Uncertainty quantification for the Hokkaido Nansei–Oki tsunami using B-splines on adaptive sparse grids

M. F. Rehme[1] S. G. Roberts[2] D. Pflüger[3]

(Received 31 December 2020; revised 14 June 2021)

Abstract

Modeling uncertainties in the input parameters of computer simulations is an established way to account for inevitably limited knowledge. To overcome long run-times and high demand for computational resources, a surrogate model can replace the original simulation. We use spatially adaptive sparse grids for the creation of this surrogate model. Sparse grids are a discretization scheme designed to mitigate the curse of dimensionality, and spatial adaptivity further decreases the necessary number of expensive simulations. We combine this with B-spline basis functions which provide gradients and are exactly integrable. We demonstrate the capability of this uncertainty quantification approach for a simulation of the Hokkaido Nansei–Oki Tsunami with ANUGA. We

DOI:10.21914/anziamj.v62i0.16121, © Austral. Mathematical Soc. 2021. Published 2021-06-29, as part of the Proceedings of the 19th Biennial Computational Techniques and Applications Conference. ISSN 1445-8810. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to the DOI for this article.

develop a better understanding of the tsunami behavior by calculating key quantities such as mean, percentiles and maximum run-up. We compare our approach to the popular Dakota toolbox and reach slightly better results for all quantities of interest.

Contents

1 Introduction

Computer experiments are commonly used to investigate real-world phenomena, and modeling uncertainties in the simulation's input parameters takes into account the inevitably limited knowledge of the researcher. Unfortunately, run-times and the necessary computational resources increase exponentially with the dimensionality of the model. This is known as the curse of dimensionality. Sparse grids [3] are an established discretization

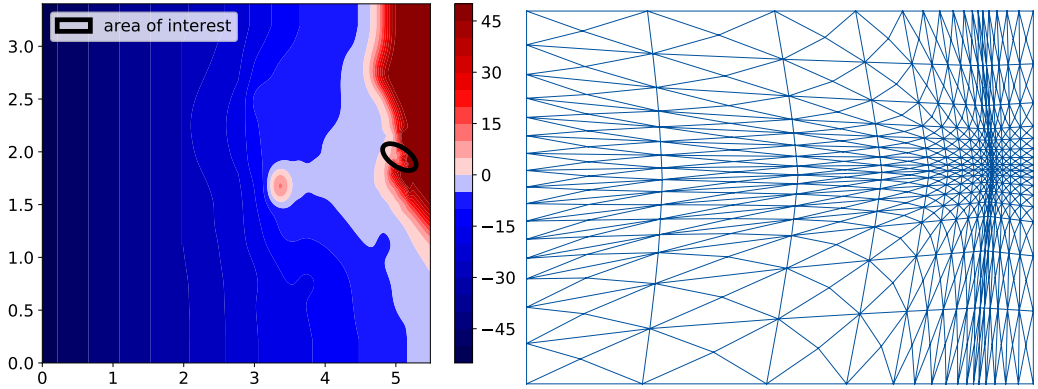

Figure 1: (left) bathymetry of the simulation domain, including part of Okushiri's coast and the Monai valley (ellipse); and (right) underlying triangulation with 16×16 triangles for illustration. In practice we use 64×64 triangles.

technique, designed to mitigate the curse of dimensionality. In this work we use spatially adaptive sparse grids [9] that automatically refine towards the quantity of interest. We apply higher-order basis functions, namely B-splines, which generalize classical hat functions [14]. With the B-splines we perform uncertainty quantification (UQ) and optimization. B-splines and spatially adaptive sparse grids have only recently been combined for UQ [10, 11]. This work demonstrates their competitiveness with popular and widely used UQ techniques for a relevant real-world application.

2 Tsunami model

In 1993 an earthquake in the Sea of Japan led to the Hokkaido Nansei–Oki tsunami whose run-up was disproportionately large in the Monai valley on the island of Okushiri. A $1/400$ scale model of the region was constructed in a $205\,\mathrm{m}$ long water tank to reproduce the run-ups [6]. The experimental data was made publicly available and the setting has become a benchmark for tsunami prediction [7, 15]. Figure 1 illustrates the bathymetry of the

benchmark.

2.1 Implementation of the tsunami benchmark

We model the benchmark with ANUGA [12], a fluid dynamics simulation framework, which uses the finite-volume method to solve the non-linear shallow water equations. The quantity of interest is the time-dependent average height of the water layer in the Monai valley. Our implementation is based on previous work [2], where it was shown that a multi-fidelity combination technique approach converges towards a high accuracy combination technique surrogate. In this work however, we investigate convergence towards the simulation itself using a fixed number of 64×64 triangles.

2.2 Wave shape parametrization

We investigate how different shapes of incoming tsunami waves influence the run-up. For this the original wave is decomposed into Gaussian bumps. The shapes of the individual Gaussian bumps are then varied, before the overall energy of the wave is normalized to the energy of the original wave.

Formally, let $W(t)$ be the original wave, depending on time $t \in [0, 22.5]$ in seconds. We set the initial residual $r_0(t) := W(t)$ and iteratively fit the dth Gaussian bump with parameters

$$\eta_d, \tau_d, \nu_d := \operatorname{argmin}_{\eta,\tau,\nu} \int_0^{22.5} \left| r_{d-1}(t) - \eta \exp\left(-\frac{(t-\tau)^2}{2\nu^2}\right)\right|^2 dt, \qquad (1)$$

where η_d, τ_d and ν_d are the height, width and center of dth Gaussian bump, respectively. The dth residual is defined as

$$r_d(t) := r_{d-1}(t) - \eta_d \exp\left(-\frac{(t-\tau_d)^2}{2\nu_d^2}\right). \qquad (2)$$

We iterate (1) and (2) for $d = 1, \ldots, D$, where $D \in \mathbb{N}_+$ is the number of

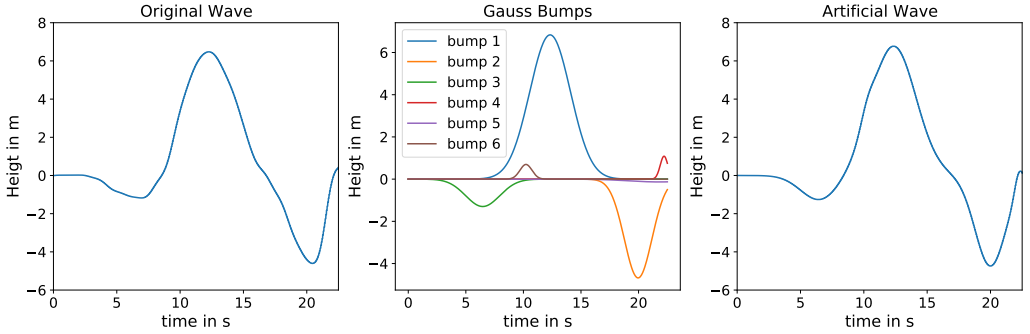

Figure 2: decomposition of the (left) original wave W into (centre) $D = 6$ Gauss bumps. To create different wave shapes, the height of the dth Gauss bump is varied, and then the energy of the wave is normalized. The L^2-difference between the original and artificial wave (right) is smaller than 0.00051 m.

bumps, and parameterize the incoming wave as

$$W_p(t,\vec{\theta}) := \sum_{d=1}^{D} \theta_d \eta_d \exp\left(-\frac{(t-\tau_d)^2}{2\nu_d^2}\right), \quad W_p^n(t,\vec{\theta}) := \frac{W_p(t,\vec{\theta})\|W(t)\|_2}{\|W_p(t,\vec{\theta})\|_2}, \tag{3}$$

where $\vec{\theta} := [\theta_1,\dots,\theta_D]$ is a vector of the uncertain parameters which determine the shape of the wave, $\|\cdot\|_2$ is the L^2-norm with respect to time t, and W_p^n is the normalized wave.

Inside the implementation the time domain is discretized with 451 uniform time steps t_m, $m = 1,\dots,451$. Our objective function is therefore

$$f : \mathbb{R}^D \to \mathbb{R}^{451}, \quad \vec{\theta} \mapsto (f_1(\theta),\dots,f_{451}(\theta))^T, \quad f_m(\vec{\theta}) = W_p^n(t_m,\vec{\theta}). \tag{4}$$

We use $D = 6$ and assume the wave shape parameters θ_d to be i.i.d., following a truncated normal distribution $\mathcal{N}(1.0, 0.125)|_{[0.5,1.5]}$ centered at the same point as the original wave. To better understand how the uncertainty in the shape parameters influences the run-up, the uncertainty must be propagated

through the model. However, a single evaluation of the model takes one CPU-hour[1], therefore we replaced the model by a surrogate.

3 Hierarchical not-a-knot B-splines

B-splines are widely used basis functions because of their local support and numerically exact differentiability and integrability. They have been applied in combination with sparse grids for interpolation, optimization and UQ [10, 11, 16]. Let $\xi := (\xi_0, \ldots, \xi_{q+n})$ be a knot sequence, that is a non-decreasing sequence of real numbers ξ_k, for $k \in \{0, \ldots, q+n\}$, $q \in \mathbb{N}_0$. The B-spline $b^n_{k,\xi}$ of degree $n \in \mathbb{N}_0$ and index $k \in \mathbb{Z}$ is defined as

$$b^n_{k,\xi}(x) := \begin{cases} \frac{x-\xi_k}{\xi_{k+n}-\xi_k} b^{n-1}_{k,\xi}(x) + \frac{\xi_{k+n+1}-x}{\xi_{k+n+1}-\xi_{k+1}} b^{n-1}_{k+1,\xi}(x) & n \geqslant 1, \\ \chi_{[\xi_k,\xi_{k+1}]}(x) & n = 0, \end{cases} \tag{5}$$

where $\chi_{[\xi_k,\xi_{k+1}]}(x)$ evaluates to one inside $[\xi_k, \xi_{k+1}]$ and zero elsewhere. The derivative of a B-spline is simply the difference of two B-splines of lower degree. Without loss of generality, we restrict ourselves to the unit interval $[0, 1]$ and define the uniform knot sequence of level $\ell \in \mathbb{N}_0$ and degree $n \in \mathbb{N}_0$ as $\xi^{n,u}_\ell := (\xi^{n,u}_{\ell,0}, \ldots, \xi^{n,u}_{\ell,2^\ell+2n})$, where $\xi^{n,u}_{\ell,k} := (k-n)h_\ell$ for grid width $h_\ell := 2^{-\ell}$. Simply using this finite uniform knot sequence does not suffice, because the resulting B-splines do not fulfill the Schoenberg–Whitney conditions [13] at the interval boundaries. However, this condition is necessary for the B-splines to form a basis of the spline space corresponding to their knot sequence.

3.1 Not-a-knot condition

One approach to revalidate the Schoenberg–Whitney conditions are not-a-knot B-splines [5]. The not-a-knot condition is motivated by requiring continuity

[1] Raijin supercomputer at NCI Australia, a Fujitsu Primergy-Lenovo NeXtScale system. Evaluations were distributed, whereby each individual evaluation was performed using one node with 2×8 core Intel Xeon E5-2670 (Sandy Bridge) 2.6 GHz

of the nth derivatives of all splines at the $\frac{n-1}{2}$ left-most and $\frac{n-1}{2}$ right-most knots. This requirement is equivalent to removing the $(n-1)$ knots from the knot sequence, but keeping them in the set of interpolation nodes. More details can be found in our previous work [11]. To significantly mitigate the curse of dimensionality the boundary points are omitted too. We apply the not-a-knot condition again, removing another two knots from the knot sequence. With this trade-off we can remove the two boundary points and their corresponding basis functions without losing approximation quality.

For the definition of sparse grids we require hierarchical basis functions. Let $I_\ell := \{0 < k < 2^\ell \mid k \text{ odd}\}$ be the hierarchical index set of level ℓ, then $\{b^n_{\ell,k} \mid k \in I_\ell\}$ are the univariate hierarchical B-splines of level ℓ. The multivariate equivalents are defined via tensor products.

4 Sparse grids

For full uniform isotropic tensor product grids the number of grid points increases exponentially with the number of dimensions D and has order $\mathcal{O}(h^{-D})$ for grid width $h < 1$. Sparse grids [3] are a discretization scheme designed to mitigate this curse of dimensionality. The number of inner grid points of a regular sparse grid of level ℓ with grid width h_ℓ only increases like $\mathcal{O}(h_\ell^{-1}(\log_2 h_\ell^{-1})^{D-1})$. At the same time, the L^2-interpolation error for B-splines of degree n still decays asymptotically like $\mathcal{O}(h_\ell^{n+1}(\log_2 h_\ell^{-1})^{D-1})$, if the objective function is sufficiently smooth [14], which is only slightly worse than the full grid error convergence rate of $\mathcal{O}(h_\ell^{n+1})$.

Regular sparse grids We define the hierarchical subgrid of level $\vec{l}$ with grid widths $h_{l_d} := 2^{-l_d}$ as $\mathcal{H}_{\vec{l}} := \{\vec{x}_{\vec{l},\vec{i}} = (x_{l_1,i_1}, \ldots, x_{l_D,i_D}) \mid \vec{i} \in I_{\vec{l}}\}$ for $x_{l_d,i_d} := i_d h_{l_d}$. The corresponding hierarchical subspace $W_{\vec{\ell}}$ of level $\vec{\ell}$ is then defined as the span of the according basis functions $W_{\vec{l}} := \operatorname{span}\{b^n_{\vec{i},\xi^{n,u}_{\vec{l}}} \mid \vec{i} \in I_{\vec{l}}\}$.

The boundary sparse grid V^b_ℓ and the non-boundary sparse grid V^s_ℓ of level ℓ

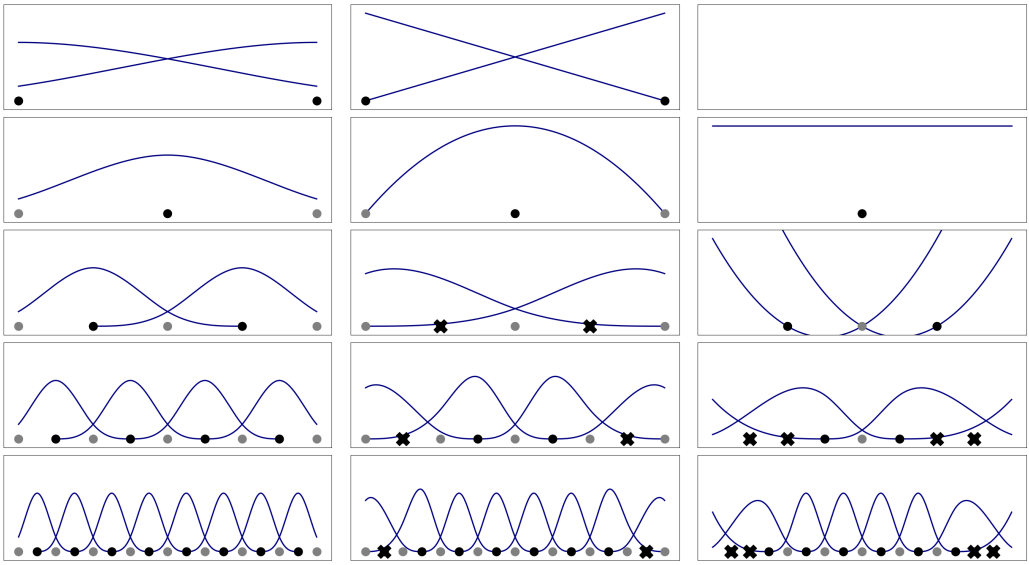

Figure 3: (left) hierarchical B-spline basis; (centre) hierarchical not-a-knot B-spline basis; and (right) hierarchical non-boundary not-a-knot B-splines of degree $n = 3$ and levels $\ell \in \{0, 1, 2, 3, 4\}$, respectively. The not-a-knot condition is indicated with crosses.

are defined as

$$V_\ell^b := \bigoplus_{|\vec{\ell}'|_1 \leqslant \ell} W_{\vec{\ell}'}, \quad V_\ell^s := \bigoplus_{|\vec{\ell}'|_1 \leqslant \ell, \ell'_d \geqslant 1, \forall d=1,\dots,D} W_{\vec{\ell}'}, \tag{6}$$

where $|\vec{\ell}'|_1 := \sum_{d=1}^{D} \ell'_d$ is the discrete L^1 norm of $\vec{\ell}'$.

Spatial adaptivity Spatially adaptive sparse grids can be refined automatically according to a specified quantity of interest, potentially significantly reducing the number of grid points required for desired approximation quality [9]. We use the standard surplus criterion, which makes use of the hierarchy of the basis, where larger interpolation coefficients indicate a worse local ap-

proximation. Thus we iteratively refine the grid point corresponding to the largest interpolation coefficient [11].

5 Numerical results

We created our sparse grid surrogates with the toolbox SG^{++} [9], a versatile open-source framework for spatially adaptive sparse grids.

5.1 Polynomial chaos expansion

To demonstrate the performance of our approach we compare it to one of UQ's most popular tools: polynomial chaos expansion (PCE) [4, 17]. Generalized polynomial chaos is based on the Wiener–Askey scheme [17]. This scheme specifies basis functions, which are orthogonal with respect to the distributions of the uncertain parameters. Our application has normally distributed parameters which correspond to Hermite polynomials. PCE generally results in good approximations and allows direct access to stochastic moments without further calculations. The downside of PCE is the strict specification of the basis functions, only depending on the distributions and not on the actual objective function. For comparison and verification of our results we used the state-of-the-art toolbox Dakota [1] and its implementation of PCE.

5.2 Accuracy of the surrogate

We measure the approximation errors based on 10 000 parameters randomly sampled from the joint parameter distribution. The results are seen in Figure 4. Because they have too many points on the boundary the not-a-knot boundary splines perform significantly worse than the alternatives. Omitting the boundary points and using the non-boundary not-a-knot B-splines instead yields the best approximation quality overall. We use cubic B-splines here because we require the derivatives of the surrogate. However, the underlying simulation is only first order accurate because of shocks, and thus none of the

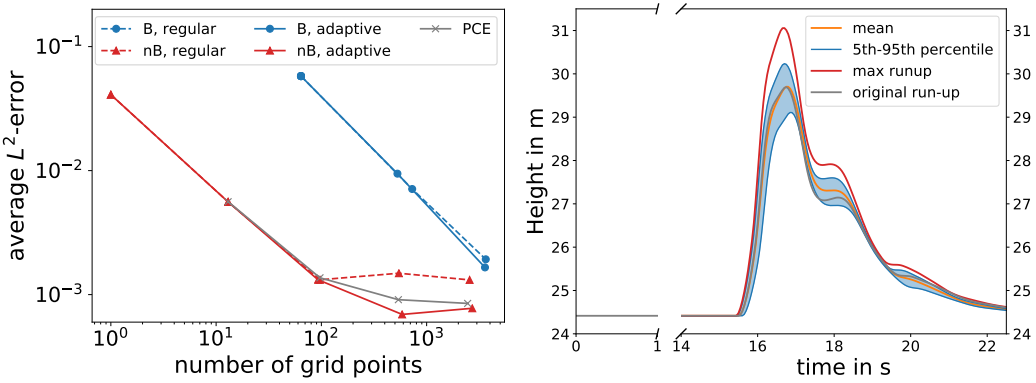

Figure 4: (left) average L^2-error in meters for regular and spatially adaptive sparse grids with boundary (B), non-boundary sparse grids (nB) and PCE; and (right) forward uncertainty quantification of the tsunami model, resulting in percentiles, the expected value and the overall maximum calculated with SG^{++}.

methods reach their optimal convergence rate. Also, all methods stagnate around 10^{-3} m. This stagnation is caused by tiny inaccuracies in the inputs. Using a six-dimensional decomposition of the original wave into Gauss bumps, the incoming wave itself is represented only up to one millimeter accuracy. This also explains the only slight improvement from spatial adaptivity. When the domain is sufficiently scanned to make use of the adaptivity, maximum accuracy is already reached.

5.3 Uncertainty quantification

By propagating the uncertainty of the input parameters through the surrogate, we better understand how they influence the quantities of interest. The results of this section are summarized in Figure 4.

Percentiles We created a set of 100 000 samples drawn from the joint parameter distribution and evaluated the surrogate in each of these parameters.

Then we derived the 5th and 95th percentile, which are the values below which 5% or 95% of the evaluations are found, respectively.

Mean By using a quadrature formula of appropriate order, B-splines can be integrated exactly. We use integration to calculate the surrogate's mean with stochastic collocation. For PCE the mean calculation is even simpler. Because the underlying basis functions are orthonormal, the mean of the surrogate is simply the expansion's first coefficient. We validate the calculated means using a reference solution based on a non-boundary not-a-knot B-spline surrogate of degree $n = 3$ on a sparse grid of level $\ell = 6$ with 10 625 grid points. Our adaptive surrogate based on 2 714 grid points has an L^2 difference of 0.0026 m to the reference mean and its worst deviation for any time step is 0.0165 m. The Dakota PCE approximation, which uses 2 465 function evaluations, has an L^2 difference of 0.0030 m and its worst deviation is 0.0191 m.

Shape optimization To find the wave shape leading to the largest run-up, we optimized the objective function over time, searching for $\theta^{\max} := \arg\max_{\vec{\theta}}[\max_m \tilde{f}_m(\vec{\theta})]$. We calculate the gradient of our surrogate using the recursive definition of B-splines. This allows us to optimize the surrogate using the gradient descent algorithm, where we determine the step sizes using the Armijo line search algorithm [8]. We optimized the surrogates from Section 5.2 and found that the maximum value is on the boundary of the domain. Thus, although its approximation is less accurate in general, the boundary sparse grid gives the best result,

$$\theta^{\max} = [1.5, 0.5, 1.5, 0.671875, 0.5, 0.5], \quad \max_m \tilde{f}_m(\theta^{\max}) = 31.057\,\text{m}. \tag{7}$$

The run-up value is larger than the values of any of the 100 000 evaluations we used to calculate the percentiles. The Dakota toolbox does not offer optimization based on its PCE surrogate. Instead it uses the mesh adaptive direct search algorithm (MADS) based on quadratic polynomials. This needs another 2 500 evaluations of the objective function, in contrast to SG^{++}, which performed all operations on the same initially constructed surrogate. Dakota's

optimization then results in 31.02 m run-up height, a little smaller than our result.

6 Conclusion and outlook

We have performed a UQ investigation of a computationally expensive tsunami benchmark using B-splines on adaptive sparse grids. Using a moderate number of evaluations of the simulation, we created an accurate surrogate. With this surrogate we investigated how different shapes of the incoming tsunami wave influence the run-up. B-splines can be evaluated very quickly, are exactly integrable and give direct access to gradients. Exploiting these properties, we calculated percentiles, the mean and the maximum of the surrogate. For comparison and verification we used the toolbox Dakota. For all quantities of interest our software SG^{++} performed slightly better than Dakota. In future work we plan to not only use incoming waves of different shapes, but to derive the tsunami directly from a simulated earthquake.

Acknowledgments The authors thank the German Research Foundation (DFG) for financial support of the project within the *Cluster of Excellence in Data-Integrated Simulation Sciences* (EXC2075).

References

[1] B. M. Adams, M. S. Ebeida, et al. *Dakota*. Sandia Technical Report, SAND2014-4633, Version 6.11 User's Manual, July 2014. 2019. URL: https://dakota.sandia.gov/content/manuals (cit. on p. C38).

[2] J. H. S. de Baar and S. G. Roberts. "Multifidelity sparse-grid-based uncertainty quantification for the Hokkaido Nansei–Oki tsunami". In: *Pure Appl. Geophys.* 174 (2017), pp. 3107–3121. DOI: 10.1007/s00024-017-1606-y (cit. on p. C33).

[3] H.-J. Bungartz and M. Griebel. "Sparse grids". In: *Acta Numer.* 13 (2004), pp. 147–269. DOI: 10.1017/S0962492904000182 (cit. on pp. C31, C36).

[4] M. Eldred and J. Burkardt. "Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification". In: *47th AIAA*. 2009. DOI: 10.2514/6.2009-976 (cit. on p. C38).

[5] K. Höllig and J. Hörner. *Approximation and modeling with B-splines*. Philadelphia: SIAM, 2013. DOI: 10.1137/1.9781611972955 (cit. on p. C35).

[6] M. Matsuyama and H. Tanaka. "An experimental study of the highest run-up height in the 1993 Hokkaido Nansei–Oki earthquake tsunami". In: *National Tsunami Hazard Mitigation Program Review and International Tsunami Symposium (ITS)*. 2001 (cit. on p. C32).

[7] O. Nielsen, S. Roberts, D. Gray, A. McPherson, and A. Hitchman. "Hydrodymamic modelling of coastal inundation". In: *MODSIM 2005*. 2005, pp. 518–523. URL: https://www.mssanz.org.au/modsim05/papers/nielsen.pdf (cit. on p. C32).

[8] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, 2006. DOI: 10.1007/978-0-387-40065-5 (cit. on p. C40).

[9] D. Pflüger. *Spatially Adaptive Sparse Grids for High-Dimensional Problems*. Dr. rer. nat., Technische Universität München, Aug. 2010. URL: https://www5.in.tum.de/pub/pflueger10spatially.pdf (cit. on pp. C32, C37, C38).

[10] M. F. Rehme, F. Franzelin, and D. Pflüger. "B-splines on sparse grids for surrogates in uncertainty quantification". In: *Reliab. Eng. Sys. Saf.* 209 (2021), p. 107430. DOI: 10.1016/j.ress.2021.107430 (cit. on pp. C32, C35).

[11] M. F. Rehme and D. Pflüger. "Stochastic collocation with hierarchical extended B-splines on Sparse Grids". In: *Approximation Theory XVI, AT 2019. Springer Proc. Math. Stats.* Vol. 336. Springer, 2020. DOI: 10.1007/978-3-030-57464-2_12 (cit. on pp. C32, C35, C36, C38).

[12] S Roberts, O. Nielsen, D. Gray, J. Sexton, and G. Davies. *ANUGA*. Geoscience Australia. 2015. DOI: 10.13140/RG.2.2.12401.99686 (cit. on p. C33).

[13] I. J. Schoenberg and A. Whitney. "On Pólya frequence functions. III. The positivity of translation determinants with an application to the interpolation problem by spline curves". In: *Trans. Am. Math. Soc.* 74.2 (1953), pp. 246–259. DOI: 10.2307/1990881 (cit. on p. C35).

[14] W. Sickel and T. Ullrich. "Spline interpolation on sparse grids". In: *Appl. Anal.* 90.3–4 (2011), pp. 337–383. DOI: 10.1080/00036811.2010.495336 (cit. on pp. C32, C36).

[15] C. E. Synolakis, E. N. Bernard, V. V. Titov, U. Kânoğlu, and F. I. González. *Standards, criteria, and procedures for NOAA evaluation of tsunami numerical models*. NOAA/Pacific Marine Environmental Laboratory. 2007. URL: https://nctr.pmel.noaa.gov/benchmark/ (cit. on p. C32).

[16] J. Valentin and D. Pflüger. "Hierarchical gradient-based optimization with B-splines on sparse grids". In: *Sparse Grids and Applications—Stuttgart 2014. Lecture Notes in Computational Science and Engineering*. Vol. 109. Springer, 2016, pp. 315–336. DOI: 10.1007/978-3-319-28262-6_13 (cit. on p. C35).

[17] D. Xiu and G. E. Karniadakis. "The Wiener–Askey polynomial chaos for stochastic differential equations". In: *SIAM J. Sci. Comput.* 24.2 (2002), pp. 619–644. DOI: 10.1137/S1064827501387826 (cit. on p. C38).

Author addresses

1. **M. F. Rehme**, IPVS, University of Stuttgart, GERMANY.
   mailto:michael.rehme@ipvs.uni-stuttgart.de
   orcid:0000-0001-7259-3501
2. **S. G. Roberts**, MSI, Australian National University, AUSTRALIA.
3. **D. Pflüger**, IPVS, University of Stuttgart, GERMANY.