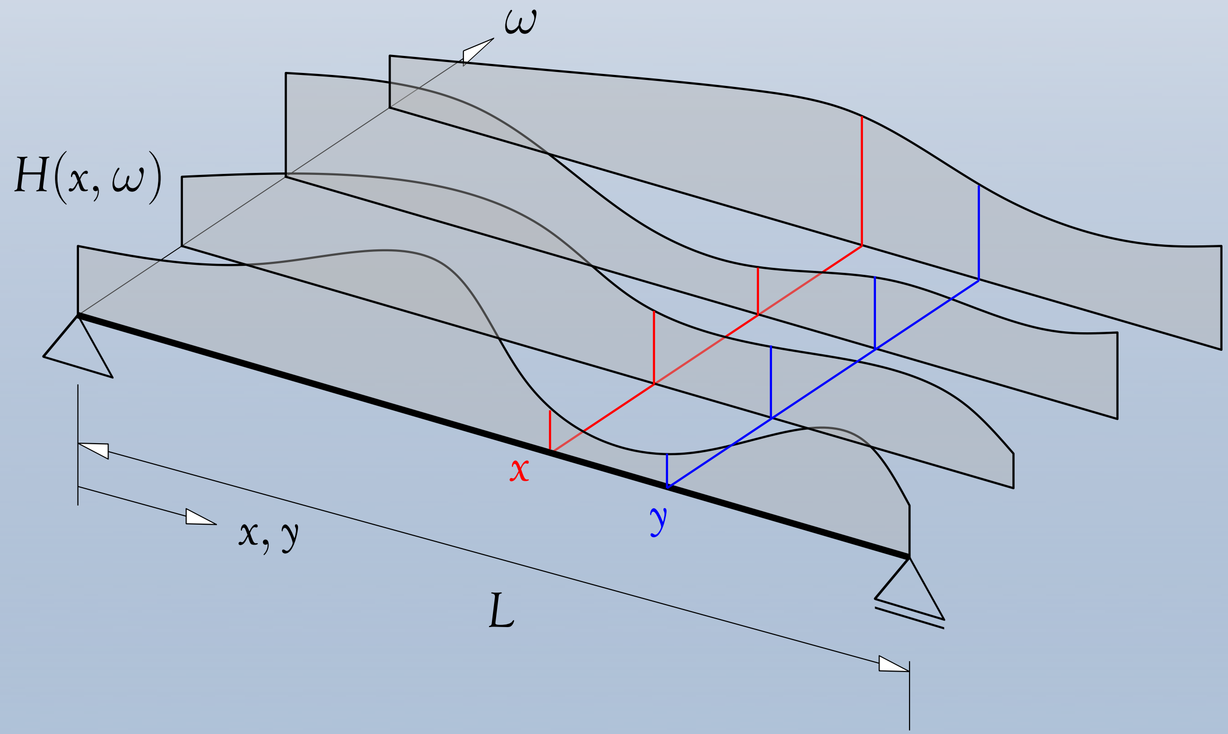
$$\bar{H}(\mathbf{x}) = \mathbf{E}[H(\mathbf{x})]$$

- Auto-covariance function

$$C_{HH}(\mathbf{x}, \mathbf{y}) = \mathbf{E}[\{H(\mathbf{x}) - \bar{H}(\mathbf{x})\}\{H(\mathbf{y}) - \bar{H}(\mathbf{y})\}]$$

Ensemble

- Different realizations of one-dimensional field



Essential properties of random fields

- Weak homogeneity

$$\bar{H}(\mathbf{x}) = \text{const.} \quad \forall \mathbf{x} \in \mathcal{D}$$

$$C_{HH}(\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}) = C_{HH}(\boldsymbol{\xi}) \quad \forall \mathbf{x}, \mathbf{x} + \boldsymbol{\xi} \in \mathcal{D}$$

- Isotropy

$$C_{HH}(\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}) = C_{HH}(\|\boldsymbol{\xi}\|) \quad \forall \mathbf{x}, \mathbf{x} + \boldsymbol{\xi} \in \mathcal{D}$$

Spectral decomposition

- Fourier-type series expansion using deterministic basis functions ϕ_k and random coefficients c_k

$$H(\mathbf{x}) = \sum_{k=1}^{\infty} c_k \phi_k(\mathbf{x}), \qquad \mathbf{x} \in \mathcal{D} \subset \mathbb{R}^n; c_k, \phi_k \in \mathbb{R}$$

- Karhunen-Loeve expansion based on eigenvalue decomposition of the auto-covariance function

$$C_{HH}(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} \lambda_k \phi_k(\mathbf{x}) \phi_k(\mathbf{y})$$

$$\int_{\mathcal{D}} C_{HH}(\mathbf{x}, \mathbf{y}) \phi_k(\mathbf{x}) \mathrm{d}\mathbf{x} = \lambda_k \phi_k(\mathbf{y})$$

- Leads to orthogonal basis functions and uncorrelated coefficients (convenient, but not required)

Spatially discrete formulation

- Discrete values of random field

$$H_i = H(\mathbf{x_i}); \quad i = 1 \dots N$$

- Spectral representation

$$H_i = \sum_{k=1}^{N} \phi_k(\mathbf{x_i}) c_k = \sum_{k=1}^{N} \phi_{ik} c_k$$

- Written as matrix-vector multiplication

$$\mathbf{H} = \mathbf{\Phi c}$$

Types of random fields

- Feature fields
  - Contain prominent features in all realizations
  - Strongly inhomogeneous
- Noise fields
  - Consist of purely random values
  - May be considered to be homogeneous
- Real field
  - Combination of feature field and noise field

Choice of basis functions

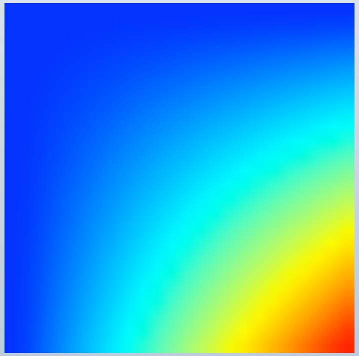
- Reduce number of random variables significantly
  - Improves statistical significance for small sample size
  - Reduces numerical effort in statistical analysis
  - Simplifies representation of input/output relations based on meta-models
- Basis functions should be orthogonal
  - Reduces computational effort for projection/reduction
- Random coefficients should be uncorrelated
  - Simplifies digital simulation of random fields

Example - JPEG Data reduction

- Basis functions are cosines with different wave lengths
- Suitable for rectangular domains
- Very efficient for smoothly varying data
- Convergence difficulties near jumps in data (Gibb's phenomenon)

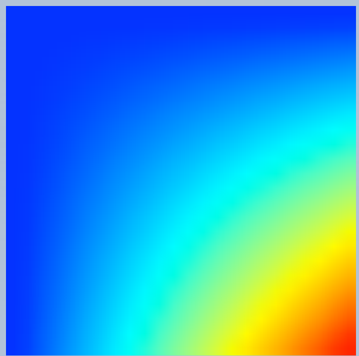
Smoothly varying data

- Original (24x24)

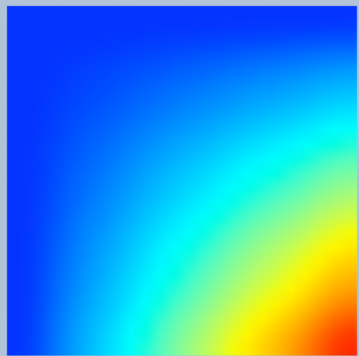


- Reduced

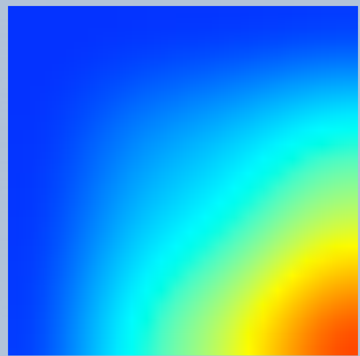
16x16



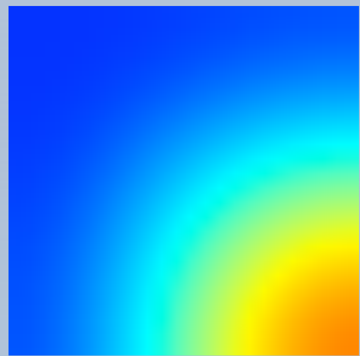
8x8



4x4

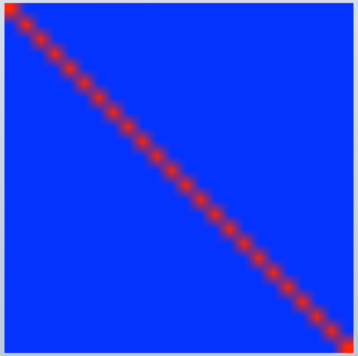


2x2



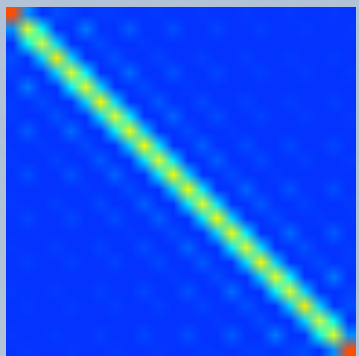
Rapidly varying data

- Original (24x24)

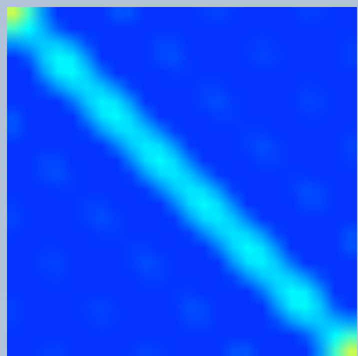


- Reduced

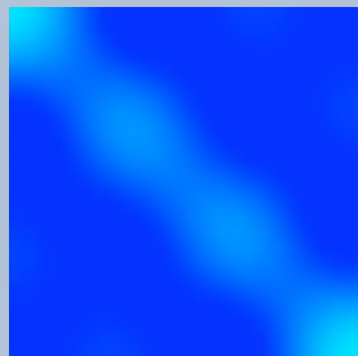
16x16



8x8



4x4



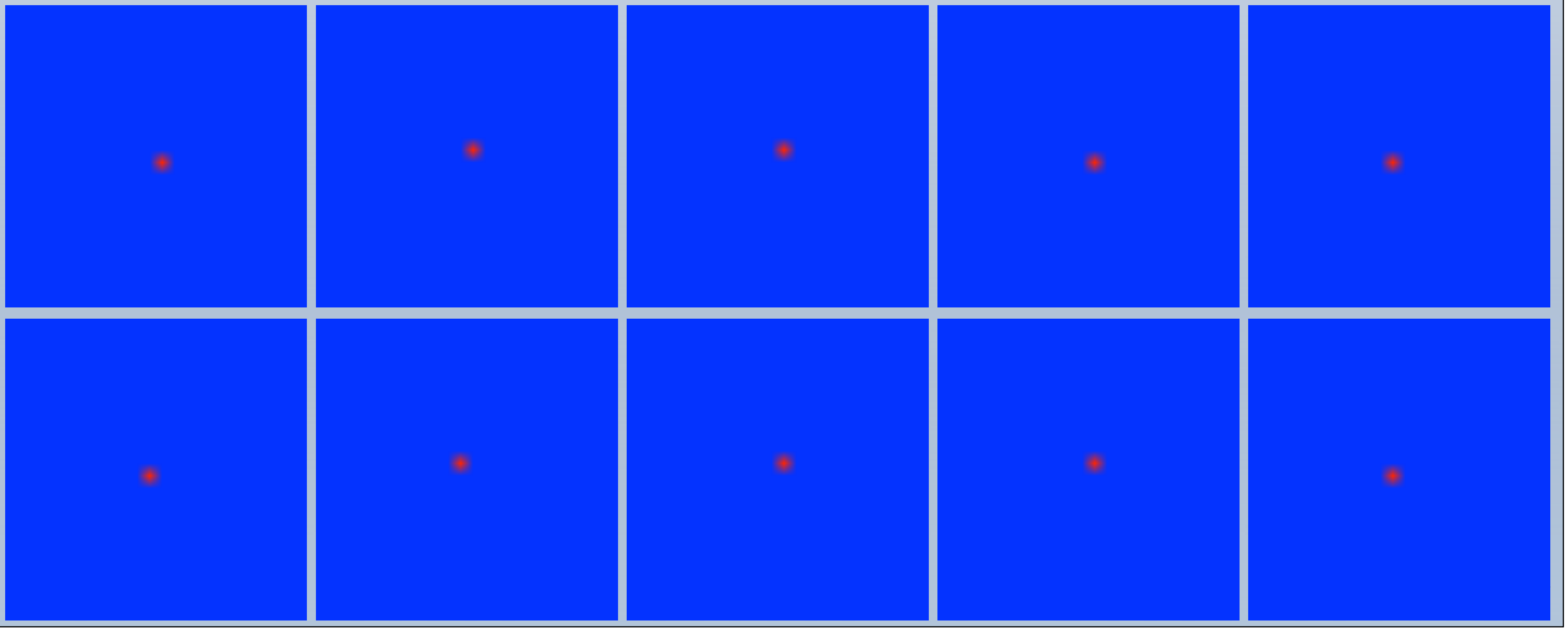
2x2

How to improve convergence?

- Describe features independently from the noise
- Map severely inhomogeneous field to a more “homogeneous” field
- Generate standardized samples
  - Subtract mean value from original sample functions
  - Divide samples by standard deviations (if non-zero), set samples to zero otherwise

Example: Hot spot

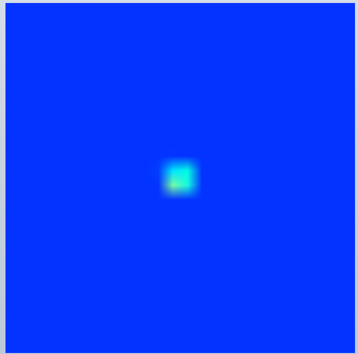
- Random background value (constant)
- One hot spot at random location in the vicinity of the center
- Sample functions

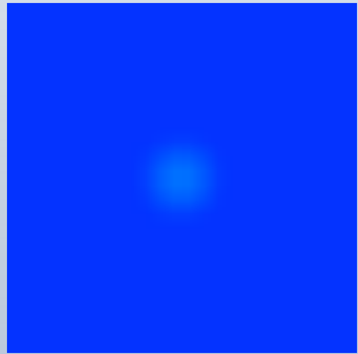


Statistics based on compressed data

- Mean value

Original

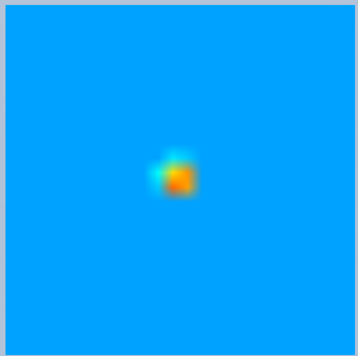


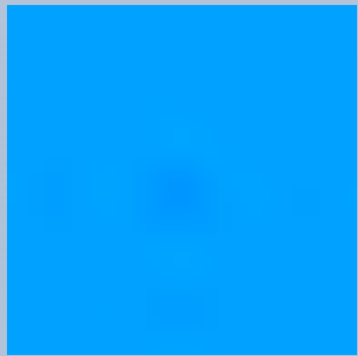


Reduced

- Standard deviation

Original





Reduced

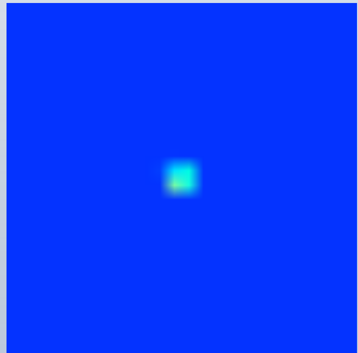
Standardization of data

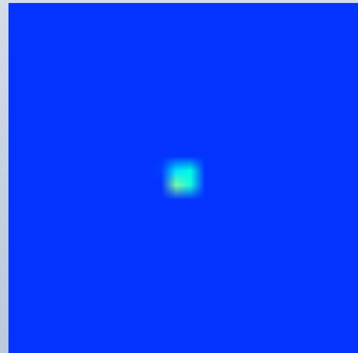
- map to a more “homogeneous” field (zero mean and constant standard deviation
- represents the deviations from the mean in terms of basis functions
- very helpful if the randomness expressed by the standard deviation is related to the mean (e.g. almost constant coefficient of variation)
- can easily represent completely deterministic areas in a structure

Standardized and reduced data

- Mean value

Original

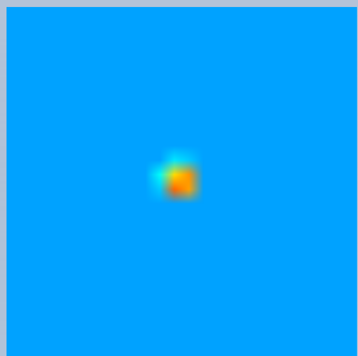


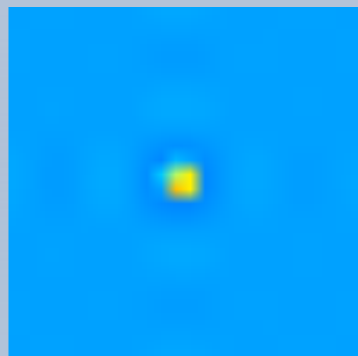


Standardized/
Reduced

- Standard deviation

Original





Standardized/
Reduced

Large structures

- Large number of elements or nodes leads to an un-manageable number of random variables
- Essential to reduce number of random variables **before** application of random field methods
- Suitable approach: represent random data by local averages
- Averaging can be achieved
  - Mesh coarsening
  - Spatial smoothing using appropriate functions
- Essential to maintain topological structure (required for physical interpretation)

Smoothing

- Original space $\mathbf{x}$
- Standardized space $$\mathbf{x}_h = \frac{\mathbf{x} - \bar{\mathbf{x}}}{\sigma_{\mathbf{x}}}$$
- Project into a smoothed space (with correlated variables **y**) $$\mathbf{y} = \boldsymbol{\Theta}^T \mathbf{x}_h$$
- Project back to standardized space $$\tilde{\mathbf{x}}_h = \boldsymbol{\Theta}\mathbf{y}$$
- Project back to original space $$\tilde{\mathbf{x}} = \sigma_{\mathbf{x}}\tilde{\mathbf{x}}_h + \bar{\mathbf{x}}$$
- Measure of loss of detail $$S = \frac{||\mathbf{x} - \tilde{\mathbf{x}}||}{||\mathbf{x}||}$$

Principal component analysis

- Further reduction of number of variables
- Operates in **smoothed** space by applying eigenvalue decomposition, choose number of eigenvalues based on representation of total variance

$$\mathbf{Q}^T \mathbf{C}_{\mathbf{yy}} \mathbf{Q} = \mathbf{\Gamma}$$

- Projection into reduced space with uncorrelated variables $\mathbf{z}$

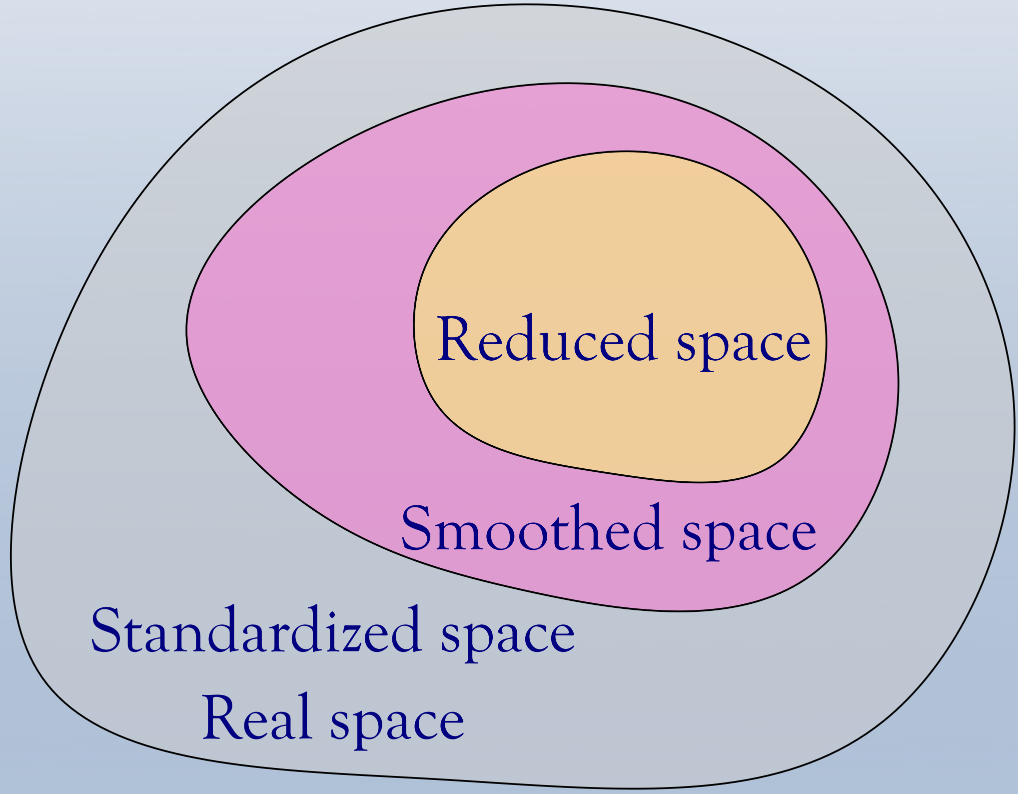
$$\mathbf{z} = \mathbf{Q}^T \mathbf{y}$$

- Projection back to smoothed space (this can also be used for Monte Carlo simulation)

$$\hat{\mathbf{y}} = \mathbf{Q}\mathbf{z}$$

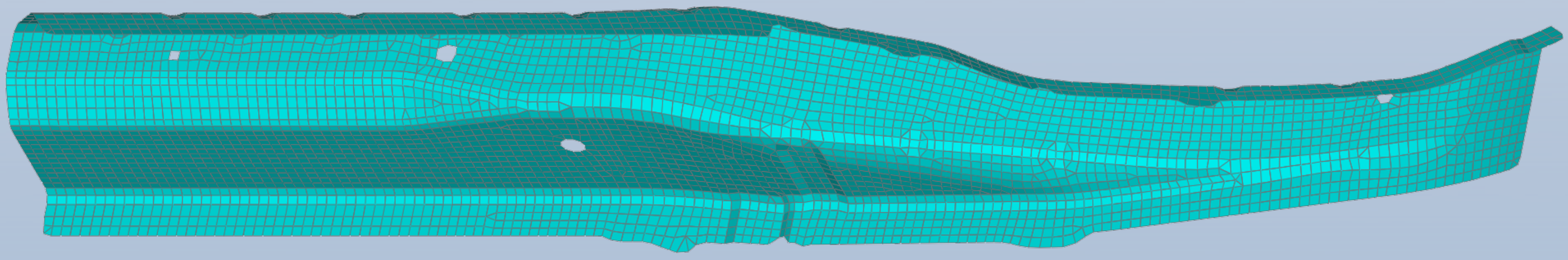
Variable spaces

- Mapping from real space to standardized space is lossless
- Mapping from standardized space to smoothed space is lossy
- Mapping from smoothed space to reduced space is lossy



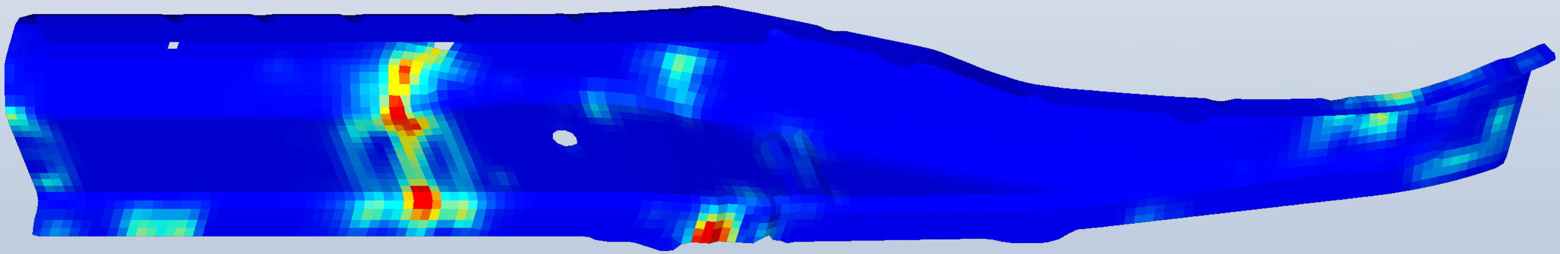
Example - small structure

- 4826 elements
- 150 samples
- Data show effective plastic strain

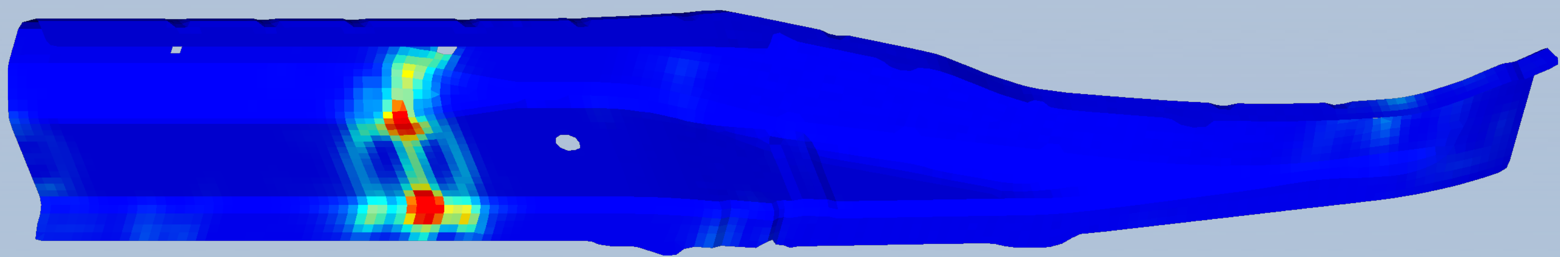


Statistics from 150 samples

- Mean value



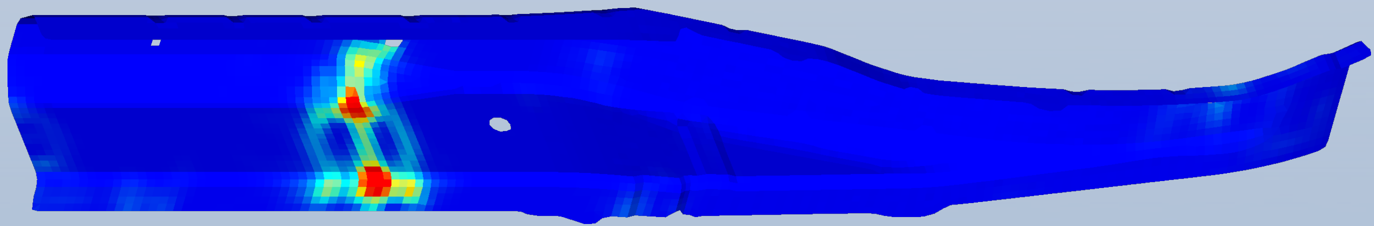
- Standard deviation (COV = 300%)



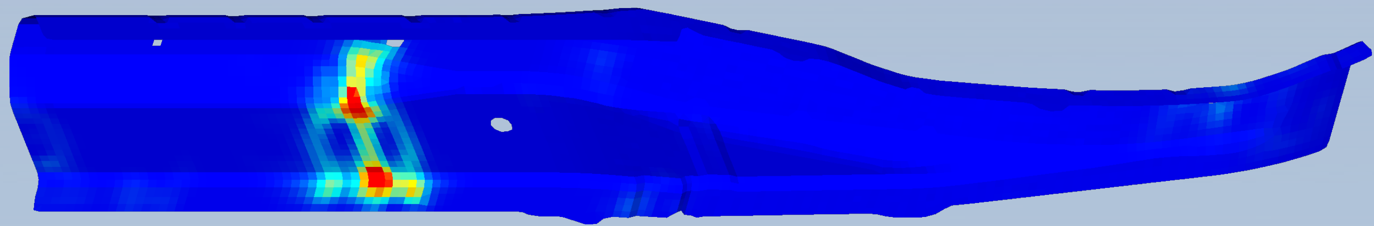
Smoothing and reduction

- 100 basis vectors for smoothing
- 9 random variables for reduction (accuracy of variance: 99%)
- Standard deviation:

Original



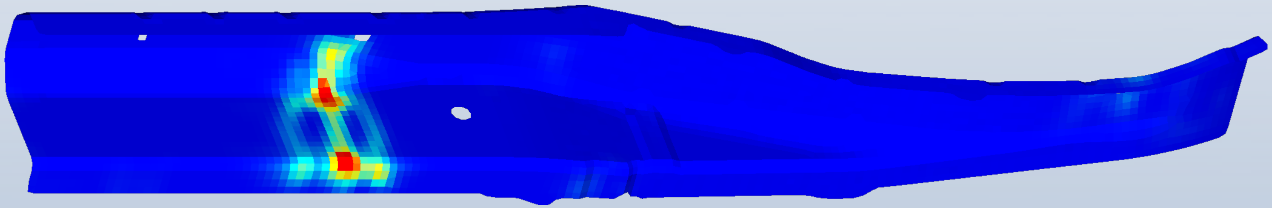
Smoothed
and reduced



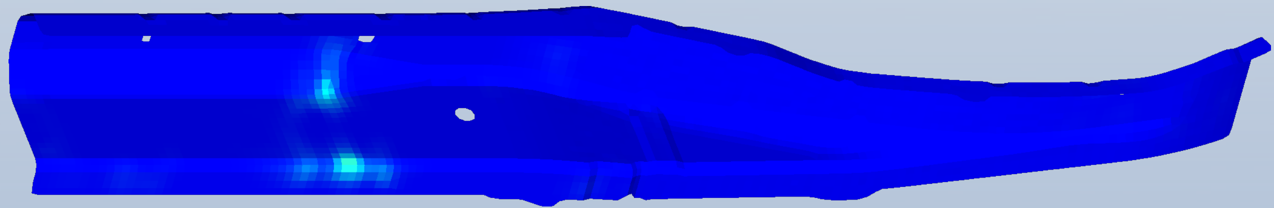
Statistics in reduced space

- Standard deviation due to individual random variables

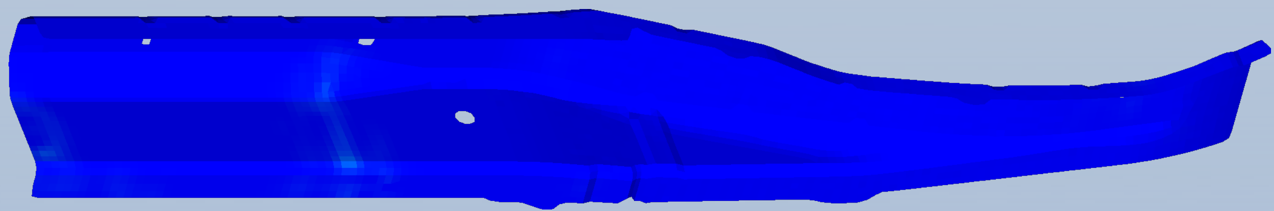
Variable 1



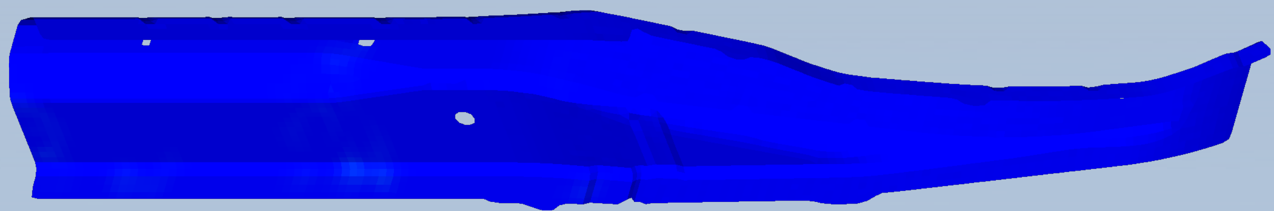
Variable 2



Variable 3

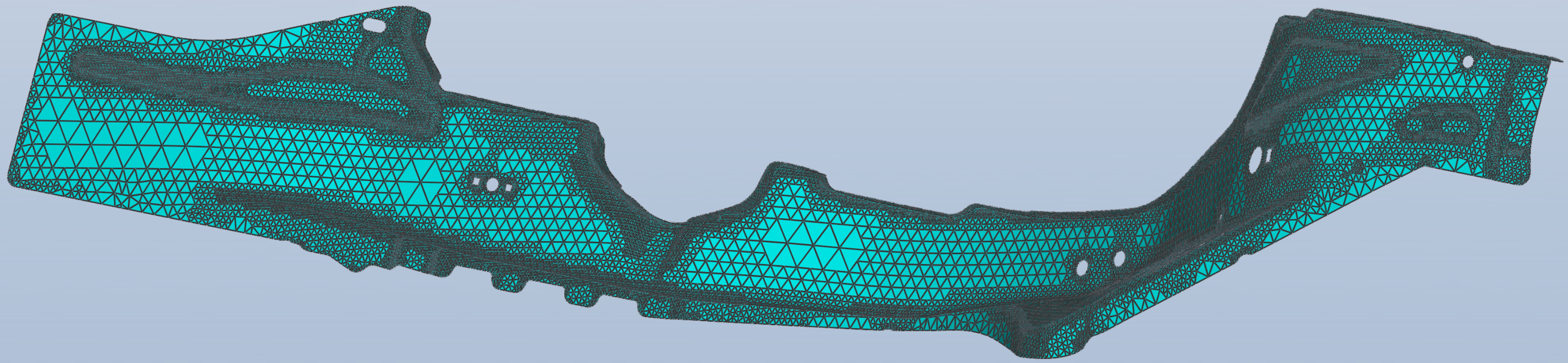


Variable 4

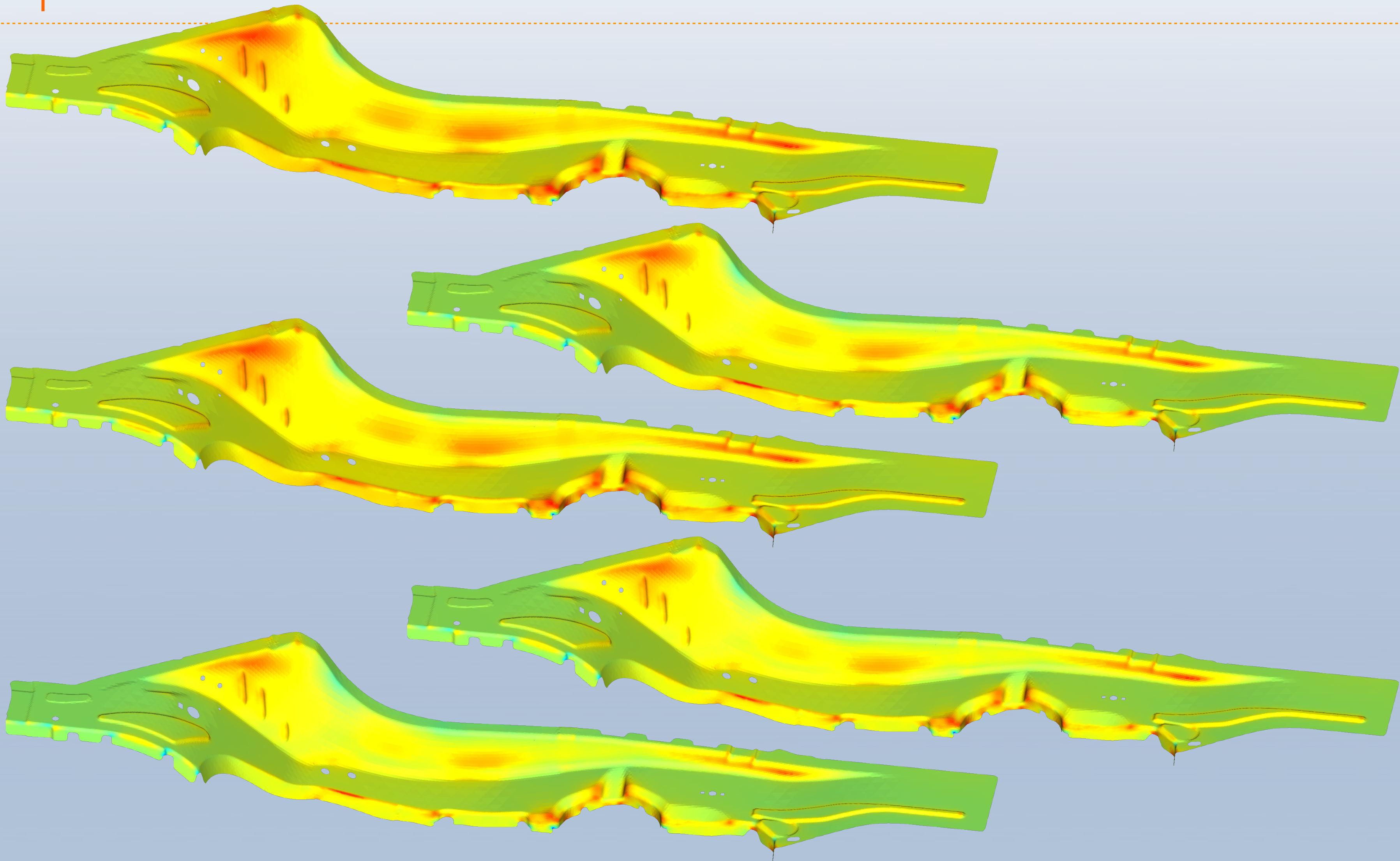


Example - larger structure

- 60.000 elements
- data show thickness variation

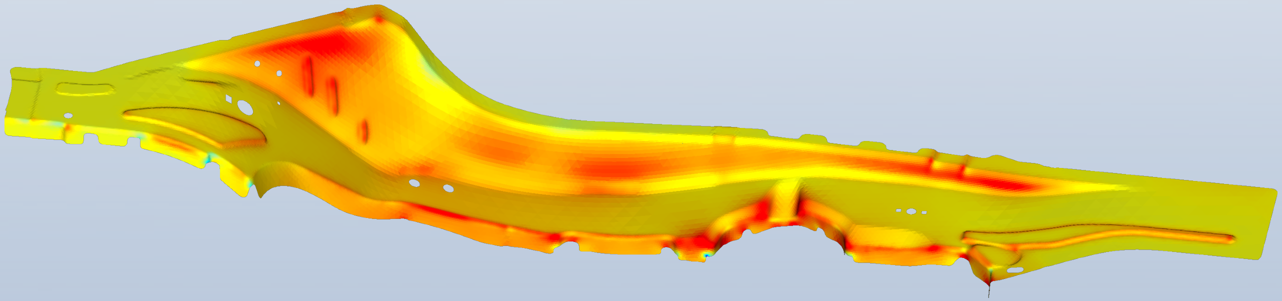


Sample functions

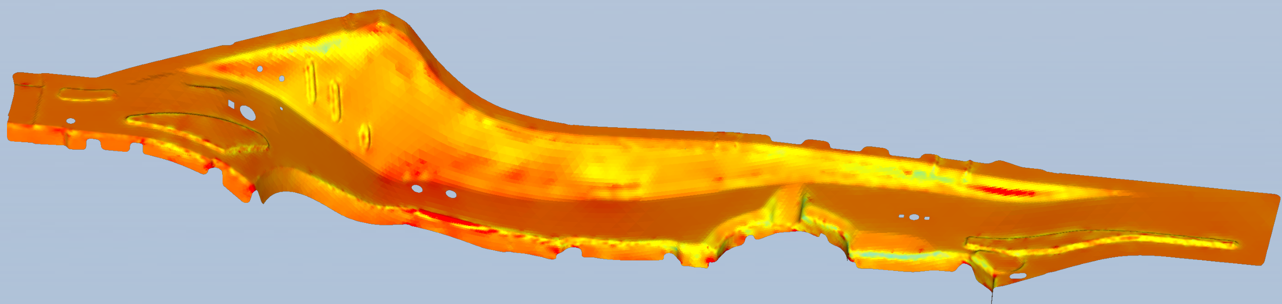


Statistics using original data

- Mean value

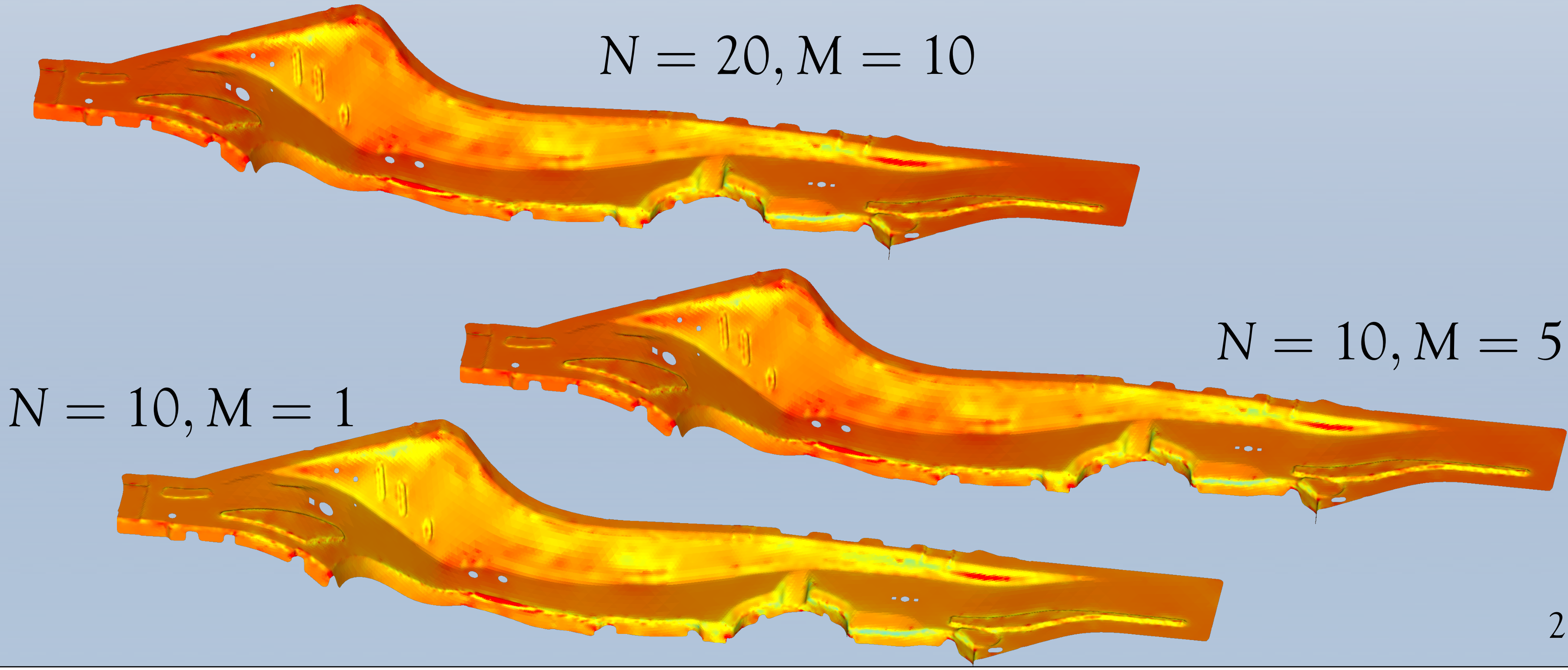


- Standard deviation



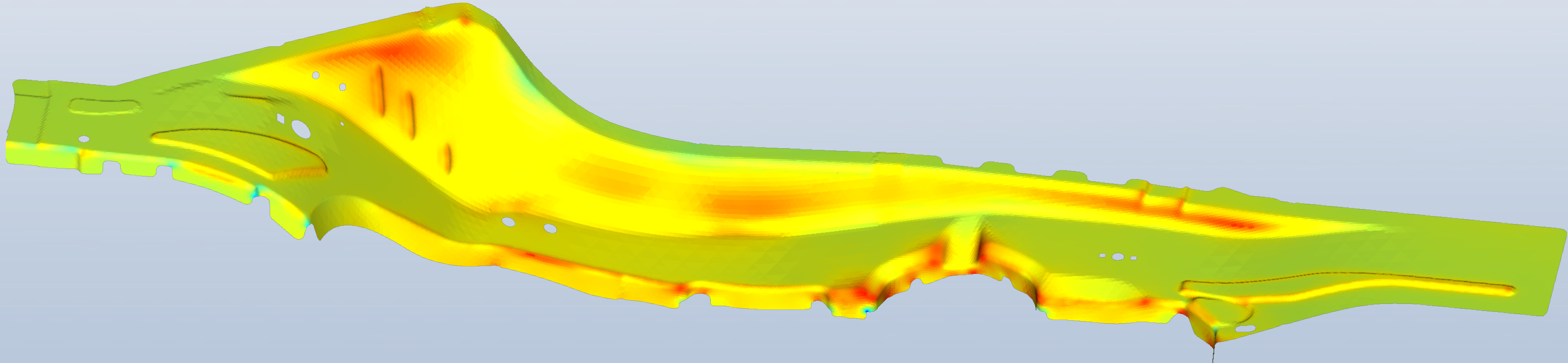
Statistics based on reduced data

- Standardized, reduced to smoothed space of dimension N, reduced to M principal components
- Mean value remains unchanged, standard deviation shows only small differences

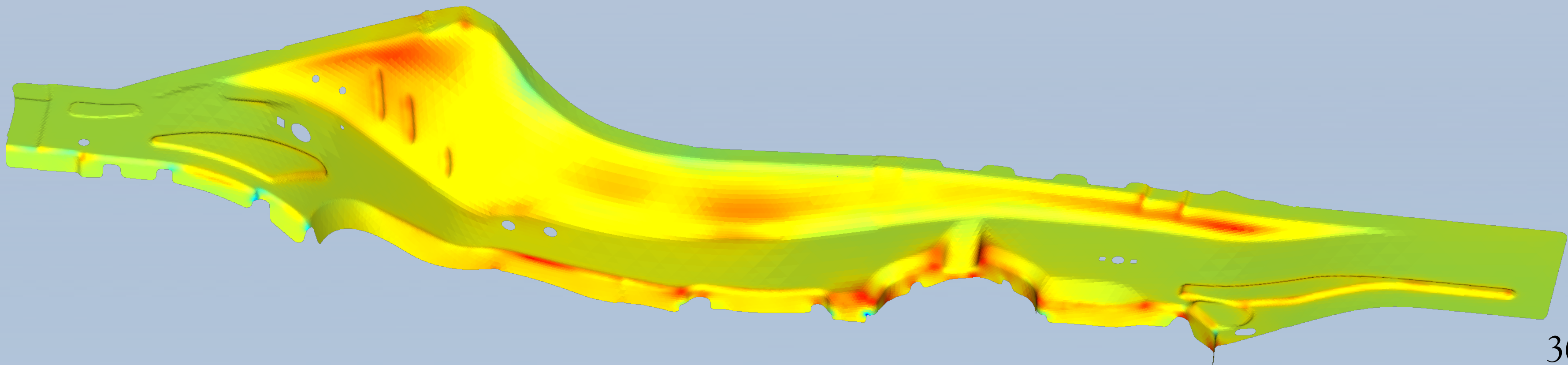


Original vs. simulation

- One real sample from FE analysis



- one virtual sample based on pure statistical analysis



Concluding remarks

- Data-driven reduction of random fields can provide high levels of accuracy
- Number of random variables can be significantly reduced
- Essential to use data-independent spatial smoothing and data-oriented principal component analysis
- New algorithms for spatial smoothing provide significant speed improvements
- Reduced representation improves statistical significance and allows for correlation analysis
- Reduced representation can be used to produce high-quality directly simulated random fields