

BioBrain Database

Héctor Maldonado de León¹, Ansh Gupta², Victor Puig I Laborda³,
Mahdi Naderibeni⁴, Cees Haringa¹

October 16, 2025

Contents

1	Background	3
2	Introduction	3
2.1	Turbulence	4
2.2	Species transport	4
2.3	List of cases	5
3	Folder structure	6
4	Case 1: 700 L reactor	7
4.1	Description	7
4.2	Mesh	9
4.3	Model development	11
4.3.1	Stopping criteria	13
4.4	Simulations	13
4.5	Data visualization	15
5	Case 2: Stavanger 22m³	16
5.1	Description	16
5.2	Mesh	17
5.3	Model development	18
5.3.1	Single-phase simulations	19

¹Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands.

²Department of Chemical Engineering, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands.

³The Novo Nordisk Foundation Center for Biosustainability, Bygning 220, Henrik Dams Allé, Kgs, Lyngby, DK-2800, Denmark.

⁴Pattern Recognition and Bio-informatics Group, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands.

⁵Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands.

5.3.2	Multi-phase simulations	20
5.3.3	Stopping criteria	20
5.4	Simulations	21
5.5	Data visualization	22
6	Case 3: Industrial Bioreactor with Cooling Coil	23
6.1	Description	23
6.1.1	Geometry	24
6.2	Mesh	26
6.3	Model Development	27
6.4	Simulations	29
6.5	Data visualization	30
7	Acknowledgments	30
	References	31
A	Instructions for Data visualization	32
B	ParaView Getting Started Guide	35

1 Background

Computational Fluid Dynamics (CFD) simulations of bioreactors, despite recent advances, remain computationally demanding, especially for assessing full bioprocesses or optimizing design due to the large combinatorial analysis of operating conditions required. Reduced Order Models (ROMs) offer a solution by reducing the complexity of the model while maintaining spatial detail affordably. Moreover, these models enable the integration of hydrodynamics and biokinetics at an affordable cost. This database aims to accelerate ROM development and promote collaboration between microbiologists and biochemical engineers by releasing CFD-derived data of some bioreactors.

However, the use of reduced-order models in bioreactor analysis is hindered by the steep learning curve of CFD software and limited access to comprehensive databases essential for training and testing these models. By openly releasing this data, we aim to overcome these barriers, facilitating rapid ROM development and enabling researchers to benchmark and compare alternative dimensionality reduction methods like compartment modeling (CM) or proper orthogonal decomposition (POD).

Therefore, this initiative democratizes access to essential bioreactor data, allowing scientists and engineers to analyze bioprocesses more efficiently without requiring access to CFD and high-performance computing, facilitating developments in model and data-driven bioprocess design and optimization.

2 Introduction

The BioBrAIn database offers curated sets of CFD simulations across various reactor geometries and operating conditions, with the goal of democratizing access to critical bioreactor data—such as velocity fields, turbulence characteristics, and gas fraction distributions. This resource enables both industrial microbiologists and machine learning practitioners to evaluate bioprocesses and develop surrogate models without the need for CFD software or high-performance computing infrastructure. In doing so, it supports advancements in bioprocess design and optimization through surrogate modeling.

Each case is introduced along with its parameterization, following a typical CFD workflow: CAD geometry generation, mesh creation, model setup for turbulence and species transport, discretization, and result visualization. Most simulations were conducted using ANSYS FLUENT™ so certain configuration details refer to features specific to its meshing and solver modules. However, similar setups can be replicated in other proprietary or open-source CFD platforms. In the following tables, we outline the general settings for the turbulence and species transport models for the cases available in this release:

2.1 Turbulence

Due to known experimental observations for the flow behavior in stirred tank reactors, along with proven comparisons of simulations, the *dispersed standard* $k - \varepsilon$ model was implemented in ANSYS FLUENTTM by setting the **viscous** model with the following parameters:

Table 1: Parameters used for the **viscous model** to account for turbulence.

Parameter	Value
Turbulence model	k-epsilon
k-epsilon model	Standard
Near-wall treatment	Standard wall functions
Turbulence multiphase model	Dispersed
<i>Model constants</i>	
Cmu	0.09
C1-epsilon	1.44
C2-epsilon	1.92
C3-epsilon	1.30
TKE Prandtl Number	1.00
TDR Prandtl Number	1.3
Energy Prandtl Number	0.85
Wall Prandtl Number	0.85
Turbulent Schmidt Number	0.70

2.2 Species transport

Subsequent validation of some design points was conducted by comparing estimations of the mixing time from CFD to the ones reported in publicly available literature, given a converged flow field that was frozen, and an inert tracer was injected. For this, the parameters described in table 2 were set for the **species transport** model.

Table 2: Parameters for the **species transport** model for determining mixing times

Parameter	Value	Units
<i>Phase properties</i>		
Phase	liquid	-
Phase material	tracer_mix	-
<i>Water-tracer mixture properties</i>		
Compounds in mixture	1. tracer	-
	2. water	-
Density	Volume weighted-mixing-law	kg m^{-3}
Viscosity	1.003×10^{-3}	$\text{kg m}^{-1} \text{s}^{-1}$
Mass diffusivity	Multicomponent	-
Diffusion coefficient for tracer	1.00×10^{-9}	$\text{m}^2 \text{s}^{-1}$

2.3 List of cases

Below is a list of the cases available per release, and an overview of the operating conditions is introduced:

Version 1.0

- **Small-scale reactor (700 L)**
 - Fixed impeller configuration: A single impeller configuration (i.e., RT-RT-RT) geometry operating at different filling levels but same stirring speed (i.e., 300 RPM).
 - Variable impeller configuration: Six different impeller configurations are provided, operating at different combinations of stirring speeds (25-133 RPM) and filling levels (0.35-1.90 m).
- **Stavanger reactor (22 m³):** An industrial-scale geometry equipped with 4 Rushton turbine impellers operating at different stirring speeds (0.45 s^{-1} to 2.22 s^{-1}), filling levels (4.72 m to 6.87 m), and aeration rates ($0 \text{ m}^3/\text{s}$ to $0.182 \text{ m}^3/\text{s}$ at normal conditions).
- **Industrial bioreactor with Cooling Coil:**

Disclaimer: The CFD simulation cases presented in this repository are purely academic and have been reproduced from publicly available studies. The authors of this repository are not liable for any misuse of the data presented herein. Users are responsible for ensuring that the data is used appropriately and in accordance with relevant regulations and guidelines. The cases presented here are based on publicly available reactor designs; hence, no patent or commercial rights can be derived from them.

3 Folder structure

Each case folder includes:

- `dp_0`
 - **case**: The necessary files for generating the simulations in the CFD package originally used for this purpose [`.cas` / `.dat` files]
 - **data**: CFD-derived data in open-source file formats [`.csv` / `.encas` files]
 - **geom**: Geometry files containing the CAD models of the reactor walls and internals [`.stl` / `.scdoc`]
 - **mesh**: Spatial discretization of the CAD models using the meshing solution chosen for each case [`.msh` files]
 - **results**: When available, a set of images displaying the flow fields for variables of interest is provided. [`.jpeg` files]
 - **config.json**: A file containing general parameters to ensure reproducibility of the CFD simulations (including: meshing, turbulence, species transport, discretization schemes, etc.)
- `dp_1`
- ...
- `operating_conditions.csv`
- `README.md`
- `CITATION.cff`

4 Case 1: 700 L reactor

4.1 Description

This case is based on the studies from Spann *et al.* [1] and Tajssoleiman *et al.* [2] on the formation of pH gradients in a 700 L batch fermentation of *Streptococcus thermophilus*. These studies employed a single-phase CFD simulation to predict the evolution of pH gradients experimentally observed, and then assessed their effect on growth when coupled to a biokinetic model. Later, Nadal-Rey *et al.* [3] utilized this reactor geometry to propose a workflow for developing a dynamic compartment model based on a library of staggered simulations, aiming to obtain a close-to-real-time representation of the hydrodynamics and respective gradients of substrate and oxygen.

Two sets of simulations based on the reactor geometry from Spann *et al.* [1] are provided in this dataset:

- **Fixed impeller configuration:** For this single-phase case, the impeller configuration consists of three Rushton Turbine impellers operating according to the conditions outlined in Table 3.
- **Variable impeller configuration:** This dataset corresponds to six reactor geometries where the impeller configuration varies between Rushton and Pitched-blade Turbines operating according to the conditions outlined in Table 3.

Table 3: Operating conditions for the 700 L cases

Parameter	Fixed geometry	Variable geometry	Unit
Stirring speed	300	25 - 133	RPM
Liquid height	1.694	0.35 - 1.90	m

DOI: 10.4121/f8d32b7c-0477-4b87-a038-39e55c8a7379

Table 4 provides the values of the dimensions of this reactor geometry (See fig. 1), including both types of impeller used: Rushton Turbine (RT) and Pitched-blade Turbine (PBT).

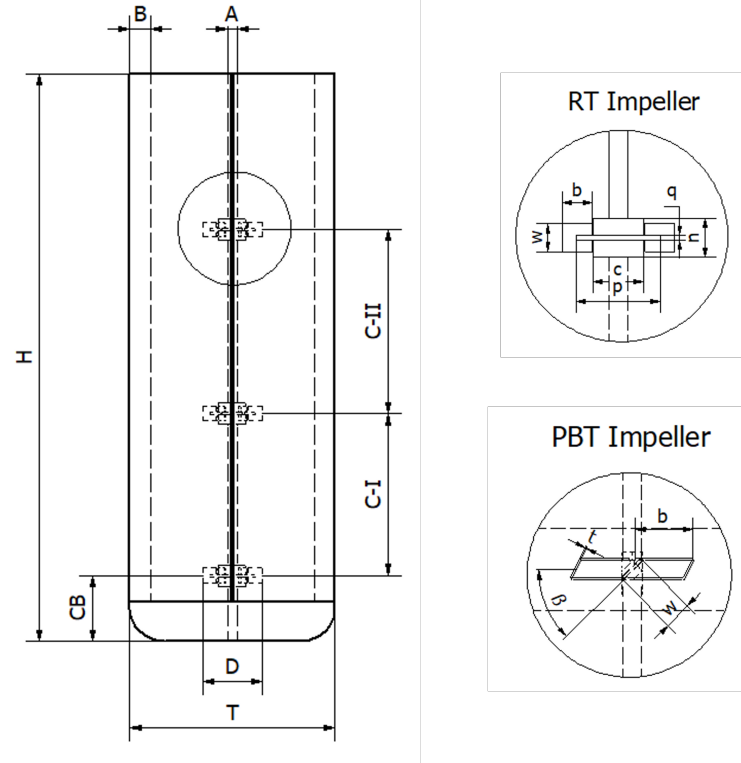


Figure 1: Sketch of the 0.7 m³ reactor geometry. Details about the dimensions are presented in Table 4.

Table 4: Reactor dimensions for 700 L case

Part	Label	Value m	Description
Tank	T	0.70	Diameter of tank
	H	1.9	Height of tank
	D	0.020	Diameter of impeller
	A	0.033	Shaft diameter
	B	0.069	Width of baffle
	CB	0.23	Bottom clearance of impeller
	C-I	0.54	Spacing between bottom and middle impeller
	C-II	0.62	Spacing between middle and top impeller
RT Impeller	b	0.053	Width of impeller blade
	w	0.050	Height of impeller blade
	c	0.090	Diameter of impeller coupling
	q	0.0080	Thickness of impeller disk
	p	0.15	Diameter of impeller disk
PBT Impeller	b	0.053	Width of impeller blade
	w	0.050	Height of impeller blade
	t	0.0040	Blade thickness
	β	45°	Pitch of blades

4.2 Mesh

Table 5 shows the parameters used for each task to generate the watertight mesh for this reactor geometry.

Table 5: Parameters used for generating a watertight mesh of the reactor geometry

Task	Parameter	Value/option
Local sizing	<i>Impellers:</i>	
	Growth rate	1.2
	Size control type	Body size
	Target mesh size [m]	0.005
	<i>Bulk:</i>	
	Growth rate	1.2
	Size control type	Body size
	Target mesh size [m]	0.02
Surface mesh	Minimum size [m]	0.005
	Maximum size [m]	0.02
	Growth rate	1.2
Describe geometry	Type	Only fluid with no voids
	Change from wall to internal	No (Fluid-fluid boundary)
	Share Topology	Yes
	Non-conformal Mesh	No
	Multizone Meshing	No
Boundary Layers	<i>Walls:</i>	
	Offset method type	smooth-transition
	Number of layers	3
	Transition ratio	0.272
	Growth rate	1.2
	Add in	fluid-regions
	Grow on	Tank_walls
	<i>Transition impeller zones (3):</i>	
	Offset method type	smooth-transition
	Number of layers	3
	Transition ratio	0.25
	Growth rate	1.15
	Add in	fluid-regions
	Grow on	impeller_zone_walls
	Fill with	poly-hexcore
	Buffer layers	2
	Peel layers	2
	Min cell length [m]	0.005
	Max cell length [m]	0.01
Improve volume mesh	Improve cell quality limit	0.3

Table 6 shows parameters for the different meshes generated based on the 700 L reactor geometry. For all of these meshes, the cell type is poly-hexcore with a minimum orthogonal quality of 0.3.

Table 6: Mesh parameters for the 700 L geometry

Configuration	Cell quantity	Faces
PBT-PBT-PBT	945192	108511
PBT-PBT-RT	891774	106662
PBT-RT-RT	836329	105160
RT-PBT-PBT	881994	106709
RT-RT-PBT	832384	105321
RT-RT-RT	783959	103524

4.3 Model development

The dataset was generated with the goal of developing surrogate models capable of handling various impeller configurations operating under different stirring speeds and filling levels. Consequently, the focus was placed on calculating the liquid flow field. Although the simulations included the headspace, they were solved exclusively for the liquid phase. This approach simplified the computations and significantly reduced run times. In the following sections, the details for setting up the different models of the simulations are provided.

Turbulence The settings for this model follow the ones described in Table 1; no further alterations were made.

Species transport Parameterization of this model followed the details shown in Table 2. Additionally, the tracer was injected using a spherical **cell-register** with a radius of 2 cm located along the x-axis 0.17 m from the centre and 5 cm below the liquid surface (H_L).

Discretization schemes Table 7 provides details about the discretization schemes and under-relaxation factors used for solving the steady simulations of the 700 L reactor geometry.

Table 7: Discretization schemes used for the solution of transient simulations for the evaluated single-phase setups in this work.

Solution methods		Solution controls
<i>Pressure-velocity coupling</i>		
Scheme	SIMPLE	
<i>Spatial discretization</i>		
Gradient	Least squares Cell Based	-
Pressure	PRESTO!	0.30
Density	Second order upwind	0.50
Momentum	Second order upwind	0.35
Volume Fraction	First order upwind	0.20
Turbulent kinetic energy	Second order upwind	0.40
Specific dissipation rate	Second order upwind	0.40
Tracer ^a	Second order upwind	0.99
Transient formulation ^a	Second order implicit	-

^a Used only when solving for the mixing-time analyses.

Convergence criteria Convergence of the simulations was considered given the values shown in Table 8 for the different equations being solved for. These were combined with a stopping criteria described in the following section (Check section 4.3.1).

Table 8: Equations being solved along with their convergence criteria

Equation	Variable(s)	Convergence criteria
Flow	continuity	1×10^{-3}
	u-liquid	
	u-gas	
	v-liquid	
	v-gas	
	w-liquid	
	w-gas	
Turbulence	k-liquid	1×10^{-3}
	eps-liquid	
Volume fraction	vf-gas	1×10^{-3}
Tracer	tracer<l>	1×10^{-3}

4.3.1 Stopping criteria

Along with the convergence criteria shown in Table 8, a stopping condition was set based on a quasi-stable *torque* report definition. This involves checking the *torque* variability being within a $\pm 2.5\%$ range over a rolling window of 200 iterations. This condition becomes active after 3000 iterations from the start.

4.4 Simulations

Table 9 provides the combinations of stirring speeds and liquid heights used per design point simulated for all cases, disregarding the impeller configuration.

Table 9: Operating parameters and dimensionless mixing times for simulations employed in this study

Design point	Stirring speed	Liquid height
	$N_s / (\text{s}^{-1})$	$H_L / (\text{m})$
0	1.9	1.7
1	2.2	0.58
2	1.8	1.5

Table 9: (continued)

Design point	Stirring speed	Liquid height
	N_s / (s^{-1})	H_L / (m)
3	0.76	0.52
4	1.3	1.2
5	0.61	1.7
6	1.2	1.6
7	0.49	0.91
8	2.1	1.0
9	2.0	1.2
10	1.5	0.35
11	2.0	1.6
12	1.0	1.5
13	1.1	1.3
14	1.9	1.9
15	1.7	0.54
16	1.5	0.93
17	1.4	1.1
18	0.66	0.49
19	0.90	0.41
20	0.73	0.86
21	1.7	1.2
22	2.2	1.3
23	1.6	1.5
24	1.4	1.7
25	1.6	1.7
26	1.7	0.65
27	0.96	0.38
28	1.9	1.1
29	1.1	0.46
30	0.69	1.8
31	0.99	1.1
32	0.85	0.67
33	0.88	0.97
34	0.81	1.8

Table 9: (continued)

Design point	Stirring speed	Liquid height
	$N_s / (\text{s}^{-1})$	$H_L / (\text{m})$
35	1.3	1.8
36	1.0	0.97
37	0.53	1.2
38	2.0	0.62
39	1.3	1.8

4.5 Data visualization

Based on the simulation results, the data for gas holdup within the reactor is shown in Figure 2, and for the magnitude of liquid velocity is shown on Figure 3, respectively. The plots shown here have been done for design point 0 of the RT-RT-RT configuration as an example.

For details and instructions regarding the generation of such plots, please refer to Appendix A.

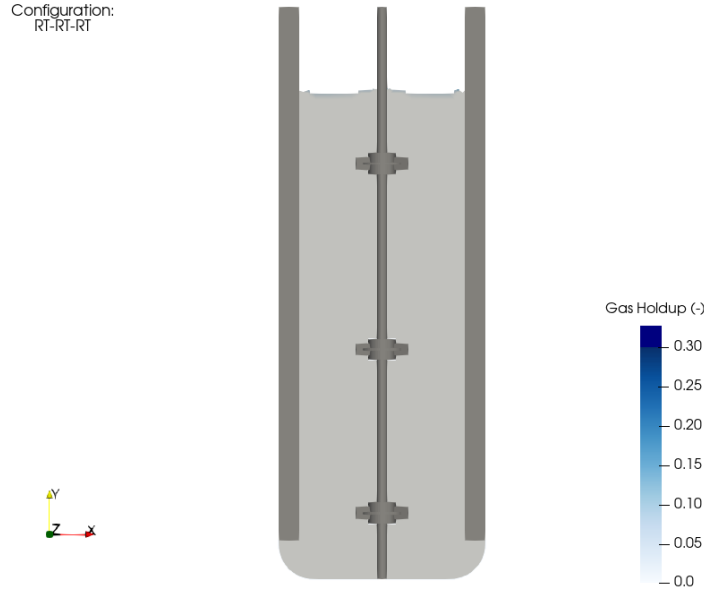


Figure 2: Gas Holdup for RT-RT-RT configuration

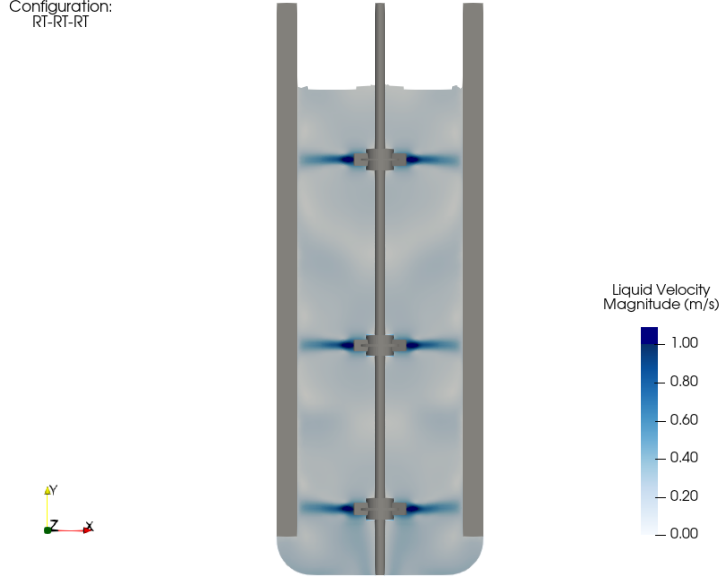


Figure 3: Liquid velocity (magnitude) for RT-RT-RT configuration

5 Case 2: Stavanger 22m³

5.1 Description

Access to publicly available detailed information on industrial-scale fermentors is limited [4]; most studies date back to the 1980s and 1990s, before the widespread availability of affordable computing power. Among these, the EU-Project *Bio-process Scale-up Strategy* analyzed a 30 m³ reactor design for a fed-batch process with a maximum working volume of 22 m³ equipped with Rushton Turbine (RT) impellers, as well as 3- and 6-blade radial Scaba or axial Scaba hydrofoil impellers. Here, we focus on the configuration using solely RT-impellers. The studies by Bylund *et al.* [5] and Vrabel *et al.* [6] provide detailed descriptions of the reactor geometry, pulse-response experiments for mixing time determination, and the process conditions. Notably, the available data spans a wide range of operating conditions—both aerated and non-aerated—resulting in distinct impeller flow regimes (i.e., flooding, loading, and transitional states). Based on these studies, simulations were conducted across the operating conditions summarized in Table 10.

Table 10: Operating conditions for the Stavanger case

Parameter	Range	Unit
Stirring rate	0.45 - 2.22	1/s
Liquid height	4.725 - 6.87	m
Aeration rate ^a	0 - 0.182	m ³ s ⁻¹

^a measured at normal conditions

DOI: 10.4121/34c5d95d-2e34-4087-b970-5f14b7463d44

5.2 Mesh

To maintain computational efficiency and ensure tractable simulations, only half of the reactor was modeled. A periodic boundary condition was applied in the xy-plane, facilitating faster solution times and improved convergence under aerated conditions. While this simplification may influence the spatial distribution of the gas phase, it is expected to preserve the essential physical behavior of the system.

Table 11: Parameters used for generating a watertight mesh of the Stavanger reactor geometry

Task	Parameter	Value/option
Local sizing	<i>Impellers:</i>	
	Growth rate	1.2
	Size control type	Body size
	Target mesh size [m]	0.01
	<i>Bulk:</i>	
	Growth rate	1.2
	Size control type	Body size
	Target mesh size [m]	0.04
	Minimum size [m]	0.01
Surface mesh	Maximum size [m]	0.04
	Growth rate	1.2
	Type	Only fluid with no voids
Describe geometry	Change from wall to internal	No (Fluid-fluid boundary)
	Share Topology	Yes
	Non-conformal Mesh	No
	Multizone Meshing	No
Boundary Layers	<i>Sparger:</i>	
	Offset method type	smooth-transition
	Number of layers	3
	Transition ratio	0.272
	Growth rate	1.15
	Add in	fluid-regions
	Grow on	sparging_face
Volume mesh	Fill with	poly-hexcore
	Buffer layers	2
	Peel layers	1
	Min cell length [m]	0.01
	Max cell length [m]	0.04 (bulk_liquid)
Improve volume mesh	Improve cell quality limit	0.3

5.3 Model development

The simulation workflow began with a single-phase setup, which served as the initial condition for the subsequent multiphase simulations. This approach ensured numerical stability and convergence once air sparging was introduced. Additionally, the top wall boundary condition was modified from a non-shear configuration to a degassing configuration to represent gas escape dynamics better. Further details on the simulation procedure are provided below:

5.3.1 Single-phase simulations

To ensure convergence of the multiphase simulations, a first batch of simulations under unaerated conditions was generated. This set considers the dispersion height expected given prior studies corresponding to this reactor geometry [6]. Under this configuration, the top wall was modeled as a non-shear surface to account for the presence of the headspace and, practically, the lack of friction exerted by the air on the liquid [7].

Turbulence The settings for this model follow the ones described in Table 1, no further alterations were made.

Species transport Parameterization of this model followed the details shown in Table 2.

Discretization schemes and under-relaxation factors Table 12 provides the discretization schemes and under-relaxation factors used for the solution of steady simulations of the 22 m³ reactor.

Table 12: Discretization schemes used for the solution of steady simulations for the evaluated single-phase setups in this work.

Solution methods		Solution controls
<i>Pressure-velocity coupling</i>		
Scheme	SIMPLE	
<i>Spatial discretization</i>		
Gradient	Least squares Cell Based	-
Pressure	Second Order	0.30
Density	First order upwind	0.50
Momentum	Second order upwind	0.35
Body forces	-	0.50
Volume Fraction	First order upwind	0.20
Turbulent kinetic energy	First order upwind	0.40
Specific dissipation rate	First order upwind	0.40
Tracer ^a	First order upwind	0.95
Transient formulation ^a	Second order implicit	-

^a Used only when solving for the mixing-time analyses.

Convergence criteria Convergence of the simulations was considered once the residuals of the simulations reached the values presented in table 13. Along with these, stopping criteria were considered (Check section 5.3.3).

Table 13: Equations being solved along with their convergence criteria

Equation	Variable(s)	Convergence criteria
Flow	continuity	1×10^{-3} or 5×10^{-3} ^a
	u-liquid	1×10^{-3}
	u-gas	1×10^{-3}
	v-liquid	1×10^{-3}
	v-gas	1×10^{-3}
	w-liquid	1×10^{-3}
	w-gas	1×10^{-3}
Turbulence	k-liquid	1×10^{-3}
	eps-liquid	
Volume fraction	vf-gas	1×10^{-3}

^a Used only when solving for multiphase simulations.

5.3.2 Multi-phase simulations

Once the single-phase simulations were stopped (See section 5.3.3) and the results used as initial conditions for the multi-phase simulations, the injection of air (NTP) via the sparging face was activated. Additionally, new report definitions were established: `gas_bal` to track the mass balance of air over 100 iterations, and `vol_avg_eg` to track the volume average of the gas fraction in the system. To simulate the expansion of bubbles as they rise, the diameter of the bubbles ($d_{b,32}$) was varied according to the model from [8].

$$d_{b,32} = 0.0076(\rho_L \varepsilon_L)^{-0.14} \quad (1)$$

In Equation (1), the reference pressure is 101.325 kPa. Additionally, a threshold for the maximum bubble diameter was set to 3 mm. Changes in bubble size are accounted for by incorporating the ideal gas law and hydrostatic pressure based on the location of the bubble.

Convergence criteria Several attempts were made to find a suitable combination of models for simulating phase interaction and ensuring convergence of the continuity equation. The lowest residual that was achieved was about 2.5×10^{-3} for just a handful of cases, hence 5×10^{-3} was deemed as a reasonable value for all the cases.

5.3.3 Stopping criteria

For the single-phase simulations, along with the convergence criteria shown in Table 13, a stopping condition was set based on a quasi-stable `torque` report

definition. This involves checking the torque variability being within a $\pm 2.5\%$ range over a rolling window of 200 iterations. This condition becomes active after 3000 iterations from the start.

The multi-phase simulations included the evaluation of the torque and the volume average of the gas fraction (`vol_avg_eg`) whose variability should also remain within a $\pm 2.5\%$ range over a rolling window of 200 iterations. The combination of these criteria was enabled after 10000 iterations counted from the last iteration of the single-phase simulation.

5.4 Simulations

Table 14 provides the combinations of stirring speeds and liquid heights used per design point simulated for all cases employing the 22 m³ reactor geometry.

Table 14: Operating parameters for simulations employed in this case study

Design point	Stirring speed N_s / (s ⁻¹)	Liquid height H_L / (m)	Aeration rate Q_g / (Nm ³ s ⁻¹)	Dispersion height H_{disp} / (m)
0	1.5	6.4	0.11	7.3
1	2.2	5.2	0.12	6.0
2	1.2	6.4	0.063	6.9
3	1.9	6.6	0.16	7.7
4	1.4	6.4	0.048	6.8
5	0.77	6.8	0.18	7.8
6	1.6	5.9	0.0049	6.1
7	1.2	5.1	0.033	5.3
8	1.6	4.8	0.13	5.4
9	1.3	5.1	0.12	5.8
10	1.9	5.8	0.068	6.4
11	0.81	6.5	0.045	6.9
12	1.7	6.7	0.018	7.0
13	1.0	5.0	0.028	5.2
14	1.8	4.9	0.061	5.3
15	2.1	6.3	0.011	6.6
16	1.3	6.7	0.074	7.4
17	0.65	5.3	0.10	5.9
18	1.1	4.8	0.055	5.2
19	0.50	6.2	0.088	6.8
20	0.48	5.0	0.14	5.6
21	2.2	6.0	0.17	7.1
22	1.6	6.8	0.089	7.6
23	1.4	5.2	0.15	6.0

Table 14: (continued)

Design point	Stirring speed $N_s / (\text{s}^{-1})$	Liquid height $H_L / (\text{m})$	Aeration rate $Q_g / (\text{Nm}^3 \text{s}^{-1})$	Dispersion height $H_{disp} / (\text{m})$
24	2.0	5.9	0.098	6.7
25	0.58	6.3	0.079	6.8
26	0.96	6.1	0.14	6.9
27	1.1	6.0	0.17	6.9

5.5 Data visualization

After performing the simulation, the results for gas holdup and magnitude of liquid velocity within the reactor are visualized in the form of contour plots. Figure 4 shows the gas holdup, and Figure 5 shows the liquid velocity magnitude for design point 0 of the Stavanger case.

For details and instructions regarding the generation of such plots, please refer to Appendix A.

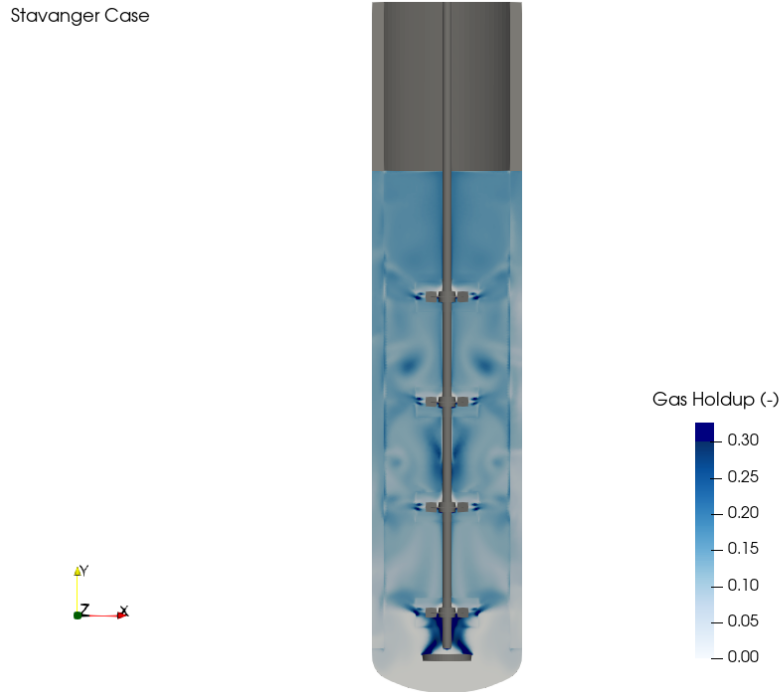


Figure 4: Gas Holdup for Stavanger case

Stavanger Case

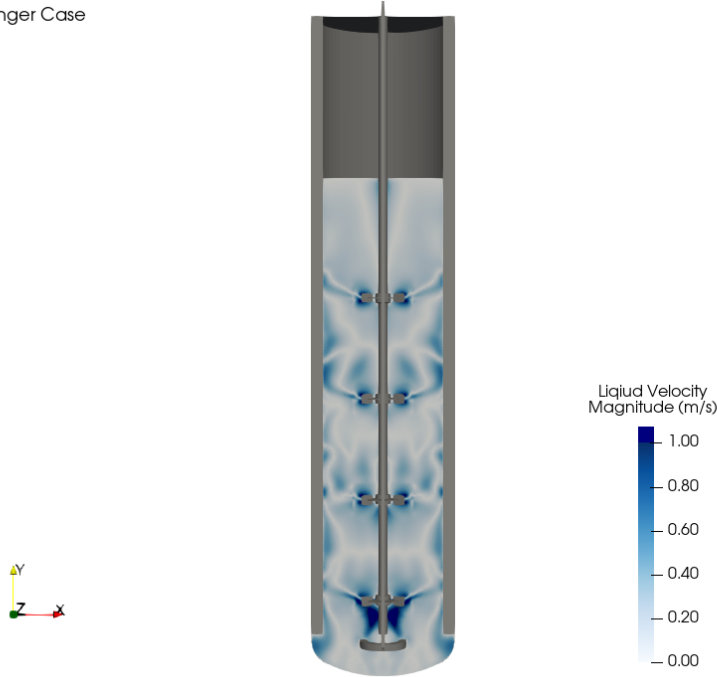


Figure 5: Liquid velocity (magnitude) for Stavanger case

6 Case 3: Industrial Bioreactor with Cooling Coil

6.1 Description

This dataset was initially generated to train supervised AI/ML models to predict spatial concentration gradients in an industrial reactor under varying operating conditions [9]. It comprises 760 steady-state CFD simulations for both fluid flow and transport of diluted species. 560 simulations are designated for model development (training and hyperparameter tuning), and 200 served as a priori held-out test set.

Each simulation is parametrized by a 10-D vector of operating conditions, biokinetic constants, and fluid properties: impeller speed, liquid height, feed-point position (sampled in cylindrical coordinates), volumetric substrate feed rate F_V , biomass concentration C_X , maximum substrate uptake rate $q_{S,\max}$, half-saturation constant K_S , density ρ , and viscosity μ . These variables define both the momentum and the reactive-transport problems solved by the CFD model and are the “global” descriptors associated with every cell of the corresponding mesh.

To cover the design space efficiently, we used Latin Hypercube Sampling (LHS), which stratifies each factor and selects one value per stratum to achieve near-uniform coverage with far fewer samples than simple random sampling [10, 11]. For the feed-point, we sampled the azimuthal angle θ_{feed} , the vertical coordinate y_{feed} , and the radial one r_{feed} in cylindrical coordinates to respect vessel symmetry, and converted to Cartesian coordinates before meshing.

The sampling ranges for all global variables are reported in Table 15.

Table 15: Operational parameter intervals for the global parameters in the CFD simulations.

Global Variable	Interval	Unit
liquid height	3 – 12.5	m
stirring speed	5 – 700	RPM
F_V	$1 \cdot 10^{-4} - 5 \cdot 10^{-3}$	$\text{kg m}^{-3} \text{s}^{-1}$
μ	0.001 – 0.02	$\text{kg m}^{-1} \text{s}^{-1}$
ρ	950 – 1250	kg m^{-3}
C_X	2 – 75	kg m^{-3}
K_S (mass)	$1 \cdot 10^{-5} - 2 \cdot 10^{-4}$	$\text{kg}_X \text{kg}_S^{-1}$
$q_{S,\text{max}}$	$6.95 \cdot 10^{-5} - 4.86 \cdot 10^{-4}$	$\text{kg}_X \text{kg}_S^{-1} \text{s}^{-1}$
θ_{feed}	0 – 2π	rad
y_{feed}	0 – 12.5	m

DOI: 10.4121/836ce408-7b40-43f2-8db0-8e207d356f73

6.1.1 Geometry

The simulations employed a Rushton-tank geometry, designed in SolidWorks[®] Figure 6 with the following cylindrical coordinates:

$$\begin{aligned}
 0 &\leq r \leq 3.1 \text{ [m]} \\
 0 &\leq \theta \leq 2\pi \text{ [rad]} \\
 0 &\leq y \leq 12.5 \text{ [m]}
 \end{aligned} \tag{2}$$

The Rushton tank was designed with four Rushton impellers and baffles. Additionally, a helical cooling coil was placed around the impellers and baffles, and an O-ring sparger at the bottom of the tank. The sizes of the parametrized Rushton Impellers (disk, blades, clearance...) and the baffles were based on a “standard” bioreactor design as described by Gogate *et al.* [12]. The total volume of the tank is 170 m³.

The different tank design dimensions, relative to the vessel height, can be found in Table 16.

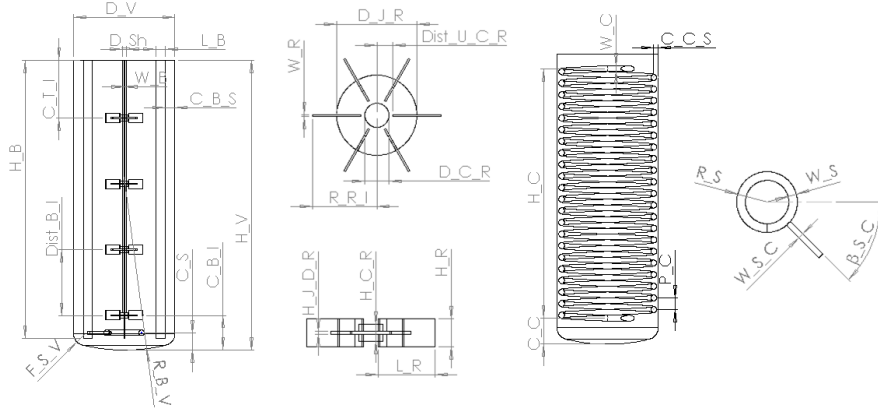


Figure 6: Rushton-tank bioreactor with helical cooling coil and O-ring sparger attached to the side walls. From left to right: General vessel geometry, top view Rushton impeller, front view Rushton impeller, front view helical cooling coil, top view sparger.

Table 16: Model parameters used for the Rushton tank geometry in the CFD simulation.

Dimensions	Description	Value
H_V	Vessel height	12.5 m
A_V	Aspect ratio vessel	0.35
D_V	Vessel diameter	$H_V \cdot A_V$
$R_{B,V}$	Radius bottom vessel	$0.6 \cdot H_V$
$F_{S,V}$	Fillet sketch skeleton vessel	$0.08 \cdot H_V$
$C_{B,I}$	Bottom impeller clearance	$0.12 \cdot H_V$
$C_{T,I}$	Top impeller clearance	$0.2 \cdot H_V$
N_I	Number of impellers	4
$Dist_{B,I}$	Distance between impellers	$\frac{(H_V - C_{T,I} - C_{B,I})}{(N_I - 1)}$
D_{shaft}	Shaft diameter	$0.016 \cdot H_V$
H_C	Height cooling coil	$0.86 \cdot H_V$
W_C	Width cooling coil	$0.02 \cdot H_V$
C_C	Bottom clearance cooling coil	$0.09 \cdot H_V$
$C_{C,S}$	Side wall tank-cooling coil clearance	$0.006 \cdot H_V$
P_C	Pitch helix cooling coil	$0.04 \cdot H_V$
H_B	Height baffle	$0.96 \cdot H_V$
L_B	Length baffle	$0.032 \cdot H_V$
W_B	Width baffle	$0.125 \cdot L_B$

Dimensions	Description	Value
$C_{B,S}$	Side tank wall–baffle clearance	$0.0332 \cdot H_V$
N_B	Number of baffles	4
R_S	Diameter sparger	$0.06 \cdot H_V$
W_S	Width sparger	$0.02 \cdot H_V$
$W_{S,C}$	Width sparger connector	$0.012 \cdot H_V$
$\beta_{S,C}$	Sparger connector plane angle	45°
S_C	Sparger clearance	$0.06 \cdot H_V$
$R_{R,I}$	Radius Rushton impeller	93 m
L_R	Length Rushton blade	$0.75 \cdot R_{R,I}$
H_R	Height Rushton blade	$0.43 \cdot R_{R,I}$
W_R	Width Rushton blade	$0.027 \cdot R_{R,I}$
–	Number of Rushton blades	6
$D_{J,R}$	Diameter joint disk	$1.24 \cdot R_{R,I}$
$H_{J,R}$	Height joint disk	$0.054 \cdot R_{R,I}$
$\text{Dist}_{U,C,R}$	Distance union Rushton blade	$0.2 \cdot D_{J,R}$
	–shaft connector disk	
$D_{C,R}$	Diameter Rushton impeller-	$0.3 \cdot D_{J,R}$
	–shaft connector disk	
$H_{C,R}$	Height Rushton impeller-	$0.27 \cdot R_{R,I}$
	–shaft connector disk	

6.2 Mesh

The meshing strategy and parameters were kept consistent across all simulation cases. Variations in element count were solely due to changes in liquid height, which affected the overall domain volume. To ensure comparability between runs, a target element density of approximately $\sim 9 \times 10^3$ cells/m³ was maintained. At lower fill levels, surface-to-volume effects led to a slightly higher effective mesh density. Across all generated meshes, the minimum orthogonal quality was 75%, and the minimum cell quality was 0.2. A detailed summary of the selected mesh parameters is provided in Table 17.

Table 17: Parameters used for generating a watertight mesh of the 54 m^3 reactor geometry

Task	Parameter	Value/option
Local sizing	Growth rate	1.2
	Size control type	Body size
	Target mesh size [m]	0.005
Surface mesh	Minimum size [m]	0.01
	Maximum size [m]	0.2
	Growth rate	1.2
Describe geometry	Type	Only fluid with no voids
	Change from wall to internal	No (Fluid-fluid boundary)
	Share Topology	Yes
	Non-conformal Mesh	No
	Multizone Meshing	No
<i>Walls:</i>		
Boundary Layers	Offset method type	smooth-transition
	Number of layers	3
	Transition ratio	0.272
	Growth rate	1.2
	Add in	fluid-regions
	Grow on	Tank.walls
Volume mesh	Fill with	poly-hexcore
	Buffer layers	2
	Peel layers	1
	Min cell length [m]	0.01
	Max cell length [m]	0.2 (stator), 0.1 (rotors)
Improve volume mesh	Improve cell quality limit	0.2

The feeding domain, different for each simulation, was then defined as a cell register and separated from the stator domain after the meshing (Figure 7).

6.3 Model Development

Due to known experimental observations for the flow behavior in stirred tank reactors, along with proven comparisons of simulations, the *realizable* $k - \varepsilon$ model was implemented in ANSYS FLUENT™ by setting the **viscous** model with the following parameters:

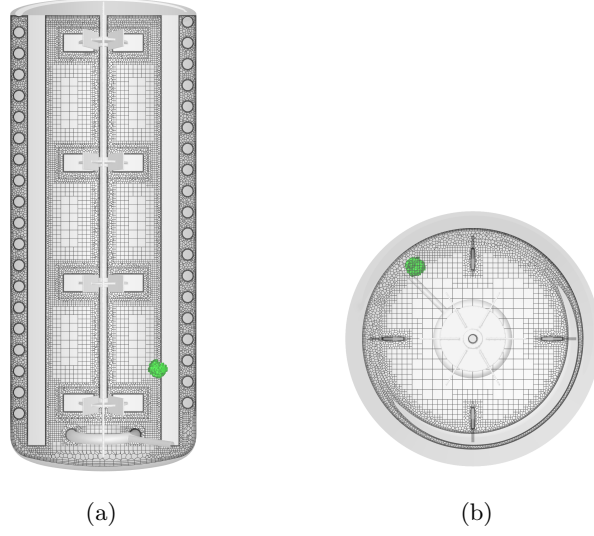


Figure 7: Case with liquid height = 10.65m, with the poly-hexcore mesh and the internal walls. The green dot indicates the feeding point location for this specific case. (a) Front view. (b) Top view.

Table 18: Parameters used for the `viscous model` to account for turbulence.

Parameter	Value
Turbulence model	k-epsilon
k-epsilon model	Realizable
Near-wall treatment	Enhanced Wall Treatment
Turbulence multiphase model	Dispersed
<i>Model constants</i>	
C2-Epsilon	1.9
TKE Prandtl Number	1
TDR Prandtl Number	1.2

Once the flow field reached a pseudo-steady state, we solved the unsteady advection–diffusion–reaction equation for the substrate concentration C_S :

$$\frac{\partial C_S}{\partial t} + \nabla \cdot (\mathbf{U}, C_S) = \nabla \cdot \left[\left(\mathcal{D} + \frac{\mu_t}{Sc_t} \right) \nabla C_S \right] + S_S, \quad (3)$$

where \mathbf{U} is the velocity field, \mathcal{D} the molecular diffusivity, μ_t the eddy viscosity, and Sc_t the turbulent Schmidt number (so the effective diffusivity is $\mathcal{D}_{\text{eff}} = \mathcal{D} + \mu_t/Sc_t$).

The total source term S_S comprises a Monod uptake sink and localized feed source:

$$S_{\text{uptake}} = -q_{S,\text{max}} \frac{C_S}{C_S + K_S} C_X, \quad S_S = S_{\text{uptake}} + S_{\text{feed}} \quad (4)$$

Here $q_{S,\text{max}}$ is the maximum specific uptake rate, K_S the affinity constant, and C_X the biomass concentration. The parameters $(q_{S,\text{max}}, K_S, C_X)$ were sampled via the LHS-DOE. Substrate addition at the feed point was implemented as the volumetric term S_{feed} , applied within the prescribed feed control volume; its magnitude (volumetric feed rate) was also a sampled variable in the DOE.

A summary of the parameters can be found at Table 19.

Table 19: Parameters for the `species transport` model for determining the substrate gradients for the different cases.

Parameter	Value	Units
<i>Phase properties</i>		
Phase	Liquid	-
Phase material	Broth-mix	-
<i>Broth-mix mixture properties</i>		
Compounds in mixture	1. Glucose	-
	2. Broth	-
Density	Volume weighted-mixing-law	kg m^{-3}
Viscosity	Variable (DoE-LHS)	$\text{kg m}^{-1} \text{s}^{-1}$
Mass diffusivity	Multicomponent	-
Diffusion coefficient for tracer	6.7×10^{-10}	$\text{m}^2 \text{s}^{-1}$

6.4 Simulations

The dataset comprises 760 CFD simulations generated from the Latin hypercube sampling (LHS) design-of-experiments described above. We split the data into a development set (560 simulations) for model selection and hyperparameter tuning, and an unseen test set (200 simulations) for final evaluation.

For transparency and reproducibility, the complete case lists—including all sampled input parameters for each simulation—are provided in the companion files `development_set_cases.csv` and `test_set_cases.csv`. Each row corresponds to a unique simulation (case ID) and can be used to reproduce or subset the experiments reported in this work.

6.5 Data visualization

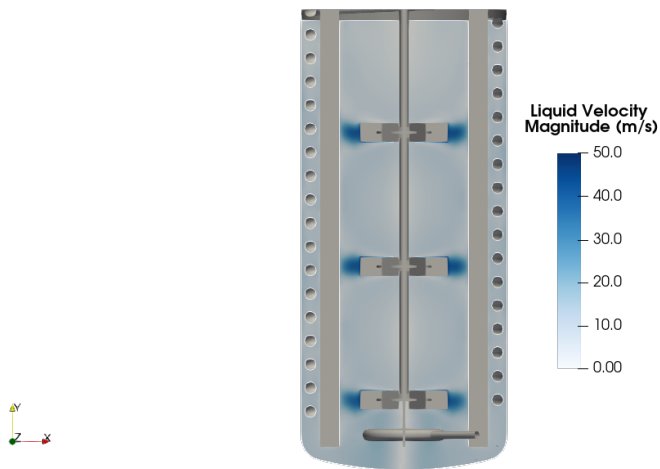


Figure 8: Liquid velocity (magnitude) for DTU case - liquid height: 9.55 m.

7 Acknowledgments

The simulations corresponding to the 700 L and 22 m³ reactor geometries used the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-11590 (Project: “CFD-derived data for surrogate models of bioreactors”).

This work was made possible through the support of the FAIR Data Fund 2024 grant awarded by the 4TU.Research Data Repository. We gratefully acknowledge the technical and financial assistance that enables the scientific community to access CFD-derived bioreactor data.

References

1. Spann, R. *et al.* CFD predicted pH gradients in lactic acid bacteria cultivations. en. *Biotechnology and Bioengineering* **116**, 769–780. ISSN: 00063592. <https://onlinelibrary.wiley.com/doi/10.1002/bit.26868> (2023) (Apr. 2019).
2. Tajssoleiman, T. *et al.* A CFD based automatic method for compartment model development. en. *Computers & Chemical Engineering* **123**, 236–245. ISSN: 00981354. <https://linkinghub.elsevier.com/retrieve/pii/S0098135418308950> (2022) (Apr. 2019).
3. Nadal-Rey, G. *et al.* Development of dynamic compartment models for industrial aerobic fed-batch fermentation processes. en. *Chemical Engineering Journal* **420**, 130402. ISSN: 13858947. <https://linkinghub.elsevier.com/retrieve/pii/S1385894721019884> (2023) (Sept. 2021).
4. Noorman, H. An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnology Journal* **6**, 934–943. <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/biot.201000406> (2011).
5. Bylund, F., Collet, E., Enfors, S.-O. & Larsson, G. Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. en. *Bioprocess Engineering* **18**, 171. ISSN: 0178515X. <http://link.springer.com/10.1007/s004490050427> (2024) (1998).
6. Vrabel, P., Van Der Lans, R. G., Luyben, K. C., Boon, L. & Nienow, A. W. Mixing in large-scale vessels stirred with multiple radial or radial and axial up-pumping impellers: modelling and measurements. en. *Chemical Engineering Science* **55**, 5881–5896. ISSN: 00092509. <https://linkinghub.elsevier.com/retrieve/pii/S0009250900001755> (2024) (Dec. 2000).
7. Haringa, C. *Computational analysis of a penicillin production process using ANSYS FLUENT* (2021), 1–44.
8. Alves, S., Maia, C., Vasconcelos, J. & Serralheiro, A. Bubble size in aerated stirred tanks. en. *Chemical Engineering Journal* **89**, 109–117 (2002).
9. Puig I Laborda, V. & Nielsen, L. *Supervised learning bioreactor regimes* In preparation.
10. McKay, M. D., Beckman, R. J. & Conover, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **42**, 55–61. ISSN: 15372723 (1 2000).
11. Helton, J. C. & Davis, F. J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety* **81**, 23–69. ISSN: 09518320 (1 2003).
12. Gogate, P. R., Beenackers, A. A. C. M. & Pandit, A. B. *Multiple-impeller systems with a special emphasis on bioreactors: a critical review* tech. rep. (2000), 109–144.

Appendix

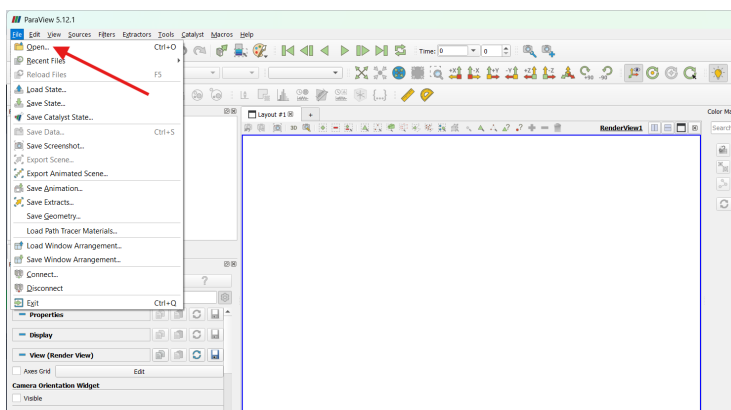
For new users of Paraview, please refer Appendix B for the Getting Started Guide.

A Instructions for Data visualization

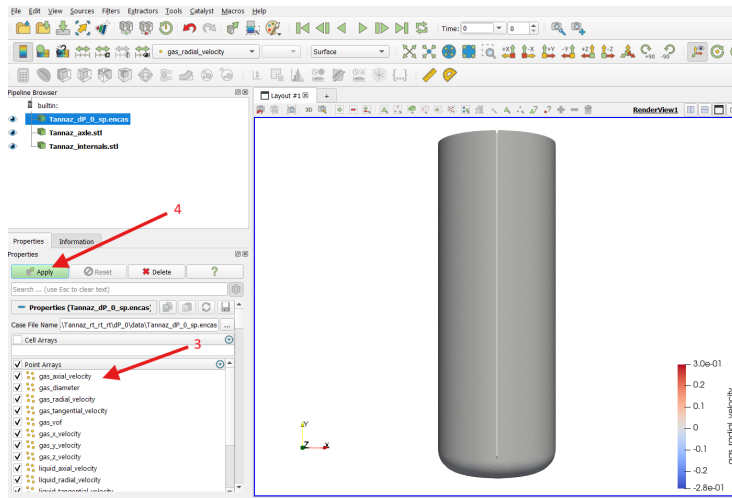
Instructions for visualizing the simulation plots in ParaView software (version 5.12.1):

Files needed: Data files (**.encas**), Reactor internals-geometry files (**.stl**)

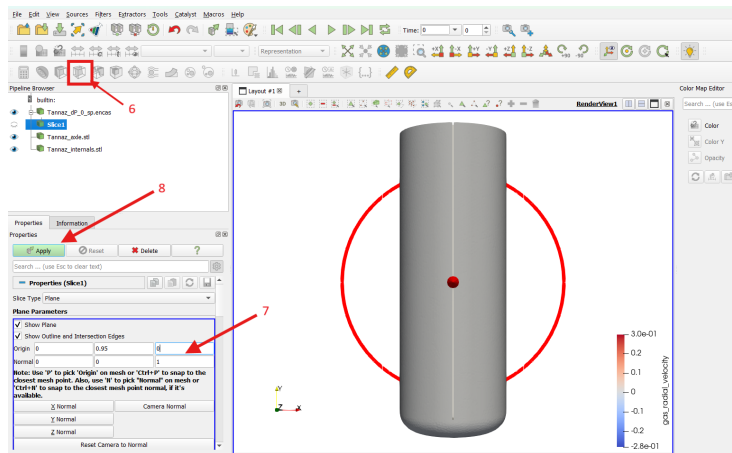
1. Open ParaView software on your system.
2. Click on **File** on the top ribbon and open the required **.encas** and **.stl** files by navigating to the correct directory on your system. (In the 'data' and 'geom' folder)



3. Once the file is loaded in ParaView, you will be able to see various variables present in the **.encas** file on the left side of your screen.

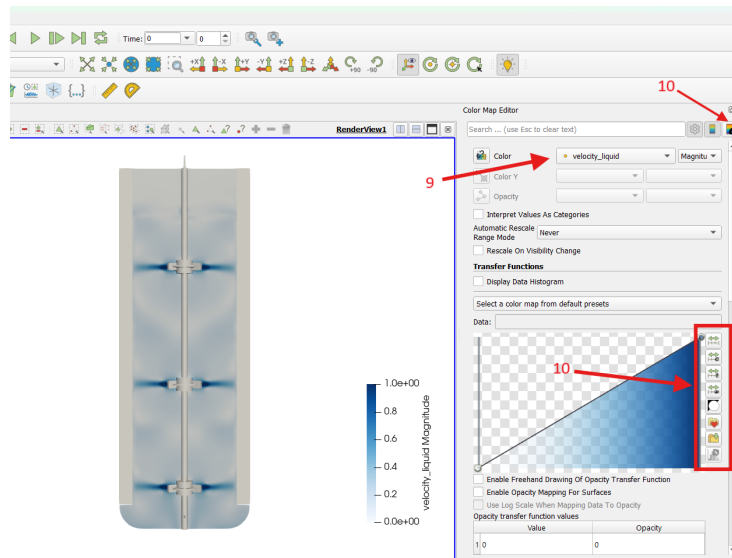


4. Select the variables that you need to work on, and click **Apply**.
5. After clicking on **Apply**, ParaView will load the geometry.
6. To view the variables in a slice (plane) of the system, select the **Slice** option in the top ribbon.



7. Set up the coordinates for the slice place changing the **Origin** and **Normal** (for normal vector of the plane) parameters in the left menu. A plane can also be seen in geometry.
8. Click on **Apply**, after which only the desired plane will be visible on the screen.

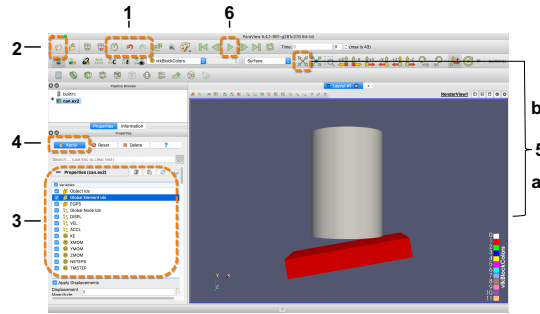
9. In the **Color** section, select the variable that needs to be displayed. A colored contour plot of the variable, with a color map, will be displayed.






10. The colors, labels, and scaling can be adjusted in the **Color Map Editor** (right side of the window).

B ParaView Getting Started Guide


ParaView Getting Started Guide

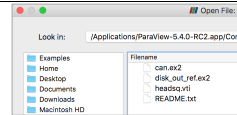


Reset Session

- 1 At any time, click  to reset ParaView to its initial state when the program is first started. The **Undo** and **Redo** buttons   are also available to undo/redos individual changes.

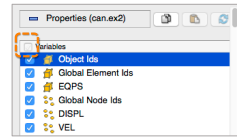
Open File

- 2 Click the folder icon  or choose **File -> Open**. From the **Examples** directory, open the file **can.ex2**. Doing so creates a **file reader** in the **Pipeline Browser**.



Select Data Variables

- 3 Before the data are loaded, you can choose which variables to load in the **Properties** panel (lower left, ParaView window). Click the checkbox next to **Variables** to load all variables. Note: many controls in ParaView are located in this panel.




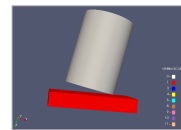
Apply Data

- 4 Click the **Apply** button to load the data. If you change any file reader properties, click **Apply** to update the visualization.



Interact with 3D View

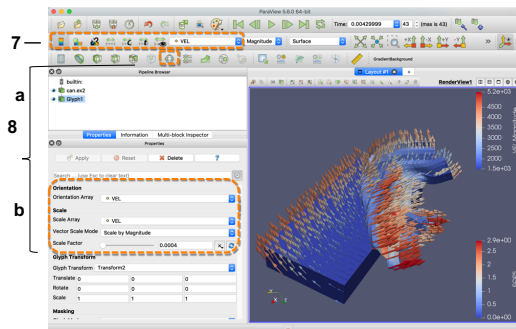
- a. Click the left mouse button and drag to rotate, the middle mouse button to translate, and the scroll wheel or right mouse button to zoom. For one- or two-button mice, hold down Shift and Control keys while clicking and dragging to tilt, translate, and zoom.
- b. To re-center the data in the view, click .



Animate the Data

- 6 Press the play button and watch the can get crushed. Other buttons enable moving to different timesteps.



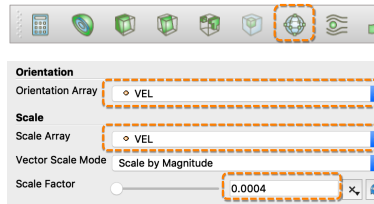


- 7 Apply a Color Map to a Variable
Click on the drop-down menu in the variables toolbar and select the **EQPS** variable.



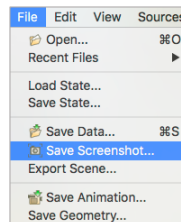
Add Vector Glyphs

- a. Ensure **can.ex2** is selected in the **Pipeline Browser**. Click the **Glyph** filter icon in the toolbar to add a **filter** to the pipeline that will add arrow glyphs to the visualization.
- 8 b. In the **Properties** panel, choose **VEL** from the **Orientation Array** menu. Change the **Scale Array** to **VEL**. Set the **Scale Factor** to 0.004 to scale the arrow glyphs to a reasonable size. Click **Apply** to update the visualization. Use the variables toolbar to color the glyphs by **VEL** (see previous step).



Save Screenshot

Save an image for presentation or publication by choosing **File -> Save Screenshot...**



Get Additional Help

Additional resources for learning about **ParaView** are available in the **Help** menu.

- [ParaView Guide](#) – comprehensive user guide for ParaView
- [Help](#) – online help for file readers and filters
- [Online Tutorials](#) – in-depth tutorials for ParaView
- [Online Blogs](#) – informative blog posts on new features in ParaView