

Main Results of the Second DA Challenge

A. Schröder^{1,5,*}, P. Godbersen¹, D. Schanz¹, A. Sciacchitano², A. Grille Guerra², S. Grauer³, B. Leclaire⁴

1: Inst. of Aerodynamics and Flow Technology, German Aerospace Center (DLR), Germany

2: Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands

3: Dept. of Mechanical Engineering, Pennsylvania State University, United States

4: DAAA, ONERA, Institut Polytechnique de Paris, 92190, Meudon, France

5: Chair of Image Based Measurement Techniques, Brandenburg University of Technology (BTU), Cottbus, Germany

*Corresponding author: andreas.schroeder@dlr.de

Keywords: Data assimilation, Benchmark testing, Lagrangian particle tracking, Flow reconstruction, Challenge

Abstract

This work presents the main results of the second Data Assimilation (DA) challenge, jointly organized by TU Delft, DLR, ONERA, and Penn State. We consider three test cases, namely three synthetic Lagrangian Particle Tracking (LPT) input cases based on advection of virtual tracers along time-series of DNS for Case 1) a turbulent rectangular duct flow at $Re_\tau = 1,000$, Case 2) a compressible turbulent boundary layer flow at a Mach 4.9 and Case 3) a Rayleigh-Bénard convection flow at $Ra = 10^7$ in a cubic domain. Time-resolved tracks are provided for all test cases at four (Case 1), two (Case 2) and three (Case 3) different particle tracking densities, translated to particle image densities (in particles per pixel, ppp) of a generic virtual camera set-up. Additionally, for Case 1 two- and four-pulse particle tracks are addressed and for Case 2 inertial particles beside flow tracers. The results are presented in terms of bias and mean RMS errors of the reconstructed velocity-, vorticity- and pressure fields with respect to the DNS data at a predefined 3D grid of points. For case 3 as well the temperature field inside the RBC flow is compared and assessed. The input data are freely available for download at the online portal, where the automatic evaluation of the submitted results is conducted and displayed as shown in the present work: https://onera-flow-benchmarks.fr/en/Second_LPT_DA_Challenge. Here we consider the status of submissions until April 14th, 2026.

1. Introduction

In many fluid dynamic applications, Eulerian fields of velocity, pressure, or related quantities are required (often on Cartesian grids) in order to compute derivatives, allowing to identify vortical- and shear flow topologies or other coherent structures using the 3D velocity gradient tensor (Meneveau 2011). In this respect, the use of LPT creates the challenge of inferring continuous or gridded Eulerian fields from discrete and randomly distributed particle trajectories. The earliest strategies for this task relied on interpolation, using, e.g., polynomials or adaptive Gaussian windowing (Agui and Jimenez 1987). While straightforward and efficient, these methods tend to be sensitive to noise and modulation of the data. To address these shortcomings, physics-based constraints have been incorporated into reconstruction schemes, re-framing the task as a data assimilation (DA) problem (He 2025).

Data assimilation already emerged in the 1950s in the field of weather prediction frameworks, which included sparse and limited measurement data in time series at discrete weather stations, based on which numerical predictions for future weather conditions have been supported (Navon 2009). With the rising computational capabilities over the past decades, various DA schemes have been developed in the field of fluid mechanics. A related method, which requires decent statistical input of (measured) flow field data is lower order modelling of PIV and PTV data using (gappy) POD, LSE or DMD methods, which allows noise filtering or identifying (quasi-) periodic flow structures and can serve as a basis for so called *super-resolution* approaches (Fukami 2021, Discetti 2023). Mathematical and numerical methods such as ensemble Kalman filtering or other variational approaches (4D- or 3DVar) (Gronskis 2013, Chandramouli 2020, Du 2023, He 2024) were implemented, mostly including adjoint CFD schemes or fully DNS/LES-based adjoint methods (Lemke 2016). Similar DA methods are dealing with hybrid numerical and experimental approaches aiming at observation induced reconstruction and state estimation (Beneddine 2017, Buchta 2021, Wang 2021), or URANS based nudging / state-observer approaches using experimental input data (Zauner 2022) or full hybrid numerical simulations using experimental observations and DNS (Suzuki 2009, Suzuki 2015). A rather different and interesting approach is using time-resolved DA results based on 3D LPT measurement data as upstream input for a subsequent DNS domain. It has been shown (Appelbaum 2021) that the time for convergence of the DNS solution with respect to large-scale content (here in a high Reynolds number turbulent boundary layer flow) can be significantly reduced.

However, in experimental fluid mechanics and aerodynamics, especially 3D LPT (Schröder and Schanz 2023) with high temporal and spatial resolution can deliver position, velocity and acceleration (material derivative) vector information along a huge number of individual particle trajectories (or 'fluid elements') inside the measured fluid flow (e.g. by Shake-The-Box processing (Schanz 2016)). Such data provide a well-suited input basis for many DA schemes allowing for a high-resolution interpolation of the time-resolved 3D velocity, velocity gradient tensor (VGT) and pressure fields by using physical constraints (Navier-Stokes-equations). A two- or four-pulse 3D LPT system (Novara 2019, 2023) can at least provide position and velocity distributions in a scattered manner from which data assimilation schemes can derive significant instantaneous flow quantities (coherent structures, vorticity, Q-values (Chong 1990) and pressure under certain assumptions) which are important for the understanding and characterization of the investigated flow field and validation of CFD methods. The resolution and accuracy of the resulting flow field interpolation is limited by the measurement accuracy and sampling of the flow features given by the mean 3D inter-particle distances of the particle trajectories in terms of Kolmogorov (or viscous) length scales, but as well by the quality and performance of the used DA algorithm, which might involve physical laws, neural networks, adjoint methods or temporal information. Such synthetic and experimental benchmark tests, e.g. performed as an international challenge with an independent jury (Sciacchitano 2021, Sciacchitano 2025) are a desired framework for comparing and discussing implementation details of the various algorithms and for inspiring further developments towards higher accuracies and processing speeds.

Related publications and online accessible benchmark test platforms demonstrate the capabilities of the various DA algorithms in comparison to each other and in outperforming simple interpolation

methods e.g. based on weighted binning or adaptive smoothing with polynomials or other functionals. Already in the past, with grid based and low-pass filtered snapshot and time-resolved Tomo PIV or sparse PTV data, many groups have developed methods that compute unsteady pressure fields (Liu and Katz 2006, Kurtulus 2007, Murai 2007, Suzuki 2009, Oudheusden 2013). By solving the pressure-Poisson-equation, 3D (or 2D) pressure fields can be estimated. The solvers are typically based on (time-series) of instantaneous PIV results and use certain assumptions, boundary conditions and various integration schemes, depending on the available flow data. An interesting report on a pressure reconstruction benchmark test using LES data of a rocket base flow for creating synthetic 2D- and 3D PIV as well as 3D LPT data can be found in van Gent 2017.

For scattered 3D LPT results, single time-step techniques for data assimilation schemes have been established using solenoidal constraints or a full incompressible Navier-Stokes-regularization. For the latter, velocity and acceleration (material derivative) of the particle tracks are often used as input values (left side of the momentum equation) for e.g. FlowFit, VIC +/-# (Gesemann 2016, Schneiders and Scarano 2016, Godbersen 2024, Sperotto 2024), or sometimes velocity only (Mons 2022, Shin 2025). Those schemes provide a continuous functional representation of the assimilated velocity vector field without additional spatial filtering on the basis of (truncated) 3D radial basis functions e.g. Gaussian shaped or B-Splines. Analytical derivatives can be computed for the VGT $A_{ij}(t)$ and 3D pressure fields are integrated by solving the Poisson equation implicitly (e.g. Huhn 2017) using non-linear optimization solvers e.g. L-BFGS (Liu and Nocedal 1989). These methods can be improved by exploiting the available time-series of LPT data by using advection models with smoothness constraints for du/dt and $d\omega/dt$ of the VGT (Jeon 2023), finite time-segments of particle trajectories (Scarano 2022) or advection of virtual particles from time-step to time-step (Ehlers et al 2020). A more recent development doing so is physics informed Gaussian process regression (GPR) (Lawson 2026), which can be implemented, similar to the previously mentioned methods, either in a (faster) linear approach by penalizing divergence of the resulting velocity field only or in a non-linear way by additionally enforcing Navier-Stokes momentum equation. Furthermore, GPR quantifies prediction uncertainty and provides physics-constrained estimates of the two-point velocity covariance, a quantity of primary interest in turbulent flow characterization.

A relatively new and fast developing field of DA which can use 3D LPT data are physics-informed neural networks (PINNs) (Raissi 2019/2020; Cai 2021, Zhou and Grauer 2023, Clark di Leonie 2023 and others) which fit a solution of the inverse augmentation problem along a full time-series of trajectories (typically > 10 time-steps) while penalizing deviations to the velocity (and acceleration) measurement data at particle track positions (u_p and a_p), to continuity ($\text{div } u = 0$) and to the Navier-Stokes-momentum equation. The augmentation results are generally smoother in space and time and have increased accuracies with respect to former and single time step methods (Zhou 2024, Shin and Schröder 2025), but on the other hand PINNs require training or better spatial-temporal fitting of the data inducing relatively high computational costs on GPUs.

Another recent study by Du 2023 compares adjoint-variational data assimilation and PINN approaches. Variational methods (such as 4D-Var) aim to adapt flow models precisely to real observational data. Here, a physical state is optimised in such a way that the deviation between the

simulated flow field and measured data is minimised, while physics constraints using Navier-Stokes equations are applied. It is found that the rigorous enforcement of physical laws and time causality stabilizes the 4D-Var solution, providing increased accuracy and a superior ability to forecast solutions to the future with respect to PINNs. On the other hand, Zhou 2024 demonstrated a superiority of PINN over FlowFit and VIC# reconstructions, especially at lower particle densities (under-sampling of flow features by LPT). However, the computational costs of a 4D-Var solution are much higher than that of a PINN, which is around two orders of magnitude higher compared to the one of FlowFit3.

Benchmark tests can deliver quantitative feedback during development steps and if being freely available on public web-portals they can build up a competitive driving force for enhancing reconstruction quality of various LPT and DA reconstruction algorithms under development in various groups and companies all over the world. In this work, we expand on the results of the previous DA challenges (Leclaire 2021; Sciacchitano 2021; Sciacchitano 2025) by considering three different test cases (three synthetic ones), where relevant parameters are varied, namely: the volumetric particle density, the timing strategies and temporal separations between synthetic particle acquisitions.

Therefore, an independent benchmark test portal for both tasks has been established (first during the H2020 EU Project HOMER (GA ID: 769237)) at the website <https://www.onera.fr/flow-benchmarks>. Input data can be downloaded freely from the online portal. Here, various flow test cases based on DNS or LES simulations are used to either create time-series of synthetic particle images based on projections of advected 3D Lagrangian tracks onto four virtual cameras, which can be downloaded together with camera calibration information for the LPT Challenge cases or the 3D particle tracks are used directly as input data for the respective DA challenge cases. Typical position and image noise is added onto the tracks and images in order to mimic experimental boundary conditions for both benchmark categories. Details on the online portal are reported in another contribution submitted to this Lisbon symposium (Leclaire 2026). The results of each benchmark test can be uploaded at the portal in a predefined format and an automatic evaluation script is comparing the uploaded reconstruction results with the “ground truth” tracks or the respective flow field data of the original DNS/LES simulations by using standard error metrics, spectra and flow field volume slices, which are provided to the respective participant. In case the participant agrees, the error report is displayed and ranked with respect to other participants at the portal. Since the 2nd Challenge (Tokyo 2025 in conjunction with the ISFV and ISPIV conferences) three DA benchmark test cases are available: Case 1 contains two-pulse (TP), four-pulse (FP) and time-resolved (TR) sequences of ideal tracers advected in an incompressible turbulent channel (rectangular duct) flow at various seeding densities. Case 2 consists of TR sequences of advected particle tracks in a compressible turbulent boundary layer flow at Mach 4.9. Parameters explored are the separation time between pulses, the seeding density, and inertial or non-inertial particles. Case 3 consists of TR sequences of advected particle tracks in Rayleigh-Bénard Convection at $Ra = 10^7$ and various seeding densities. More detailed descriptions of the test cases can be found at the portal website, where more test cases will be uploaded in the future, and in Leclaire 2026.

1.1. Participants

The list of research groups who submitted their contributions to the different test cases of the 2nd DA challenge until April 14th, 2026, together with the references to their algorithms, is summarized in Table 1.

Table 1. List of research groups that contributed to each DA Case 01 with references to their algorithms.

Test Case	Research Group	Algorithm	Reference
DA Case 01 -TP	Dantec Dynamics	PII	Sperotto et al. (2024)
	DLR Göttingen	FlowFit3-linear	Godbersen et al. (2024)
	LaVision GmbH	FlowFit3-linear	Godbersen et al. (2024)
	Southampton Univ.	PIGLET-linear	Lawson (2026)
	Southampton Univ.	PIGLET	Lawson (2026)
DA Case 01 -FP	DLR Göttingen	FlowFit3-linear	Godbersen et al. (2024)
	DLR Göttingen	FlowFit3-non-linear	Godbersen et al. (2024)
	LaVision GmbH	FlowFit3-non-linear	Godbersen et al. (2024)
	Penn State Uni	NIPA (PINN)	Zhou and Grauer (2023)
	Southampton Univ.	PIGLET-linear	Lawson (2026)
	Southampton Univ.	PIGLET-non-linear	Lawson (2026)
DA Case 01 -TR	Auburn Univ.	PINN	Clark di Leoni (2023) + more
	Dantec Dynamics	MTA / PII	Sperotto et al. (2024)
	DLR Göttingen	FlowFit3-linear	Godbersen et al. (2024)
	DLR Göttingen	FlowFit3-non-linear	Godbersen et al. (2024)
	DLR Göttingen	PINN	Shin and Schröder (2025)
	LaVision GmbH	FlowFit3-non-linear	Godbersen et al. (2024)
	ONERA	State Observer/Nudging	Zauner et al. (2022)
	Southampton Univ.	PIGLET-linear	Lawson (2026)
	Southampton Univ.	PIGLET-non-linear	Lawson (2026)
	TU Berlin	PINN	Steinfurth et al (2024)
DA Case 02 -TR	No contribution (yet)		
DA Case 03 -TR	DLR Göttingen	PINN	Mommert et al. (2024)
	Penn State Uni	NIPA(PINN)	Zhou and Grauer (2023)

2. 2nd DA challenge Case 01

2.1. Test case description of Turbulent Channel Flow

The first test case, provided by the group of Delft University of Technology, consists of two-pulse (TP), four-pulse (FP) and time-resolved (TR) synthetic particle track data advected along a time-series of DNS of a turbulent flow in a square channel. This test case considers the square duct DNS simulations of Pirozzoli et al. (2018) to study a fully developed turbulent flow. The Figure 2-1 below illustrates the simulation results, conducted at a bulk Reynolds number of $Re_b = 2hu_b/\nu = 40,000$ and a friction Reynolds number Re_τ approximately equal to 1,000, where h denotes the duct half-width, u_b is the bulk velocity and ν the fluid's kinematic viscosity. This test case of the 2nd DA challenge is requesting the reconstruction of the velocity and pressure fields from scattered positions of particles

along tracks. Simulated 3D track results from TR-, FP- and TP- recordings are considered. For each timing strategy, seeding density is varied over 0.001 ppp, 0.01 ppp, 0.05 ppp and 0.2 ppp according to 1,000, 10,000, 50,000 and 200,000 particles advected inside DNS of the turbulent channel flow sub-volume indicated in Figure 2-1 by advanced temporal 4th-order Runge-Kutta schemes and higher-order spatial interpolation along the 100-, 4- and 2- time steps each. Details of some parameters of the synthetic LPT experiment as input for the respective DA test cases clarifying the meaning of ppp-levels for particle densities inside the flow are shown in Table 2.

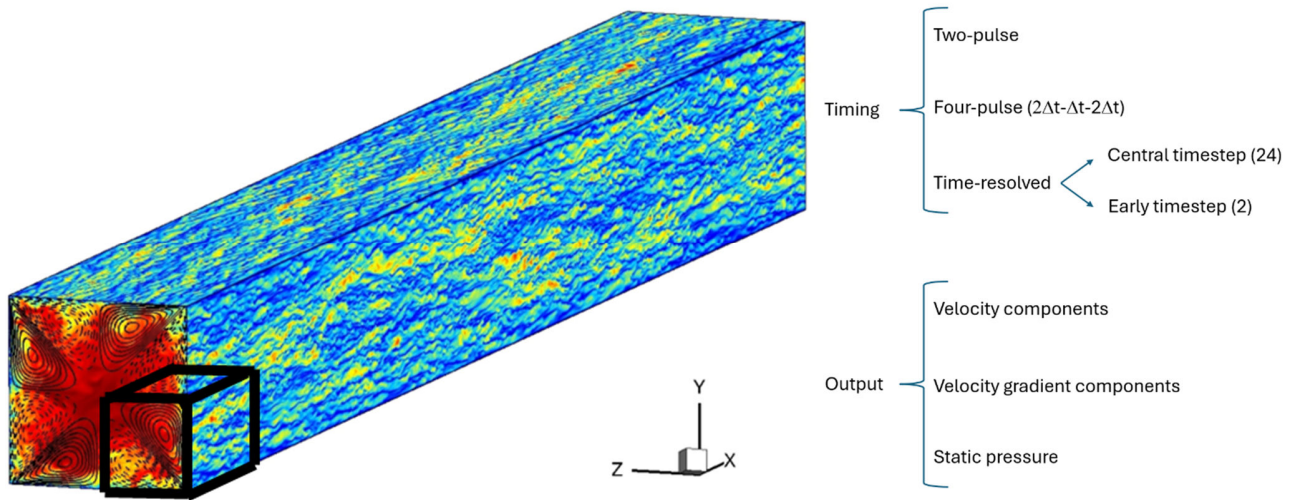


Figure 2-1. Visualization of the square duct flow DNS simulations from Pirozzoli et al. (2018) with indicated sub-volume for DA case01(left) and three timing cases for provided LPT input data and requested output variables (right)

Table 2. Main parameters of the DA Case 01 synthetic experiment setup.

Parameter	Value
Friction Reynolds number	1,000
Number of cameras	4
Sensor size [px]	1024×1024
Pixel pitch [μm]	10
Focal length [mm]	105
$f\#$	22
Imaged volume [mm]	100×100×100
Number of particles	1,000; 10,000; 50,000; 200,000
ppp	0.001, 0.01, 0.05, 0.2
Position noise	Non-disclosed

For the challenge, a domain of $h \times h \times h$ in the x , y and z directions is considered, where x denotes the streamwise direction, and $y = 0$ and $z = 0$ are the duct walls. A virtual experiment is conducted in water ($\rho = 1000 \text{ kg/m}^3$, $\nu = 10^{-6} \text{ m}^2/\text{s}$), with $h = 100 \text{ mm}$ and $u_b = 0.2 \text{ m/s}$. A cloud of virtual particles, hypothesized as fully passive tracers, has been propagated within the DNS simulation.

2.1.1. *TR cases*

A sequence of 50 consecutive frames (starting at 0) is provided for each seeding density case. The time separation between snapshots is 5.916 ms.

2.1.2. *FP cases*

The particles belonging to each of the four pulses of the sequence are provided for every seeding density case. The timing template of the four-pulse acquisition is 2-1-2, with the unitary separation of 5.916 ms. This means that the time separation between the first and second pulse is 11.832 ms, the separation between the second and third pulses is 5.916 ms, and finally between the third and fourth pulses is again 11.832 ms.

2.1.3. *TP cases*

The particles belonging to each of the two pulses of the sequence are provided for every seeding density case. The time separation between pulses is 5.916 ms.

2.2. Requested output and formatting rules

The results need to be prescribed on a Cartesian grid with constant spacing of 1 mm in each direction. The output variables required on this grid are the following (total of 13 variables):

- Three velocity components: VX , VY , VZ in m/s
- Nine velocity gradient components: $dVXdX$, $dVXdY$, $dVXdZ$, $dVYdX$, $dVYdY$, $dVYdZ$, $dVZdX$, $dVZdY$, $dVZdZ$ in 1/s
- Static pressure: p in Pa, where the pressure variable refers to the relative pressure with respect to the point $(x,y,z) = (0,0.1,0.1)$ mm.

2.2.1. *TR cases*

The flow field at timestep 24 (from the sequence starting at 0) needs to be supplied for each seeding density situation. Providing the output also at time-step 2 is optional.

2.3. Results DA Case 01

Here we restrict our presentation of the assessment of DA Case 01 to the time-resolved (TR) case only, because this attracted the largest number of participants resp various DA codes (in total 12) with the possibility to compare their performances. In the following short descriptions written by the individual DA code developers before the Tokyo 2nd LPT and DA challenge event took place in June 2025 are summarized (with some editorial changes) in order to give some insights into the respective DA architectures and parameter choices. More details can be found in the respective literature given in Table 1 or the ones indicated at the web-portal.

The **Auburn University PINN** follows a conventional framework with locations in space-time where velocities are known (x, y, z, t) as input and the corresponding fluid velocity and pressure (u, v, w, p)

as output. A fully connected neural network connects the input to the output nodes. Auto differentiation is used to calculate the gradients of these output variables with respect to the input variables to be used in the calculation of the loss function. The loss function consists of PDE loss, data loss, and boundary loss. The overall loss is backpropagated to optimize the network parameters over multiple iterations. Given the particle positions for 50-time instances for each case of data, a second-order central difference scheme is used to determine the velocities. This forms our input and output training data points. Table 3 shows the hyperparameters used in the PINNs to process the data. Auburn PINN uses min-max normalization for the inputs, which takes the minimum and maximum value for each coordinate and rescales the coordinate between -1 and 1. For output, they used z-score normalization, where in we normalize the output velocity with the mean and standard deviation of velocity data of each component. To enhance the network's ability to learn high-frequency features, we employ a multi-scale fourier feature encoding for transformation of the input coordinates. A set of frequency matrices is generated using normally distributed random values at seven predefined scales ranging from 0.1 to 25.0, with each scale containing 32 frequencies. The input vectors are projected through these matrices and transformed using sine and cosine functions to produce periodic features, which are then concatenated and fed into the neural network.

Table 3. Hyperparameters used to train the Auburn Uni PINN model

Hyperparameter	Value	Remark
Network size and depth	100 * 8	
Activation	tanh	
Number of epoch	1200	
Batch size	300, 500, 1000, 3000	Different for each case
Optimizer	Adam	
Normalization	minmax and z-score	
Fourier encoding	32*7*2 fourier features	

The neural network architecture utilized in this model consists of 8 fully connected (dense) layers, each comprising 100 neurons. To introduce nonlinearity and enable the network to learn complex functional mappings, each layer employs the hyperbolic tangent (tanh) activation function. The training was conducted using the Adam optimizer with a cosine decay restart learning rate scheduler. The initial learning rate was set to 0.001 and followed a warm restart strategy. This approach, with a decay multiplier of 0.90 and a cycle stretch factor of 1.2, ensured a gradual reduction of the learning rate while periodically allowing exploration with higher rates. The network was trained over 1200 epochs using mini-batch stochastic gradient descent with a different batch size for different cases with varying particle density. They divide the PDE loss into individual terms such that the final loss term is the sum of the individual ones including scaled weight coefficients in order to mitigate gradient imbalance between multiple loss terms. Because the order of magnitude of the gradients of the loss terms is different for individual loss terms, they scale the weights of the other loss terms with respect to the continuity loss. To ensure full coverage of the domain, residual points are randomly sampled from a continuous uniform distribution in both space and time within the measurement domain for the calculation of the physics loss. We use dense sampling in our model, assigning a different number

of collocation points than data points, with the number of collocation points for each case varied according to the local particle density.

Dantec Dynamics used Meshless Track Assimilation (MTA), a Data Assimilation (DA) method that aims to find the sweet spot between measured data and underlying physics, like continuity and momentum conservation (see more details in Sperotto 2024). For what concerns the specific computation strategy, no smoothing or other preprocessing is applied to the tracks as the method has a built-in time smoothing penalization. For all the seeding densities the 30 central-most time-step were optimized together with our DA method. The typical number of bases per particle is 5 for all the test cases. Continuity (or divergence-free) is no longer penalized but hard constrained in the computational grid points. This allows us to get rid of the penalization regarding the time derivative of the divergence. Furthermore, also the Poisson equation is regarded as not necessary anymore. The gradient is properly projected into a divergence-free space. Due to the divergence free optimization, the variables concerning the velocity per grid point are reduced from three to two. Moreover, a similar reduction can be applied to the pressure, the pressure appears only in the momentum equation having a linear relationship. This means that at any gradient query the pressure can be optimized directly and independently exploiting the linear relationship. As the pressure is optimal at any given gradient query it can be shown that the pressure does not influence the Jacobian of the velocities, but it does influence the residuals. Meaning that the final optimized variables are just 2 arbitrary velocity directions, for the given case Z and Y were selected. This approach stabilizes the whole process, avoiding the pressure getting stuck in sub-optimal solutions, which is common for small static pressure flows.

DLR uses **FlowFit3** linear and non-linear as described in Godbersen 2024 and a **PINN** described by Shin and Schröder 2025 for the TR DA Case 01. DLR **FlowFit** consists of B-Spline grid-based representation of the domain which is then optimized to fit the particle data using L-BFGS subject to incompressible Navier-Stokes and smoothness constraints. The method applies to individual time-steps (not a 4D method) and expects position and velocity values of the particles, or additionally their acceleration depending on its mode of operation. Nonlinear mode additionally incorporates the acceleration data via the momentum equation and is able to compute pressure as well. For the two-pulse data a simple linear fit was performed. Due to the lack of acceleration data only the linear mode FlowFit was then applied. The four-pulse tracks were subjected to a second order polynomial fit also giving acceleration data. The time-resolved data was processed using TrackFit and evaluated at the desired time step. In order to enforce the no-slip wall boundary conditions on two sides of the domain we additionally added artificial stationary particles at those walls in all variants. FlowFits were conducted using a B-Spline grid spacing of 0.6 mm, a curvature penalization factor of 0.1, and a L-BFGS termination criterion of 0.005 relative change in the loss. For DLR **PINN** first, particle trajectories were reconstructed using the (local) TrackFit method, which employs B-spline-based curve fitting to reduce trajectory noise. Second, Cartesian Grid Reduction (CGR) was performed using a Physics-Informed Neural Network approach. The PINN utilized a Multi-Layer Perceptron (MLP) architecture consisting of 15 hidden layers, each comprised of 300 neurons with hyperbolic tangent (tanh) activation functions. Network weights were initialized using the Glorot uniform

initializer, and biases were initialized to zero. To enhance training stability, weight normalization was applied. Before feeding data into the network, inputs underwent standardization followed by a Fourier feature mapping using randomly selected, fixed frequencies. Throughout optimization, additional collocation points were adaptively introduced into regions of the domain where physical constraints exhibited large residuals, evaluated every 10,000 iterations. The batch size was consistently set to 10,000 points, covering data points, governing equations, and boundary conditions. Optimization was conducted using the Shampoo with Adam in the Preconditioner's eigenbasis (SOAP) optimizer with an exponential moving average coefficient of 0.95 and preconditioning frequency set to 5. The initial learning rate was 0.001, decaying by a factor of 0.9 every 20,000 iterations. Loss balancing factors are 1, $1e^{-6}$, $1e^{-6}$, $1e^{-7}$ for data loss, boundary loss, divergence free loss and incompressible Navier-Stokes loss respectively. Training took 8 hrs with Nvidia A100 GPUs.

LaVision implemented FlowFit3 (Godbersen 2024) in a homemade code, while linear and non-linear modes are available. All weighting factors should be given in advance since they have not been automatically decided.

Boundary penalization:

- Cost functions are introduced at grid meshes crossing the wall using the following terms:
 - Velocity components at the wall position
 - Wall-normal pressure gradient at the wall position

Coarse grid approximation into FF3:

- It shares the same concept implemented in VIC# (Jeon 2022).
- FF3 is conducted for multiple grid spacings, from large to final, where the grid spacing is halved between each reconstruction. After sampling and divergence-free filtering, a velocity field from the coarse grid is used as an initial condition for FF3 on the finer grid.

Vorticity time-marching into FF3 only for the TR case, $t = 24$:

- It shares the same concept implemented with VIC# (Jeon 2023)
- Cost functions regarding the vorticity vector are introduced at grid points.
- The vorticity vector at neighbouring time instances is evaluated using the Taylor series with its first and second-time derivatives.
- Spatial derivatives of the velocity for the vorticity time derivatives are evaluated using the basis function without discretization.

The computation grid points are identical to those where a result is requested. One grid is extended to the y- and z-axis wall direction, i.e., $101 \times (100 + 1) \times (100 + 1)$. Therefore, the computation domain contains the wall. The extended grids beyond the wall were not included in the submitted result.

ONERA used a data-assimilation procedure that may be classified as a state observer/nudging technique (Zauner 2022) and consists in adding a forcing/feedback term $\mathbf{f}$ in the momentum equations to the (incompressible) Navier-Stokes equations. This forcing term is built in two main

steps. We start from pointwise feedback term $\mathbf{f}_p$ that would only act at the particle locations. An adjustable coefficient that allows us to adjust the intensity of the feedback term is tuned to find a compromise between numerical stability and reaching low discrepancy levels with respect to the data. To obtain the LPT velocity data from the particle positions provided as input to the DA Challenge, we used a third-order temporal polynomial fit (Savitzky-Golay), over 7 instants. Feedback term $\mathbf{f}_p$ cannot be used as such in the nudging technique, as it would introduce discontinuities in the numerical simulation. In order to prevent the feedback term from directly affecting the pressure field, we then remove its potential part which provides the (divergence-free) actual feedback term $\mathbf{f}$. The numerical resolution of equations relies on a 3D incompressible Navier-Stokes solver which has been developed based on the software *FreeFEM++*. The total number of grid points in the streamwise direction and at the wall boundaries is here $228 \times 224 \times 224$, versus $2048 \times 512 \times 512$. The time step of the computations is taken equal to the time interval between two consecutive LPT snapshots. The simulations are initialized from an arbitrary flow solution that was obtained from a previous calculation of the same configuration that reached statistical convergence. They are performed over a time horizon that coincides with the time interval between the 0th and 24th LPT snapshots.

Penn State uses neural-implicit particle advection (NIPA) as a PINN to recover flow states from unfiltered, noisy LPT tracks (Zhou and Grauer 2023; Zhou2025) (not submitted to the TR DA Case 01, but to the FP Case (not shown here) and the TR DA Case 03 (RBC) before April 14th, 2026). NIPA has two modules: neural flow states and kinematics constraint tracks (KCTs), which represent the flow fields and particle states, in turn. Flow states (e.g., velocity, pressure, density, etc.) are parameterized by one or more neural networks. Fourier encoding is adopted to combat the low-frequency bias of gradient descent-type optimizers. Partial derivatives of the flow with respect to space-time coordinates, obtained by automatic differentiation, are used to evaluate the governing equations; the equation residuals are integrated over the flow domain and measurement interval and added up to form a “flow physics loss.” A KCT is a trainable function, e.g., a linear combination of radial basis functions or similar, that describes a particle’s velocity over time. The function is constrained to integrate to the tracked particle locations. Integration and differentiation of the KCT yield the particle’s position and acceleration history, respectively. A “particle physics loss” is computed by adding integrated residuals from a suitable equation of motion, such as the Maxey–Riley equation, for each particle. For ideal tracers, the Maxey–Riley equation reduces to equality between flow and particle velocity. Simultaneously minimizing the flow and particle physics losses yields flow states that approximately solve the governing equations, even when there is particle slip (Zhou and Grauer 2023). Position variables in the KCTs are included as trainable parameters in the optimization, and a “data loss” is added to the flow and particle physics losses. The data loss may be based on weighted residuals in 3D space, or it can be formulated in terms of image residuals. Optimizing all three losses, together, amounts to synchronous flow reconstruction and physics-enhanced particle tracking. For the turbulent channel flow case, we employ two distinct neural networks: a flow network to infer the three velocity components (u, v, w), and a pressure network to infer pressure p . Under higher particle densities (0.2 and 0.05 ppp), each velocity network comprises 25 hidden layers with 500 neurons per layer, while each pressure network has 25 hidden layers with 250 neurons apiece. In scenarios with

lower particle densities (0.01 and 0.001 ppp), network complexity is reduced to 15 hidden layers with 400 neurons and 10 hidden layers with 300 neurons for velocity, and 15 hidden layers with 200 neurons and 10 hidden layers with 150 neurons for pressure. All networks are augmented with a Fourier encoding with frequencies drawn from zero-mean Gaussian distributions (standard deviations of unity for spatial and 0.2 for temporal features, using 1024 frequencies) to enhance expressivity. Particle tracks are split into fixed-length segments of 17 positions to form KCTs. A non-slip boundary loss is added along the bottom and vertical wall. Training utilizes the Adam optimizer, employing a learning rate of 10^{-3} . Training continues until convergence, generally requiring approximately 30,000 epochs and the computational cost is around 15 hours.

Southampton University did not participate at the Tokyo event 2025 but now at the recent call for submission until April 14th, 2026 considered here for presentation. Therefore, there is no short description from the author(s) of the code available. We have to refer here to the introduction and more detailed descriptions in Lawson 2026.

TU Berlin used a Fourier-embedded physics-informed neural networks for assimilation of particle tracking velocimetry data. The employed algorithm represents an extension of the method introduced in Steinfurth 2024. The provided synthetic particle tracks were linearized to extract the velocity at each timestep. For the submitted results, a fully-connected neural network (4 layers with 256 neurons each) is optimized to map the coordinates (x, y, z, t) to the requested output parameters (u, v, w, p) . The training is constrained by using three timesteps of the extracted synthetic PTV data on the one hand and residuals of the three-dimensional incompressible Navier-Stokes equations on the other. Furthermore, a no-slip condition was enforced by penalizing non-vanishing velocity at the channel walls. The three loss terms (i.e., the data loss, the physics loss and the boundary loss) were balanced by applying a weight $\lambda = 0.01$ to the physics loss so that the overall loss becomes a weighted sum of the individual losses. The minimization of the overall loss was handled by Adam optimization (10000 epochs) before refining the network with LBFGS optimization (4000 epochs). It is important to mention that the data loss was formulated in such a way that the absolute error associated with each of the three velocity components is minimized; alternatively, the minimization of relative errors would also be possible with the same framework. In an extension of the method introduced in Steinfurth 2024 Fourier features were embedded to map the input vector $\mathbf{x}$ to a higher-dimensional space using a sinusoidal transformation, where a range of random frequency components sampled from a Gaussian distribution, which was defined differently for space and time coordinates based on the expected frequency content. Specifically, 256 Fourier features were incorporated, sampled with standard deviations of $\sigma\mathbf{x} = 0.5$ and $\sigma t = 0.05$ for space and time, respectively. The transformed input layer features higher-frequency components compared to the raw input data which enables the network to learn more complex functions whereas standard networks favor learning low frequency components.

In the following we present the results of the various DA schemes from the participants for the time-resolved (TR) DA test case 1. First, in Figure 2-2 we show the random velocity magnitude error (mean RMS) of the various submitted results with respect to the ground truth DNS at a large number of predefined grid points. One can see that if there are very sparse data available at $\text{ppp} = 0.001$ the

resulting interpolations do have the widest spreading of errors, while the two PIGLET versions tackle this massive under-sampling problem best, a bit better than the TUB_PINN solution.

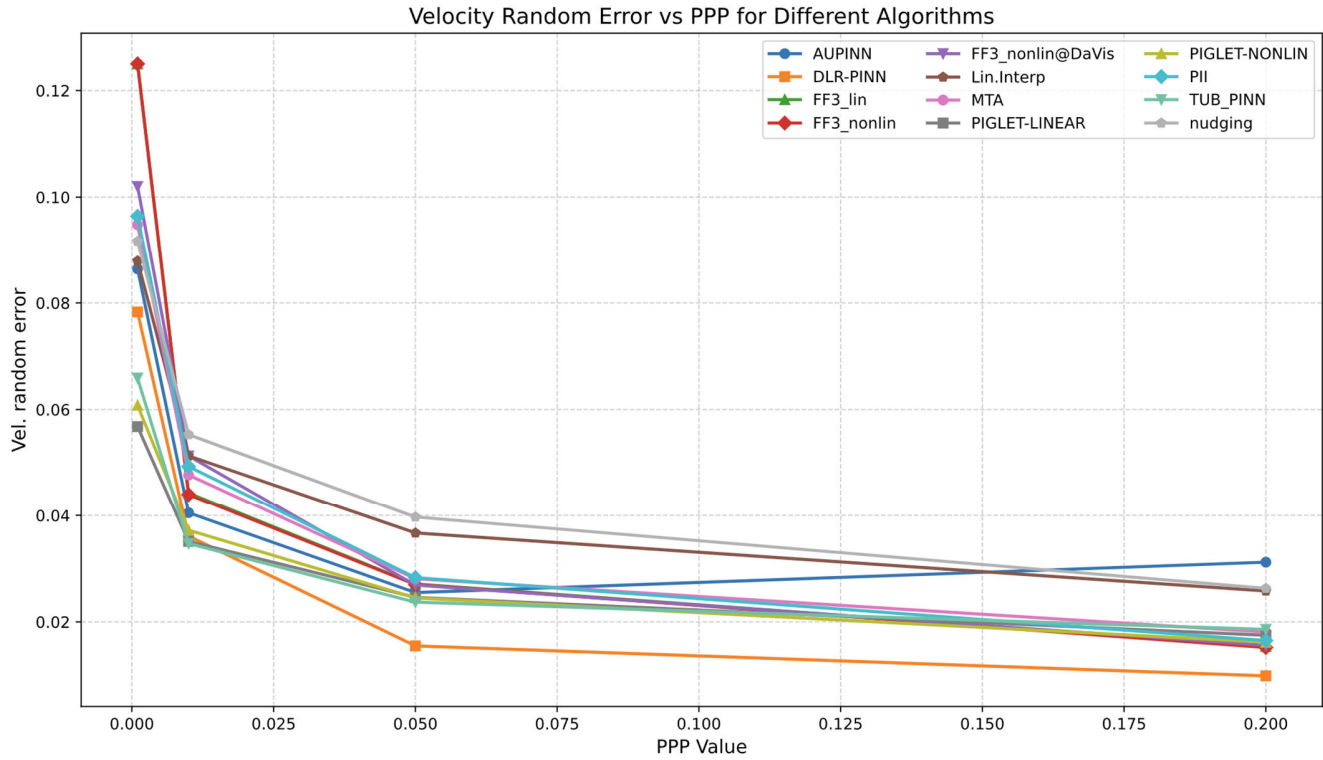


Figure 2-2. Random error of reconstructed velocity magnitudes of the various DA schemes with respect to the ground truth over the four volumetric particle track densities given in ppp (x-axis). For visibility the 10 curves (one being the linear interpolation of GT data) are colour coded by algorithm

Beyond values of $ppp > 0.01$, which represent typical measurement conditions with LPT systems, the variation of errors in between the various DA algorithms is narrowed and quite similar. Nevertheless, beside the fact that most algorithms benefit significantly from the increase of provided volumetric LPT data up to the value of $ppp = 0.2$, some are resulting towards larger improvements with respect to the velocity random error. For 0.05 ppp and 0.2 ppp DLR PINN is a visible step ahead all others in reducing the error. The majority of DA algorithms perform quite similar for these two higher ppp-values with small discrepancies visible in the graphs of Figure 2-2. Probably, some artifacts lead to more deviations for single algorithms or further fitting regarding the choice of (hyper-)parameters need to be performed, which might improve individual results. For the high $ppp = 0.2$ values the differences between PINNs and other Constraint Cost Minimization algorithms diminishes almost, beside the DLR PINN. In a former benchmark test (Zhou 2024) using a boundary layer flow the NIPA PINN approach performed always better at low ppp-values with respect to random velocity- and vorticity errors than the two DA methods FlowFit and VIC#. Typically, always a bunch of parameter settings need to be optimized for all DA methods, which are not known a priori and should be a controllable set of parameters for experimental data. A way to consider the statistics of the LPT input

data e.g. the velocity co-variances for optimizing such DA parameters is paved with the PIGLET methodology (Lawson 2026), which performed quite stable over the range of ppp values in the present test TR DA Case01.

In general, for statistically stationary flows the measured 3D LPT data can be used to calculate one- and two-point statistics via high-resolution binning methods, which can provide the mean velocity field and its gradients, corresponding turbulent kinetic energy / RMS fields, space-(time) correlations and related local wave-number spectra. Furthermore, velocity- (acceleration) structure-functions enable estimating local Kolmogorov- and Taylor space- and time scales, which in turn can be used for optimizing DA parameters, like grid resolution, curvature penalisation, local anisotropies of the VGT, or advection methods for exploiting the temporal domain etc..

Figure 2-3 shows the instantaneous distributions of the error velocity magnitude in a transversal cut through the individual DA results of TR Case01 with colour coded deviations to the DNS ground truth at ppp = 0.05. In the instantaneous deviation plots it is clearly visible that almost all DA codes can handle the bulk region of the turbulent duct flow at the given volumetric track density, because only limited velocity gradients are present. The various methods differ specially in the instantaneous shear regions with local high values of vorticity, specifically close to the two walls. While the interpolation on LPT ground truth data suffers from missing information along all boundaries of the domain, all other methods used flow assumptions, like no-slip boundary conditions at the wall and smooth boundary conditions inside the flow, to cope with the lack of particle tracking data at the borders. Additionally, the constraints of incompressible Navier-Stokes equations used for all other methods overall improves the instantaneous velocity field interpolation.

One can see that even for the relative high level at ppp = 0.05 or relatively dense sampling of the flow field by particle trajectories, the small scales of the velocity (gradient) field, especially close to the wall, and of the vorticity field in the shown cut cannot be fully recovered by the various DA methods, including those additionally incorporating the temporal domain for their augmentation results. For the cases at lower ppp-levels the low-pass filtering effects on the reconstructed flow results become even more prominent due to the larger 3D interparticle spacings with respect to the local Kolmogorov length scale. Resolving all scales of the flow at high Reynolds numbers is always a challenge due to the increasing temporal- and spatial scale separation with increasing Reynolds number. Therefore, the provided test cases are shaped such that they mimic state-of-the-art hard- and software applied to typical experimental 3D LPT conditions. Nevertheless, some differences are still visible regarding the reached spatial resolution: the URANS based nudging show a clear trend of spatial modulation towards low wave numbers, while most of the various DA methods chose similar grid resolutions. Some show better performance in adapting the local velocity maxima and minima, while suffering a bit from spatial modulation, while others show a trend vice versa. However, the method with the lowest error overall in the bulk and close to the strong shear regions and the local vortices is provided by the DLR PINN, while there is still a reasonable low deviation to all other contributors. Especially, with respect to the computational effort one need to decide for a trade-off between desired accuracy and spatial resolution and the computational effort related to the various DA codes.

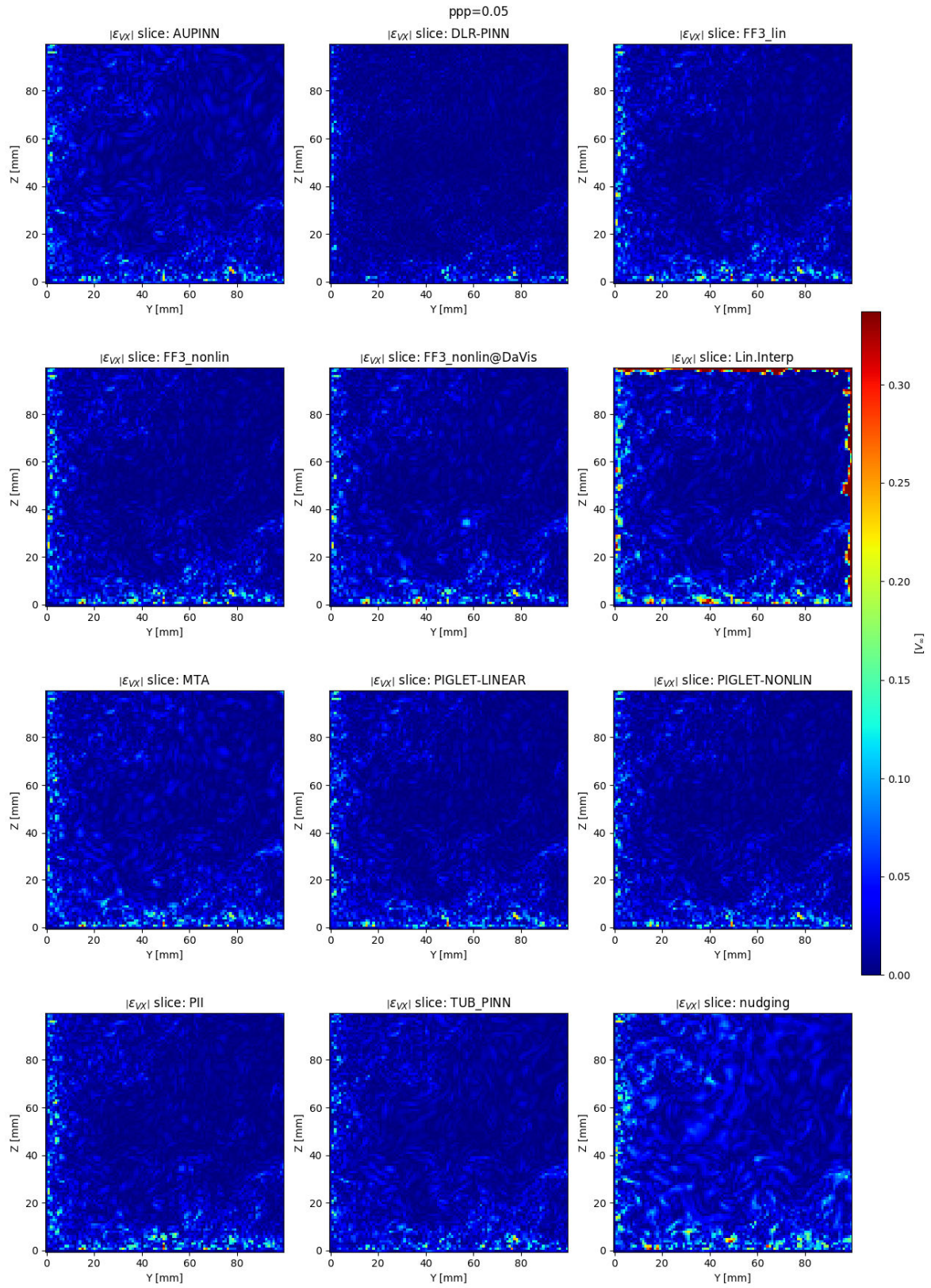


Figure 2-3. Transversal cut of velocity errors of the individual DA results through a y-z-plane, x- in flow direction, left and bottom are the channel walls, showing deviations to DNS in colour coded velocity magnitude at ppp = 0.05.

Considering the random error of the vorticity magnitudes with respect to DNS is a more sensitive indicator of how well the spatial resolution of the interpolated velocity field has been provided by the DA algorithm. 3D vorticity estimation requires that local velocity gradients in all directions can be derived accurately. The respective graphs of the results provided by the individual algorithms are displayed in Figure 2-4. The spreading of the various DA methods is more pronounced here as solutions with slight low pass filtering effects or data including high-frequency noise are penalized due to the amplification of the errors by computing spatial derivatives.

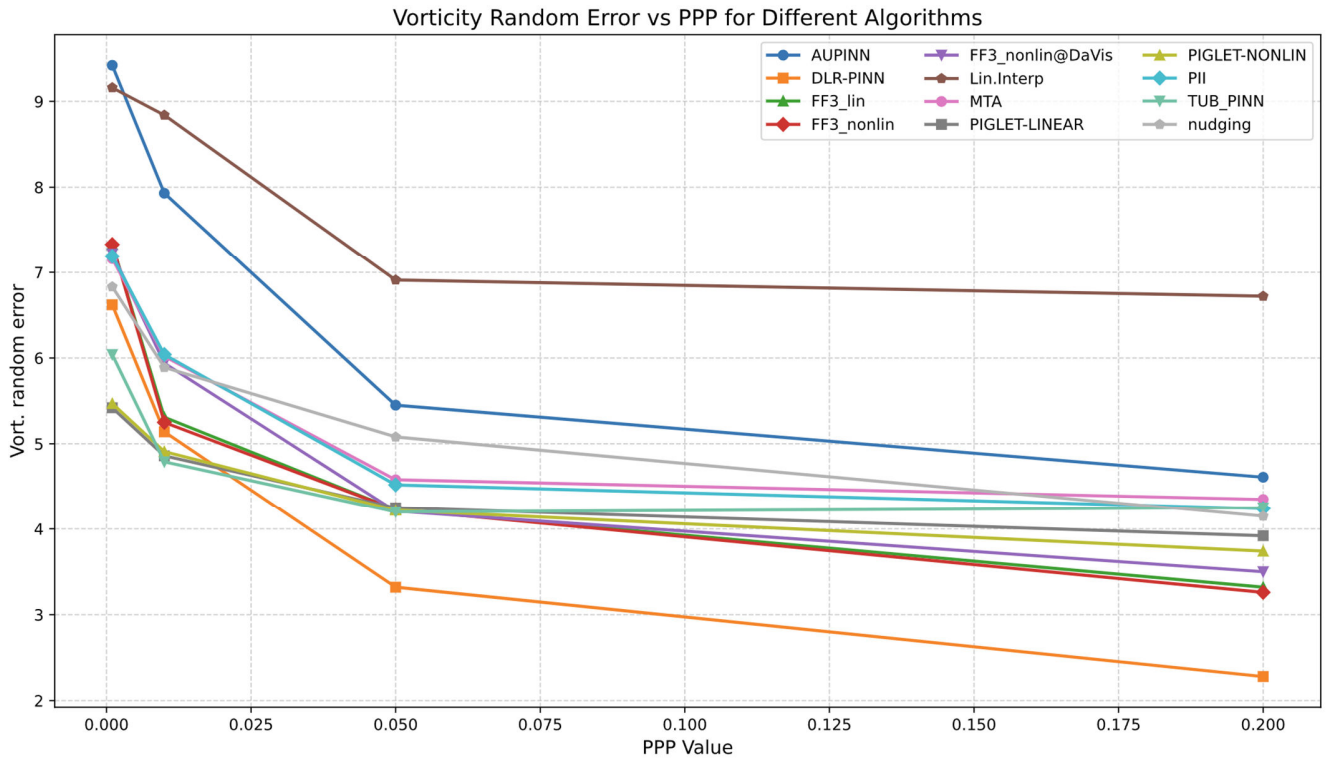


Figure 2-4. Random error of reconstructed vorticity magnitudes of the various DA schemes with respect to the ground truth over the four volumetric particle track densities given in ppp (x-axis). For visibility the 10 curves (one being the linear interpolation of GT data, which is a kind of “hacker solution”) are colour coded by algorithm

One can see that for the two high values of $ppp = 0.05$ and 0.2 the DLR PINN solution show even smaller vorticity errors with respect to the other methods (including other PINNs) than for the velocity field. Despite the fact that most other DA methods perform quite similar at $ppp = 0.05$, the second-best solution at $ppp = 0.2$ is provided by the FlowFit family (similar to results shown in Zhou 2024). While other algorithms pretty close to each other at the velocity random error plots for this ppp -value (see Figure 2-2) seem to suffer from lower spatial resolution or random noise here at the vorticity random error plot. This might be improvable with adapted parameters and is not necessarily a problem of the algorithmic concept. The potential of the individual algorithms will hopefully be clearer in the future as many DA and parameter optimization methods are still under development.

From the three plots in Figure 2-2, Figure 2-3 and Figure 2-4 it is not yet very conclusive which DA strategy is best suited for gaining accurate and high resolution velocity vector volumes from (sparse) LPT data, but overall PINNs seem to have a greater potential to further enhance the DA capabilities

based on LPT input. On the other hand, long time series of 3D LPT measurement data (typically needed for converged flow statistics) require realistic (short) processing times for subsequent DA applications. A measure which should be amended in the current table (at least with respect to given (standard) hardware components (CPUs or GPUs)) in order to complement the assessment.

An important field for the assessment of DA codes in fluid mechanics are the pressure fields estimated from the embedded physical constraints given with the momentum equation of Navier-Stokes. Pressure is of particular interest due to its non-locality especially for characterization of turbulent flows, but as well as for estimating aerodynamic forces and moments on (wing) models inside the flow. In Figure 2-5 the random error of the pressure of the TR-Case 1 for the various DA methods in comparison to the DNS ground truth are given at various ppp levels. Interpolation results are computed directly on DNS pressure data, which represents a kind of “hacker-solution”.

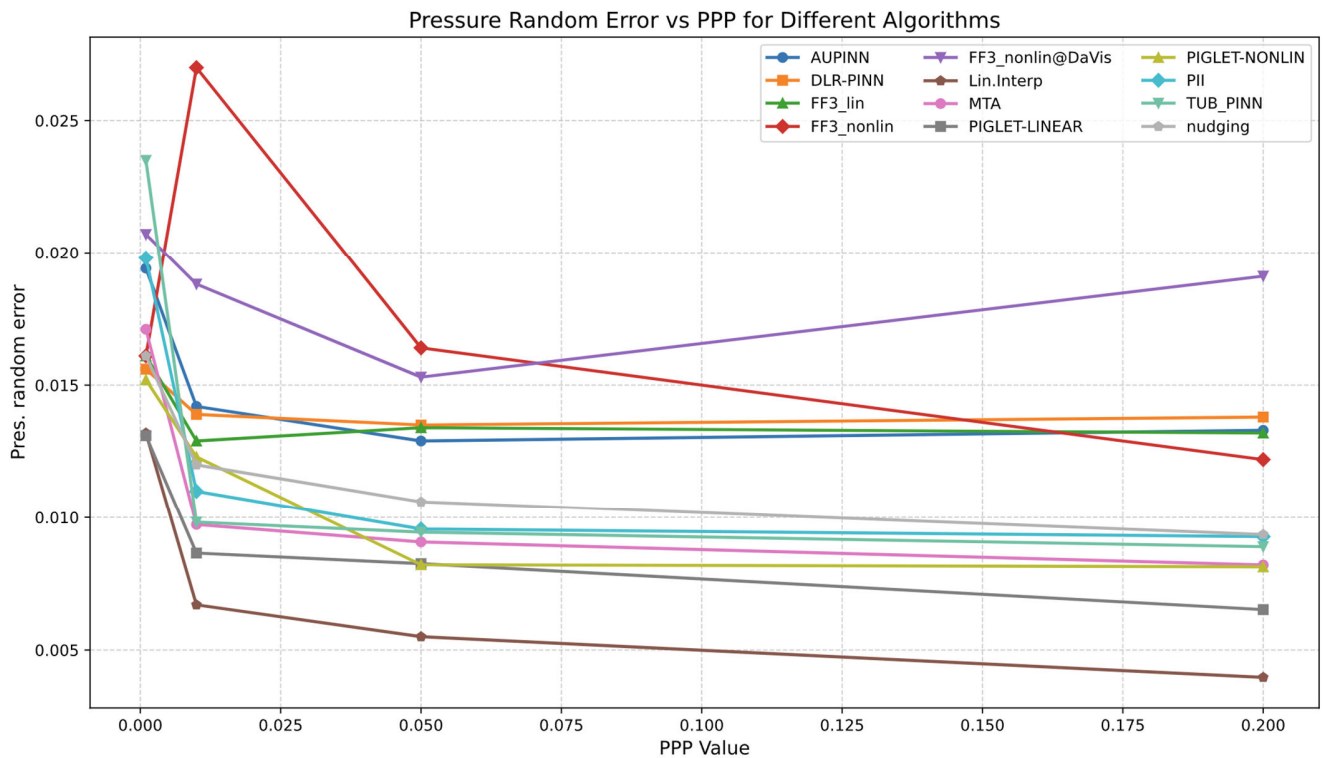


Figure 2-5. Pressure random error of the TR-Case 1 for the various DA methods in comparison to the DNS ground truth at various ppp levels. Interpolation results are computed directly on DNS pressure data, which represents a kind of “hacker-solution”.

Interestingly, those codes, like the FlowFit family or the DLR PINN which performed well in velocity- and vorticity- field estimates at high ppp-levels (see Figure 2-2 and Figure 2-4) produce poor results with respect to pressure reconstruction accuracy, which is a novelty compared to previous pressure benchmark tests (van Gent 2017, Zhou 2024). Overall PIGLET, MTA and TU Berlin PINN achieved the best results for the pressure reconstructions at all relevant ppp-levels. The specific parameter choices or implementation strategies for physics constraints regarding pressure seem to be advantageous for this test case. But still the pure interpolation of pressure on the GT DNS data at the particle positions provides the best pressure representation, which was the worst performing e.g. for

vorticity (see Figure 2-4) and is a kind “hacker solution” as it is based on direct DNS sampling. Furthermore, the correlation coefficients of the pressure fields with DNS are quite low for most approaches, but even for the successful ones not as high as one would expect especially for the high ppp-levels. A further study about the specific difficulties of reconstructing pressure using the present benchmark test TR DA case 01 for almost all applied DA methods need to be performed in detail in the near future.

One can imagine from the other results that maybe more iterations, a better non-linear optimizer, a higher (or lower) grid resolution, better fitted weighting factors for the various cost functions and other parameters could still improve the results on the basis of the given concepts. We could imagine that a more physics driven parameter optimization procedure seem to be required, assessed and described in more detail in order to fully exploit the potentials of the individual DA codes. The shown results represent the current status-quo during a phase of steep development for various codes. This is reflected by a large amount of new journal articles in this field. We hope the LPT and DA Challenge web-portal can help addressing remaining questions, how to enhance the results for velocity-, velocity gradient tensor-, pressure- or even density- and temperature field augmentation of a certain DA framework using the given and maybe further customized benchmark tests.

The turbulent channel flow TR DA Case 1 of the 2nd DA Challenge has shown that high accuracies and spatial resolutions for velocity and vorticity can be reached with PINN approaches at least as a tendency with respect to the classical approaches. However, this time-resolved test case attracted the highest number of participants amongst the three provided test cases (so far), most likely due to its incompressible nature excluding density and temperature variations. The overall task in solving optimization algorithms with physics constraints for the flow field augmentation of incompressible fluid flows is in general easier to achieve with respect to the other two test cases. Nevertheless, non-linear schemes are required for simultaneously reconstructing instantaneous velocity vector- and pressure field distributions, which still remains a challenging task as can be seen in Figure 2-5.

3. 2nd DA challenge Case 02

3.1. Test case description of Compressible turbulent boundary layer at $Ma \sim 4.9$

The second test case, provided by Penn State University, consists of TR sequences of advected particle tracks in a compressible turbulent boundary layer flow at a Mach ~ 4.9 . Parameters explored are the separation time between pulses, the seeding density, and inertial or non-inertial particles.

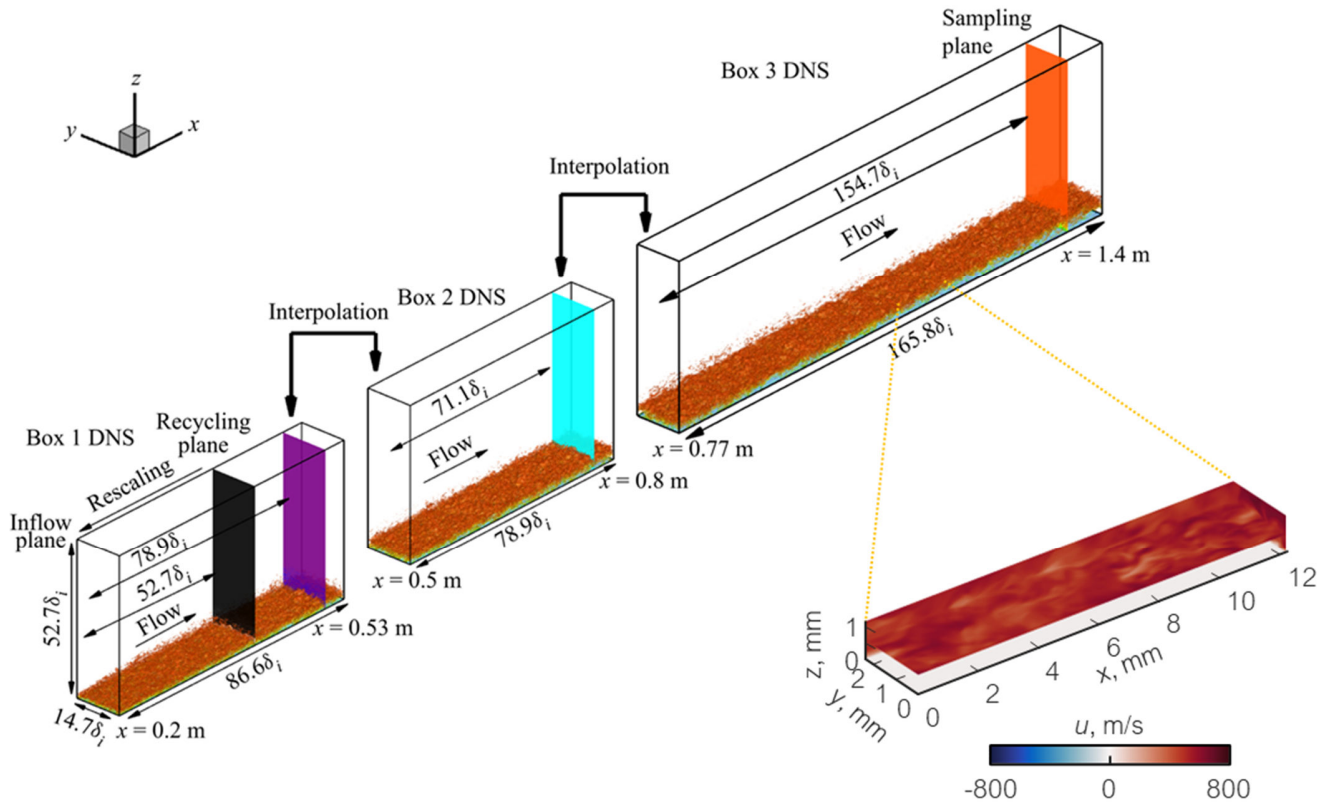


Figure 3-1: Visualization of DNS computation strategy using three boxes as sub-domains with recycling and rescaling in the first box and interpolation towards the second and third box in order to reach a fully-developed compressible TBL flow. Bottom-right: example of subdomain for advection of inertia- and non-inertia particles providing LPT input data for the DA test Case 02

This test case examines a compressible turbulent boundary layer (TBL) using data from a direct numerical simulation (DNS), while the code solves the full three-dimensional (3D) compressible Navier–Stokes equations along with the energy equation. The dataset corresponds to the Mach 4.9 case from Huang et al. (2022), which numerically replicates a high-speed wind tunnel experiment conducted at the National Aerothermochemistry Laboratory (NAL) at Texas A&M University. The simulation employs a quasi-adiabatic wall condition with a wall-to-recovery temperature ratio of $T_w/T_r = 0.91$ and sustains a friction Reynolds number of up to $Re_\tau = 1244$. Figure 3-1 illustrates computational domains from Huang et al. (2022), including the subset of the TBL selected for this data assimilation (DA) challenge (see bottom-right). The DNS results contain time-series of three velocity components, along with pressure, density, and temperature fields within the simulated measurement domain for this challenge. In the following some parameters of the flow properties and DNS of the compressible TBL at $Ma = 4.9$ are listed. The resulting time-series of instantaneous velocity-, pressure- and density fields are used to advect inertial- and ideal tracer particles along the flow following the Maxey-Riley equation with a high order temporal and spatial interpolation scheme.

The particle seeding and advection were performed using two mean interparticle spacings of $7\ell_v$ and $15\ell_v$ which corresponds to 120,000 and 12,000 particles or in a corresponding synthetic LPT experimental set-up to 0.12 ppp and 0.012 ppp at 0.1 MPx camera resolution. Ideal and inertial tracers are advected, while the inertial tracers are solid spheres with $d_p = 1.5 \mu\text{m}$ and $\rho_p = 800 \text{ kg/m}^3$ computed with one-way coupling, which is an eligible approach due to the low particle volume fraction and mass loading.

Direct numerical simulation (DNS) of a Mach 4.9 turbulent boundary layer (TBL), replicating conditions in the National Aerothermochemistry Laboratory (NAL) at Texas A&M University. The free stream flow velocity was set to $U_\infty = 794 \text{ m/s}$ at a fluid density of $\rho_\infty = 0.272 \text{ kg/m}^3$ and a temperature of $T_\infty = 66.2 \text{ K}$ with quasi-adiabatic wall conditions at $T_w = 317 \text{ K}$ and a recovery temperature ratio of $T_w/T_r = 0.91$. For the challenge, a domain of $12.24 \times 2.28 \times 1.16 \text{ mm}^3$ (streamwise length by width by height) is considered, located approximately 0.74 m downstream within the DNS domain (see Figure 3-1, bottom-right). At this location, the boundary layer thickness is $\delta = 11 \text{ mm}$, the viscous length scale is $\ell_v = 9.4 \mu\text{m}$, and the viscous time scale is $\tau_v = 0.26 \mu\text{s}$. In boundary layer units, the measurement domain size is $1.11 \times 0.21 \times 0.11$, while in viscous units, it is $1300 \times 240 \times 125$. The computational domain for this region consists of $200 \times 50 \times 110 \text{ px}$, and the DNS time step is $\Delta t = 0.08 \mu\text{s}$, which is 0.31 in viscous time units. This flow case is designed to explore the limits of DA in the presence of compressibility effects and inertial particle dynamics. Despite significant advances in high-speed imaging technology and high-power, high-repetition laser sources, conducting time-resolved 3D Lagrangian Particle Tracking (LPT) experiments in high-Mach-number flows remains extremely challenging.

Furthermore, particles in compressible flows may experience strong inertial effects, causing them to deviate from the fluid velocity field. In this challenge, we consider both inertial tracer particles and idealized tracers. Ideal tracers are assumed to be massless particles that perfectly follow the flow, providing a baseline for evaluating tracking accuracy and DA performance. Real particles are not massless and may exhibit ballistic motion in unsteady flows. The ability of inertial particles to follow the flow can be quantified using the Stokes number, $St = \tau_p/\tau$, where τ_p and τ are characteristic time scales of the particle and flow, respectively.

As output values velocity fields, the VGT elements and thermodynamic properties pressure and density are requested. Unfortunately, so far, no group has contributed to this challenging benchmark test case of a compressible TBL flow with and without inertial particle tracks as 3D LPT input data.

4. 2nd DA challenge Case 03

4.1. Test case description of Rayleigh-Bénard convection (RBC)

DA Case 03 is a DNS flow simulation of Rayleigh-Bénard convection in a closed cube with a heated floor, cooled top and adiabatic walls at a Rayleigh number of $Ra = 1 \times 10^7$ and a Prandtl number at $Pr = 0.7$ (air) driven by temperature gradient (see Figure 4-1), provided gratefully by Mohammad Shah Emran and Olga Shishkina at MPI-DS Göttingen. Its dimensions are scaled to resemble a real particle

tracking experiment in air at 1 bar 20°C ($\alpha = 0.00342086/\text{K}$, $\nu = 1.5314 \cdot 10^{-5} \text{ m}^2/\text{s}$, $\rho = 1.204 \text{ kg}/\text{m}^3$). The size of the cube is $0.2\text{m} \times 0.2\text{m} \times 0.2\text{m}$ with the coordinate origin in one of the corners oriented such that the applied temperature difference of 13K between bottom and top is in the z-direction. This difference is applied symmetrically (+6.5 K bottom, -6.5 K top with respect to ambient).

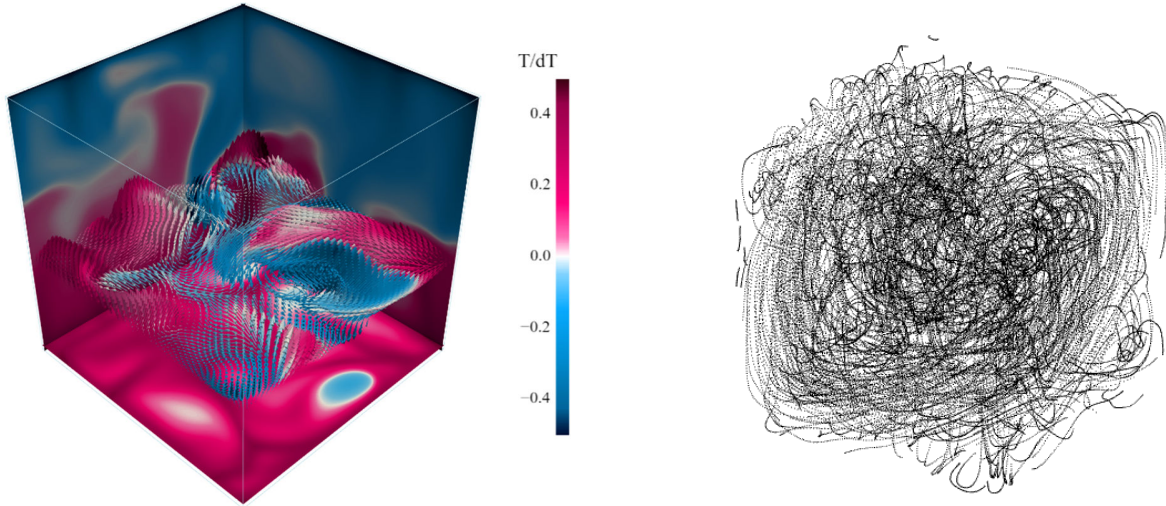


Figure 4-1: Visualization of underlying DNS of RBC at $Ra = 10^7$ (left) and example of advected tracer particles along tracks driven by turbulent motion and buoyancy forces inside cubic domain (right)

Test cases were generated at various fictitious seeding densities (0.001, 0.01, and 0.05 ppp) each resulting in a certain number of simulated particles in the domain (see Table 4) advected over a time series of 100- time steps using a 4th-order Runge-Kutta scheme and thus different mean inter particle distances. This TR- LPT case was created using massless perfect tracer particles, advected inside the DNS grid using cubic spline interpolation in space and with a fixed time stepping matched to DNS time steps. The particle positions have been initialized in a random and uniform distribution inside the RBC domain volume before three different ppp-levels and respective numbers of particles in the domain are finally advected over 100 time-steps along each track: 0.05 ppp corresponding to 40,500 particles; 0.01 ppp corresponding to 8,100 particles and 0.001 ppp corresponding to 810 particles. Due to the enclosed nature of the measurement, no particles enter or exit the domain so the number of particles across all timesteps or all tracks is constant. We provide perfect tracking in the sense that no tracks are disconnected, end prematurely, or are falsely connected. Measurement error of 0.1px is simulated by adding noise to the advected particle positions drawn from a normal distribution with a corresponding standard deviation. The added error is doubled in the y-direction to represent nonuniform error in a “depth” direction. A series of 100 timesteps is provided for each seeding density at a simulated acquisition frequency of approximately 33 Hz ($dt = 30.3228 \text{ ms}$).

We provide simulated tracks at three different seeding densities (see Figure 4-1). Similar to DA Case 01, participants submitting entries for this case must submit results for all densities. Beyond the velocity and pressure reconstruction also targeted by the other DA cases we also use this case to test out temperature reconstruction with provision of LPT data only (e.g. no temperature information at

the particle positions) as well. Participants attempting this shall include temperature in their submission data but this is an optional component of this case.

Based on the described input data participants shall provide their results on an output grid covering the interior of the domain with a fixed 2 mm spacing. A cube grid with lower corner at (2mm, 2mm, 2mm), upper corner at (198mm, 198mm, 198mm) and (99, 99, 99) points. The submitted pressure values shall be offset corrected by subtracting the pressure value at the central grid point (0.1 m, 0.1 m, 0.1 m). Temperature shall be given without the ambient, (e.g. between -6.5 K and 6.5 K).

Only the result for the central timestep (number 49 for the sequence starting at 0) is requested. Participants may optionally include the result for timestep number 1 as well.

Table 4. Number of particles per *ppp* value along 100 time steps supplied to the participants for DA Case 03.

Acquisition type	<i>ppp</i>	Number of particles	Number of images
TR	0.001; 0.01; 0.05	810; 8,100; 40,500	100

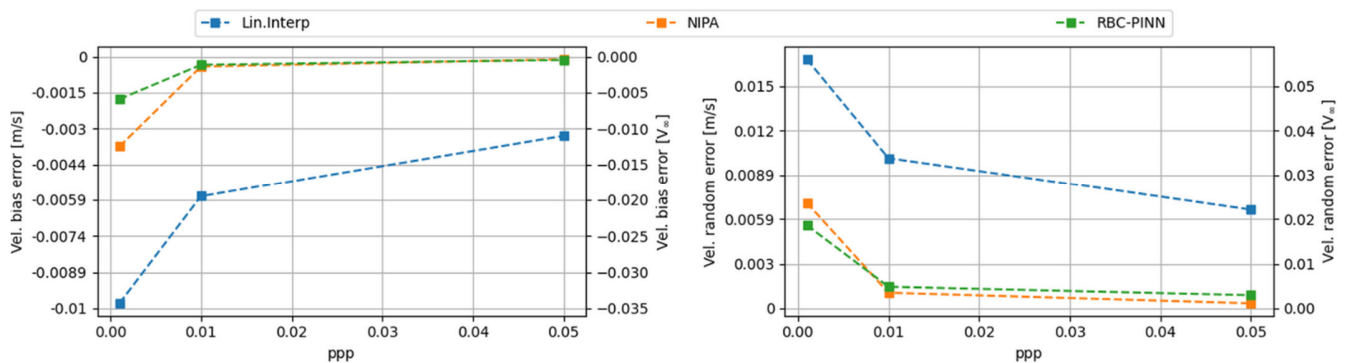
4.2. Results DA Case 03

The **DLR-PINN** developed and applied by the department of BOA at the DLR Göttingen used a neural network that consists of a multilayer perceptron (MLP) with $N_L = 10$ densely connected layers, each containing $N_N = 256$ sine activated neurons, except for the output layer which uses linear activation. The PINN used periodically activated neural networks because they have been shown to outperform monotonically activated neural networks in turbulent flow assimilation tasks. Using sine activation functions, the initialization of the weights has a non-negligible effect on the accuracy and convergence speed of the network. The weights are randomly initialized by a uniform distribution in a certain range. The Adam optimizer with its default configuration is used for training of PINNs for $\sim 10^4$ epochs using a learning rate scheduler that gradually reduces the learning rate by a factor of 0.8 from an initial value of $\eta_{\max} = 1 \times 10^{-3}$ to a minimum of $\eta_{\min} = 1 \times 10^{-4}$. Furthermore, the DLR-PINN used static weights for the loss components of the PINN. The collocation points, where the residuals of the PDEs are calculated, are randomly sampled at each iteration using the Latin Hypercube Sampling (LHS) scheme. Since PINNs generally have difficulty resolving high gradients, a higher density of collocation points is allocated in the boundary layers in comparison to the bulk (for a more detailed description see Mommert 2024). Furthermore, since the v component of the velocity is more susceptible to noise in the challenge dataset, we reduce the weight of the data loss for this component by a factor of 0.7.

The **NIPA** code, the general method already described above, employed two distinct neural networks for the Rayleigh–Benard Convection Case 03: A flow network to infer the three velocity components (u , v , w), and a pressure–temperature network to infer pressure p and temperature T . Across all particle densities (0.001, 0.01, and 0.05 *ppp*), both networks share the same architecture, consisting of

8 hidden layers with 150 neurons per layer. Time-resolved particle tracks are split into fixed-length segments of 20 positions to feed into the KCT model. The total loss involves a hybrid boundary loss, comprised of (1) a non-slip boundary loss enforcing zero velocity at all walls, (2) an adiabatic temperature loss (Neumann condition) at the vertical walls, and (3) a Dirichlet temperature loss at the top and bottom walls. Training is performed using the Adam optimizer with a learning rate of 10^{-3} , requiring approximately 50,000 epochs and 12 hours of computational cost to reach convergence.

Bias- and random error of velocity and vorticity and for pressure and temperature as well the correlation coefficient of respective reconstructed fields by the three DA schemes (one being interpolation of GT data at track positions) in comparison to the RBC DNS ground truth over the three volumetric particle track densities given in ppp-levels are given in Figure 4-2, Figure 4-3, Figure 4-4 and Figure 4-5 top rows. Visualizations of the deviation of the three reconstructed instantaneous flow fields by the respective DA methods in comparison to the underlying DNS results show color-coded iso-contours of the velocity-, vorticity-, pressure- and temperature fields in one plane of the RBC flow in the bottom rows of Figure 4-2, Figure 4-3, Figure 4-4 and Figure 4-5. While only minor differences between the DNS velocity- and vorticity field and the augmentations of both PINN based DA solutions are visible at the relatively dense sampling of the flow field at a ppp-value of 0.05, with a slightly better resolution of the velocity gradients for the NIPA solution. On the other hand, a higher accuracy for the PINN reconstruction from the DLR side are visible for the pressure- and temperature fields. The interpolation method on GT data (DNS values at velocity, pressure and temperature) at track positions fails to resolve the high temperature and velocity gradients close the RBC bottom and top plates due to under sampling of the strong gradients and missing information at all RBC borders, where the two PINN methods applied reasonable physical assumptions. Therefore, in the middle of the RBC domain the interpolation “hacker solution” performs still quite well. Overall, both PINN showed very good performance in delivering velocity- and vorticity- fields, while DLR PINN performed especially well for the “hidden parameters” pressure and temperature. All measures need to be adapted by physics constraints and making use of the temporal domain of the Lagrangian track input data. The two submissions to the DA case 03 represent the status of April 14th, 2026.



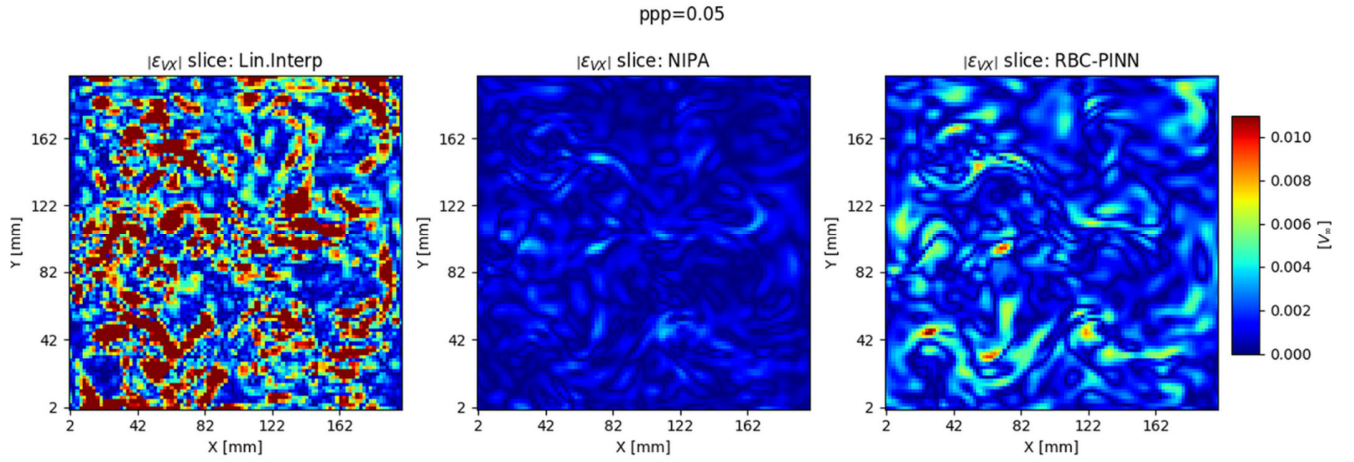


Figure 4-2: Bias- and random error of reconstructed velocity magnitudes of DA schemes with respect to the RBC DNS ground truth over the three volumetric particle track densities given in ppp (x-axis). For visibility the 3 curves (one being the linear interpolation of GT track data) are colour coded by algorithm (top) and cut of velocity errors of the individual DA results through a x-y-plane, z- in gravitation direction, showing deviations to DNS in colour coded velocity magnitude at ppp = 0.05 (bottom).

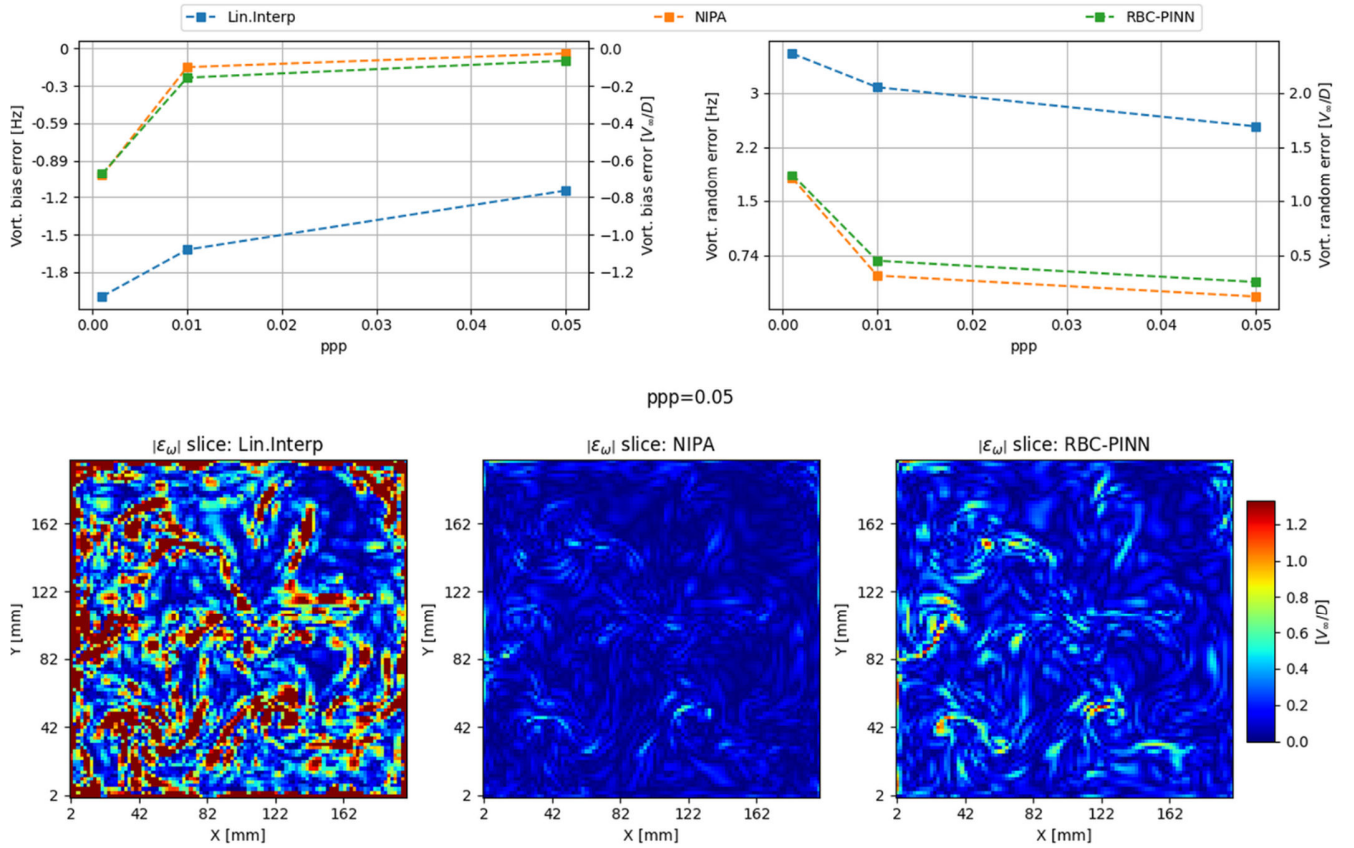


Figure 4-3: Bias- and random error of reconstructed vorticity magnitudes of DA schemes with respect to the RBC DNS ground truth over the three volumetric particle track densities given in ppp (x-axis). For visibility the 3 curves (one being the linear interpolation of GT track data) are colour coded by algorithm (top) and cut of vorticity errors of the individual DA results through a x-y-plane, z- in gravitation direction, showing deviations to DNS in colour coded vorticity magnitude at ppp = 0.05 (bottom).

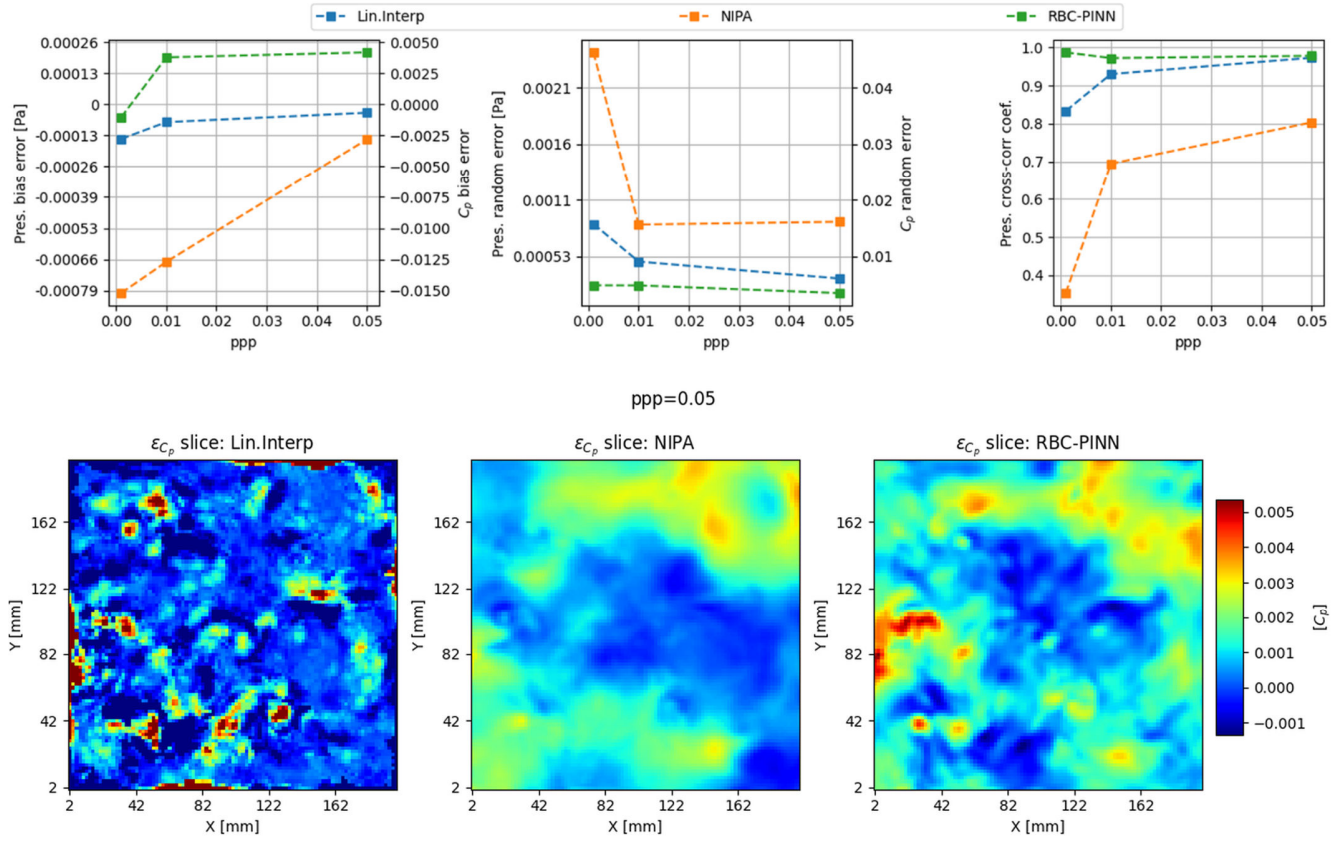
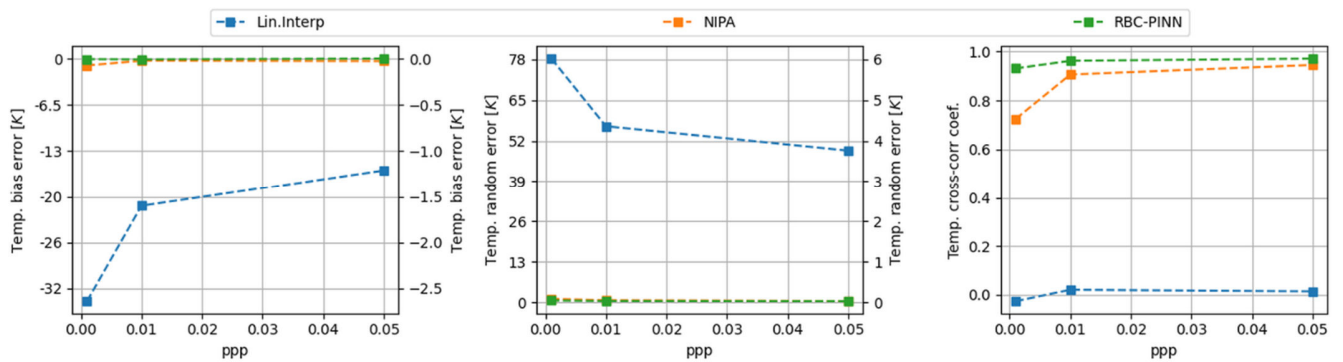


Figure 4-4: Bias- and random error and correlation coefficient of reconstructed pressure magnitudes of DA schemes with respect to the RBC DNS ground truth over the three volumetric particle track densities given in ppp (x-axis). For visibility the 3 curves (one being the linear interpolation of GT pressure data at track positions) are colour coded by algorithm (top) and cut of pressure errors of the individual DA results through a x-y-plane, z- in gravitation direction, showing deviations to DNS in colour coded pressure coefficients at ppp = 0.05 (bottom).



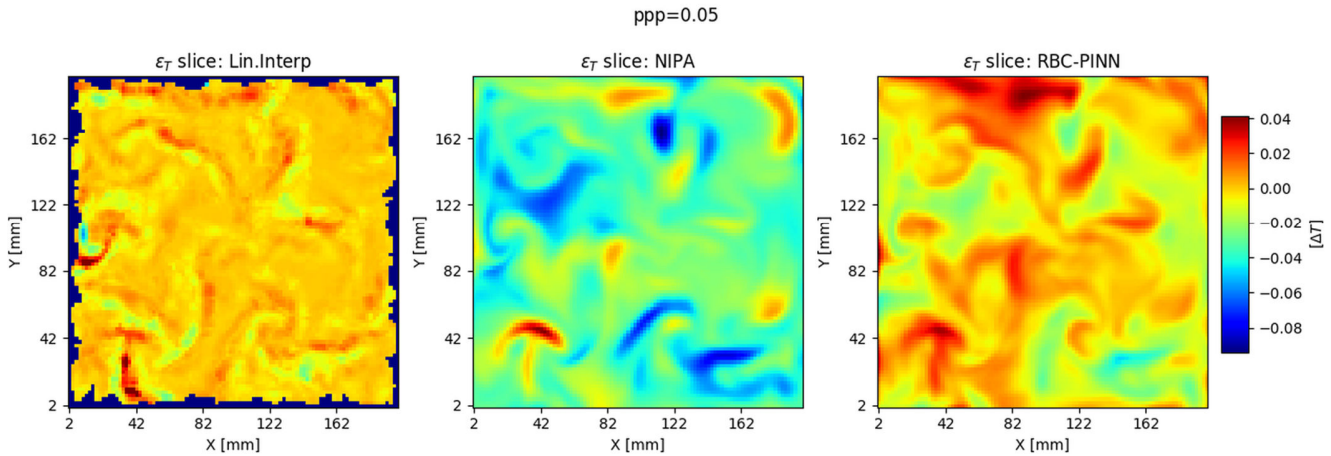


Figure 4-5: Bias- and random error and correlation coefficient of reconstructed temperatures of DA schemes with respect to the RBC DNS ground truth over the three volumetric particle track densities given in ppp (x-axis). For visibility the 3 curves (one being the linear interpolation of GT temperature data at track positions) are colour coded by algorithm (top) and cut of temperature errors of the individual DA results through a x-y-plane, z- in gravitation direction, showing deviations to DNS in colour coded temperature differences at $ppp = 0.05$ (bottom).

5. Conclusions

We present the main results of the second Data Assimilation (DA) challenge, jointly organized by TU Delft, DLR, ONERA, and Penn State University. In this challenge, three test cases were considered, all based on synthetic benchmark tests with time-series of DNS as ground truth data, namely a turbulent channel flow at $Re_\tau = 1,000$, a compressible turbulent boundary layer flow at $Ma = 4.9$ and a Rayleigh-Bénard Convection flow at $Ra = 10^7$. The effects of relevant parameters of experimental flow conditions were addressed, including the interparticle distances or particle image densities (ppp) of mimicked camera system set-ups as well as the timing strategies of corresponding image recordings (two-pulse, four-pulse or time-resolved). The input data were processed by research groups from Auburn University, Dantec, DLR, LaVision, Penn State University, ONERA, Southampton University, TU Berlin and TU Delft. The submitted data was analysed in terms of RMS errors and deviations of velocity-, vorticity- fields as well as of pressure and temperature distributions with respect to the underlying DNS. The presentation at the Lisbon conference will show a detailed analysis of the results of the two DA test cases with submissions, focusing on the difference in performance among the various algorithms based on the results submitted to the portal before April 14th, 2026 (https://onera-flow-benchmarks.fr/en/Second_LPT_DA_Challenge).

References

Agarwal K, Ram O, Wang J, Lu Y, Katz J (2021) Reconstructing velocity and pressure from noisy sparse particle tracks using constrained cost minimization. *Exp Fluids* 62:75

- Agüí JC, Jimenez J (1987) On the performance of particle tracking. *Journal of fluid mechanics*, 185, 447-468.
- Appelbaum J, Ohno D, Rist U, Wenzel C (2021) DNS of a turbulent boundary layer using inflow conditions derived from 4D-PTV data. *Exp Fluids* 62:1-21
- Beneddine S, Yegavian R, Sipp D, Leclaire B (2017) Unsteady flow dynamics reconstruction from mean flow and point sensors: An experimental study. *J Fluid Mech* 824:174-201
- Buchta DA, Zaki TA (2021) Observation-infused simulations of high-speed boundary-layer transition. *J Fluid Mech* 916:A44
- Cai S, Mao Z, Wang Z, Yin M, Karniadakis GE (2021) Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mechanica Sinica*. 2021 Dec;37(12):1727-38
- Cakir BO, Saiz GG, Sciacchitano A, van Oudheusden B (2022) Dense interpolations of LPT data in the presence of generic solid objects. *Meas Sci Technol* 33:124009
- Chong MS, Perry AE, Cantwell BJ (1990) A general classification of three-dimensional flow fields, *Phys. Fluids* A2(5), 765-777
- Chandramouli P, Mémin E, Heitz D (2020) 4D large scale variational data assimilation of a turbulent flow with a dynamics error model. *J Comput Phys* 412:109446
- Clark Di Leoni P, Agarwal K, Zaki TA, Meneveau C and Katz J (2023) Reconstructing turbulent velocity and pressure fields from under-resolved noisy particle tracks using physics-informed neural networks, *Exp. Fluids* 64:95
- Discetti S, Liu Y (2022) Machine learning for flow field measurements: a perspective. *Meas Sci Technol* 34:021001
- Du Y, Wang M, Zaki TA (2023) State estimation in minimal turbulent channel flow: A comparative study of 4DVar and PINN. *International Journal of Heat and Fluid Flow*. 1; 99:109073.
- Duraisamy K, Iaccarino G, Xiao H (2019) Turbulence modeling in the age of data. *Annu Rev Fluid Mech* 51:357-377
- Ehlers F, Schröder A, Gesemann S (2020) Enforcing temporal consistency in physically constrained flow field reconstruction with FlowFit by use of virtual tracer particles. *Meas Sci Technol* 31:094013
- Fukami K, Fukagata K, Taira K (2021) Machine learning based spatio-temporal super resolution reconstruction of turbulent flows. *J Fluid Mech* 909:A9
- Gesemann S, Huhn F, Schanz D, Schröder A (2016) From noisy particle tracks to velocity, acceleration and pressure fields using B-splines and penalties. 18th Int. Symp. On Appl. of Laser and Imaging Tech. to Fluid Mech. Lisbon, Portugal, July 4 – 7 2016
- Gesemann S (2021) Trackfit: Uncertainty quantification, optimal filtering and interpolation of tracks for time-resolved Lagrangian particle tracking. in 14th International Symposium on Particle Image Velocimetry -- ISPIV 2021, August 1-5, 2021

- Godbersen P, Gesemann S, Schanz D, Schröder A (2024) FlowFit3: Efficient data assimilation of LPT measurements, Paper 216, 21th International Symposium on Application of Laser and Imaging Techniques to Fluid Mechanics, Lisbon, Portugal, July 8-11, 2024
- Gronskis A, Heitz D, Mémin E (2013) Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. *J Comput Phys* 242:480–497
- Hanrahan S, Kozul M, Sandberg RD (2023) Studying turbulent flows with physics-informed neural networks and sparse data. *Int J Heat Fluid Flow* 104:109232
- He C, Wang P, Liu Y, Gan L (2022) Flow enhancement of tomographic particle image velocimetry measurements using sequential data assimilation. *Phys Fluids* 34:035101
- He C, Zeng X, Wang P, Wen X, Liu Y (2024) Four-dimensional variational data assimilation of a turbulent jet for super-temporal-resolution reconstruction. *J Fluid Mech* 978:A14
- He C, Li S, Liu Y (2025) Data assimilation: new impetus in experimental fluid dynamics. *Experiments in Fluids*, 66(5), 94.
- Huang J, Duan L, Choudhari MM (2022) Direct numerical simulation of hypersonic turbulent boundary layers: effect of spatial evolution and Reynolds number. *Journal of Fluid Mechanics*, 937, p.A3.
- Huhn F, Schanz D, Gesemann S, Dierksheide U, van de Meerendonk R, Schröder A (2017) Large-scale volumetric flow measurement in a pure thermal plume by dense tracking of helium-filled soap bubbles. *Exp Fluids* 58:116
- Huhn F, Schanz D, Manovski P, Gesemann S, Schröder A (2018), Time-resolved largescale volumetric pressure fields of an impinging jet from dense Lagrangian particle tracking, *Exp Fluids*, 59, 5, 1-16
- Jeon YJ, Müller M, Michaelis D (2022) Fine scale reconstruction (VIC#) by implementing additional constraints and coarse-grid approximation into VIC+. *Exp Fluids* 63:70
- Jeon YJ (2023) Vorticity time-marching method in Fine scale reconstruction (VIC#) for describing 4-D space-time. In *Proc. 15th International Symposium on Particle Image Velocimetry*, San Diego
- Kurtulus DF, Scarano F, David L (2007) Unsteady aerodynamic forces estimation on a square cylinder by TR-PIV, *Exp. Fluids* 42:185-196
- Lawson JM (2026) Physics-informed Gaussian process regression for particle-tracking data assimilation. *Physical Review Fluids*, 11(1), 014902.
- Leclaire B, Mary I, Liauzun C, Peron S, Sciacchitano A, Schröder A, ... and Champagnat (2021) First Lagrangian Particle Tracking and Data Assimilation challenge: datasets description and evolution to an open online benchmark. In *Proceedings of 14th International Symposium on Particle Image Velocimetry* (Vol. 1, No. 1, pp. 1–2). ILLINOIS Tech/Paul V. Galvin Library.
- Leclaire B, Cornic P, Champagnat F, Godbersen P, Grauer S, Grille Guerra A, Schanz D, Schröder A, Sciacchitano A (2026) LPT and DA challenges anytime: an open web portal for continuous

- development and benchmarking of algorithms, 22nd International Symposium on Application of Laser and Imaging Techniques to Fluid Mechanics, Lisbon, Portugal, June 29 – July 2, 2026
- Lemke M, Sesterhenn J (2016) Adjoint-based pressure determination from PIV data in compressible flows—validation and assessment based on synthetic data. *EuR J Mech B Fluid* 58:29-38
- Li L, Pan Z (2024) Three-dimensional time-resolved Lagrangian flow field reconstruction based on constrained least squares and stable radial basis function. *Exp Fluids* 65:57
- Liu X, Katz J (2006) Instantaneous pressure and material acceleration measurements using a four-exposure PIV system. *Experiments in fluids*, 41(2), 227-240.
- Liu DC, Nocedal J (1989) On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1–3), 503–528.
- Lourenco L, Krothapalli A (1995), On the accuracy of velocity and vorticity measurements with PIV, *Exp. Fluids* 18:421-428
- Mommert M, Barta R, Bauer C, Volk MC, Wagner C (2024). Periodically activated physics-informed neural networks for assimilation tasks for three-dimensional Rayleigh–Bénard convection. *Computers & Fluids*, 283, 106419.
- Mons V, Marquet O, Leclaire B, Cornic P, Champagnat F (2022) Dense velocity, pressure and Eulerian acceleration fields from single-instant scattered velocities through Navier–Stokes based data assimilation. *Meas Sci Technol* 33:124004
- Murai Y, Nakada T, Suzuki T, Yamamoto F (2007) Particle tracking velocimetry applied to estimate the pressure field around a Savonius turbine. *Meas Sci Technol* 18:2491–2503
- Navon IM (2009) Data assimilation for numerical weather prediction: A review. In: Park, S.K., Xu, L. (eds) *Data assimilation for atmospheric, oceanic and hydrologic applications*. Springer Publishing Company, Incorporated, Springer, Berlin, Heidelberg
- Novara M, Schanz D, Geisler R, Gesemann S, Voss C, Schröder A (2019) Multi-exposed recordings for 3D Lagrangian particle tracking with multi-pulse Shake-The-Box. *Experiments in Fluids*, 60(3)
- Novara M, Schanz D, Schröder A (2023). Two-pulse 3D particle tracking with shake-the-box. *Experiments in Fluids*, 64(5), 93.
- Pirozzoli S, Modesti D, Orlandi P, Grasso F (2018). Turbulence and secondary motions in square duct flow. *Journal of Fluid Mechanics*, 840, 631–655.
- Raissi M, Perdikaris P, Karniadakis GE (2019) Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys* 378:686–707
- Raissi M, Yazdani A, Karniadakis GE (2020) Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. *Science* 367:1026-1030

- Scarano F, Schneiders JFG, Saiz GG, Sciacchitano A (2022) Dense velocity reconstruction with VIC-based time-segment assimilation. *Exp Fluids* 63:96
- Schanz D, Schröder A, Gesemann S, Michaelis D, Wieneke B (2013), Shake-The-Box: a highly efficient and accurate Tomographic Particle Tracking Velocimetry (TOMO-PTV) method using prediction of particle position, 10th Symp. on PIV, Delft, The Netherlands
- Schanz, D., Gesemann, S, Schröder A (2016), Shake-The-Box: Lagrangian particle tracking at high particle image densities. *Exp. Fluids*, 57(5), 1-27. Doi: 10.1007/s00348-016- 2157-1
- Schanz D, Huhn F, Gesemann S, Dierksheide U, van de Meerendonk R, Manovski P, Schröder A (2016b), Towards high-resolution 3D flow field measurements at cubic meter scales, 18th Int. Symp. on Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal
- Schneiders J.F.G. & Scarano, F. (2016), Dense velocity reconstruction from tomographic PTV with material derivatives. *Exp Fluids* 57, 139 (2016). Doi:10.1007/s00348-016-2225-6
- Schröder A, Schanz D (2023) 3D Lagrangian Particle Tracking in Fluid Mechanics, *Annu. Rev. Fluid Mech.* 2023. Bd. 55:511–540, <https://doi.org/10.1146/annurev-fluid-031822-041721>
- Sciacchitano A, Leclaire B and Schröder A (2021) Main results of the first Data Assimilation Challenge. In 14th International Symposium on Particle Image Velocimetry – ISPIV 2021, August 1–5, 2021
- Sciacchitano, A., Leclaire, B. and Schröder, A. (2025) On the accuracy of data assimilation algorithms for dense flow field reconstructions. *Experiments in Fluids*, 66(2), 42.
- Sperotto P, Pieraccini S, Mendez MA (2022) A meshless method to compute pressure fields from image velocimetry. *Meas Sci Technol* 33:094005
- Sperotto P, Watz B, Hess D (2024). Meshless track assimilation (MTA) of 3D PTV data. *Measurement Science and Technology*, 35(8), 086005.
- Steinfurth B, Hassanein A, Doan NAK, Scarano F (2024) Physics-informed neural networks for dense reconstruction of vortex rings from particle tracking velocimetry *Phys. Fluids* 36(9), <https://doi.org/10.1063/5.0212585>.
- Suzuki T, Ji H, Yamamoto F (2009a) Unsteady PTV velocity field past an airfoil solved with DNS: Part 1. Algorithm of hybrid simulation and hybrid velocity field at $Re \approx 103$. *Exp Fluids* 47:957–976
- Suzuki T, Sanse A, Mizushima T, Yamamoto F (2009b) Unsteady PTV velocity field past an airfoil solved with DNS: Part 2. Validation and application at Reynolds numbers up to $Re \leq 10^4$. *Exp Fluids* 47:977–994
- Suzuki T, Yamamoto F (2015) Hierarchy of hybrid unsteady-flow simulations integrating time-resolved PTV with DNS and their data-assimilation capabilities. *Fluid Dyn Res* 47:051407
- van Gent PL, Michaelis D, van Oudheusden BW, Weiss P-É, De Kat R, Laskari A, Jeon Y, David L, Schanz D, Huhn F, Gesemann S, Novara M, McPhaden C, Neeteson NJ, Rival DE, Schneiders JFG, Schrijer FFJ (2017) Comparative assessment of pressure field reconstructions from PIV measurements and Lagrangian particle tracking, *Exp Fluids* 58:33

- van Oudheusden BW (2013) PIV-based pressure measurement, *Meas. Sci. Technol.* 24 032001
- Wang M, Zaki TA (2021) State estimation in turbulent channel flow from limited observations. *J Fluid Mech* 917:1–29
- Zauner M, Mons V, Marquet O, Leclaire B (2022) Nudging-based data assimilation of the turbulent flow around a square cylinder. *J Fluid Mech* 937:A38
- Zhou K, Grauer SJ (2023) Flow reconstruction and particle characterization from inertial Lagrangian tracks. *arXiv:2311.09076*
- Zhou K, Grauer S, Schanz D, Godbersen P, Schröder A, Rockstroh T, Jeon YJ, Wieneke B (2024). Benchmarking data assimilation algorithms for 3D Lagrangian particle tracking. In 21st International Symposium on the Application of Laser and Imaging Techniques to Fluid Mechanics, Lisbon, Portugal
- Zhou K, Tang R, Ke G, and Grauer SJ (2025) Neural-implicit particle advection. in 17th International Symposium on Particle Image Velocimetry (ISPIV 2025) June