Application of recurrence CFD (rCFD) to species transport in turbulent vortex shedding

S. Abbasi\textsuperscript{a,b,c}, S. Pirker\textsuperscript{a,b}, T. Lichtenegger\textsuperscript{a,h,*}

\textsuperscript{a} Department of Particulate Flow Modelling, Johannes Kepler University, Linz, Austria
\textsuperscript{b} Linz Institute of Technology (LIT), Johannes Kepler University, Linz, Austria
\textsuperscript{c} K1-MET Metallurgical Competence Center, Linz, Austria

\textbf{A R T I C L E   I N F O}

Article history:
Received 21 March 2019
Revised 24 September 2019
Accepted 23 October 2019
Available online 24 October 2019

Keywords:
Large eddy simulations
Recurrence statistics
Multi-scale problems
Time-extrapolation
Recurrence CFD
Data-assisted simulation

\textbf{A B S T R A C T}

The functionality of computational fluid dynamics (CFD) for turbulent flows is limited by huge computational demands which prevent any detailed long-term studies. In this publication, we apply the recently introduced, data-assisted method “recurrence CFD” (rCFD) to turbulent vortex shedding after a circular cylinder at Reynolds number $Re = 3900$. Using a database of flow fields from short, conventional simulations, we time-extrapolate their behavior to arbitrary durations and obtain promising results for passive species transport with speed-up factors of more than 120 at 1/20 of the required computer power compared to the underlying large eddy simulation (LES). Besides this massive run-time reduction, we focus on data efficiency. For cases with strong, spatial scale separation, rCFD’s resilience towards grid coarsening allows us to carry out calculations at lower mesh resolution provided one retains meso-scale velocity fluctuations as a contribution to diffusivity. This reduces database size which would otherwise become a bottleneck in the methodology.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

The drastic increase of available computer power has extended the scope of applicability of CFD from simple two-dimensional single-phase flows to complex, transient, three-dimensional multi-phase flows even under highly turbulent conditions. To solve such problems, there are several well-established techniques \cite{1} including direct numerical simulation (DNS), large eddy simulations (LES) and Reynolds-averaged Navier-Stokes simulations (RANS). Performing DNS is highly expensive since they solve the Navier-Stokes equations directly in the entire domain and resolve the whole temporal and spatial scales of the flow down to the Kolmogorov length \cite{2–4}. The computational cost can be reduced by solving just the large and intermediate scales and modeling the influence of the smaller ones, which is done in LES \cite{5}. However, still the essential range of time-scales and length-scales for LES limits their use to low and moderate Reynolds numbers. RANS models require lower spatial and temporal resolution, however they bring an additional loss of information due to statistical averaging \cite{6}.

Put differently, with increasing cell size, the employed sub-grid corrections need to capture more of the ongoing physics on smaller scales by modeling the effect of the corresponding degrees of freedom without any explicit calculation. An alternative approach to eliminate inessential degrees of freedom was inspired by the experimental observation of coherent structures in turbulent flows \cite{7–9}. These are characteristic spatial features which reappear in an irregular fashion but exhibit a distinctive temporal evolution. A huge body of research emerged trying to extract such structures and obtain their equations of motion using techniques like proper orthogonal decomposition (POD) or dynamic mode decomposition (DMD) (cf. the reviews \cite{10,11}).

While the high computational costs per time step and their large number prevent long-term investigations of transient problems using LES or unsteady RANS despite intense research activity \cite{12–17}, reduced-order Galerkin models built from POD modes are numerically much cheaper but suffer from long-time \cite{18–25} and other types of instabilities \cite{26–31} that require additional handling. In this work, we want to apply the recent approach rCFD \cite{32} to turbulent problems to overcome the limitation of costly computations. Based on the characteristic temporal life-cycle of dominant flow structures, rCFD aims to time-extrapolate the phase-space tra-
Nomenclature

Greek letters
\[ \Delta \] filter width
\[ \Delta t \] time step
\[ \theta \] angle
\[ \nu \] viscosity
\[ \rho \] density
\[ \tau \] time range
\[ \tau^* \] residual scalar flux
\[ \tau^{*\prime} \] residual stress tensor
\[ \Omega \] vorticity

Latin symbols
\[ C \] species, constant, coefficient
\[ D \] diameter, distance norm/matrix
\[ D' \] diffusivity
\[ E \] mean interval steps
\[ F/G \] filter function
\[ F \] force
\[ f \] frequency
\[ k \] residual kinetic energy
\[ L \] length
\[ N' \] distance norm’s normalization constant
\[ P \] jump probability
\[ p \] pressure
\[ R \] recurrence norm/matrix
\[ Re \] Reynolds number
\[ r \] radius
\[ r/x \] position
\[ S \] rate of strain tensor
\[ Sc \] Schmidt number
\[ St \] Strouhal number
\[ S \] source term
\[ t \] time
\[ U_\infty \] free stream velocity
\[ \mathbf{u} \] velocity
\[ x \] stream-wise direction
\[ y \] transverse direction
\[ z \] span-wise direction

Subscripts/Superscripts
\[ b/e \] begin/end
\[ \text{bin.} \] binary
\[ \text{c} \] cells
\[ D/L \] drag/lift
\[ F/G \] filter
\[ i \] interval i
\[ \text{rec} \] recurrence
\[ \text{rms} \] root mean square
\[ \text{sep} \] separation
\[ \text{sgs} \] sub-grid scale
\[ \text{sim} \] similar state
\[ \text{vs} \] vortex scale
\[ \infty \] free stream
\[ \text{–} \] filtered
\[ (\) \] averaged

In the context of systems exhibiting recurring patterns, e.g., vortices behind blunt bodies, according to recurrence plots introduced by Eckmann et al. [33] and later formalized in the recurrence quantification analysis (RQA) [34,35]. Using fields from an initial, short, detailed simulation (regardless if DNS, LES or RANS), the long-time evolution is approximated assuming that similar configurations will show similar short-term behavior. Essentially, a large number of flow field sequences are chained in a smooth fashion, which amounts in an iterated method of analogues [36].

Although the fields can have the same high level of spatial resolution as the underlying simulation, time-extrapolation is extremely cheap. Not even a set of model equations of motion needs to be solved. At the same time, no stability problems occur since only information from the reliable, initial simulation is employed. These properties make the method very interesting for long-time studies of passive or weakly coupled processes on highly dynamic backgrounds, e.g., species transport in a turbulent flow.

Here, to examine the proper functionality of rCFD in turbulent flows, we want to answer two questions: a) Is it possible to capture turbulent long-time transport behavior based on small initial data sets? b) Without the need to solve the fluid momentum equation and to take care of pressure-velocity coupling, how much can we reduce temporal and spatial resolution and still get accurate results with respect to species distribution? To this end, we test our methodology on the well-studied case of vortex shedding behind a circular cylinder with available experimental data to validate our detailed simulation results.

We first review the fluid equations of motion and describe the rCFD procedure in Section 2. Next, Section 3 illustrates main parameters and the setup of the full and the rCFD simulations. Then, we show data validation of the LES calculation, investigate the flow’s recurrent behavior, present simulation results of species transport using rCFD and carry out a performance analysis in Section 4. The final Section 5 points to conclusions of the work and gives an outlook on future activities.

2. Theoretical background

Here, we briefly review the set of equations governing a single-phase, incompressible flow from the LES perspective and describe the rCFD approach. While rCFD can be applied under compressible conditions in a straight-forward fashion, we restrict ourselves to the incompressible, divergence-free case in the present work for the sake of simplicity.

2.1. Fluid equations of motion

The system of equations considered for an incompressible, single-phase fluid along with passive transport of a scalar concentration \( C \) are

\[
\nabla \cdot \mathbf{u} = 0 \tag{1}
\]

\[
\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho} \nabla p + \nabla \cdot \nu \nabla \mathbf{u} \tag{2}
\]

\[
\frac{\partial C}{\partial t} + (\mathbf{u} \cdot \nabla)C = \nabla \cdot D \nabla C + S \tag{3}
\]

where diffusivity

\[
D = \frac{\nu}{Sc} \tag{4}
\]

is connected to viscosity \( \nu \) via the Schmidt number \( Sc \), and \( S \) is a constant source term. As mentioned in Section 1, LES is one of the methods to numerically simplify the Navier-Stokes equations through a filtering operation

\[
\tilde{u}(x, t) = \int d\mathbf{r} \mathbf{G}(\mathbf{r}, x) \mathbf{u}(\mathbf{x} - \mathbf{r}, t). \tag{5}
\]

Applying the kernel \( \mathbf{G}(\mathbf{r}, x) \) to Eqs. (1) – (3), the set of filtered equations can be written as

\[
\nabla \cdot \tilde{\mathbf{u}} = 0 \tag{6}
\]
\[ \frac{\partial \hat{u}}{\partial t} + (\hat{u} \cdot \nabla) \hat{u} = -\frac{1}{\rho} \nabla \hat{p} + \nabla \cdot \nu \nabla \hat{u} - \nabla \cdot \mathbf{r} \]  
(7)

\[ \frac{\partial \hat{C}}{\partial t} + (\hat{u} \cdot \nabla) \hat{C} = \nabla \cdot D \nabla \hat{C} - \nabla \cdot \mathbf{r}^e + \hat{S} \]  
(8)

where the residual stress tensor \( \mathbf{r}^e = \mathbf{U} - \mathbf{u} \) and the residual scalar flux \( \mathbf{r}^e = \mathbf{U} - \mathbf{u} \) require closure \([1]\). Introducing the sub-grid scale viscosity \( \nu_{sgs} \), the Boussinesq approach postulates the simple relationship

\[ \mathbf{r}^e = -2\nu_{sgs} \hat{S} \]  
(9)

between the residual stresses and the filtered rate-of-strain tensor

\[ \hat{S} \equiv \frac{1}{2} \left( \nabla \hat{u} + (\nabla \hat{u})^T \right) . \]  
(10)

Based on dimensional arguments, one can make an ansatz

\[ \nu_{sgs} = C \nu \sqrt{\Delta} \]  
(11)

in terms of the residual kinetic energy \( k \equiv \frac{1}{2}(\mathbf{U} \cdot \mathbf{u} - \mathbf{u} \cdot \mathbf{u}) \), the filter width \( \Delta \) which is the cube root of cell volume and a model coefficient \( C \). Among the large number of models to obtain \( k \) from available quantities, we chose a simple, algebraic approach following Smagorinsky \([37]\), which is based on the assumption of local equilibrium between the production and dissipation of turbulent kinetic energy and leads to \([1]\)

\[ k = \frac{2C_k}{C_t} \Delta^2 \hat{S} : \hat{S} \]  
(12)

\[ C_k = 0.094 \]  
(13)

\[ C_t = 1.048. \]  
(14)

Analogous to the residual stress tensor, we can model the residual scalar flux with the gradient diffusion hypothesis

\[ \mathbf{r}^e = D_{sgs} \nabla \hat{C} . \]  
(15)

Although most generally, sub-grid diffusivity can show some degree of anisotropy and hence \( D_{sgs} \) is a tensor \([38–41]\), recent studies \([42,43]\) have pointed out that LES predict transport mechanisms with a scalar sub-grid diffusivity \( D_{sgs} = \nu_{sgs}/Sc_{sgs} \) accurately, whereas RANS require an anisotropic treatment.

We stress that the methodology presented in the following sections is completely independent of the specific choice of \( D_{sgs} \) and, more generally, may be applied with any turbulence model to generate the velocity fields for the database.

### 2.2. Recurrence-based time extrapolation

The main objective of rCFD is time-extrapolation of a system's behavior based on information from a short-time, detailed simulation while maintaining a high degree of accuracy. Therefore, to start with rCFD, it is essential to record a sufficient amount of data comprising the required flow fields over a certain time range \( \tau_{rec} \) during which the fields are stored with a specified sampling time interval \( \Delta t_{rec} \). The latter needs to be chosen such that the flow does not change considerably within \( \Delta t_{rec} \) in order to solve the transport equation. \( \tau_{rec} \) is rarely known a priori, e.g. if the flow is driven by a single dominant frequency. Otherwise, it needs to be determined from the recurrence statistics (see below).

#### 2.2.1. Recurrence statistics

To quantify a system's degree of recurrence, a distance norm \([32]\), e.g.

\[ D(t, t') \equiv \frac{1}{N} \int d^3 r (u(r, t) - u(r, t'))^2 \]  
(16)

is normalized by

\[ N \equiv \max t_{t}, \int d^3 r (u(r, t) - u(r, t'))^2 \]  
(17)

which compares velocity fields at two times at all \( N = \tau_{rec}/\Delta t_{rec} \) times stored in the database giving rise to the distance matrix

\[ D_{m,n} = D(m\Delta t_{rec}, n\Delta t_{rec}) . \]  
(18)

The lower values in the distance matrix represent the similarity between two states, and vice versa. Moreover, one can obtain various characteristic numbers of the system from the distance matrix, e.g. the mean nearest-neighbor distance

\[ D_{sim} \equiv \frac{1}{N} \sum_m \min_n D_{m,n} . \]  
(19)

The saturation or rather the appearance of plateaus of \( D_{sim} \) or of RQA parameters with database size can be used to estimate sensible values of \( \tau_{rec} \).

Alternatively, Eqs. (16) and (18) may be reformulated in terms of a recurrence norm and matrix,

\[ R(t, t') \equiv 1 - D(t, t') \]  
(20)

\[ R_{m,n} \equiv 1 - D_{m,n} . \]  
(21)

respectively. The binary version of Eq. (21), \( R_{m,n}^{(\text{bin})}(\epsilon) \equiv \Theta(\epsilon - D_{m,n}) \), has been used extensively in literature. Among various target systems, recurrence plots and statistics have also been employed to analyze turbulent flows \([44–50]\).

In the following section, we show how one can utilize them to approximate their temporal evolution.

#### 2.2.2. Recurrence path

Time evolution of a recurrent flow is obtained from a recurrence path on the recurrence statistics. As explained in previous work \([32]\), one starts at a given begin time with the corresponding original flow data. After generating a time interval of random length (see below for details) with consecutive fields, the state of the interval end is identified with the most similar past one in the distance matrix. Likewise, after a certain number of steps in the database from that state on, the flow field is updated based on the minimum value of distance norm in the distance matrix for the end of the second interval and so on.

For some cases, it is advantageous to generate a recurrence path from two or more separate databases, e.g. because of various possible process states \([51,52]\) or to enforce certain constraints, which we model as a Markov process. For example, the reflection symmetry of time-averaged vortex shedding in geometries with the same symmetry is not fully recovered from a limited number of pseudo-periods due to turbulent fluctuations. However, based on this knowledge, we can record a database \( D^{(1)} \) for a small number of cycles and guarantee symmetry of the final results by extending it with its mirrored counterpart \( D^{(2)} \) to give \( D = D^{(1)} \cup D^{(2)} \). Compared to the case of a single database, the only difference is that the continuous time intervals cannot range from one into another. One can set a "jump probability" \( P \) to obtain a desired mean interval length \( E(\Delta t_{rec}) = (1 - P)/P \) well below the database size, e.g. 1/4 of it, but still large enough to take intervals of multiple consecutive steps. We go from one state to that \( \Delta t_{rec} \) later
with probability \( 1 - P \) unless it corresponds to the end of the current database \( D^{(m)} \). Otherwise, a jump in the recurrence statistics is performed. First, based on their predefined, relative weights, either the same or a new database \( D^{(n)} \) is selected and then, the most similar flow configuration is looked up in the database. Ultimately, this procedure which is depicted in Figs. 1 and 2 gives a sequence of intervals \([t_i^{(b)}, t_i^{(e)}] = t_i^{(b)} + \Delta t_i\), \([t_i^{(b)}, t_{i+1}^{(b)}] = t_i^{(b)} + \Delta t_{i+1}, \ldots \), with mean length \( E(\Delta t_i / \Delta t_{\text{rec}}) \). According to the recurrence path generated from these intervals, the corresponding original flow data is used to go on with the process resulting in time series of the form

\[
\mathbf{u}(\mathbf{r}, t_i^{(b)}), \mathbf{u}(\mathbf{r}, t_i^{(b)} + \Delta t_{\text{rec}}) \ldots ,
\]

\[
\mathbf{u}(\mathbf{r}, t_i^{(e)}), \mathbf{u}(\mathbf{r}, t_{i+1}^{(b)}), \mathbf{u}(\mathbf{r}, t_{i+1}^{(b)} + \Delta t_{\text{rec}}) \ldots .
\]

(22)

### 2.2.3. Recurrence CFD

After acquiring the recurrence path, we can study any passive process like species transport by solving only a single convection-diffusion equation

\[
\frac{\partial \mathbf{c}}{\partial t} + (\mathbf{u}_{\text{rec}} \cdot \nabla) \mathbf{c} = \nabla \cdot \mathbf{D}_{\text{rec}} \mathbf{c} + S
\]

(23)

with the information recorded in the database (\( \mathbf{u}_{\text{rec}} \) and \( \mathbf{D}_{\text{rec}} \)). The diffusivity \( \mathbf{D}_{\text{rec}} \) contains contributions from the molecular and the sub-grid scale diffusivities \( \mathbf{D} \) and \( \mathbf{D}_{\text{sgs}} \). If the mesh used for rCFD is the same as the one built for LES, one can write

\[
\mathbf{D}_{\text{rec}} = \mathbf{D} + \mathbf{D}_{\text{sgs}},
\]

(24)

and directly save and use \( \mathbf{D}_{\text{sgs}} \) in the database. However, since here only species transport is of importance and there is no need to solve the momentum and pressure equations, we may use a much coarser grid onto which we map the recorded flow fields before solving Eq. (23). While computationally much cheaper, the transition from the fine LES to the coarse rCFD mesh introduces a new closure problem. We need not only include sub-grid scale fluctuations already taken care of by \( \mathbf{D}_{\text{sgs}} \) but also retain fine-grid scale ones due to the second filtering operation.

There are at least two strategies to obtain \( \mathbf{D}_{\text{sgs}} \) for the coarser grid. In the paper, we chose to use an LES closure model like

\[
\text{Eqs. (11) and (15) on the new mesh. Following Germano [53,54], “given two filters F and G, the residual stress tensor at FG level,}
\]

\[
\tau_{FG}(u_i, u_j) = \langle \bar{u}_i \bar{u}_j \rangle_{FG} - \langle \bar{u}_i \rangle_{FG} \langle \bar{u}_j \rangle_{FG} ,
\]

\[
\text{is equal to the G-averaged value of the residual stress at the F level plus the resolved turbulent stress extracted from the resolved scale F,}
\]

\[
\tau_{FG}(u_i, u_j) = \langle \bar{u}_i \rangle_{FG} \langle \bar{u}_j \rangle_{FG} + \tau_C \langle \bar{u}_i \rangle_{FG} \langle \bar{u}_j \rangle_{FG}.
\]

(25)

If we consider our two filters as the LES one and the grid coarsening (C) for rCFD, the residual kinetic energy of the coarser mesh can be computed as

\[
k_c = \left( k + \frac{1}{2} \mathbf{u} \cdot \mathbf{u} \right)_{FG} - \frac{1}{2} \mathbf{u} \cdot \mathbf{u} \right)_{FG} - \frac{1}{2} \mathbf{u} \cdot \mathbf{u} \right)_{FG}
\]

(26)

in terms of the LES sub-grid kinetic energy \( k \) and the fine grid velocity \( \mathbf{u} \). We can use an LES model and calculate the sub-grid diffusivity for the coarsened mesh based on the sub- and fine-grid
kinetic energy $k$ provided by the recurrence database, e.g.

$$D_{SGS} = \frac{C_k \sqrt{\nu \Delta}}{S_{CGS}}.$$  

(27)

Clearly, this method eventually becomes unreliable for very coarse grids. As an alternative to our approach employing Eq. (27), one could directly save the sub- and fine-grid diffusivity of passive transport from the LES simulation with respect to the coarser grid [43]. Such a strategy would also allow to use other turbulence models than those employing eddy viscosities, e.g. deconvolution-based ones [55,56], for the generation of the database. However, we stress that the resulting sub-grid diffusivity needs to be obtained from the initial simulation and not during the rCFD run with grids being far too coarse for any deconvolution operations. It is worth mentioning that for the first approach – provided we utterly trust rCFD –, there is no requirement of solving the passive transport equation (8) during LES, whereas in the second strategy, we would be forced to include species transport in the data-generating calculations.

3. Simulation setup

Our case of study to apply the method to a turbulent flow was vortex shedding after cylinder with diameter $D$ in a sub-critical flow regime [57,58] at $Re = \frac{U_{max} D}{\nu} = 3900$. In this section, we describe the setups of the LES and the rCFD simulations. In several aspects, the former is aligned with other studies [59–61] with respect to the flow configuration and comparing the results.

3.1. LES

The set of equations was solved using the PISO algorithm [62,63] with the OpenFOAM toolbox. The computational grid built for this case (in the following referred to as “L1”) is presented in Fig. 3, and has about $3.85 \times 10^6$ cells. The dimensions of the domain are listed in Table 1. The mesh was resolved near the cylinder to avoid using wall functions ($\Delta r/D = 0.001$), and the inlet boundary condition was considered to be laminar [59–61]. A cell set at $x/D = 2.85$ with the length of 0.3$D$ was created for the purpose of adding a constant source of species. The sub-grid and laminar Schmidt numbers were set to 0.7 and 0.9, respectively. A time step of $\Delta t_{LES} = 10^{-3} D/ U_{max}$ was sufficiently small to ensure a Courant number of less than 0.3. The discretization scheme of the transient term was first-order implicit Euler, those applied for pressure gradient, convection and diffusion term were second-order central difference [63].

3.2. rCFD

As mentioned in Section 2.2, if one is not interested in smallest-scale details, it is possible to perform rCFD on a coarser mesh because only the passive transport equation needs to be solved. We built three different grids referred to as “R1”, “R2” displayed in Fig. 4 and “R3” by proportionally decreasing the number of nodes relative to L1 with the aim of studying the method’s efficiency and its limit in grid coarsening.

In addition to grid size, another factor affecting the performance was the recurrence sampling time $\Delta t_{rec}$. As explained in Section 2.2, the flow information should not change too strongly from one time step to the next. In other words, the recording frequency $f_{rec}$ for which we tested three different values had to be significantly larger than the vortex shedding frequency $f_{vs}$ [32]. On the other hand, too high values would have increased database size unnecessarily. A summary of the number of cells for each mesh along with the ratio between the recurrence sampling frequencies and the shedding frequency $f_{vs}$ is provided in Table 2.

Another matter to consider was the time range of the database $\tau_{rec}$ which should cover a suitable number of shedding cycles with period $\tau_{vs}$. In this work, we chose a range of five cycles. The reason for this choice is discussed in detail in Section 4.2.

For generating the recurrence path, we mirrored the databases with respect to the $y = 0$ plane and assigned them equal weights to guarantee symmetric time-averaged results. We chose a mean

Table 1: Geometric dimensions of the computational domain.

<table>
<thead>
<tr>
<th>Domain parameter</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_x$</td>
<td>30$D$</td>
</tr>
<tr>
<td>$L_y$</td>
<td>10$D$</td>
</tr>
<tr>
<td>$L_z$</td>
<td>$\pi D$</td>
</tr>
</tbody>
</table>

Table 2: Computational grids and sampling times employed in rCFD. The number of nodes are reduced proportionally from R1 to R3 with reference to L1.

<table>
<thead>
<tr>
<th>Spatial resolution</th>
<th>$N_x$</th>
<th>$N_y$</th>
<th>$N_z$</th>
<th>$f_{rec}/f_{vs}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{L1}$</td>
<td>3.85 $\times 10^6$</td>
<td></td>
<td></td>
<td>450</td>
</tr>
<tr>
<td>$R_{R1}$</td>
<td>7.68 $\times 10^5$</td>
<td></td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>$R_{R2}$</td>
<td>4.75 $\times 10^5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_{R3}$</td>
<td>2.65 $\times 10^5$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
interval size of consecutive fields of $E[Δt/Δt_{rec}]^1 = 0.2t_{rec}/Δt_{rec}$ from which the corresponding jump probabilities were derived.

Discretization schemes used for rCFD were as the same as the ones in LES, but the generally coarser mesh allowed for a larger time step of $Δt_{rCFD} = 5 \times 10^{-3}D/U_∞$ for the passive scalar equation.

4. Results

In this section, we present a thorough validation of our LES results and our study on species transport with LES as well as rCFD. The performance improvement by rCFD is highlighted from various perspectives.

4.1. LES results

Varying numbers of shedding periods considered to acquire mean flow statistics can be found in the literature. Franke and Frank [64] mentioned that at least 40 shedding periods are necessary. On the other hand, Lysenko et. al. [60] and D’Alessandro et. al. [61] have gathered statistics over more than 150 shedding cycles for their works. We carried out time averaging over approximately 75 vortex shedding cycles after a duration of $125D/U_∞$ after which we considered the turbulent flow as fully established.

4.1.1. Wall data

One can define lift and drag coefficients $C_L = \frac{F_L}{\frac{1}{2}\rho_∞U_∞^2L_2D}$ and $C_D = \frac{F_D}{\frac{1}{2}\rho_∞U_∞^2L_2D}$ where $F_L$ and $F_D$ are normal and tangential forces acting on the cylinder surface, respectively [58]. Fig. 5 depicts time histories of $C_L$ and $C_D$ over the time range of approximately 75 shedding cycles. We found the root mean square of our lift coefficient, $C_{Lrms} = 0.13$, is in the range of $[0.05, 0.6]$. Similarly, the mean drag coefficient, $C_D = 1.1$, is consistent with the results available in literature from both LES and experimental data [60,61,64–69].

The Strouhal number $St = f_∞D/U_∞$ describes the periodicity of the vortex shedding and can be obtained through fast Fourier transform of the lift coefficient. We determined a value of $St = 0.224$ in close agreement with Zhang et al. [66] and Ouvrard et al. [70].

Fig. 6 a displays the distribution of the time-averaged pressure coefficient $\langle Cp \rangle = 2(\bar{p} - p_∞)/\rho U_∞^2$ on the cylinder surface. The resulting mean base suction coefficient of $\langle C_pθ=180° \rangle = -0.875$ is in very good agreement with the range $[-0.97, -0.85]$ from other studies [60,61,64,65,68,70]. The normalized magnitude of mean vorticity $⟨Ω⟩/2√Re$ on the cylinder surface, with $Ω = \nabla \times \mathbf{u}$, depicted in Fig. 6b shows a separation angle of $θ_{sep} = 87°$ which lies within the bounds $[86°, 89°]$ acquired from experimental and LES data [60,61,64,67,68,71].
4.1.2. Flow statistics

We validated our results for the flow statistics by comparing them with experimental data [59,67]. The time-averaged streamwise velocity component along the center-line of the wake region presented in Fig. 7 is in excellent agreement with the work of Parnaudeau et al. [59]. Fig. 8 illustrates the mean velocity profiles at different sections in the wake of the cylinder compared to measurements. Likewise, mean velocity fluctuations are sketched along with the experimental information in Fig. 9. It is evident that the present simulation was very close to the PIV measurements [59].

Table 3 provides a summary of our LES data compared to other studies.

4.2. Recurrence analysis

The rCFD procedure relies on a database of adequate size which is generally not known a priori. Since both its creation and manipulation are computationally expensive, one will look for it to be as small as possible but still contain the relevant physics of the flow it was obtained from. If any symmetries of the system are known, one can use them to extend it with hardly any additional costs. Therefore, we expanded the database containing approximately five shedding cycles (see below for the choice of size) with velocity fields mirrored in y-direction as explained in Section 3.

Next, we assessed the flow's degree of recurrence after the cylinder by calculating the distance norms for both databases to obtain the combined one, $D(t, t')$, shown in Fig. 10 for R2 with $\Delta t_{\text{rec}}/\Delta t_{\text{LES}} = 50$. One can intuitively perceive some amount of periodicity corresponding to $\tau_{\text{vs}}$. Line plots of $D(t, t')$ with one time argument fixed in Fig. 11 underline that while minima are located at multiples of $\tau_{\text{vs}}$, the flow is clearly not fully periodic. Although velocity fields recur approximately after $\tau_{\text{vs}}$, they are not identical due to turbulent fluctuations. This explains the need to retain several shedding cycles for a proper description.
Fig. 10. The combined distance plot calculated for R2 corresponding to $\tau_{en} = 10$ shedding cycles. The dashed line separates the primary database and the mirrored one. Parallel to the main diagonal, weaker local minima (similar configurations) are located, which are separated by pronounced maxima (different states).

Fig. 11. The distance norm value for R2 at the times of 2 (---) and 3 (----) cycles for the primary database. Though recurrent within a certain tolerance, the flow is not fully periodic.

Fig. 12. Average nearest-neighbor distance over recurrence matrix size. After a fast drop, the curve decreases slowly and even shows a very weak increment for large times. For each duration, the recurrence matrix contained the same number of original and of mirrored flow fields.

We stress that such a (pseudo-)periodicity is not necessary for rCFD which can handle a completely irregular reappearance of patterns as long as the database contains their recurrence. Consequently, the average degree of recurrence may be used as an indicator for sufficient database size by looking at the evolution of the mean nearest neighbor distance $D_{\sim}(t)$ defined in Eq. (19). In a recent study on a gas-solid fluidized bed, one of us [73] found a pronounced initial drop of $D_{\sim}(t)$ that was followed by a slow descent connected to the complexity of the system. Once within the second regime which may even show slight, temporary increments of $D_{\sim}(t)$, extending the database added only little further information. Fig. 12 illustrates $D_{\sim}(t)$ for the combined, i.e. symmetrized, database for up to ten cycles. Unsurprisingly, the characteristic, fast drop is followed by a rather constant plateau after one to two shedding cycles. After this duration, new information accumulates only slowly. As a compromise between memory demands and sufficient amounts of information, we generally use database sizes of a few multiples of the dropping time as a rule of thumb. In the present case, we chose approximately five cycles, which resulted in a total of ten cycles because of the symmetrization procedure.

4.3. Time-extrapolation

Using the recurrence database, we could time-extrapolate the evolution of $u(r, t)$. In order to verify if the recording duration based on $D_{\sim}(t)$ had been sufficiently long to reproduce statistically valid long-term properties, we investigated the time-averaged velocity components and its fluctuations over 75 shedding cycles acquired from rCFD simulations with different numbers of cycles in the database and compared them with mean profiles calculated from the database itself according to different number of shedding cycles. Therefore, we defined error norms

$$ E_1 = \frac{1}{VU_{\infty}} \int d^3r |\bar{u} - \bar{u}_{LES}| $$

$$ E_2 = \frac{1}{VU_{\infty}^2} \int d^3r |\overline{\mathbf{u}} - \overline{\mathbf{u}}_{LES}| $$
which are displayed in Fig. 13. The errors from the recurrence process after a few shedding cycles were in the range of one or two percent and below for the velocity average and fluctuations, respectively, which we deemed acceptable for our purposes. Interestingly, the errors kept decreasing when the mean recurrence degree in Fig. 12 had already leveled out. Consequently, one should not blindly rely on a single characteristic number of the recurrence statistics, but carry out a posteriori analyses, too.

Additionally, we underline the purpose of rCFD to study long-term processes like species transport by time-extrapolating the underlying velocity field. However, it does not provide new information on velocity itself. Hence, the error in the time-extrapolated velocity field cannot be lower than that of the evenly time-averaged database in Fig. 13. We stress that the situation is utterly different for species transport (cf. Section 4.4).

4.4. Species transport

We studied passive scalar transport by adding a source in the wake region (cf. Fig. 3). Instantaneous mass concentrations in the wake region at two selected times for LES on L1 and rCFD on R1, R2 and R3 are displayed in Figs. 14 and 15. In the beginning of the rCFD simulations, the generated recurrence paths were the same, therefore species profiles demonstrated similar patterns at $t = 3\tau_{vs}$. After some time at $t = 30\tau_{vs}$ with arising slight changes
in recurrence paths in each rCFD simulation and depending on the level of fluctuations in these periods, the instantaneous profiles showed some minor deviations particularly further downstream of the source. However, the time-averaged species profiles after 75 shedding cycles presented in Fig. 16 agreed extremely well and different recurrence paths did not affect the mean species profile due to statistical averaging. Fig. 17 compares time-averaged concentration profiles over 5 and 75 shedding cycles and verifies the necessity of time averaging over long time periods (as mentioned in Section 4.1) in order to acquire statistically valid data for turbulent flows.

We note that the results for the first two grids were in very good agreement with the LES simulation with reduced cell numbers by factors of 5 and 8, and even the coarsest mesh with 1/15 of the original cell number showed only little deviations. The effect of grid coarsening could be distinguished for time-averaged species fluctuations in Fig. 18. However, one can observe that the profiles were qualitatively comparable.

For the sake of a more detailed comparison, we provide three groups of species line plots based on different values of $\Delta t_{rec}$ for the computational grids specified in Table 2. Fig. 19a illustrates the mean profiles at different sections in the wake region of the cylinder studied with LES on L1 compared to rCFD on R1, R2 and R3 with the smallest $\Delta t_{rec}$. rCFD on R1 and R2 was in very good agreement with LES, but for the coarsest mesh R3, we see the onset of deviations. The same behavior can be observed for larger time steps, i.e. smaller recurrence frequencies in Fig. 19b and c, which indicates that one may be quite generous with the choice of sampling steps.

Additionally, we provide two quantitative measures for the obtained accuracy. The relative error for the system-wide, total amount of time-averaged species calculated from rCFD on R1, R2 and R3 in comparison to LES on L1 was well below one percent for all investigated cases, cf. Table 4. Furthermore, the error of the spatial species distribution is displayed in Fig. 20 for the various meshes. Although the number of cells was decreased 5, 8 and 15 times, the error remained in the same order of magnitude, which demonstrates that rCFD is capable of dealing with rather coarse grids. However, we believe that an intuitive notion for the method’s accuracy may be drawn more easily from Figs. 14–19.

Furthermore, Fig. 21 shows a comparison between the errors of species distribution time-averaged over 75 shedding cycles from rCFD and obtained from considering only databases with different number of cycles without extrapolation. Here, one can observe the importance of methods like rCFD since even for an almost periodic process like vortex shedding and even after 10 cycles, it makes a difference for species profiles if we time-extrapolate the dynamics or not.

![Fig. 16. Snapshots of mean species after 75 cycles for (a) LES on L1 and rCFD on (b) R1, (c) R2 and (d) R3. Contour lines are sketched at values of 0.01, 0.03, 0.05, 0.07, 0.09 and 0.11.](image1)

![Fig. 17. Time-averaged species transport profile at different locations in the wake region of the cylinder over 5 and 75 shedding cycles. A large number of shedding cycles is required for statistical averaging because of the high-level of fluctuations. The data are shifted to be visible in one plot.](image2)

| Table 4
<table>
<thead>
<tr>
<th>Mesh</th>
<th>Relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>0.0064</td>
</tr>
<tr>
<td>R2</td>
<td>0.0066</td>
</tr>
<tr>
<td>R3</td>
<td>0.007</td>
</tr>
</tbody>
</table>

$E_c = \int d^3 r |\vec{C} - \vec{C}_{LES}|$
The resilience of rCFD towards grid coarsening is underlined by the anticipated, unsatisfactory performance of LES on equally coarse meshes. We carried out an LES on R1 and compare time-averaged species profiles on L1 and R1 over 75 shedding cycles in Fig. 22. As expected, without very fine grids and small time steps (associated with much higher computational costs), LES led to poor results, while rCFD worked well with a considerably reduced number of cells as well as greater time steps.

4.5. Performance analysis

One of the main purposes and advantages of the rCFD method is the reduction of computational costs. We analyzed its performance for turbulent vortex shedding for various selected cases specified in Table 5. To draw a reasonable comparison, we tried to keep the number of cells per number of processors approximately constant. Therefore, the last column of Table 5 corresponds to number of cells per one core.

Fig. 23 depicts the total run-time of the LES and rCFD cases separated into contributions from the solution of the passive scalar transport equation and other processes. The drastic reduction of the overall computation time had two reasons. First, about 90% of run-time in LES was spent with the velocity and pressure equations which need not to be solved in rCFD. Second, due to the larger permissible value of $\Delta t_{rCFD}$ compared to $\Delta t_{LES}$, the passive transport equation was solved more than 14 times faster in rCFD than in LES.

On the other hand, rCFD had to read the database and calculate the distance matrix, for which the loading time is presented in
The resilience of rCFD towards mesh coarsening can be deduced from the relatively weak dependence of $E_c$ from the cell number.

The difference of 1.5% even after 10 shedding cycles demonstrates the advantage of using rCFD.

The error $E_c$ of the spatial species distribution for different number of cycles in the database. We note that the simulation run-times for the case with the smallest sampling step, it became negligible for that with the largest one.

CPU times for updating the fields along with solving the passive transport equation are indicated in Fig. 24b. The simulation times were close to each other because we used approximately the same number of cells per processor. We note that the lower simulation time for R1 corresponds to the lower number of cells per one CPU in comparison to other grids.

Altogether, the LES simulation was performed on 40 processors with an approximate total time of 200 h while our rCFD simulation on the coarsest mesh and with the largest sampling step took only about 6000 sec on 2 processors. This amounts in a speed-up of more than 120 at 1/20 of required CPU power.

As mentioned above, since finer grids increase loading and hence total run-time, we recommend to use as coarse meshes as compatible with the resolution of the meso- and macroscopic phenomena of interest. Furthermore, they are much less demanding with regard to memory consumption than geometries with huge...
cell numbers as illustrated in Fig. 25. In such cases, RAM could easily become the limiting factor.

5. Conclusion and outlook

In this work, we have demonstrated the application of rCFD to species transport in the turbulent flow after a circular cylinder at Re = 3900 and compared it with a detailed LES which we validated against experimental and numerical data. The good agreement of time-resolved and -averaged species profiles between rCFD and LES shows that already short time series of the velocity field contain most of the relevant information required for long-term studies if reappearing patterns prevail and smallest-scale structures are not of interest. We stress that recurrences within some tolerance (which will ultimately affect the obtained accuracy) are sufficient and no strict periodicity is needed, which would make the whole procedure trivial. Elsewhere [74], one of us demonstrated that not even a clearly dominating frequency as \( f_{\text{ev}} \) is necessary to construct a proper recurrence process.

The obtained speed-ups of more than 120 at 1/20 of the required computer power originated from various sources. Most importantly, rCFD only solves the passive transport equation but neither that for velocity nor that for pressure. Consequently, meshes may be much coarser and time steps larger than for LES. More specifically, we employed three different computational grids with 5, 8 and 15 times less cells as well as a time step 5 times greater than for LES. The database loading time and memory consumption of rCFD, which could have otherwise easily become bottlenecks, were reduced by increasing the sampling time step and lowering spatial resolution, respectively.

Since we used a relatively simple, idealized test case for our novel method, there are several open questions and tasks for future investigations. Among various others, we stress two current restrictions. Firstly, similarity of states depends on the underlying norm which is chosen quite subjectively. In the case of transport in single-phase flow, it is reasonable to compare velocity fields with each other, but the situation becomes more complicated for multiple phases. Is velocity alone still a good indicator or should one rather contrast local volume fractions or combinations thereof? Secondly, we considered completely passive species transport that did not affect the velocity field. However, many slow, long-lasting processes alter the fluid dynamics gradually or abruptly. Heat transfer or chemical conversion can change density or viscosity, which in turn will impact the flow. For such scenarios, it will be necessary to create more than a single database corresponding to different parameter settings and interpolate between them [51].

In summary, we can state that rCFD significantly reduced the computational costs for the investigated case. For the finer two grids, the results agreed very well with the full simulation, and only for the coarsest mesh, we found slight differences, which we regard as an indicator for the lower resolution limit of our simulation. Taken together, this makes us optimistic that rCFD can be a useful method for fast long-term studies of turbulent flows. We are convinced that eventually, it will be applicable not only to passive, single-phase but to coupled, multi-phase problems.

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.

Acknowledgement

This work was partly funded by the Linz Institute of Technology (LIT), Johannes Kepler University (project LIT-2016-1-YOU-007). We furthermore acknowledge support from K1–MET GmbH metallurgical competence center.

References


Lourenco L, Shi B. Characteristics of the plane turbulent near wake of a circular cylinder. A particle image velocimetry study. Published in Ref[68][1993].


