Efficient Uncertainty Quantification and Global Time-Varying Sensitivity Analysis of Conceptual Hydrological Model

Ivana Jovanovic Buha[§], Florian Künzner[§], Tobias Neckel[§], Hans-Joachim Bungartz[§]
[§] Technical University of Munich, **ivana.jovanovic@tum.de**, florian.kuenzner@tum.de, neckel@in.tum.de, bungartz@in.tum.de

Motivation - Determine the (Un)influential parameters of a Hydrological Model

Application

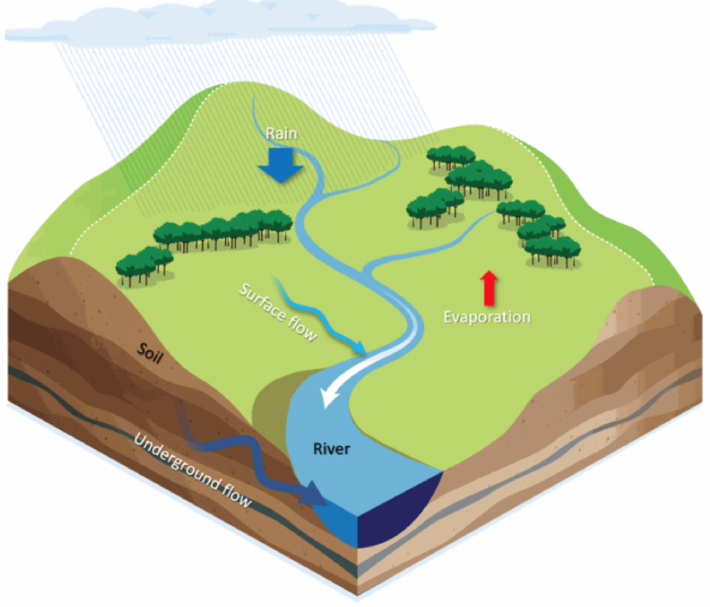
Conceptual Hydrological Model Larsim

Key pillars of our work

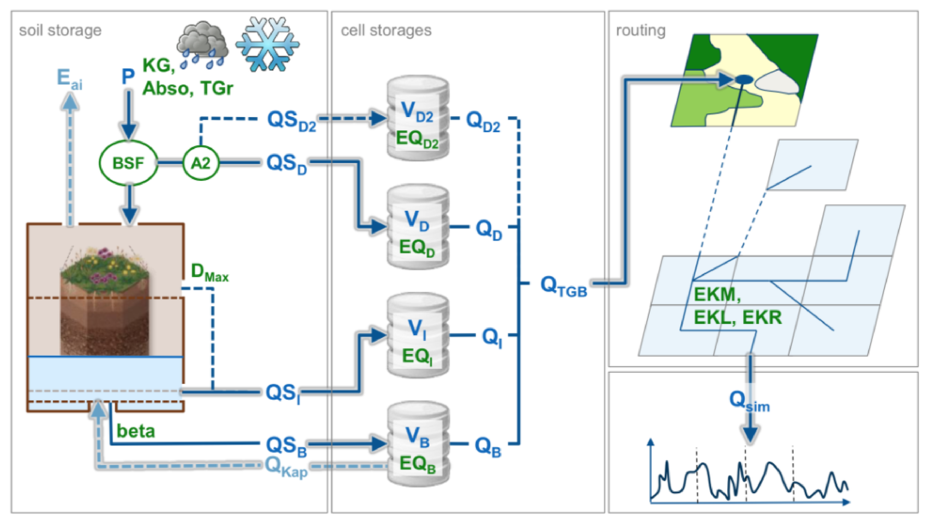
- Challenging (Real-world) application
- UQ Algorithms
- Unifying (Parallel) Software Solution

Bridging the gap between the theoretical work on UQ and more complex real-world problems

A sketch of hydrological processes occurring in the catchment



Processed in the water balance model LARSIM used in operational flood forecasting on Bayern, DE



Challenges

- Complex, computationally expensive model
- Quantity of interest is a time signal
- Many different 'modes' of the systemś behavior
- Physical dependency between parameters
- Other sources of uncertainty (e.g. forcing data)

Benefit

- Investigating dimensionality reduction strategies using Sobolíndices (e.g., speed up the following calibration process)
- Selecting an appropriate model structure for a given application

Uncertainty propagation

Stochastic input is composed of **independent** random variables

$$\rho(\boldsymbol{\theta}) = \prod_{i=1}^{d} \rho_i(\theta_i), \quad \boldsymbol{\theta} = (\theta_i)_{i=1}^{d}$$

We draw samples from $\rho(\boldsymbol{\theta})$ and solve the complex forward model $f(x,t,\boldsymbol{\theta})$, for each input sample.

Zero step - mapping of the range Γ to some standard *convenient* range (e.g., $[0,1]^d$ or $[-1,1]^d$) - transformation $\boldsymbol{\theta} = \boldsymbol{T}(u)$.

An **observation operator**, $\mathcal{O}$, is employed to determine the output of interest.

Ensemble of simulations are used to assess quantities:

$$E[\mathcal{O}] = \int_{\Gamma} \mathcal{O}(f(x,t,\boldsymbol{\theta}))\boldsymbol{\rho}(\boldsymbol{\theta})d\boldsymbol{\theta}; \quad Var[\mathcal{O}] = E[\mathcal{O}^2] - (E[\mathcal{O}])^2$$

Variance based (Sobol') sensitivity indices[1]

$$S_i^T = \frac{Var(f) - Var(E(f|\boldsymbol{\theta}_{-i}))}{Var(f)} = \frac{E(Var(f|\boldsymbol{\theta}_{-i}))}{Var(f)}$$

Monte Carlo methods for UQ and SA

- MC and **Quasi-MC** sampling approaches[2]
- **Sobol' indices** approximated based on the ensemble of simulations[1]

Polynomial Projection[3]

$$f(x,t,\boldsymbol{\theta}) \approx \boldsymbol{\mathcal{U}_p^N} = f_N(x,t,\boldsymbol{\theta}) = \sum_{p=0}^{N-1} m_p(x,t)\boldsymbol{\Phi}_p(\boldsymbol{\theta})$$

Where $\boldsymbol{\Phi}_n(\boldsymbol{\theta})$ are orthonormal multivariate polynomials
Coefficients:

$$m_p(x,t) = \mathbb{E}[f(x,t,\boldsymbol{\theta})\boldsymbol{\Phi}_p(\boldsymbol{\theta})] = \int_{\Gamma} f(x,t,\boldsymbol{\theta})\boldsymbol{\Phi}_p(\boldsymbol{\theta})\boldsymbol{\rho}(\boldsymbol{\theta})d\boldsymbol{\theta}$$

Pseudospectral approximation - $\{m_p\}_{p=0}^{N-1}$ evaluated via tensor product (Gaussian-Legendre) quadrature rule;[4] 1D quadrature rule

$$U_q^n = \sum_{q=0}^{n} f((\theta_i)_q)\phi_p((\theta_i)_q)\omega_q, \quad n = floor[\frac{DE}{2}]$$

Expectation:

$$\mathbb{E}[f_N(x,\boldsymbol{\theta})] = m_0(x,t)$$

Variance:

$$Var[f_N(x,t,\boldsymbol{\theta})] = \sum_{p=1}^{N-1} m_p^2(x,t)$$

Sobol' indices[5] for sensitivity analysis:

$$S_i^T = \frac{\sum_{p \in A_i} m_p^2(x,t)}{Var[f_N(x,t,\boldsymbol{\theta})]}, \qquad i = 1,\dots,d$$

Spatially Adaptive Sparse Grid[6]

- SG Interpolation of $f(x,t,\boldsymbol{\theta})$

$$f(x,t,\boldsymbol{\theta}) \approx \boldsymbol{\mathcal{U}_{SGI}} = f_{SGI}(x,t,\boldsymbol{\theta}) = \sum_{\boldsymbol{l} \in j, \boldsymbol{i} \in I} \alpha_{\boldsymbol{l},\boldsymbol{i}}(x,t)\boldsymbol{\varphi}_{l,i}(\boldsymbol{\theta})$$

Where $\alpha_{\boldsymbol{l},\boldsymbol{i}}(x,t)$ are hierarchical surpluses, and $\varphi_{l_j,i_j}(\boldsymbol{\theta}) = \varphi_{l,i}(\boldsymbol{\theta}) = \prod_j^d \varphi_{l_j,i_j}(\theta_j)$ are d-variate hierarchical basis functions constructed as a tensor product of 1D hierarchical basis functions

Compute the PC coefficients in two ways:

- SG interpolation followed by pseudo-spectral projection[7]

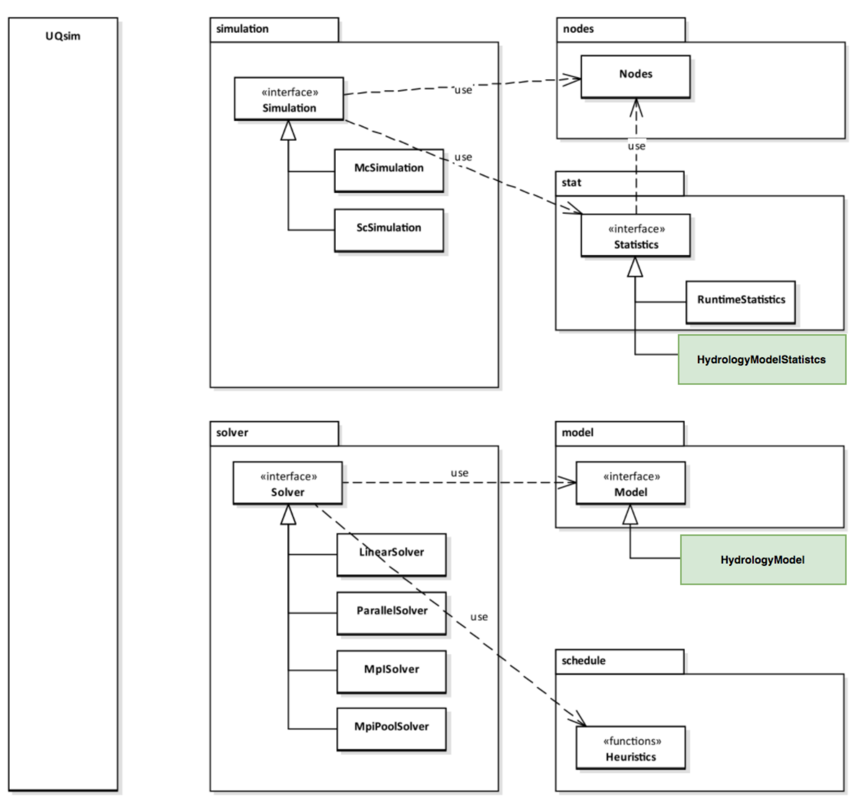
$$m_n(x) = \int_{\Gamma} f_{\mathcal{I}}(x,\boldsymbol{\theta})\boldsymbol{\Phi}_n(\boldsymbol{\theta})\boldsymbol{\rho}(\boldsymbol{\theta})d\boldsymbol{\theta}$$
$$= \sum_{\boldsymbol{l},\boldsymbol{i}} \alpha_{\boldsymbol{l},\boldsymbol{i}}(x) \prod_j \int_{[0,1]} \boldsymbol{\Phi}_j(T(\theta_j))\varphi_{l_j,i_j}(\theta_j)dT(\theta_j)$$

- Approximate all the weighted integrals of f via some standard interpolatory quadrature schemes (e.g., SG Gauss-Legendre)

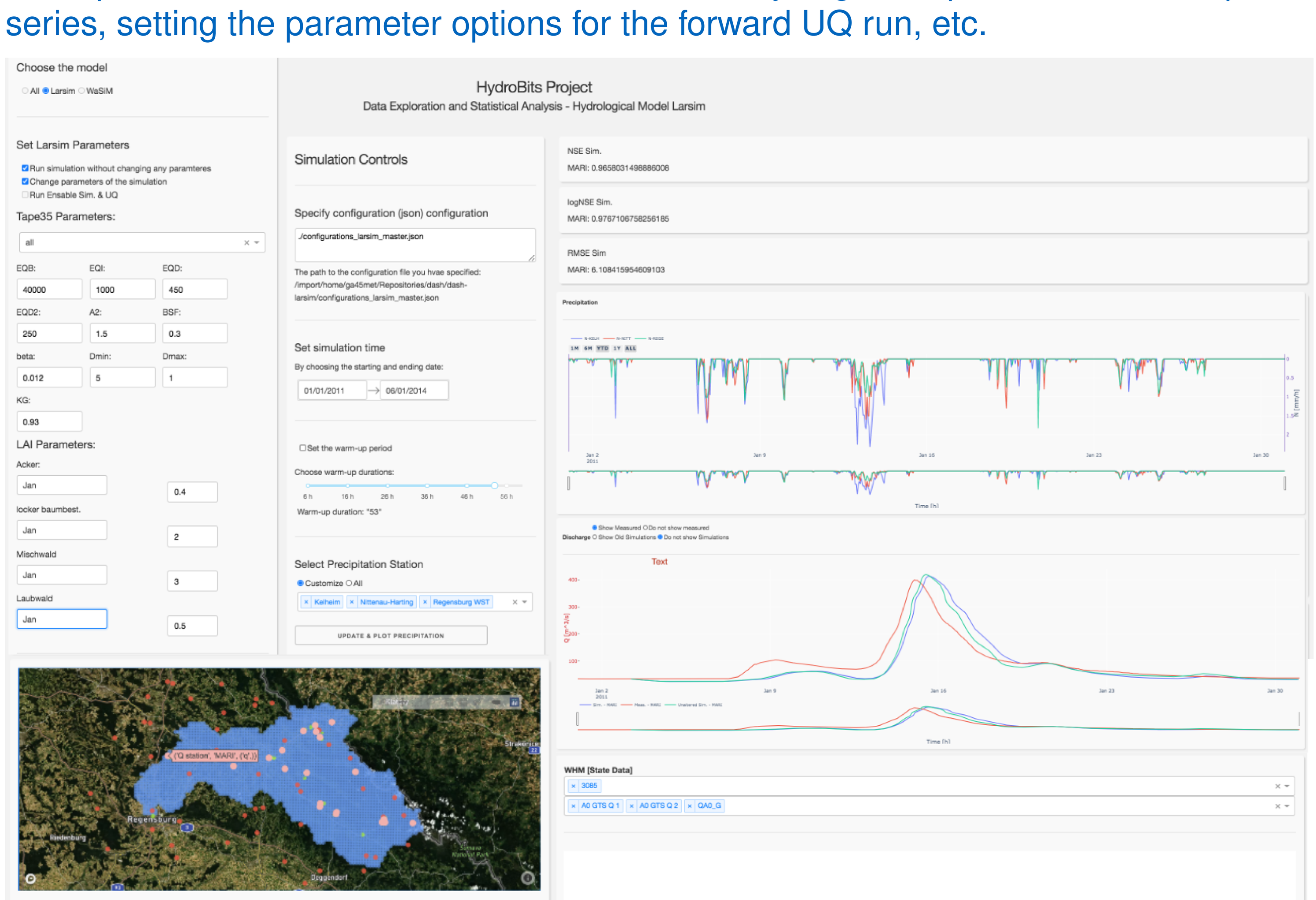
Combination Technique approach
linearly combine full lower-order anisotropic grids Dimension-wise

Spatial Refinement with SG refinement strategy with a maximum surplus refinement[8]

Software solutions & Results



- `Software Tools:`
  - Pandas Library - Time series data representation
  - ChaosPy
  - **UQEF**[9] - UQ parallel execution framework
  - **SparseSpACE**[10] - the sparse grid spatially adaptive combination environment
- `Hardware - Leibniz Supercomputing Centre (LRZ) CooLMUC-2 Cluster`
- Necessity for the parallelization
  - parallel execution of model runs (with all methods)
  - parallel time-wise post processing
- Dash panel - visualization and control tool for analyzing the input, state, and output time series, setting the parameter options for the forward UQ run, etc.

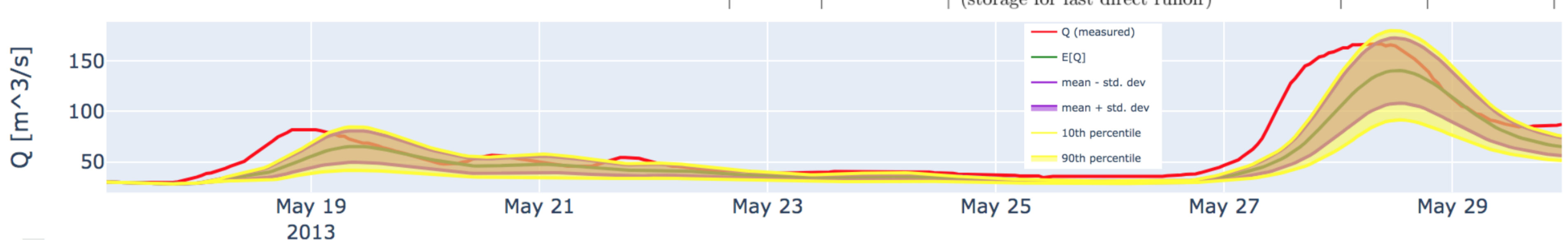


Case 1.

Impact of 5 Uncertain Parameters on LARSIM Model predictions under High Flow conditions

- **QoI:** Hourly discharge values
- **Method:** gPCE (Gauss-Hermite quadrature) $N = n = 7$
- **#model evaluations:** 32768
- **Computation time:** $\sim$ 24 hours on 4 compute nodes (28CPU+55GB RAM per node)

	Parameter	Description	Unit	Rang
	KG	correction factor for precipitation	[-]	0.9-1.1
High Flow	$A2$	threshold for division of fast and slow direct runoff	[mm]	0.5-4.0
	b (BSF)	form parameter of BSF (soil moisture saturation area function)	[-]	0.01-0.5
	EQD	calibration parameter for retention constant (storage for slow direct runoff)	[-]	50-5000
	$EQD2$	calibration parameter for retention constant (storage for fast direct runoff)	[-]	10-1000

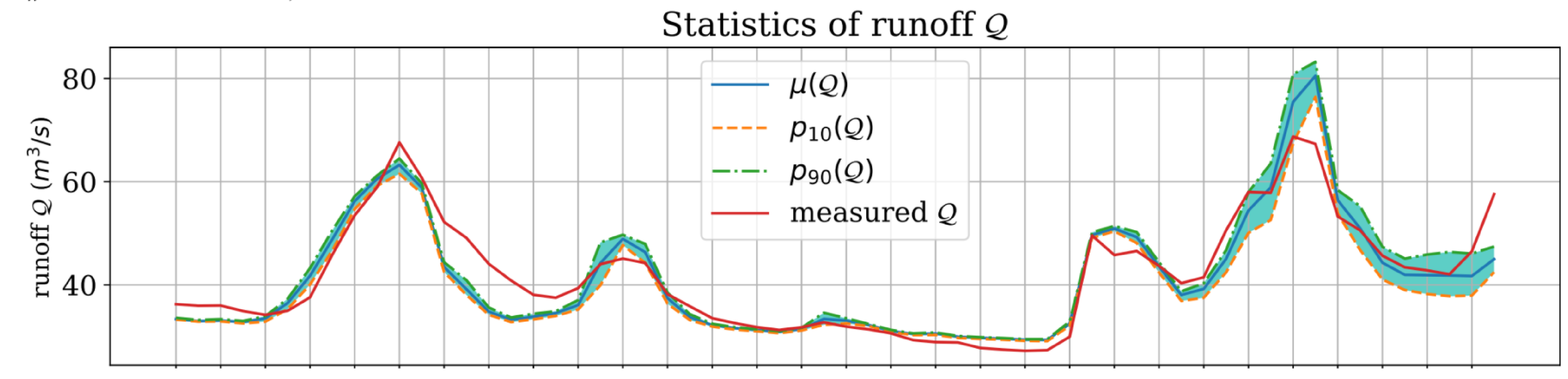


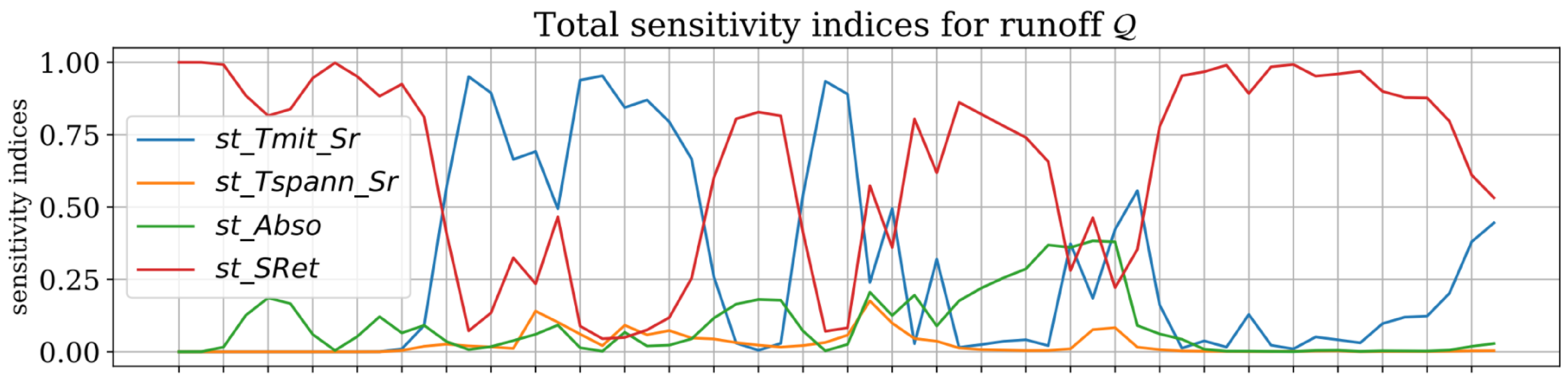
Case 2.

Impact of 4 Uncertain Parameters on LARSIM Model predictions under Snow Melting conditions

- **QoI:** Hourly discharge values over 2 months
- **Method:** gPCE (Gauss-Legendre quadrature) $n = 9; N = 7$
- **#model evaluations:** 10,000

Snow	T_{mit}	Average limit temperature of snow/rain mixture	[°C]	U[-2,4]
	T_{spann}	Temperature range around T_{mit} [°C]	[°C]	U[-0.08,0.1]
	ε (Abso)	Absorption coefficient for snow layer	[-]	U[-25,17]
	S_{Ret}	Maximum snow retention capacity	[%]	U[-2,3]





References

[1] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, "Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index," *Computer Physics Communications*, vol. 181, no. 2, pp. 259–270, 2010.

[2] R. E. Caflisch *et al.*, "Monte carlo and quasi-monte carlo methods," *Acta numerica*, vol. 1998, pp. 1–49, 1998.

[3] D. Xiu and G. E. Karniadakis, "The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations," *SIAM Journal of Scientific Computing*, vol. 24, pp. 619–644, 2002.

[4] P. G. Constantine, M. S. Eldred, and E. T. Phipps, "Sparse pseudospectral approximation method," *Computer Methods in Applied Mechanics and Engineering*, vol. 229, pp. 1–12, 2012.

[5] B. Sudret, "Global sensitivity analysis using polynomial chaos expansions," *Reliability Engineering and System Safety*, vol. 93, pp. 964–979, 2008.

[6] D. Pflüger, *Spatially Adaptive Sparse Grids for High-Dimensional Problems.* München: Verlag Dr. Hut, Aug. 2010.

[7] I.-G. Farcaș, B. Uekermann, T. Neckel, and H.-J. Bungartz, "Nonintrusive uncertainty analysis of fluid-structure interaction with spatially adaptive sparse grids and polynomial chaos expansion," *SIAM Journal on Scientific Computing*, vol. 40, no. 2, pp. B457–B482, 2018.

[8] H.-J. Bungartz and M. Obersteiner, "A generalized spatially adaptive sparse grid combination technique with dimension-wise refinement." 2020.

[9] F. Künzner, *Efficient non-intrusive uncertainty quantification for large-scale simulation scenarios.* Dissertation, Technische Universität München, München, 2021.

[10] M. Obersteiner, "Sparsespace - the sparse grid spatially adaptive combination environment." `https://github.com/obersteiner/sparseSpACE`, 2020.