Computers and Fluids 268 (2024) 106108

Contents lists available at ScienceDirect

Computers and Fluids

journal homepage: www.elsevier.com/locate/compfluid

Quantifying uncertainties in direct numerical simulations of a turbulent channel flow

Joseph O'Connor [a,*], Sylvain Laizet [a], Andrew Wynn [a], Wouter Edeling [b], Peter V. Coveney [c]

[a] Department of Aeronautics, Imperial College London, London, UK
[b] Department of Scientific Computing, Centrum Wiskunde and Informatica, Amsterdam, Netherlands
[c] Department of Chemistry, University College London, London, UK

ARTICLE INFO

Dataset link: https://doi.org/10.5281/zenodo.8038943

Keywords:
Direct numerical simulation
Sensitivity analysis
Turbulent channel flow
Uncertainty quantification
Xcompact3d

ABSTRACT

Direct numerical simulation (DNS) provides unrivalled levels of detail and accuracy for simulating turbulent flows. However, like all numerical methods, DNS is subject to uncertainties arising from the numerical scheme and input parameters (e.g. mesh resolution). While uncertainty quantification (UQ) techniques are being employed more and more to provide a systematic analysis of uncertainty for lower-fidelity models, their application to DNS is still relatively rare. In light of this, the aim of this work is to apply UQ and sensitivity analysis to the DNS of a canonical wall-bounded turbulent channel flow at low Reynolds number ($Re_\tau = 180$). To compute the DNS, `Incompact3d` – a highly scalable open-source framework based on high-order compact finite differences and a spectral Poisson solver – is used as a black-box solver. Stochastic collocation is used to propagate the input uncertainties through `Incompact3d` to the output quantities of interest (QOIs). To facilitate the non-intrusive forward UQ analysis, the open-source EasyVVUQ package is used to provide integrated capability for sampling, pre-processing, execution, post-processing, and analysis of the computational campaign. Three separate UQ campaigns are conducted. The first two examine the effect of domain size and the numerical parameters (e.g. mesh resolution, time step, sample time), respectively, and adopt Gaussian quadrature rules combined via tensor products to sample the multi-dimensional input space. Finally, the third campaign investigates the performance of a dimension-adaptive sampling strategy that significantly reduces the computational cost compared to the full tensor product approach. The analysis focuses on the cross-channel statistical moments of the QOIs, as well as local and global sensitivity analyses to assess the sensitivity of each QOI with respect to each individual input. This enables an assessment of the robustness and sensitivity of DNS to the user-defined numerical parameters for wall-bounded turbulent flows, and provides an indication of suitable ranges for defining the values of these parameters.

1. Introduction

Turbulent flows govern the behaviour of many natural and engineering systems, from arterial blood flow to aircraft aerodynamics and industrial process mixing, to name a few. Such flows are characterised by complex and chaotic dynamics that span disparate spatial and temporal scales [1]. These characteristics make it difficult to model turbulent flows using traditional analytical approaches. While the Navier–Stokes equations constitute a broadly accepted mathematical model to describe the dynamics of a turbulent flow, they can be extremely challenging to solve, due to the chaotic and inherently multi-scale nature of turbulence. The smallest scales influence the largest scales, which are typically separated by several orders of magnitude. Furthermore, small changes to initial conditions, boundary conditions, or mesh resolution, for example, can have a dramatic impact on the final solution. Direct numerical simulation (DNS) aims to resolve all of these scales and, as a result, is the gold standard for providing in-depth analysis of flow physics, developing new theory, and validating lower-fidelity models [2]. However, even with today's state-of-the-art algorithms and petascale high-performance computing (HPC) systems, DNS is only feasible for a small class of problems, namely, those at low-moderate Reynolds numbers with simple geometries. While this is a major limitation of DNS, significant effort has been dedicated to this area, and advances in algorithms and hardware are continuously pushing the frontier of what is achievable (e.g. see Lee and Moser [3]).

Another major challenge with DNS is assessing the accuracy and reliability of a simulation [4]. Numerical simulations are inherently uncertain due to various sources of error and variability, such as modelling assumptions, initial and boundary conditions, and choice

* Corresponding author.
E-mail address: joseph.oconnor@imperial.ac.uk (J. O'Connor).

https://doi.org/10.1016/j.compfluid.2023.106108
Received 14 June 2023; Received in revised form 7 October 2023; Accepted 8 November 2023
Available online 10 November 2023

of discretisation scheme, to name a few [5]. This uncertainty makes it difficult to interpret the results of a simulation, particularly for real-world decision-making. To mitigate against this issue, verification, validation, and uncertainty quantification (VVUQ) techniques provide a way to assess a model's reliability and ensure its outputs are actionable in the real world [6,7]. Verification and validation are the foundation of most computational fluid dynamics (CFD) studies and are often used to demonstrate the credibility of the results. Uncertainty quantification (UQ), on the other hand, is an area that is typically overlooked in most CFD studies. Beyond performing mesh-independence studies, it is rare to see rigorous sensitivity analyses of input parameters or UQ studies demonstrating forward propagation of uncertainties to build a statistical picture of a model's performance. This is most likely due to the computational cost and added complexity associated with performing such studies. However, these days, with advances in hardware/algorithms and the availability of open-source libraries that can handle such workloads automatically (e.g. see Richardson et al. [8] and Rezaeiravesh et al. [9]), this is an area that should be considered more often to provide a measure of confidence in a model. This is especially true if a model is to be translated from academic to real-world applications, where engineers require a measure of the uncertainty to make informed decisions (particularly in safety-critical applications).

Lucor et al. [10] were one of the first to apply systematic UQ and sensitivity analysis to scale-resolving simulations of turbulent flows. In particular, they applied non-intrusive generalised polynomial chaos (gPC) to study the effect of the Smagorinsky constant in large eddy simulations (LES) of homogeneous isotropic turbulence. Their results showed that sensitivity to the Smagorinsky constant mainly manifests in the small-scale behaviour. Similar techniques have also been applied to investigate the behaviour of the mixing layer instability. For example, Ko et al. [11] combined gPC with DNS to investigate the sensitivity of a 2D mixing layer to uncertainties in the inlet boundary conditions. On the other hand, Meldi et al. [12] examined the effect of mesh resolution and Smagorinsky constant for the LES of a 3D mixing layer, showing that the Smagorinsky constant is the dominant parameter responsible for the majority of the observed variance. Similar approaches have also been adopted to examine the sensitivity of LES in turbulent combustion [13], and to help explain the dispersion in previously reported LES results for the flow around a rectangular cylinder [14]. In addition to this, Congedo et al. [15] compared both Reynolds-averaged Navier–Stokes (RANS) and LES for turbulent pipe flow with an axisymmetric expansion, while also taking into account the uncertainty in the experimental measurements used to define the inlet conditions. The results showed that, under certain flow configurations, the RANS approach does not accurately capture the mean flow features. Furthermore, the variance associated with the uncertain inputs can be high. On the other hand, LES showed good agreement and small variance for all considered flow conditions.

The canonical channel flow test case, constructed by driving the flow between two infinitely long parallel plates, is a popular benchmark for validating CFD codes. As a result, it has also become a popular case for UQ studies of scale-resolving simulations. Meyers and Sagaut [16] applied under-resolved DNS to investigate the grid convergence behaviour in a channel flow. They showed that there is a line of combinations of streamwise and spanwise mesh resolution where the error in the skin friction is zero, due to error balancing in the discretisation. Furthermore, they also showed that these errors mainly manifest in the second-order velocity moments. More recently, Oliver et al. [17] developed a Bayesian extension of Richardson extrapolation to estimate discretisation error, while also accounting for the sampling error (due to finite time-averaging). Safta et al. [18] estimated the probability distribution of LES subgrid-scale model parameters by applying Bayesian inference to DNS data of forced isotropic turbulence. Following this, forward UQ was applied to LES of a channel flow to propagate the uncertainty through to the flow solution. Two recent studies have examined the uncertainties arising from LES of a channel flow within OpenFOAM, including the sensitivity to grid resolution, wall model, and subgrid model, to name a few [19,20]. More recently, this work was extended and applied to Nek5000 [21].

Based on the above, it is clear that so far the majority of the work on UQ of scale-resolving simulations of wall-bounded turbulence has focussed on LES, with DNS usually providing a reference for benchmarking. In light of this, the purpose of this work is to apply VVUQ techniques to DNS, with the aim of propagating uncertainties and assessing the sensitivity of certain quantities of interest (QOIs) with respect to various input parameters. The high-order compact finite difference flow solver `Incompact3d` is used to perform DNS of a turbulent channel flow at a friction Reynolds number of $Re_\tau = 180$. To perform the UQ analysis, the EasyVVUQ package – a Python library designed to facilitate VVUQ workflows on HPC systems – is adopted. Three separate UQ campaigns are carried out. The first investigates the effect of the domain length along the two periodic (streamwise and spanwise) directions. The second campaign investigates the effect of five numerical parameters: mesh spacing in each direction, time step size, and the length of time the statistics are collected for. Finally, the third campaign repeats the second campaign using a dimension-adaptive sampling strategy to significantly reduce the number of samples required. In the following sections the flow solver and UQ approach are introduced, followed by a description of the case and campaign setup. Finally, the results are presented and discussed for each campaign, before concluding with suggestions for future work.

2. Methods

2.1. Flow solver

The numerical simulations presented in this work are performed with the high-order compact finite-difference flow solver `Incompact3d` [22], which is part of the open-source framework of flow solvers `Xcompact3d` [23]. The governing equations are the unsteady three-dimensional incompressible Navier–Stokes equations, given by:

$$\nabla \cdot \mathbf{u} = 0 \tag{1}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{1}{2}\left[\nabla(\mathbf{u}\otimes\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{u}\right] = -\frac{1}{\rho}\nabla p + \nu\nabla^2\mathbf{u} + \mathbf{F} \tag{2}$$

where $\mathbf{u}$ is the velocity vector, t is time, ρ is mass density, p is pressure, ν is the kinematic viscosity, and $\mathbf{F}$ accounts for any external forcing. Note that Eq. (2) is given in skew-symmetric form to reduce aliasing errors [24].

Eqs. (1) and (2) are discretised using sixth-order compact finite-difference stencils. For time integration, an explicit third-order Adams–Bashforth scheme is adopted. This is combined with an implicit Crank–Nicolson scheme for the diffusive terms in the wall-normal direction to circumvent the stability constraints imposed by the non-uniform mesh resolution used to properly resolve the near-wall effects. The pressure Poisson equation (PPE), which enforces incompressibility, is solved entirely in spectral space via the use of relevant three-dimensional fast Fourier transforms (FFTs). Through the use of a modified wavenumber [25], the divergence-free condition is ensured up to machine accuracy. To avoid the spurious pressure oscillations observed in fully-collocated approaches [22], the pressure field is defined offset (by half a mesh width) with respect to the velocity field. The simplicity of the structured mesh allows easy implementation of a two-dimensional domain decomposition strategy, based on pencils, using the message passing interface (MPI) [26]. The computational domain is split into several subdomains (pencils), each of which are assigned to an MPI process. The derivatives and interpolations in the x, y, and z direction are performed from within the X, Y, and Z pencils, respectively. The three-dimensional FFTs required by the PPE solver are performed as a series of one-dimensional FFTs, computed in one direction at a time. Global transpositions to switch from one pencil to another are performed via the MPI command `MPI_ALLTOALL(V)`. `Incompact3d`

has been extensively validated on a variety of turbulent flows [27–29] and shown to scale well on up to hundreds of thousands cores on CPU-based supercomputers [23,26].

2.2. Uncertainty quantification

The focus of this study is the forward UQ problem, where uncertainties in the input parameters are propagated through the model to generate statistical moments of the output QOIs. As a complement to UQ, sensitivity analysis is also performed, where the sensitivity (i.e. the relative contribution of the input parameters to the observed variance) of the QOIs are quantified with respect to each uncertain input parameter. There are a number of methods for this type of analysis. However, a requirement for this work is the approach should be non-intrusive, so that no modifications to the underlying code are required. Such techniques usually work by collecting samples (running simulations) at various points in the uncertain input space and computing the required statistics as a post-processing step. Response surface methods (e.g. stochastic collocation, gPC) are a subset of these methods that work by constructing a relatively cheap surrogate model of the QOIs with respect to the uncertain inputs. Through a judicious choice of the sample locations (usually via quadrature rules) exponential convergence is possible [30] and, as a result, these methods are usually more sample-efficient than brute-force statistical approaches (e.g. Monte Carlo) [31]. This work adopts the stochastic collocation method, which is similar in many regards to the popular gPC approach, with the main difference being in the choice of polynomial used to construct the underlying surrogate model [30].

All of the UQ analyses in this work are carried out using the EasyVVUQ library [8], a Python package designed to facilitate VVUQ workflows. EasyVVUQ automatically handles sampling, pre-processing, execution, post-processing, and analysis, and has been successfully applied to problems involving nuclear fusion [32], population migration [33], and COVID epidemiology [34], to name a few. EasyVVUQ is part of the Verified Exascale Computing for Multiscale Applications (VECMA) toolkit [35], which is a suite of tools to facilitate VVUQ on HPC systems and includes packages for code coupling and job submission and management, all of which are interoperable with each other. This is exploited in the present work, where the dimension-adaptive campaign, which is sequential in nature, combines EasyVVUQ with the FabSim3 library [36] to enable an automated workflow on HPC systems. This allows the EasyVVUQ campaign to be executed locally, while individual samples (simulations) are offloaded to a remote HPC system for execution with `Incompact3d`.

2.2.1. Stochastic collocation

The main idea behind stochastic collocation is to construct a polynomial approximation to the output QOIs as a function of the uncertain inputs. This is generated by sampling the input space at various locations and then building 1D Lagrange polynomials that can be used to interpolate the QOIs onto new locations within the input space. For multi-dimensional problems, the polynomial approximation is constructed via a tensor product of the 1D Lagrange polynomials. Given a generic QOI q in the stochastic (uncertain) input space $\boldsymbol{\xi} \in \mathbb{R}^d$, the polynomial approximation at any point in the input space is given by [30]:

$$q(\boldsymbol{\xi}) \approx \sum_{j_1=1}^{m_1} \cdots \sum_{j_d=1}^{m_d} q\left(\xi_1^{(j_1)}, \ldots, \xi_d^{(j_d)}\right) L_1^{(j_1)} \otimes \cdots \otimes L_d^{(j_d)} \tag{3}$$

where m_i is the number of collocation points required to generate the 1D Lagrange polynomial for the ith input ξ_i (which is determined by the polynomial order), and $L_i^{(j)}$ is the jth Lagrange basis polynomial for the ith input:

$$L_i^{(j)} = \prod_{\substack{k=1 \\ k \neq j}}^{m_i} \frac{\xi_i - \xi_i^{(k)}}{\xi_i^{(j)} - \xi_i^{(k)}} \tag{4}$$

Each input is assigned its own independent probability distribution, which, if available, should be informed by prior knowledge. The goal of UQ is to propagate these probability distributions through the model to get the corresponding statistical moments of the output QOIs.

The stochastic collocation expansion provides three main uses: (1) a surrogate model to interpolate QOIs onto unsampled locations in the input space; (2) estimates of the first two statistical moments (mean and variance) of the QOIs; (3) variance-based global sensitivity analysis (e.g. Sobol indices). The Sobol indices provide an indication of the relative influence of an input (or combination of inputs) for a given QOI [37]. These are calculated by computing the normalised partial variance for a given multi-index:

$$S_u = \frac{D_u}{D} \tag{5}$$

where $D = \mathrm{Var}(q)$ is the total variance of q, and D_u is the partial variance associated with the multi-index u, which can be any subset of $\{0, \ldots, d\}$. For more details on how the statistical moments and Sobol indices are computed, the reader is referred to Gerstner and Griebel [38] and Tang et al. [39].

An important part of UQ is providing an assessment of the model's robustness with respect to the uncertain inputs. For example, for a CFD practitioner, it is useful to know if small changes in the inputs lead to large changes in the outputs, as this would raise concerns around the robustness of the CFD model. This type of analysis requires quantifying the degree to which uncertainty in the input parameters is amplified/dampened by the model and manifested in the output QOIs. To aid with this assessment, the present work adopts the coefficient of variation ratio (CVR) metric, defined in Edeling et al. [34]. The CVR is defined as the ratio between the coefficients of variation (CV) for a given QOI and the input parameters. The CV itself is the ratio between the standard deviation and mean of a random variable and provides a dimensionless measure of its variability. Therefore, the CVR provides a relative measure of the output-to-input variability and is given by:

$$\mathrm{CVR}(q) = \mathrm{CV}(q) \;/\; \overline{\mathrm{CV}(\boldsymbol{\xi})} = \left|\frac{\sigma(q)}{\mu(q)}\right| \Big/ \left(\frac{1}{d}\sum_{i=1}^{d}\left|\frac{\sigma(\xi_i)}{\mu(\xi_i)}\right|\right) \tag{6}$$

where $\mu(q)$ and $\sigma(q)$ represent the mean and standard deviation of q, which are provided by the stochastic collocation expansion. For the input parameters, the statistical moments can be calculated from the user-defined probability distributions. A CVR value greater than one indicates amplification of uncertainty, whereas a CVR value smaller than one indicates damping.

The sampling strategy for stochastic collocation should be chosen based on the defined probability distributions of the uncertain inputs. Assuming appropriate Gaussian quadrature is adopted, exponential convergence of the polynomial approximation is possible [30]. For multi-dimensional problems, a tensor grid of quadrature points is required to sample the input space. Therefore, in general, the number of required samples is given by $N = \Pi_{i=1}^{d} m_i$. Based on this, it is clear to see that stochastic collocation still suffers from the curse of dimensionality and, as a result, typically becomes intractable for moderate/high-dimensional problems (e.g. $d > 5$). One approach to mitigating this is to adopt a dimensionality-reducing strategy, as discussed in the following section.

2.2.2. Dimension-adaptive sparse-grid sampling

In general, for real-world problems, it is unlikely that all uncertain inputs will have equal importance with respect to their effect on the output QOIs. Therefore, a 'smart' sampling strategy that exploits this effective low-dimensionality can significantly reduce the number of samples required. In addition to the full tensor grid approach described in the previous section, this work also adopts a dimension-adaptive scheme to investigate the potential of such methods for UQ of scale-resolving turbulent simulations. The basic idea behind the dimension-adaptive approach is to combine nested anisotropic sparse

grids with a ranking scheme for each uncertain input (dimension). Each dimension initially starts with a zeroth-order quadrature rule. Following this, the most effective dimensions are identified and successively refined to high-order quadratures in an iterative manner, while keeping the unimportant inputs at low-order quadratures. A brief overview of the dimension-adaptive approach is given here. However, for more details, the reader is referred to the tutorial of Edeling [40] and other works [34,38,41].

For the dimension-adaptive scheme, the stochastic collocation expansion is extended to a linear combination of separate tensor products of 1D Lagrange polynomials. Here, each tensor product can be represented by a quadrature-order multi-index, given by $\mathbf{l} = (l_1, \dots, l_d)$, where l_i is the quadrature order associated with the ith input. For example, in 2D, the multi-index $\mathbf{l} = (2, 1)$ assigns second and first-order quadrature rules to the first and second inputs, respectively. Let Λ be the set containing all currently selected quadrature-order multi-indices. Initially, a zeroth-order quadrature rule is applied to all inputs, so that the initial set of multi-indices is $\Lambda_0 = \{(0, \dots, 0)\}$. At each iteration of the dimension-adaptive scheme, a new multi-index is added to the set according to a chosen ranking metric designed to identify the most important inputs. However, a given multi-index is only admissible if all of its backward neighbours – defined by $\{\mathbf{l} - \mathbf{e}_i \mid l_i > 0, 1 \leq i \leq d\}$, where $\mathbf{e}_i$ is the elementary basis vector for the ith input – are in the current multi-index set. Referring back to the 2D example, assuming the current multi-index set is $\Lambda = \{(0, 0), (1, 0)\}$, then the set of admissible forward neighbours is $M = \{(0, 1), (2, 0)\}$.

To adaptively refine along each input dimension, a look-ahead step is performed to select the new multi-index that is to be added to the multi-index set. This requires drawing samples from all of the new quadrature collocation points associated with each admissible forward neighbour of the current multi-index set. After the samples are drawn, each admissible forward neighbour is ranked according to a chosen metric. This work follows Edeling et al. [34] and adopts the hierarchical surplus error, which is defined for each admissible forward neighbour as the difference between the code output at the new sample locations associated with that forward neighbour and the current polynomial approximation at those same sample locations. This can be thought of as a measure of the accuracy of the current polynomial approximation and at each iteration the admissible forward neighbour with the largest surplus error is included in the new multi-index set. After adding the new multi-index to the set, this may cause new forward neighbours to become admissible, and the cycle repeats until a user-defined stopping criterion is met.

The polynomial interpolation given by Eq. (3) must be modified to work with the dimension-adaptive algorithm, since each index within the index set Λ constitutes a separate tensor product of 1D quadrature rules [34,40]:

$$q(\boldsymbol{\xi}) \approx \sum_{\mathbf{l} \in \Lambda} c_{\mathbf{l}} \sum_{j_1=1}^{m_{l_1}} \cdots \sum_{j_d=1}^{m_{l_d}} q\left(\xi_1^{(j_1)}, \dots, \xi_d^{(j_d)}\right) L_1^{(j_1)} \otimes \cdots \otimes L_d^{(j_d)} \tag{7}$$

where m_{l_i} is the number of points generated by a 1D quadrature rule of order l_i, and the coefficients $c_{\mathbf{l}}$ are given by [38]:

$$c_{\mathbf{l}} = \sum_{z_1=0}^{1} \cdots \sum_{z_d=0}^{1} (-1)^{\|\mathbf{z}\|_1} \chi^{(\mathbf{l}+\mathbf{z})} \tag{8}$$

where $\mathbf{z} = (z_1, \dots, z_d)$ and:

$$\chi^{(\mathbf{l})} = \begin{cases} 1 & \text{if } \mathbf{l} \in \Lambda \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Note that the statistical moments and Sobol indices are computed slightly differently in the dimension-adaptive scheme, compared with the tensor grid approach described in the previous section. Here, following previous works [34,42,43], the adaptive stochastic collocation expansion is transformed into a polynomial chaos expansion to facilitate the calculation of these quantities. For more details, the reader is referred to Buzzard [42], Jakeman et al. [43], Edeling et al. [34] and Edeling [40].

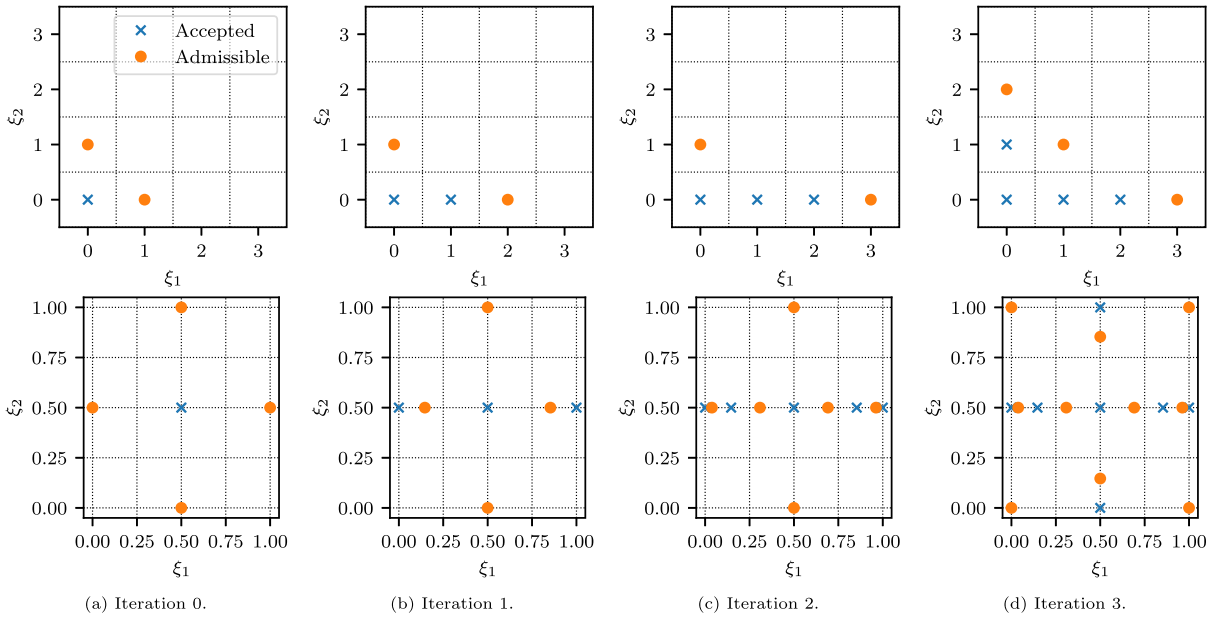
It is important to highlight that the choice of quadrature rule can have a large impact on the number of samples required. This work adopts the Clenshaw-Curtis quadrature rule for the dimension-adaptive scheme since it exhibits similar performance to Gaussian quadrature in terms of interpolation accuracy, while also offering the advantage of being nested [44]. This is a highly desirable property as it ensures that the sample points of a given quadrature order contain all the points generated by the same rule at all lower orders. As a result, many points overlap when taking linear combinations of tensor products built from nested 1D rules of different orders, thus leading to an efficient sparse sampling strategy. When combined with the dimension-adaptive scheme described above, this enables the study of high-dimensional problems beyond what is currently achievable with the full tensor grid approach. As an example, the reader is referred to Edeling et al. [34], where this method is applied to an epidemiological model with a 19-dimensional input space.

An illustration of the dimension-adaptive algorithm for a generic (example) 2D problem is given in Fig. 1. The initial index set is given by $\Lambda_0 = \{(0, 0)\}$, for which the set of admissible forward neighbours is $M_0 = \{(1, 0), (0, 1)\}$. To select which admissible forward neighbour should be included in the index set at the next iteration, the hierarchical surplus error must be calculated for each admissible forward neighbour. This requires drawing all new and unique samples associated with the admissible forward neighbours (given by the orange circles in the bottom row of Fig. 1). Once the hierarchical surplus has been calculated for each admissible forward neighbour, the one with the largest error is added to the index set. In this case, suppose the new index set becomes $\Lambda_1 = \{(0, 0), (1, 0)\}$. This leads to a new admissible forward neighbour set $M_1 = \{(0, 1), (2, 0)\}$, which requires additional model evaluations at the new sample locations associated with the new multi-index. However, as can be seen from the bottom row of Fig. 1(b), some of the new sample locations overlap with previously sampled locations (due to the nested quadrature rule). Therefore, only two additional model evaluations are required at this iteration. Following this, the algorithm repeats for a number of iterations until a user-defined stopping criterion is met. Note that in the third iteration, with $\Lambda_3 = \{(0, 0), (1, 0), (2, 0), (0, 1)\}$, the set of admissible forward neighbours is given by $M_3 = \{(3, 0), (0, 2), (1, 1)\}$. This contains a cross-coupling multi-index, given by $l = (1, 1)$, which will capture any variance associated with high-order cross-coupling between the two inputs.

3. Simulation setup

This work focusses on a canonical turbulent channel flow at a target friction Reynolds number of $Re_\tau = 180$, as shown in Fig. 2. The streamwise, wall-normal, and spanwise directions are denoted by x, y, and z, respectively. Equivalently, u, v, and w denote the instantaneous velocity components in the three spatial directions, each of which can be decomposed into a time-averaged ($\overline{u}$) and fluctuating component (u'), such that $u = \overline{u} + u'$. Note that in the following sections all quantities are non-dimensionalised with respect to the channel half-height (δ). However, where it is more convenient to use inner-scaled quantities (denoted by the plus superscript notation), these are non-dimensionalised with respect to the target friction velocity ($u_\tau = 180\nu/\delta$) and friction length ($\delta_\tau = \nu/u_\tau$). Furthermore, an additional time scale (denoted by the star superscript notation) is also defined ($t^* = tu_\tau/\delta$).

The boundary conditions are composed of periodic conditions in the streamwise and spanwise directions and a no-slip condition on the top and bottom walls. The mesh is uniformly spaced in the streamwise and spanwise directions, with non-uniform spacing in the wall-normal direction to properly resolve the near-wall effects. This is achieved



(a) Iteration 0. (b) Iteration 1. (c) Iteration 2. (d) Iteration 3.

Fig. 1. Sparse-grid quadrature order (top) and sample locations (bottom) for an illustrative 2D problem using the dimension-adaptive sampling approach. Note that this is just an example illustration and the exact sequence will be dependent on the specific problem and algorithmic settings.

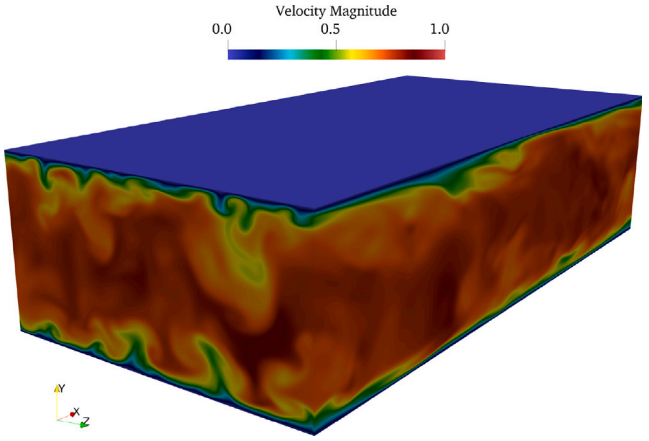


Fig. 2. Illustrative snapshot of the turbulent channel flow case.

via a carefully designed mapping function that preserves the strict physical/spectral equivalence of the discretisation scheme, while also avoiding an expensive convolution in spectral space when solving the PPE (see Laizet and Lamballais [22] for more details). For all cases, the mesh resolution at the wall was chosen to be $\Delta y_w^+ \approx 0.5$ to ensure that the turbulence at the wall is fully resolved. The flow is initialised to a laminar Poiseuille solution through the entire domain, and is subsequently driven by a dynamically-adjusted body force to preserve a constant bulk flow rate (u_b) throughout the entire simulation. The kinematic viscosity is set to match a bulk Reynolds number of $Re_b = u_b 2\delta/\nu = 5600$, which has been empirically shown to lead to a friction Reynolds number of approximately $Re_\tau \approx 180$ [45]. To accelerate the transition to turbulence, an initial noise level of 12.5% is prescribed on the velocity field, in addition to a rotational forcing which is applied until $t^* = 1$ ($t \approx 23.3$). After this, the flow is allowed to develop until $t^* = 10$ ($t \approx 233.3$) before the recording of the statistics begins. For the analysis, the main QOIs are the measured friction Reynolds number as well as cross-channel profiles of the first and second-order velocity moments, all of which are averaged in time as well as along the two homogeneous (streamwise and spanwise) directions.

The main sources of uncertainty/variability that should be considered for this specific case and configuration are:

- **Domain Size**: the length of the domain in the two periodic directions (streamwise and spanwise).
- **Parameter Uncertainty**: the bulk Reynolds number (i.e. kinematic viscosity) required to achieve the target friction Reynolds number.
- **Spatial Discretisation**: including both the discretisation scheme and mesh resolution.
- **Temporal Discretisation**: including both the discretisation scheme and time step size.
- **Statistical Uncertainty**: due to the finite sampling time (controlled by specifying the length of time over which the statistics are time-averaged).

As mentioned above, the present work neglects the uncertainty associated with the kinematic viscosity by selecting an empirical value obtained from previous works. Furthermore, to reduce the complexity of the present study, the effects of the spatial and temporal discretisation schemes are neglected. For the mesh resolution, the wall-normal stretching in `Incompact3d` is typically controlled by the number of mesh points in the wall-normal direction (N_y) and a dilation parameter (β). For practical purposes, in this work the wall-normal mesh resolution is instead parametrised in terms of the mesh spacing at the wall (Δy_w^+) and at the centre of the channel (Δy_c^+). As already stated above, the mesh resolution at the wall is chosen to be $\Delta y_w^+ \approx 0.5$ to ensure that the near-wall effects are fully resolved. Therefore, in this work Δy_c^+ is treated as the uncertain parameter for the wall-normal mesh resolution. Accordingly, for each individual case, N_y and β are set in order to match the desired Δy_c^+ while also ensuring $\Delta y_w^+ \approx 0.5$. Finally, the statistical uncertainty is reformulated in terms of the length of time over which the statistics are time-averaged. Statistical uncertainty is a form of aleatoric uncertainty and is therefore typically treated differently to the other types of (epistemic) uncertainty, either through ensembles [46] or by direct estimation [17]. Since scale-resolving simulations of turbulent flows typically control this type of uncertainty by increasing the length of time over which the statistics are time-averaged, the present work treats this user-defined parameter as a proxy for the statistical uncertainty. This simplifies the problem by allowing each source of uncertainty to be treated in the same

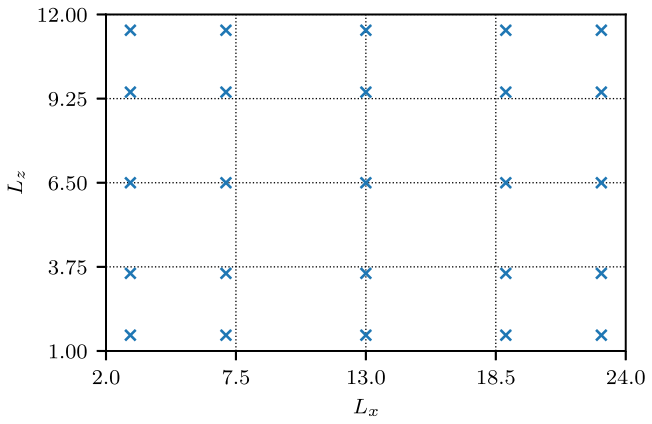


Fig. 3. Uncertain input space and sample locations for Campaign I. The sample scheme is constructed via a tensor product of fourth-order 1D Gauss–Legendre quadrature points.

Table 1
Ranges for the uncertain inputs in Campaign I. For comparison, equivalent values from selected previous works are also shown.

	L_x	L_z
Present	2–24	1–12
Kim et al. [45]	12.6	6.3
Moser et al. [47]	12.6	4.2
Oliver et al. [17]	12.6–37.7	6.3–12.6
Vreman and Kuerten [48]	12.6	4.2
Lee and Moser [3]	25.1	9.4

Table 2
Fixed parameters for Campaign I.

Parameter	Value
Δx^+	6
Δy^+_w	0.5
Δy^+_c	4
Δz^+	6
Δt^*	1×10^{-4}
t^*_{stat}	50

way (via the stochastic collocation expansion). Note that, due to the fact the QOIs are spatially-averaged in both homogeneous directions, in addition to the time-averaging, the statistical uncertainty in the present setup is expected to be small. Therefore, the effect of this subtle distinction in the treatment of the statistical uncertainty is expected to be negligible.

After the above simplifications, the remaining sources of uncertainty are the effects of domain size (two directions), mesh resolution (three directions), time step size, and the length of time over which the statistics are time-averaged. Even with the above simplifications, the size of the UQ problem is still too large to be practically feasible. Therefore, the present work splits the sources of uncertainty into two groups and performs a separate UQ campaign for each group. The first campaign focusses on just the domain size, whereas the second campaign focusses on the remaining numerical parameters.

3.1. Campaign I: Domain size

The aim of the first campaign is to investigate the effect of the domain size on the DNS of a turbulent channel flow. Specifically, the lengths of the domain along the two periodic directions – streamwise (L_x) and spanwise (L_z) – are selected as the uncertain inputs. The case is set up as described in the introduction to Section 3. The ranges for the two inputs are chosen based on previous works. These are given by $L_x \in [2, 24]$ and $L_z \in [1, 12]$, which covers most of the previously selected values in the literature for this Reynolds number, as shown in Table 1. Following Rezaeiravesh et al. [21], uniform probability distributions are applied to each of the two inputs and a tensor grid of fourth-order 1D Gauss–Legendre quadrature points is constructed to sample the input space. This results in a total of $5^2 = 25$ separate DNS, as shown in Fig. 3. All other parameters are fixed throughout the UQ campaign, as shown in Table 2, and are chosen by balancing accuracy with the cost of the campaign. Note that for each simulation it was necessary to slightly adjust the streamwise and spanwise mesh spacing (Δx^+ and Δz^+) to exactly match the required lengths (L_x and L_z) while still maintaining an integer number of mesh points.

3.2. Campaign II: Numerical parameters

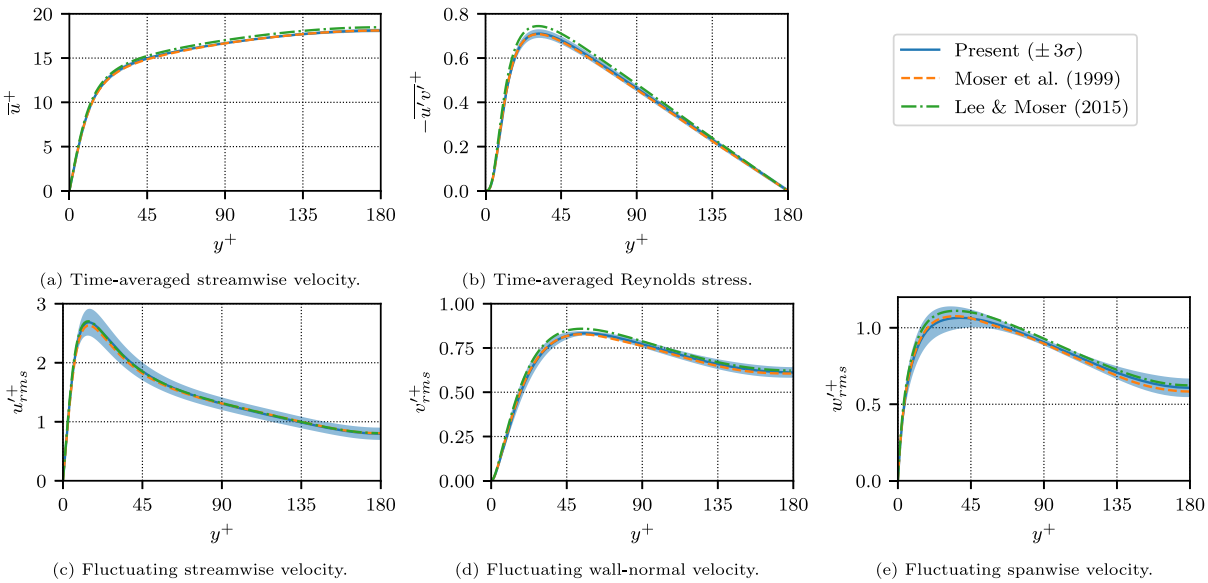
The aim of the second campaign is to investigate the effect of various numerical parameters – mesh resolution in each direction (Δx^+, Δy^+_c, and Δz^+), time step size (Δt^*), and the length of time the statistics are collected for (t^*_{stat}) – on the DNS of a turbulent channel flow. Note that Δy^+_c refers to the wall-normal mesh spacing at the centre of the channel (since the mesh spacing is non-uniform in the wall-normal direction). Like the first campaign, the case is set up as described in Section 3 and the ranges for the uncertain inputs are selected based on previous works in the literature, as shown in Table 3. These ranges are given by $\Delta x^+ \in [4, 30]$, $\Delta y^+_c \in [4, 30]$, $\Delta z^+ \in [4, 30]$, $\Delta t^* \in [0.25, 1.25]\times 10^{-4}$, and $t^*_{stat} \in [5, 50]$. Also like the first campaign, uniform probability distributions are applied to each of the uncertain inputs and a tensor grid of fourth-order 1D Gauss–Legendre quadrature points is constructed to sample the input space, resulting in a total of $5^5 = 3125$ separate DNS. All other parameters are fixed throughout the campaign and are given in Table 4. Note that for each simulation it was necessary to slightly adjust the streamwise and spanwise lengths (L_x and L_z) to exactly match the required mesh spacing (Δx^+ and Δz^+) while still maintaining an integer number of mesh points. For the wall-normal mesh spacing, this was set by adjusting both the number of mesh points in the wall-normal direction (N_y) and the stretching parameter (β). For each sample, the combination of these two parameters that led to a wall-adjacent mesh spacing closest to the target value ($\Delta y^+_w \approx 0.5$), while still exactly matching the target mesh spacing at the centre of the channel, was selected.

4. Results

This section presents the results from the two main campaigns, as well as the results from the dimension-adaptive sampling campaign. All of the simulations were performed on either the Cirrus UK National Tier-2 HPC service or the ARCHER2 UK national supercomputer. Cirrus consists of 280 nodes housing two Intel Xeon E5-2695 (Broadwell) CPUs (2×18 total cores @ 2.1 GHz) in a dual non-uniform memory access arrangement (NUMA). Network communication on Cirrus is via the InfiniBand FDR interconnect, with 54.5 Gbit/s of bandwidth. ARCHER2, on the other hand, is made up of 5860 nodes housing two AMD EPYC 7742 CPUs (2×64 total cores @ 2.25 GHz), with eight NUMA regions per node. Network communication on ARCHER2 is via the HPE Slingshot interconnect, which provides 2×100 Gbit/s of bidirectional bandwidth. Owing to the nature of the present study, there is a large variation in total execution time across all individual simulations, from approximately 20 min to 4 days.

4.1. Campaign I: Domain size

The purpose of this section is to investigate the effect of the domain size. As discussed in Section 3.1, this campaign consisted of 25 separate



(a) Time-averaged streamwise velocity.
(b) Time-averaged Reynolds stress.
(c) Fluctuating streamwise velocity.
(d) Fluctuating wall-normal velocity.
(e) Fluctuating spanwise velocity.

Fig. 4. Cross-channel profiles of first and second-order velocity moments for Campaign I. Note that the inner-scaling is with respect to the target friction velocity.

Table 3
Ranges for the uncertain inputs in Campaign II. For comparison, equivalent values from selected previous works are also shown.

	Δx^+	Δy_c^+	Δz^+	$\Delta t^*(\times 10^{-4})$	t^*_{stat}
Present	4–30	4–30	4–30	0.25–1.25	5–50
Kim et al. [45]	12	4.4	7	–	10
Moser et al. [47]	17.7	4.4	5.9	–	–
Oliver et al. [17]	6.1–27.4	2.3–10.1	3.0–12.2	3.21–12.86	48.5–2141.6
Vreman and Kuerten [48]	4.4–17.7	2.2–4.4	2.9–5.9	2.5–10	161–1300
Lee and Moser [3]	4.5	3.4	3.1	–	31.9

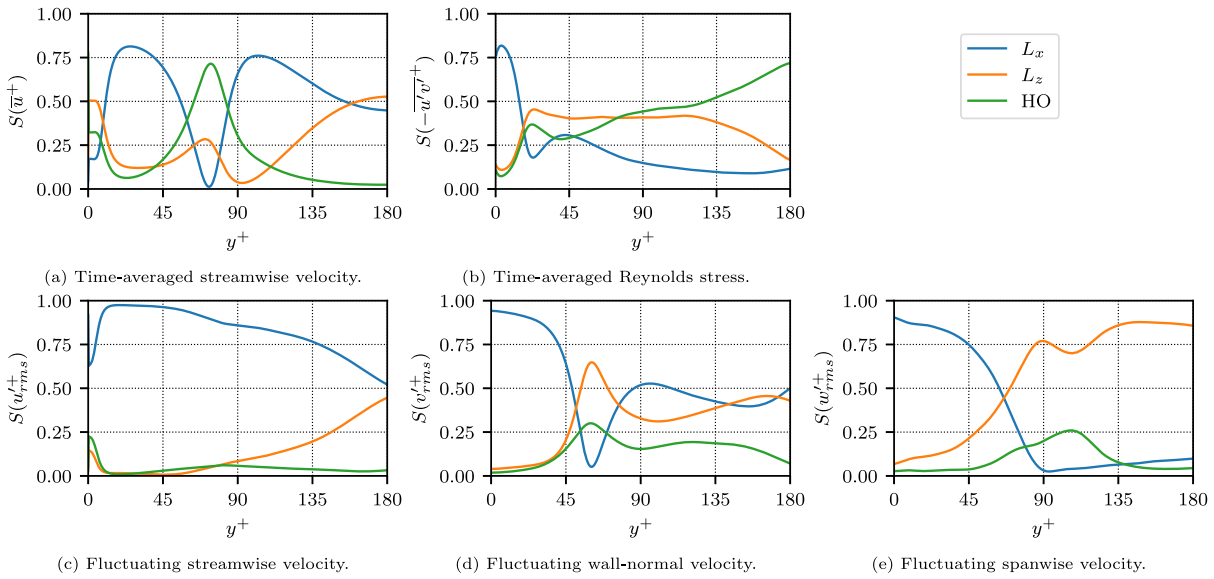
Table 4
Fixed parameters for Campaign II.

Parameter	Value
L_x	8
L_z	4
Δy_w^+	0.5

DNS, requiring approximately 350,000 core-hours of compute time. Fig. 4 shows the statistical moments of the cross-channel profiles for various QOIs, where the mean is given by the solid line and the shaded region indicates plus/minus three standard deviations. For reference, the benchmark data of Moser et al. [47] and Lee and Moser [3] are also shown. The mean profiles generally agree well with the benchmark data. However, significant variance can also be observed, depending on the QOI and wall-normal location. In particular, the streamwise and spanwise components of the fluctuating velocity show significant relative variance across the channel, but especially around the peaks in the buffer layer ($y^+ \approx 20$–25). On the other hand, the variance is somewhat lower for the Reynolds stress and wall-normal component of the fluctuating velocity, and is negligible for the mean streamwise velocity.

Fig. 4 clearly demonstrates that the domain size has a significant effect on the observed QOIs, particularly the second-order velocity moments. Fig. 5 shows how sensitive this response is with respect to each input, as a function of the wall-normal location. Overall, there is a mix of behaviours. However, focussing on the streamwise and spanwise components of the fluctuating velocity, since they exhibit the largest relative variance, it is clear that L_x dominates the variance observed in the streamwise velocity across most of the channel height. In fact, only at the channel centre does the sensitivity associated with L_z approach that of L_x. For the spanwise velocity, the variance is dominated by L_x close to the wall. However, at approximately $y^+ \approx 70$ this behaviour reverses and instead L_z becomes the most influential input. Generally, the sensitivity to high-order cross-coupling terms is low, apart from in the mean streamwise velocity and Reynolds stress. However, it should be noted that the variance associated with these QOIs is relatively low, which will affect the robustness of the calculation of the Sobol indices.

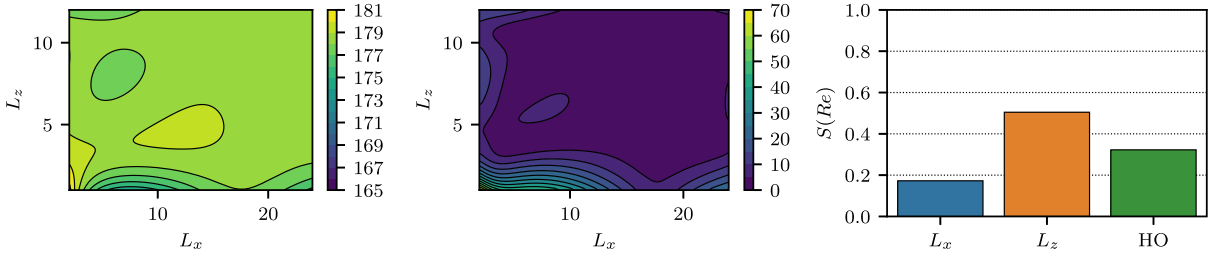
Although the target friction Reynolds number in all of the simulations is $Re_\tau = 180$, in practice this will not be achieved exactly. Therefore, the measured friction Reynolds number is also an interesting QOI to study. Fig. 6(a) shows the underlying surrogate model for the measured friction Reynolds number, as a function of the two inputs. This is constructed as part of the stochastic collocation procedure and is built via a tensor product of Lagrange polynomials using the measured Re_τ at the sample locations. Note that this figure cannot be used to locate the 'correct' region of the parameter space that will exactly match the target Reynolds number. This is because, as discussed in Section 3, the viscosity is set according to empirical data [45] to approximately match the target Reynolds number. Therefore, even in the case of zero modelling error, there is no guarantee that the measured Reynolds number will exactly match the target. Instead, Fig. 6(a) provides useful information about the behaviour of the measured Reynolds number with respect to the uncertain inputs. In particular, the alignment and spacing of the contour lines indicate the sensitivity of the measured Reynolds number with respect to each input at a particular point in the input space. Fig. 6(b) helps with this analysis by providing a local sensitivity analysis, which is calculated by taking the magnitude of the gradient of the surrogate model. Note that, to facilitate comparison between different inputs, the input space is normalised to a unit square before computing the gradient. Fig. 6(b) displays regions of the parameter space where the measured Reynolds numbers is especially sensitive to these two inputs. Clearly, for a CFD practitioner, it would be beneficial to choose a domain size in the upper right quadrant of



(a) Time-averaged streamwise velocity. (b) Time-averaged Reynolds stress.

(c) Fluctuating streamwise velocity. (d) Fluctuating wall-normal velocity. (e) Fluctuating spanwise velocity.

Fig. 5. Cross-channel Sobol indices of first and second-order velocity moments for Campaign I. HO refers to high-order cross-coupling terms. Note that the inner-scaling is with respect to the target friction velocity.



(a) Surrogate model of Re_τ. (b) Local sensitivity analysis of Re_τ. (c) Global sensitivity analysis of Re_τ.

Fig. 6. Analysis of measured Reynolds number (Re_τ) for Campaign I.

this figure (e.g. $L_x > 10$ and $L_z > 7$) to ensure robust results that are insensitive to the domain size. This is an intuitive result, since it is expected that the results of a simulation would converge as the domain length along a periodic direction increases. Finally, Fig. 6(c) displays the global sensitivity of the measured Reynolds number with respect to each input. Clearly, L_z is the dominant input for this QOI. However, L_x is also influential and there is also some cross-coupling. This is supported by examining Fig. 6(a), where it can be seen that most of the variance is with respect to L_z, whereas the influence from L_x is somewhat smaller.

To assess the robustness of `Incompact3d` for this specific setup, Fig. 7 shows the cross-channel profiles of CVR, as given by Eq. (6), for various QOIs. For the most part, it is clear that `Incompact3d` dampens the effect of the uncertain inputs on the output QOIs. In fact, for almost all QOIs and channel locations, the CVR is much lower than one. The only instance where this is not true is in the Reynolds stress close to the channel centre, where the CVR increases significantly to an extreme value of approximately 7 (not shown). However, care should be taken when interpreting the CVR in this region, since the mean of the Reynolds stress approaches zero here (see Fig. 4(b)), and thus the CV will be amplified in this region, even if the variance is low in absolute terms. Table 5 also shows the CVR for the measured Reynolds number, as well as the mean of the cross-channel CVR profiles displayed in Fig. 7. Again, the damping effect of `Incompact3d` with respect to the input uncertainty is clear. Furthermore, it can be seen that the QOIs that display the most variability, in addition to the Reynolds stress, are the streamwise and spanwise components of fluctuating velocity, which can be qualitatively confirmed by examining Fig. 4.

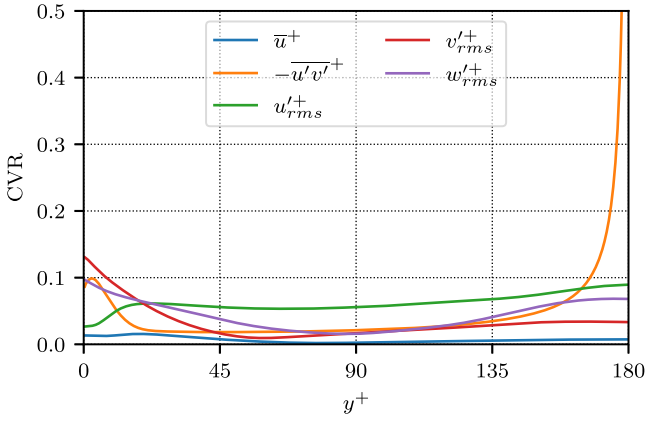
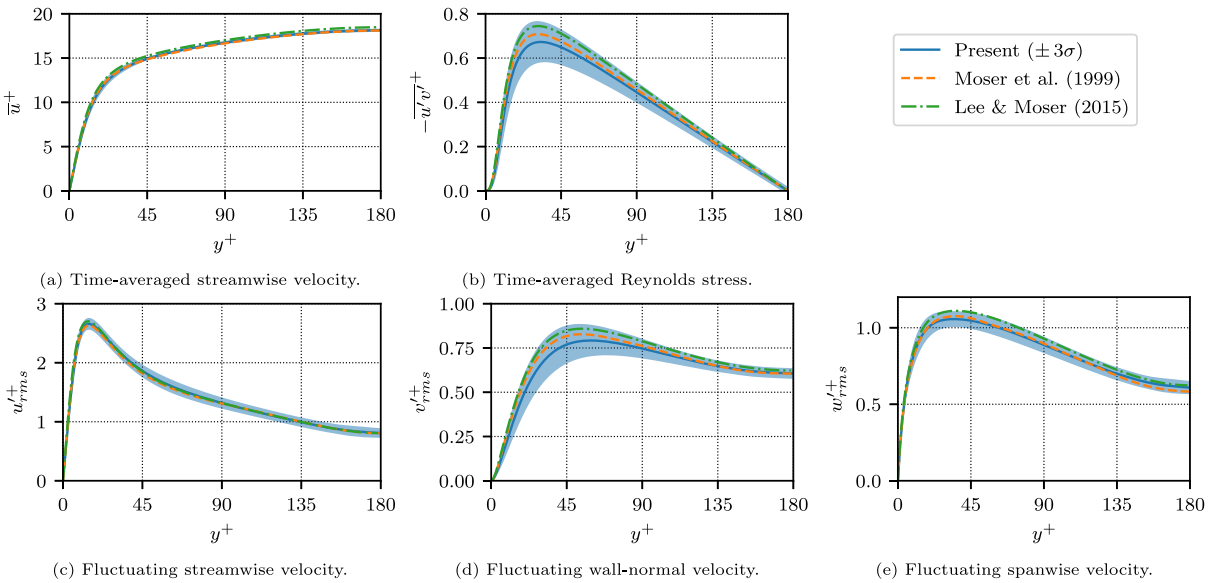


Fig. 7. Cross-channel profiles of CVR for the first and second-order velocity moments for Campaign I. Note that the limits of the CVR axis have been clipped for clarity and the inner-scaling is with respect to the target friction velocity.

4.2. Campaign II: Numerical parameters

The purpose of this section is to investigate the effect of the numerical parameters. As discussed in Section 3.2, this campaign consisted of 3125 separate DNS, requiring approximately 450,000 core-hours of compute time. Fig. 8 shows the statistical moments of the cross-channel



(a) Time-averaged streamwise velocity. (b) Time-averaged Reynolds stress.

(c) Fluctuating streamwise velocity. (d) Fluctuating wall-normal velocity. (e) Fluctuating spanwise velocity.

Fig. 8. Cross-channel profiles of first and second-order velocity moments for Campaign II. Note that the inner-scaling is with respect to the target friction velocity.

Table 5
Mean CVR for all QOIs for Campaign I. For this campaign the CV for the inputs is 0.489.

QOI	CV	CVR
Re_τ	0.003	0.007
$\overline{u}^+$	0.003	0.007
$-\overline{u'v'}^+$	0.033	0.069
u'^+_{rms}	0.030	0.062
v'^+_{rms}	0.016	0.032
w'^+_{rms}	0.021	0.043

profiles for various QOIs. The findings are broadly similar to Fig. 4, with good agreement between the mean profiles and the benchmark data. Furthermore, the variance is almost exclusively manifested in the second-order velocity moments, with negligible variance in the mean streamwise velocity. The main contrast with the first campaign is that while the variance in the Reynolds stress and wall-normal fluctuating velocity in the first campaign was relatively small, for this campaign these QOIs display the largest variance. Again, the peak in the variance generally occurs in the buffer layer, around $y^+ \approx 20$–25.

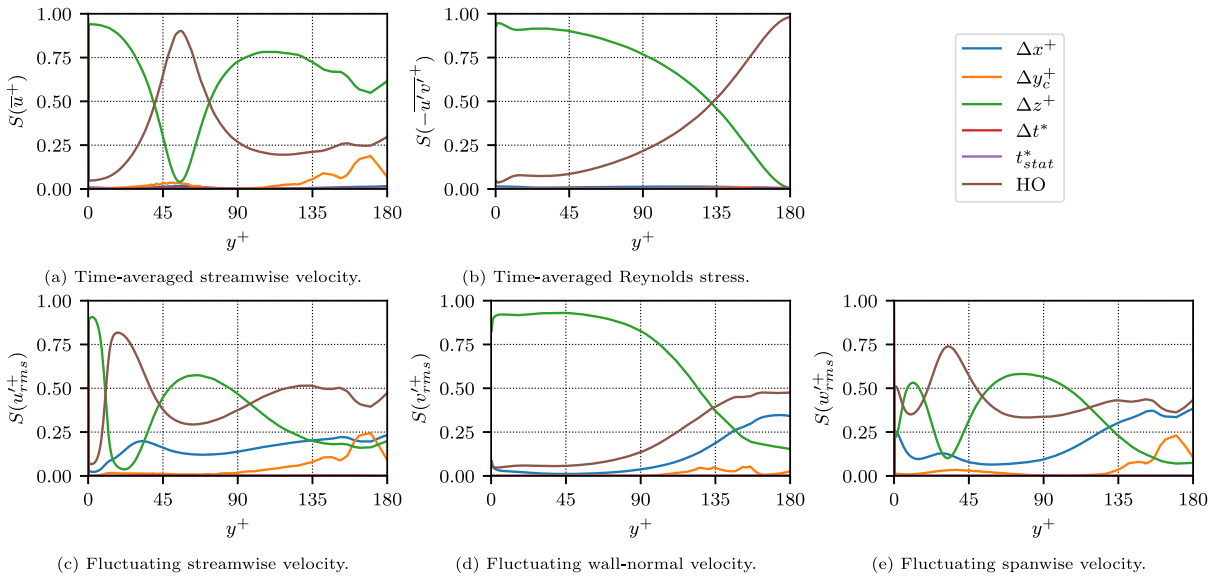
Fig. 9 shows the cross-channel Sobol indices for each uncertain input and QOI. Again, there is a mix of behaviours. However, it is clear that the spanwise mesh resolution (Δz^+) is the most influential input. This is especially true close to the wall, whereas towards the centre of the channel the streamwise mesh resolution (Δx^+) and high-order cross-couplings become more influential. It is interesting to note that there is almost no sensitivity associated with the time step size (Δt^*) or the length of time the statistics are collected for (t^*_{stat}). In the case of the time step size, this can be explained by the fact that the range for this parameter was chosen primarily based on stability constraints. In particular, to ensure a stable solution across all sampled locations in the input space, it was necessary to choose a range where the maximum time step size would still satisfy the stability requirements for the most highly resolved mesh. By accommodating for these extreme values it meant the majority of the simulations in this campaign were 'over-resolved' in the temporal domain. With regards to the length of time the statistics were collected for, in addition to time-averaging, all QOIs were also spatially-averaged along the two homogeneous directions. This accelerates the convergence of the statistics and is therefore the most likely reason why the QOIs are not especially sensitive to the length of the time-averaging window.

Fig. 10(a) shows a 2D slice of the underlying surrogate model for the measured friction Reynolds number, as a function of the streamwise and spanwise mesh resolution. These inputs were chosen since, from Fig. 9, they were deemed to be the most influential inputs with respect to the QOIs. For this slice, the remaining inputs were set to the centre of their respective ranges. Based on the alignment of the contour lines, Δz^+ is clearly the dominant input responsible for the majority of the observed variance in the measured Reynolds number. This is perhaps not surprising, given that the coherent structures associated with wall-bounded turbulence are known to align and stretch with the flow along the streamwise direction. As a result, the mesh resolution requirements are not as strict in the streamwise direction as they are in the spanwise direction. Examining the spacing of the contour lines, it is clear that the measured Reynolds number is more sensitive when Δz^+ is large. This is reinforced in Fig. 10(b), which displays the local sensitivity by taking the magnitude of the gradient of the surrogate model (after normalising the input space to a unit hypercube). Similar to the previous section, this figure can be used to determine appropriate values for the input parameters to ensure robust results. In this case, for the spanwise mesh resolution, it is seen that a value of $\Delta z^+ < 10$ is necessary to minimise the sensitivity. On the other hand, the measured Reynolds number is generally insensitive to the streamwise mesh resolution over the examined range, although there is a small region of slightly increased sensitivity at approximately $10 < \Delta x^+ < 20$. Finally, Fig. 10(c) displays the global sensitivity of the measured Reynolds number with respect to each uncertain input. Clearly, Δz^+ dominates and is responsible for approximately 95% of the total variance, whereas the effect of each of the other uncertain inputs is negligible.

Fig. 11 shows the cross-channel profiles of CVR for various QOIs. The results are broadly similar to the first campaign. Again, it is evident that `Incompact3d` dampens the effect of the uncertain inputs on the output QOIs. However, the CVR values are somewhat larger than the first campaign, indicating that the uncertainty is damped to a lesser degree for this set of uncertain inputs. Table 6 also shows that, for this campaign, the Reynolds stress and wall-normal component of fluctuating velocity exhibit the most variability, which can also be confirmed by examining Fig. 8.

4.3. Campaign III: Dimension-adaptive sampling

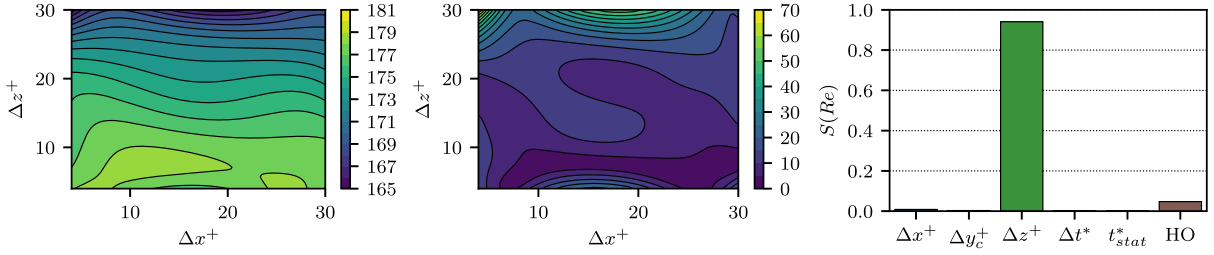
While the results presented in the previous sections are interesting and informative, they are expensive to obtain and require significant



(a) Time-averaged streamwise velocity. (b) Time-averaged Reynolds stress.

(c) Fluctuating streamwise velocity. (d) Fluctuating wall-normal velocity. (e) Fluctuating spanwise velocity.

Fig. 9. Cross-channel Sobol indices of first and second-order velocity moments for Campaign II. HO refers to high-order cross-coupling terms. Note that the inner-scaling is with respect to the target friction velocity.



(a) Surrogate model of Re_τ. (b) Local sensitivity analysis of Re_τ. (c) Global sensitivity analysis of Re_τ.

Fig. 10. Analysis of measured Reynolds number (Re_τ) for Campaign II. Note that the inner-scaling is with respect to the target friction velocity.

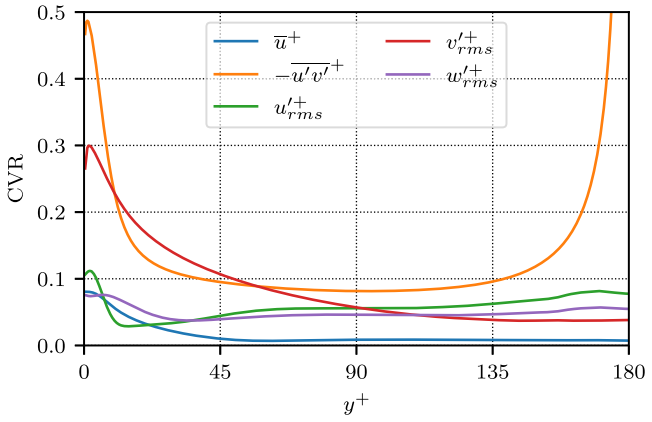


Fig. 11. Cross-channel profiles of CVR for the first and second-order velocity moments for Campaign II. Note that the limits of the CVR axis have been clipped for clarity and the inner-scaling is with respect to the target friction velocity.

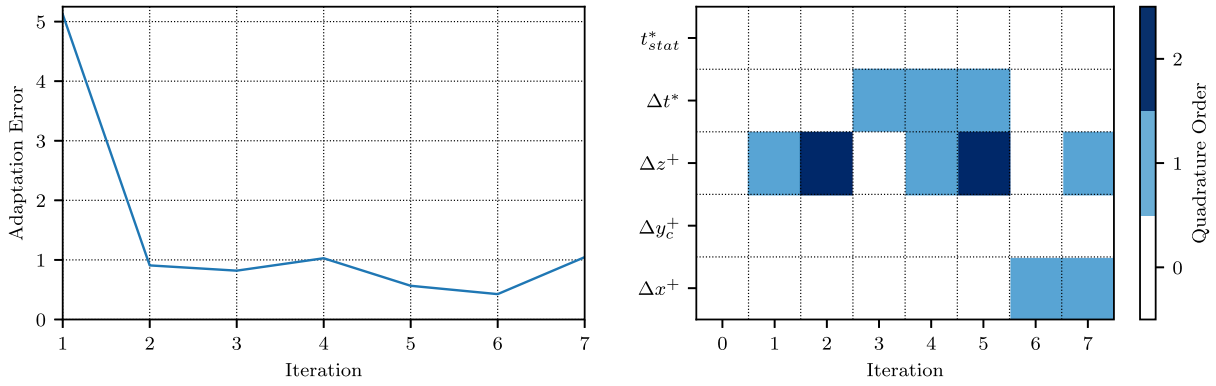
computational resources. Furthermore, including even just one more dimension in the uncertain input space would make this type of analysis intractable. Therefore, the purpose of this section is to investigate the potential benefits of adopting a 'smart' sampling strategy to reduce the number of samples (simulations) required to perform the types of analyses presented in the previous sections. Specifically, the dimension-adaptive approach described in Section 2.2.2 is used to repeat the same

Table 6
Mean CVR for all QOIs for Campaign II. For this campaign the CV for the inputs is 0.436.

QOI	CV	CVR
Re_τ	0.018	0.041
$\overline{u}^+$	0.007	0.015
$-\overline{u'v'}^+$	0.112	0.257
u'^+_{rms}	0.025	0.057
v'^+_{rms}	0.037	0.085
w'^+_{rms}	0.021	0.049

study conducted in the second campaign. The setup for this campaign is therefore the same as Campaign II, which has five uncertain inputs, given by the mesh resolution in each direction, the time step size, and the length of time the statistics are collected for. Based on the results from the second campaign (Section 4.2), this problem should be ideally suited to the dimension-adaptive approach, since the sensitivities of the QOIs are essentially dominated by just one of the inputs (the spanwise mesh resolution). Therefore, the dimension-adaptive scheme should identify this input as the most important and prioritise resources towards resolving its effect.

Table 7 displays an overview of the progression of the dimension-adaptive campaign at each iteration. Here, the measured friction Reynolds number is selected as the QOI to direct the sampler at each iteration. As described in Section 2.2.2, this is done by including the admissible forward neighbour with the largest surplus error into the accepted set of quadrature orders at each iteration. The campaign is



(a) Reynolds number (Re_τ) surplus error at each iteration.

(b) Quadrature order applied to each input at each iteration.

Fig. 12. Summary of Campaign III progression for each iteration.

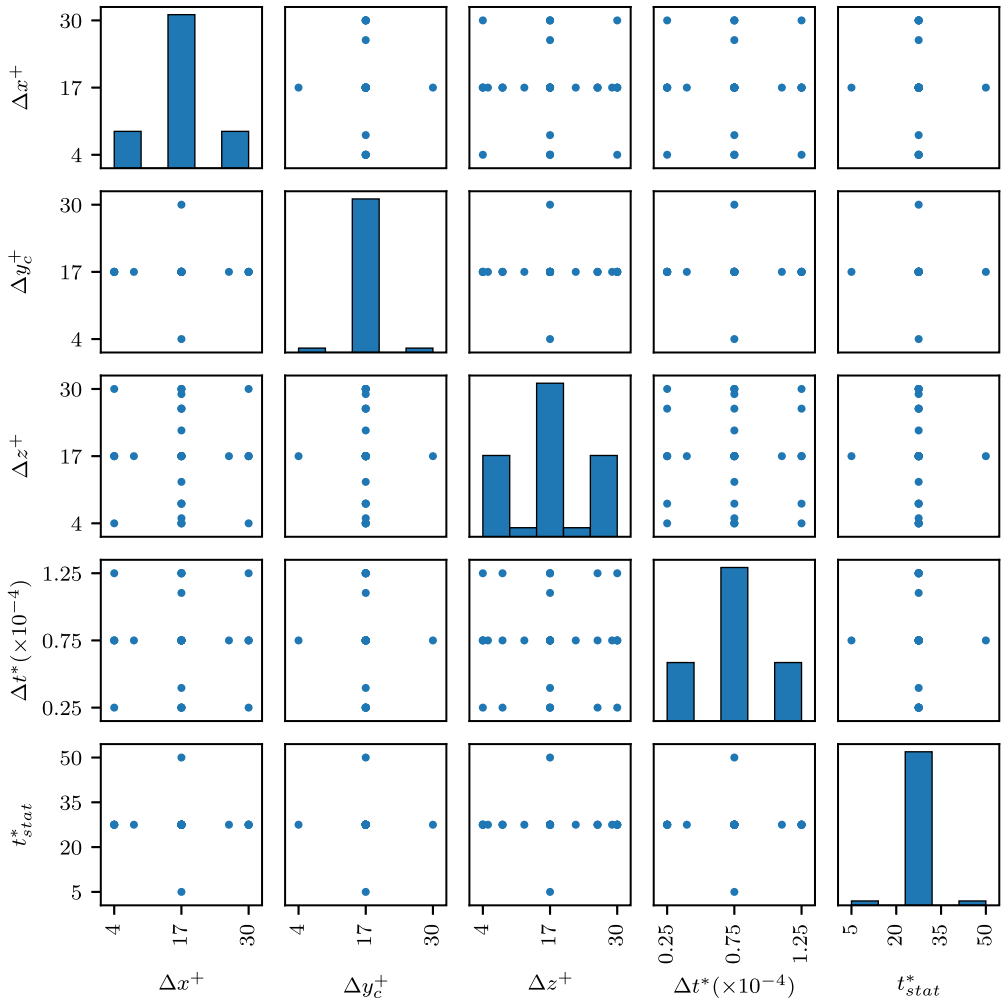
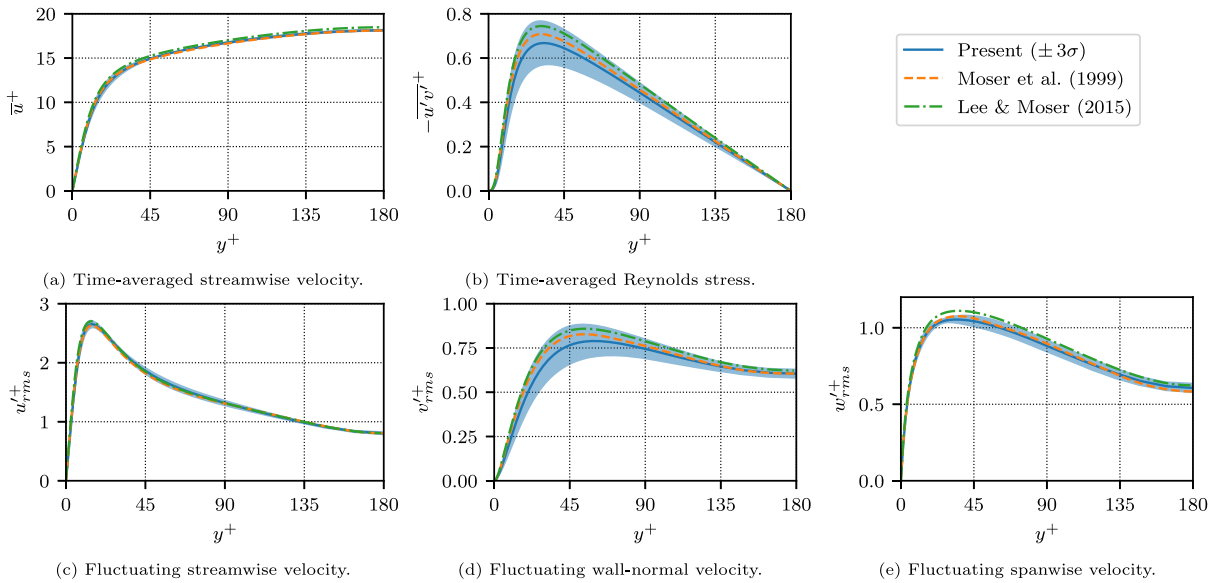


Fig. 13. 2D projection of uncertain input space and sample locations for each pair of inputs in Campaign III. The sample scheme is constructed by successively refining along important dimensions via nested anisotropic sparse grids of varying quadrature orders. Histograms for each input are shown along the diagonal. Note that the inner-scaling is with respect to the target friction velocity.

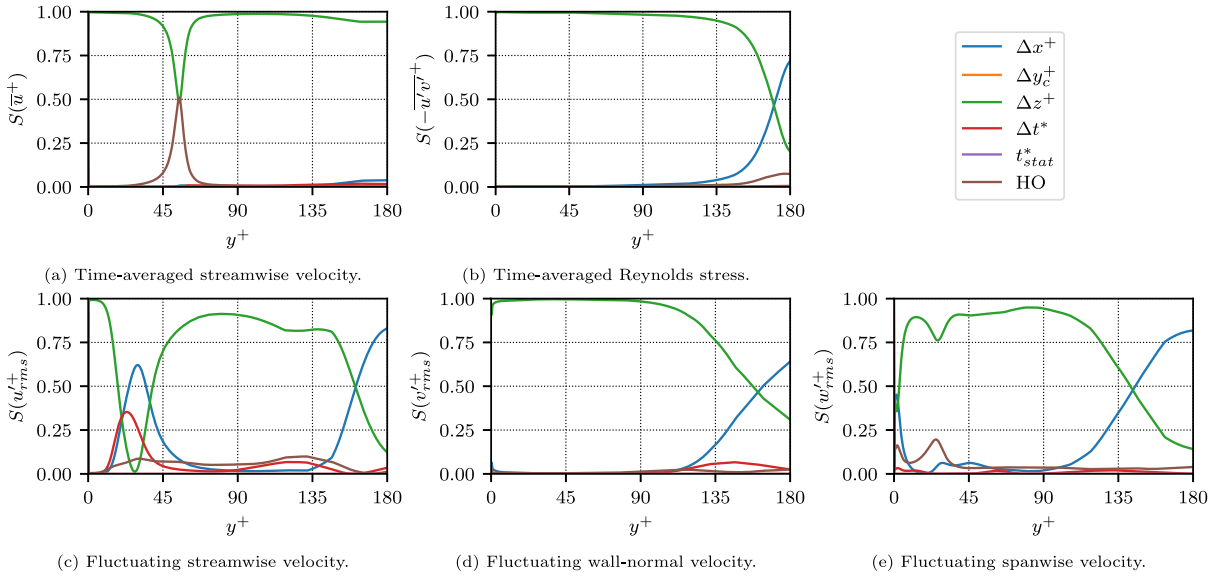
run for seven iterations, requiring 37 separate DNS, at which point the adaptation error has clearly plateaued. Compared to the full tensor grid approach in Campaign II, which required 3125 samples, this is a reduction of approximately two orders of magnitude. Furthermore, the actual computational cost was approximately 14,000 core-hours, which is just over thirty times smaller than the full tensor grid approach. That being said, this does not necessarily translate into equivalent savings in total execution time. This is because, while each sample in the tensor grid approach can be computed simultaneously, each iteration of the dimension-adaptive scheme is sequential (although the simulations within an iteration can be executed in parallel). Depending on available HPC resources, if the entire tensor grid campaign can be executed simultaneously, then this could lead to a faster total execution time compared to the dimension-adaptive approach. Having said that, this



(a) Time-averaged streamwise velocity. (b) Time-averaged Reynolds stress.

(c) Fluctuating streamwise velocity. (d) Fluctuating wall-normal velocity. (e) Fluctuating spanwise velocity.

Fig. 14. Cross-channel profiles of first and second-order velocity moments for Campaign III. Note that the inner-scaling is with respect to the target friction velocity.



(a) Time-averaged streamwise velocity. (b) Time-averaged Reynolds stress.

(c) Fluctuating streamwise velocity. (d) Fluctuating wall-normal velocity. (e) Fluctuating spanwise velocity.

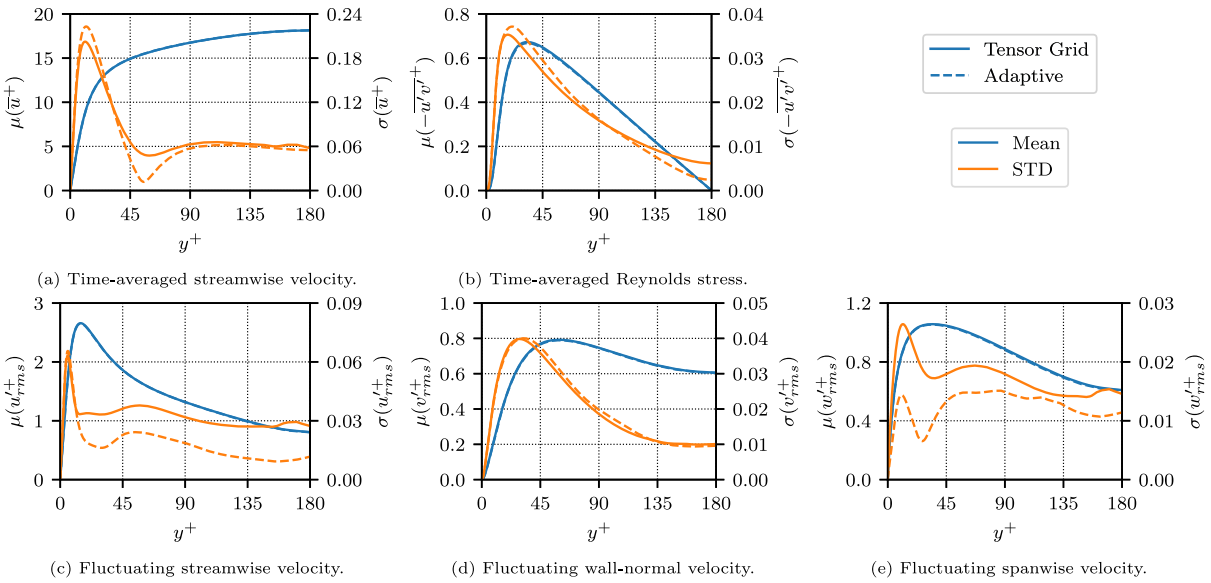
Fig. 15. Cross-channel Sobol indices of first and second-order velocity moments for Campaign III. HO refers to high-order cross-coupling terms. Note that the inner-scaling is with respect to the target friction velocity.

Table 7
Overview of Campaign III progression for each iteration. The mapping of quadrature order to each individual input is given by (Δx^+, Δy_c^+, Δz^+, Δt^*, t^*_{stat}).

Iteration	Simulations	Adaptation error	Quadrature order
0	1	–	(0, 0, 0, 0, 0)
1	10	5.12	(0, 0, 1, 0, 0)
2	2	0.91	(0, 0, 2, 0, 0)
3	4	0.82	(0, 0, 0, 1, 0)
4	6	1.03	(0, 0, 1, 1, 0)
5	4	0.57	(0, 0, 2, 1, 0)
6	0	0.43	(1, 0, 0, 0, 0)
7	10	1.05	(1, 0, 1, 0, 0)

is unlikely to be the case for large campaigns. Furthermore, the actual compute cost (e.g. in core-hours) would still be significantly higher than with the dimension-adaptive approach.

Fig. 12(a) shows the adaptation (surplus) error at each iteration. The error reduces quite rapidly after the first iteration and then settles for the remaining iterations, indicating convergence of the sampling scheme. Fig. 12(b) shows the quadrature orders accepted by the sampling scheme at each iteration. Clearly, the sampler immediately identifies the spanwise mesh resolution (Δz^+) as the dominant input and dedicates most of the resources throughout the campaign towards resolving its effect. This is in line with the findings from Section 4.2, which also identified Δz^+ as the most dominant input. During the course of the dimension-adaptive campaign, the sampler also dedicates some resources towards Δt^* and Δx^+, while completely neglecting Δy^+ and t^*_{stat}. Again, this is in line with the findings from Section 4.2, which showed that the contributions from the inputs other than Δz^+ are almost negligible. Finally, Fig. 13 provides an illustration of the multi-dimensional sample space by showing the 2D projection of each sample (simulation) on each pair of input dimensions. As can be seen, most of the resources are dedicated towards resolving the effect of Δz^+,



(a) Time-averaged streamwise velocity. (b) Time-averaged Reynolds stress.

(c) Fluctuating streamwise velocity. (d) Fluctuating wall-normal velocity. (e) Fluctuating spanwise velocity.

Fig. 16. Comparison of cross-channel profiles of first and second-order velocity moments for the tensor grid and dimension-adaptive sampling schemes. Note that the inner-scaling is with respect to the target friction velocity.

with some resources dedicated to Δx^+ and Δt^*, and no resources, other than the initial samples, dedicated to Δy^+ and t^*_{stat}.

Fig. 14 shows the statistical moments of the cross-channel profiles for various QOIs. The findings are very similar to Fig. 8, and support the use of the dimension-adaptive scheme, as an alternative to the full tensor grid approach, for providing the same type of analysis at a fraction of cost. Overall, there is a broad agreement with the benchmark data. Furthermore, like the previous campaigns, most of the variance manifests in the second-order velocity moments. This is especially true for the Reynolds stress and the wall-normal component of the fluctuating velocity. Fig. 15 shows the cross-channel Sobol indices for each uncertain input and QOI. Although there are clearly quantitative differences with respect to Fig. 9, the overall trends are very similar. In particular, Δz^+ is the dominant input across most of the channel height, especially close to the wall, whereas Δx^+ becomes more dominant towards the centre of the channel. In the tensor grid approach it was shown that high-order cross-coupling terms were typically responsible for the variance observed at the centre of the channel. However, this is not the case for the dimension-adaptive scheme. A possible explanation for this is that only one iteration (the final iteration) attempts to resolve the cross-coupling effect between Δx^+ and Δz^+. Therefore, it is possible that more iterations are required to properly capture this effect. It is also worth highlighting that it is the surplus error in the measured friction Reynolds number that directs the sampler at each iteration. Since the friction velocity is computed at the wall, then it is possible that the sampler is favouring choices that resolve the effects at the wall more than the effects towards the centre of the channel.

To properly assess the validity of the dimension-adaptive scheme, Fig. 16 provides a quantitative comparison of the cross-channel statistical moments (mean and variance), with respect to the tensor grid approach, for each QOI. For the mean profiles, the agreement is excellent and almost indistinguishable across all QOIs and wall-normal locations. On the other hand, the agreement in the variance is mixed and depends heavily on the QOI and wall-normal location. Specifically, the agreement is excellent for the Reynolds stress and wall-normal component of the fluctuating velocity. Furthermore, it is very good for the mean streamwise velocity, with some discrepancy in certain locations across the channel height. However, there are notable quantitative differences across most the channel height for the streamwise and spanwise components of the fluctuating velocity, although the overall trends are somewhat similar. The reason for these mixed results across

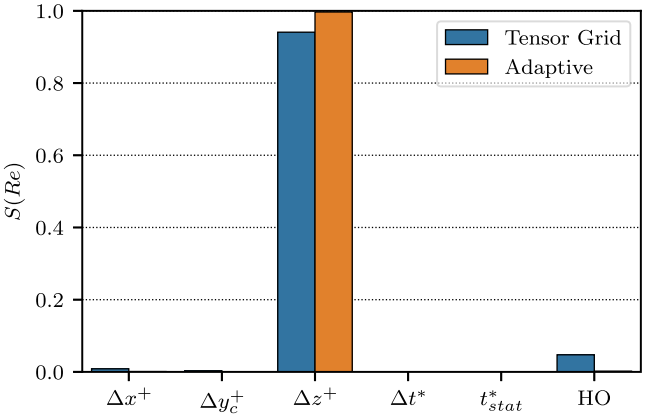


Fig. 17. Comparison of global sensitivity analysis of Re_τ for the tensor grid and dimension-adaptive sampling schemes.

different QOIs is an interesting finding, and is left for future work. Finally, Fig. 17 compares the Sobol indices for each uncertain input with respect to the measured Reynolds number. The values are very similar between the tensor grid and dimension-adaptive schemes and show that the sensitivity is dominated by Δz^+, while the effect from the remaining inputs is negligible.

5. Summary & conclusions

This work presents a systematic forward UQ study and sensitivity analysis for the DNS of a canonical turbulent channel flow at a friction Reynolds number of $Re_\tau = 180$. The DNS is computed using the high-order compact finite difference solver `Incompact3d`, whereas the UQ analysis is performed using stochastic collocation. The overall workflow is handled by EasyVVUQ, which automatically manages the preprocessing, execution and postprocessing of the simulation data on HPC systems. Three separate UQ campaigns are conducted. The first campaign investigates the effect of the domain size along the two periodic directions (L_x and L_z). The second campaign, on the other hand, investigates the effect of various numerical parameters, including the mesh resolution in each direction (Δx^+, Δy^+ and Δz^+), as well as the time step size (Δt^*) and the length of time the statistics

are collected for (t^*_{stat}). For the first two campaigns, the input space is sampled by constructing a tensor grid of fourth-order 1D Gauss–Legendre quadrature points, resulting in a total of 25 and 3125 separate DNS for the first and second campaigns, respectively. Finally, the third campaign reproduces the second campaign, but adopts a dimension-adaptive sparse-grid sampling strategy to drastically reduce the number of samples (simulations) required.

The results for the first campaign show that the variance associated with the domain size mainly manifests in the second-order velocity moments, especially the streamwise and spanwise components of the fluctuating velocity. Furthermore, the sensitivity to a given input is highly dependent on the particular QOI and wall-normal location. For the second campaign, again, the variance is mainly manifested in the second-order velocity moments. However, this time it is the Reynolds stress and the wall-normal component of the fluctuating velocity that exhibit the most variance. This sensitivity is found to be mostly affected by the spanwise mesh resolution, especially in the region close to the wall. However, some variation in the relative influence of each input can be observed across the channel height, depending on the particular QOI.

The results for the third campaign show that it is possible to reproduce the analysis from the second campaign using a 'smart' sampling strategy that significantly reduces the number of required samples. By adopting a dimension-adaptive approach that automatically determines the most influential inputs and prioritises resources towards resolving their effect, the number of required samples (simulations) was reduced by two orders of magnitude, leading to significant computational savings. A comparison of the statistical moments and sensitivity indices show good agreement between the tensor grid and dimension-adaptive approaches, with some slight quantitative discrepancies in the variance, depending on the QOI and wall-normal location. This demonstrates the potential of the dimension-adaptive approach as an affordable alternative to full tensor grid sampling that can be readily incorporated into CFD workflows, which will ultimately enable more UQ studies of scale-resolving simulations. However, it is important to highlight that this approach would not provide any benefit over the tensor grid approach for problems where the influence of each input parameter is approximately equal. Nevertheless, this is rarely the case in practice.

The restriction to a single Reynolds number is one of the main limitations of the present study. Nevertheless, it was a necessary one in order to ensure that the objectives of this work were computationally feasible. Furthermore, previous works have shown that propagated uncertainties and sensitivities are approximately invariant over small-moderate Reynolds numbers for these types of flows [21]. Therefore, it is expected that the conclusions of this work can be extrapolated to a (limited) range of flow regimes beyond what is considered here. However, further work is required to verify this assumption. A second limitation of this work is the absence of a rigorous and systematic analysis of the performance of the dimension-adaptive scheme. For example, in Section 4.3, the campaign is stopped after seven iterations based on the behaviour of the adaptation error. While this number of iterations is shown to provide good agreement with the full tensor grid approach for this setup, a more detailed investigation into the effects of the stopping criteria, the choice of QOI used to guide the adaptive sampling, and the overall convergence behaviour, among others, would be useful. However, since the purpose of this work is to demonstrate the potential of the dimension-adaptive approach for scale-resolving simulations of turbulent flows, this type of analysis is beyond the scope of the present work, as it would require its own separate study.

There are several potential directions for future research. Firstly, the logical next step following this work is to combine the input spaces of the first and second campaigns. This would allow any cross-coupling effects between the domain size and numerical parameters to be properly captured. This type of analysis would only be feasible using the dimension-adaptive scheme, since the full tensor grid approach would require $5^7 = 78125$ separate DNS. Secondly, extending this work to different flow problems – such as different geometries (e.g. turbulent boundary layer) and flow conditions (e.g. Reynolds number) – would provide further insights into the UQ and sensitivity analysis of DNS. Finally, it would also be informative to investigate and compare other dimensionality-reducing strategies for UQ, such as the deep active-subspace method [49] (also available in EasyVVUQ), which would further enable high-dimensional UQ studies of scale-resolving CFD.

CRediT authorship contribution statement

Joseph O'Connor: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Sylvain Laizet:** Conceptualization, Software, Validation, Formal analysis, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Andrew Wynn:** Formal analysis, Writing – review & editing, Supervision, Funding acquisition. **Wouter Edeling:** Methodology, Software, Validation, Formal analysis, Writing – review & editing. **Peter V. Coveney:** Conceptualization, Formal analysis, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data were obtained with the open-source codes Incompact3d and EasyVVUQ. Also, the data and scripts to reproduce the figures are available at https://doi.org/10.5281/zenodo.8038943.

Acknowledgements

The authors would like to thank the Engineering and Physical Sciences Research Council (EPSRC), United Kingdom for funding this work under grant numbers EP/T021144/1 and EP/W007711/1. The authors would also like to acknowledge the use of the Cirrus UK National Tier-2 HPC Service at EPCC (https://www.cirrus.ac.uk), funded by the University of Edinburgh and EPSRC, United Kingdom (EP/P020267/1), and the ARCHER2 UK National Supercomputing Service (https://www.archer2.ac.uk).

References

[1] Duraisamy K, Iaccarino G, Xiao H. Turbulence modeling in the age of data. Annu Rev Fluid Mech 2019;51(1):357–77.
[2] Moin P, Mahesh K. Direct numerical simulation: a tool in turbulence research. Annu Rev Fluid Mech 1998;30(1):539–78.
[3] Lee M, Moser RD. Direct numerical simulation of turbulent channel flow up to $Re_\tau \approx 5200$. J Fluid Mech 2015;774:395–415.
[4] Xiao H, Cinnella P. Quantification of model uncertainty in RANS simulations: A review. Prog Aerosp Sci 2019;108:1–31.
[5] Najm HN. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. Annu Rev Fluid Mech 2009;41(1):35–52.
[6] Roy CJ, Oberkampf WL. A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. Comput Methods Appl Mech Engrg 2011;200(25):2131–44.
[7] Coveney PV, Groen D, Hoekstra AG. Reliability and reproducibility in computational science: Implementing validation, verification and uncertainty quantification in silico. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 2021;379(2197):20200409.
[8] Richardson RA, Wright DW, Edeling W, Jancauskas V, Lakhlili J, Coveney PV. EasyVVUQ: a library for verification, validation and uncertainty quantification in high performance computing. J Open Res Softw 2020;8(1):11.
[9] Rezaeiravesh S, Vinuesa R, Schlatter P. Uqit: a python package for uncertainty quantification (UQ) in computational fluid dynamics (CFD). J Open Source Softw 2021;6(60):2871.

[10] Lucor D, Meyers J, Sagaut P. Sensitivity analysis of large-eddy simulations to subgrid-scale-model parametric uncertainty using polynomial chaos. J Fluid Mech 2007;585:255–79.

[11] Ko J, Lucor D, Sagaut P. Sensitivity of two-dimensional spatially developing mixing layers with respect to uncertain inflow conditions. Phys Fluids 2008;20(7):077102.

[12] Meldi M, Salvetti MV, Sagaut P. Quantification of errors in large-eddy simulations of a spatially evolving mixing layer using polynomial chaos. Phys Fluids 2012;24(3):035101.

[13] Khalil M, Lacaze G, Oefelein JC, Najm HN. Uncertainty quantification in LES of a turbulent bluff-body stabilized flame. Proc Combust Inst 2015;35(2):1147–56.

[14] Mariotti A, Siconolfi L, Salvetti MV. Stochastic sensitivity analysis of large-eddy simulation predictions of the flow around a 5:1 rectangular cylinder. Eur J Mech B 2017;62:149–65.

[15] Congedo PM, Duprat C, Balarac G, Corre C. Numerical prediction of turbulent flows using Reynolds-averaged Navier–Stokes and large-eddy simulation with uncertain inflow conditions. Internat J Numer Methods Fluids 2013;72(3):341–58.

[16] Meyers J, Sagaut P. Is plane-channel flow a friendly case for the testing of large-eddy simulation subgrid-scale models? Phys Fluids 2007;19(4):048105.

[17] Oliver TA, Malaya N, Ulerich R, Moser RD. Estimating uncertainties in statistics computed from direct numerical simulation. Phys Fluids 2014;26(3):035101.

[18] Safta C, Blaylock M, Templeton J, Domino S, Sargsyan K, Najm H. Uncertainty quantification in LES of channel flow. Internat J Numer Methods Fluids 2017;83(4):376–401.

[19] Rezaeiravesh S, Liefvendahl M. Effect of grid resolution on large eddy simulation of wall-bounded turbulence. Phys Fluids 2018;30(5):055106.

[20] Rezaeiravesh S, Mukha T, Liefvendahl M. Systematic study of accuracy of wall-modeled large eddy simulation using uncertainty quantification techniques. Comput & Fluids 2019;185:34–58.

[21] Rezaeiravesh S, Vinuesa R, Schlatter P. On numerical uncertainties in scale-resolving simulations of canonical wall turbulence. Comput & Fluids 2021;227:105024.

[22] Laizet S, Lamballais E. High-order compact schemes for incompressible flows: A simple and efficient method with quasi-spectral accuracy. J Comput Phys 2009;228(16):5989–6015.

[23] Bartholomew P, Deskos G, Frantz RAS, Schuch FN, Lamballais E, Laizet S. Xcompact3D: an open-source framework for solving turbulence problems on a Cartesian mesh. SoftwareX 2020;12:100550.

[24] Kravchenko AG, Moin P. On the effect of numerical errors in large eddy simulations of turbulent flows. J Comput Phys 1997;131(2):310–22.

[25] Lele SK. Compact finite difference schemes with spectral-like resolution. J Comput Phys 1992;103(1):16–42.

[26] Laizet S, Li N. Incompact3d: A powerful tool to tackle turbulence problems with up to $O(10^5)$ computational cores. Internat J Numer Methods Fluids 2011;67(11):1735–57.

[27] Diaz-Daniel C, Laizet S, Vassilicos JC. Wall shear stress fluctuations: Mixed scaling and their effects on velocity fluctuations in a turbulent boundary layer. Phys Fluids 2017;29(5):055102.

[28] Mahfoze O, Laizet S. Skin-friction drag reduction in a channel flow with streamwise-aligned plasma actuators. Int J Heat Fluid Flow 2017;66:83–94.

[29] Deskos G, Laizet S, Piggott MD. Turbulence-resolving simulations of wind turbine wakes. Renew Energy 2019;134:989–1002.

[30] Eldred M, Burkardt J. Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification. In: 47th AIAA aerosp. sci. meet. new horiz. forum aerosp. expo.. Aerospace sciences meetings, American Institute of Aeronautics and Astronautics; 2009.

[31] Sudret B. Global sensitivity analysis using polynomial chaos expansions. Reliab Eng Syst Saf 2008;93(7):964–79.

[32] Luk OO, Lakhlili J, Hoenen O, von Toussaint U, Scott BD, Coster DP. Towards validated multiscale simulations for fusion. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 2021;379(2197):20200074.

[33] Suleimenova D, Arabnejad H, Edeling WN, Groen D. Sensitivity-driven simulation development: A case study in forced migration. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 2021;379(2197):20200077.

[34] Edeling W, Arabnejad H, Sinclair R, Suleimenova D, Gopalakrishnan K, Bosak B, Groen D, Mahmood I, Crommelin D, Coveney PV. The impact of uncertainty on predictions of the CovidSim epidemiological code. Nat Comput Sci 2021;1(2):128–35.

[35] Groen D, Arabnejad H, Jancauskas V, Edeling WN, Jansson F, Richardson RA, Lakhlili J, Veen L, Bosak B, Kopta P, Wright DW, Monnier N, Karlshoefer P, Suleimenova D, Sinclair R, Vassaux M, Nikishova A, Bieniek M, Luk OO, Kulczewski M, Raffin E, Crommelin D, Hoenen O, Coster DP, Piontek T, Coveney PV. Vecmatk: A scalable verification, validation and uncertainty quantification toolkit for scientific simulations. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 2021;379(2197):20200221.

[36] Groen D, Arabnejad H, Suleimenova D, Edeling W, Raffin E, Xue Y, Bronik K, Monnier N, Coveney PV. FabSim3: an automation toolkit for verified simulations using high performance computing. Comput Phys Comm 2023;283:108596.

[37] Sobol' IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simulation 2001;55(1):271–80.

[38] Gerstner T, Griebel M. Numerical integration using sparse grids. Numer Algorithms 1998;18(3):209–32.

[39] Tang G, Iaccarino G, Eldred M. Global sensitivity analysis for stochastic collocation. In: 51st AIAA/ASME/ASCE/AHS/ASC Struct. struct. dyn. mater. conf.. Structures, structural dynamics, and materials and co-located conferences, American Institute of Aeronautics and Astronautics; 2010.

[40] Edeling W. Adaptive sparse-grid tutorial. 2022, https://ir.cwi.nl/pub/31588.

[41] Gerstner T, Griebel M. Dimension–adaptive tensor–product quadrature. Computing 2003;71(1):65–87.

[42] Buzzard GT. Global sensitivity analysis using sparse grid interpolation and polynomial chaos. Reliab Eng Syst Saf 2012;107:82–9.

[43] Jakeman JD, Eldred MS, Geraci G, Gorodetsky A. Adaptive multi-index collocation for uncertainty quantification and sensitivity analysis. Internat J Numer Methods Engrg 2020;121(6):1314–43.

[44] Trefethen LN. Is Gauss quadrature better than Clenshaw–Curtis? SIAM Rev 2008;50(1):67–87.

[45] Kim J, Moin P, Moser R. Turbulence statistics in fully developed channel flow at low Reynolds number. J Fluid Mech 1987;177:133–66.

[46] Jansson F, Edeling W, Attema J, Crommelin D. Assessing uncertainties from physical parameters and modelling choices in an atmospheric large eddy simulation model. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 2021;379(2197):20200073.

[47] Moser RD, Kim J, Mansour NN. Direct numerical simulation of turbulent channel flow up to Re_τ=590. Phys Fluids 1999;11(4):943–5.

[48] Vreman AW, Kuerten JGM. Comparison of direct numerical simulation databases of turbulent channel flow at Re_τ = 180. Phys Fluids 2014;26(1):015102.

[49] Edeling W. On the deep active-subspace method. SIAMASA J Uncertain Quantif 2023;62–90.