DOI: 10.13009/EUCASS2025-646
11TH EUROPEAN CONFERENCE FOR AERONAUTICS AND SPACE SCIENCES (EUCASS)

FLOW FIELD RECONSTRUCTION APPLIED ON FLOW PHENOMENA DETECTION IN A SPARSE SENSED TURBOPUMP MANIFOLD

Sarah KRAMER [1,2], Louis SOUVEREIN [1], Lionel AGOSTINI [3],
Günther WAXENEGGER-WILFING [4,5], Stefan SCHLECHTRIEM[2,4]

[1]*ArianeGroup GmbH, Munich, Germany*
[2]*Institute of Space Systems, University of Stuttgart, Germany*
[3]*Pprime Institute, CNRS, University of Poitiers, Poitiers, France*
[4]*Institute of Space Propulsion, German Aerospace Center (DLR), Lampoldshausen, Germany*
[5]*Institute of Computer Science, Julius-Maximilians-University, Würzburg, Germany*
sarah.kramer@ariane.group , https://ariane.group/en/

Abstract
Facing the tough conditions of turbomachinery instrumentation, such as limited accessibility, a restricted number of sensors, not directly measurable variables, and harsh operating conditions, it is challenging to capture the global flow organization and investigate the behaviour of dynamic flow phenomena. [1]

Machine Learning (ML) offers a promising approach to derive more information from the acquired measurement data to enable a better understanding of lifetime degrading effects, an optimization of performance and using data driven control strategies to avoid critical operational conditions and prevent anomalies. [2][3] This paper presents a Proper Orthogonal Decomposition (POD) for reconstructing the flow field within a turbine inlet manifold, trained on validated computational fluid dynamic (CFD) simulation data to detect flow separation.

1. Introduction

To ensure a safe operation of technical products a detailed understanding of the dynamic behaviour within the performance envelope and the therein occurring loads is fundamental. Knowing maximum loads of critical conditions and operation dynamics drives requirements for the component design and defines limits for the control system. Nowadays, computing technologies and capacities enables detailed analysis by simulations and are not only used for few final analyses, but also in the early development process to perform automated design optimizations. Nonetheless, a residual inaccuracy remains and highly dynamic system responses as well as interactions of multiple components in complex systems are difficult to simulate. In respect of this, tests are essential and relevant to verify the real behaviour.

Since rocket engines are complex and highly dynamic, as well as operating under challenging conditions close to material limits, tests are mandatory. The spatial sparse instrumentation, as a consequence of poor accessibility in these compact engines and the overall limited number of sensors, restricts the obtainable measurements. The restricted information gain is further intensified by a low test cadence due to the high costs of performing engine tests, which makes detailed analysis of the operational behaviour of each component challenging. The geometrically small turbopumps of rocket engines are highly loaded by temperature gradients, supersonic flow phenomena and compared to the other rocket engine components by additional rotational forces and rotor stator interactions. For a better understanding of these phenomena and their complex interaction, as well as their lifetime and performance degrading effect, it is essential to extract more information from tests.

Machine Learning (ML) has the potential to simplify complex dynamic systems by reducing their dimension, extend more information out of big data sets by recognize patterns and correlations, and transfer information between different data sets. For sparse sensed rocket engine components such as the turbopumps, ML techniques can be used to predict values by learned correlations, recognize anomalies by an unexpected deviation and transfer the dense information from simulations to a small measurement data set. This paper presents how the flow field of simulations, can be reconstructed in the test anchored on few sensors. The reconstructed flow field can be used to analyse the global flow behaviour and detect the propagation of dynamic flow phenomena such as flow separation. Detecting flow phenomena

DOI: 10.13009/EUCASS2025-646

in an early stage paves the way for data driven control strategies to mitigate their critical growth and extend the components life time as well as prevent anomalies.

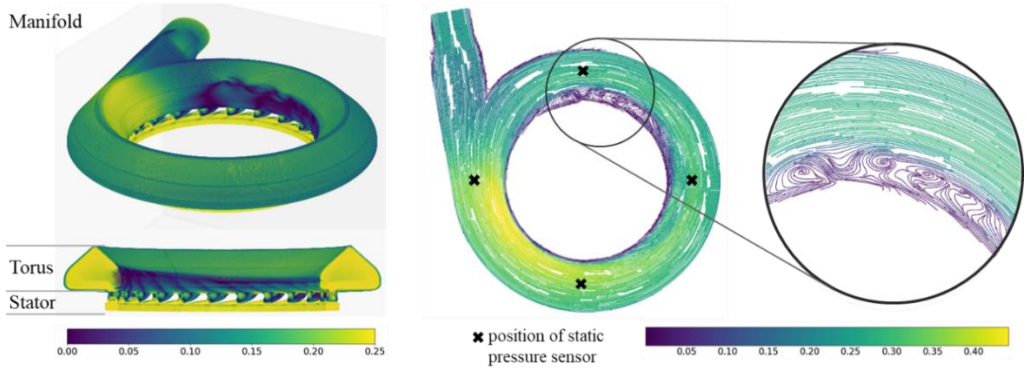

Figure 1: Turbine manifold, left: 3D velocity flow field, middle: 2D velocity streamlines, right: zoom - flow separation

This paper presents the method of flow field reconstruction applied on a turbopump inlet manifold to detect the flow separation, as visible in Figure 1. The flow separation in the manifold torus, significantly affects the inflow into the downstream convergent-divergent stator channels, ultimately increases the pressure losses and impacting the turbine performance. As this phenomenon is spatially dynamic, it cannot be determined exactly at every operational point by a few sensors, and its detailed global propagation cannot be fully captured without an optical access.

Reduced Order Models (ROM) compress the information of an application and capture its most dominant components to reconstruct the information in a reduced dimension. The here used ROM, the Proper Orthogonal Decomposition (POD) is a common used tool to catch coherent structures in flow fields. With the most dominant determined spatial modes, which contain the most flow energy, and their temporal behaviour, flow fields can be reconstructed in reduced dimension. The determined coherent structures can further be exploited to analyse the most sensitive sensor positions, to increase the reconstruction accuracy. This paper shows the reconstruction of the manifold flow field and discusses the therefore best senor placement to capture the organization of the flow

2. Theoretical Background

This chapter introduces the theoretical background of the here used POD, the method of flow field reconstruction, the QR decomposition to determine best sensor placements, the metrics to measure the accuracy of ML models, and the turbine inlet manifold where the method is applied.

2.1 Proper Orthogonal Decomposition (POD)

POD is a widely used method to determine spatial coherent structures in a given flow field, by using the technique of the Singular Value Decomposition (SVD). The SVD decomposes a specific matrix A, by performing a matrix factorization, in three simpler matrices U, Σ, and V^T, as given in equation **(1)** and illustrated in Figure 2.

$$A = \mathrm{U}\,\Sigma\,V^T \tag{1}$$

The orthogonal Matrix U contains the left singular vectors of A and forms an orthonormal basis for the column space of A. The rows of V^T contains the right singular vectors of A, which are forming an orthonormal basis for the row space of A. The diagonal matrix Σ gives the singular values of A, which are describing the importance of each corresponding singular vector.

DOI: 10.13009/EUCASS2025-646

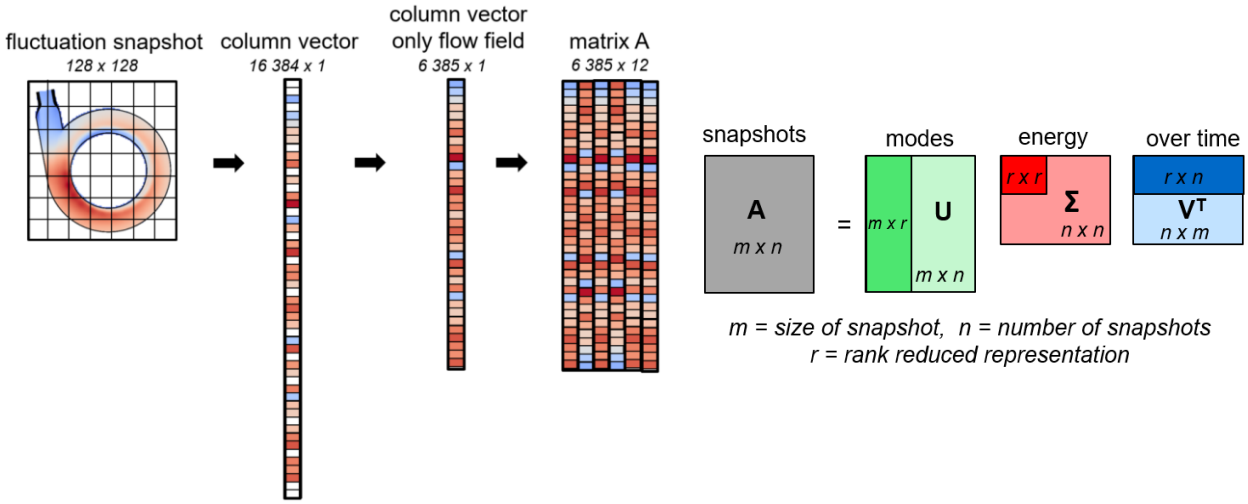

Figure 2: left: preparation of matrix A from snapshots, right: SVD illustration used for POD

Applied on fluid flows, the matrix A will contain a time series of snapshots from a given flow field. Therefore, each snapshot will be reshaped in a column vector and lined up in increasing time, as shown in figure **(3)**. Matrix A has the size of $m \times n$, where m is the size of snapshots and n is the number of snapshots. The size of snapshot is the number of data points in horizontal direction times the number of data points in vertical direction. By applying the POD to snapshots of a fluid flow, the matrix U contains the coherent structures or modes sorted by their dominance. Σ gives the energy level of these modes and V^T represents their amplitudes over time. Since V^T is transposed, one row of it describes the amplitude of the column modes in U.

The dimension can be reduced by selecting the first most dominant modes by defining a rank r. The flow field can be reconstructed in reduced representation by solving equation **(1)** with the ranked matrices.

2.3 Tailored sensing

This section describes how a flow field can be reconstructed by spatially sparse sensor measurements and how their positions can be optimized to increase the accuracy of the reconstruction. The system that has to be solved here is illustrated in Figure 3. The sensor positions are given by matrix C. The information at these spatial positions from each dominant mode can be extracted from the tailored U_r matrix by taking the rows and are saved in matrix Θ, as illustrated in Figure 3. To determine the amplitudes over time of the ranked modes based on sensor positions, contained in vector a, the system can be solved with measurements from the original snapshots at the sensor positions, given by the vector y, and the inverse of Θ.

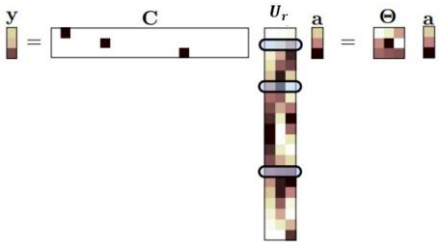

Figure 3: illustration of system equation, adapted by [3]

$$y = C\, U_r\, a = \Theta\, a \qquad \textbf{(2)}$$

The flow field can then be reconstructed by the inner product of the estimated amplitudes in vector a and the ranked U_r matrix. To increase the accuracy of this reconstruction the most sensitive positions for each mode can be determined by performing a QR decomposition of the tailored mode matrix U_r. The QR factorization decomposes a matrix into a unitary matrix Q, an upper-triangular matrix R and a column permutation matrix C. The provided sparse measurement matrix C contains r pivots that best sample the r basis modes. The flow field can be reconstructed as with random

DOI: 10.13009/EUCASS2025-646

sensors, by determining vector a with solving the equation **(2)** for given measurements y. The so estimated amplitude will provide a reconstructed flow field of higher accuracy compared to random sensor positions.

2.4 ML metrics

Common used metrics to analyse the accuracy of ML models are the Mean Squared Error (MSE), the coefficient of determination (R^2), the Mean Absolute Percentage Error, and especially for images the Peak Signal to Noise Ratio (PSNR). To enable a comparison to a wide range of similar applications all of these quality parameters will be determined here. The mathematical expressions of these metrics are given in the following equations **(3)**, **(4)**, **(5)** and **(6)(3)**.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \quad \textbf{(3)}$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)\right)^2} \quad \textbf{(4)}$$

$$MAPE = 100\frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \quad \textbf{(5)}$$

$$PSNR = 20 \cdot log_{10}\left(\frac{MAX}{\sqrt{\frac{1}{m\,n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2}}\right) \quad \textbf{(6)}$$

All of these parameters calculating the difference between the target values Y and the ML generated values $\hat{Y}$. n is giving the number of samples, excluding in PSNR. The MAX in PSNR contains the maximum value of the pixels in an image with a size of $m \times n$ pixels. The target image I is compared to the generated image K. A higher value of PSNR identifies a higher accuracy of the reconstructed image. MSE is the most used quality parameter and describes the mean of the squares of the errors. MAPE gives the mean error in percentage. R^2, the coefficient of determination describes the proportion of variance that is represented in the model, where the residual sum of squares SS_{res} is divided by the total sum of squares SS_{tot}. A value of 0 identifies a worse accuracy and a value of 1 stands for a high accuracy.

2.5 Turbine inlet manifold

The inlet manifold as shown in Figure 1, is a 3D printed component and consists of the torus and the first stator vane row. The torus works like a diffusor and decelerates the incoming flow to increase the potential energy and to generate a circumferentially homogeneous inlet flow to the first stator row and into the turbine, with minimum total pressure losses. A flow separation in the torus, as visible in Figure 1 has a bad impact on these requirements, as it increases the pressure losses and affects the inlet flow as well as the turbine performance. A good understanding of the flow separation dynamic within the operation envelope is important to define control strategies that ensure stable operation.

3. Methodology

This section reports on the generation, analysis and preparation of the here used data set, as well as the construction of the POD model and the flow field reconstruction, analysis and evaluation.

3.1 Data generation

The here used data set was generated by twelve steady state computational fluid dynamic (CFD) simulations over the range of operation. A mesh study was performed as defined in [4] and the simulation results were validated with few test data. The 2D snapshots are created by interpolating the data of the unstructured 3D CFD mesh on 2D horizontal planes through the manifold torus with a structured grid of size 128 x 128. A good indicator of flow separation is the circumferential velocity, since flow separation can be directly identified by negative values of it. The circumferential velocity was calculated by the Cartesian velocities u, v w. The simulated steady states are within the range of a cold-flow test campaign and represent operational points with the same pressure ratio over the manifold under hot-gas conditions. Lining up these operational points by an increasing pressure ratio represents a temporal behaviour of a turbine spin up.

DOI: 10.13009/EUCASS2025-646

3.2 Analysis of flow separation over operational range

The flow field of the circumferential velocity on the horizontal planes of the twelve operational points are shown in Figure 4. The flow separation, visible as the dark blue region, starts on the inner diameter at the bottom right region and increase along the inner diameter in counter clockwise direction to the region in the top. It is also recognizable that the thickness of separation increases over the operational points. In operational point nine and ten the separation disappears and comes back in point eleven and twelve. Further, an overall increase in velocity is visible.

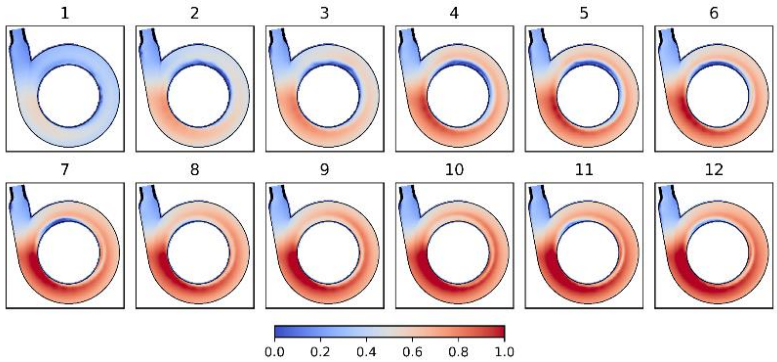

Figure 4: Original snapshots – circumferential velocity over operational points

3.3 Data Preparation

To increase the strength of the POD to capture flow dynamics, the mean flow field of the twelve snapshots was subtracted to focus the model on the pure fluctuation. The mean flow field is shown in Figure 5. The final step in the reconstruction process involves adding the mean flow field to the reconstructed fluctuation again. Figure 5 on the right side shows the snapshots of pure fluctuation.

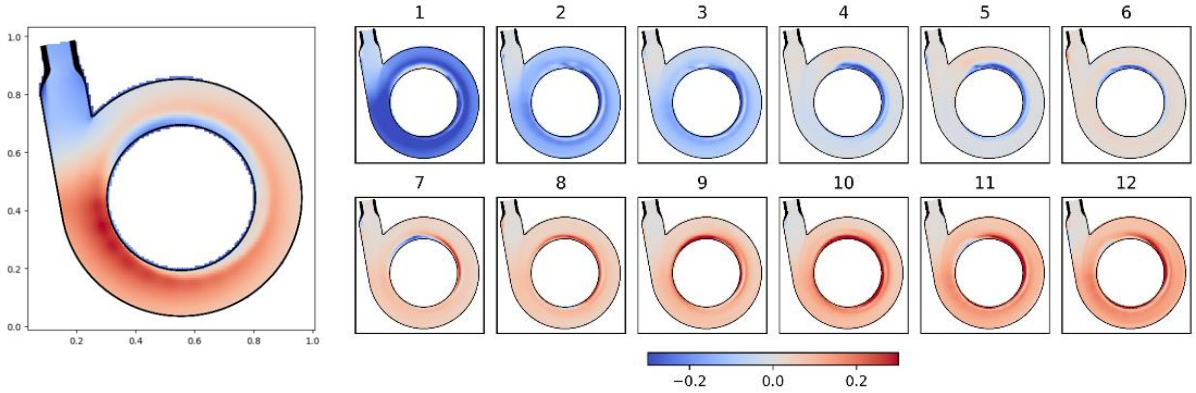

Figure 5: left: mean flow field; right: fluctuation of circumferential velocity flow field

Since the values of fluctuation are smaller than the original flow field, the colour scale of the fluctuation snapshots in Figure 5 is zoomed in. To feed in the snapshots over time into the POD, each snapshot of size 128 x 128 data points is reshaped to a column vector of size 16384 x 1 and lined up in increasing pressure ratio in the matrix A, as shown in Figure 2. The size of the matrix A is reduced to a 6385 x 12, by deleting the rows of data points outside the flow field.

3.4 Construction of POD model

Since the model is coded in the python environment, the calculation of the SVD can be directly called by the numpy command “linalg.svd”. It has to be mentioned that this command delivers the transposed V matrix. In the next step the

DOI: 10.13009/EUCASS2025-646

identified modes, their energies and amplitudes over time are analysed, as shown in chapter 4.1. The modes which capture the most flow field energy are used to reconstruct the fluctuation of the flow field by the related columns of U, diagonal elements of Σ and rows of V^T.

3.5 Flow field reconstruction

The creation of the ROM model is achieved through the dimension reduction of the POD model by tailor the POD matrices to the first dominant modes with a defined rank. The so reconstructed flow fields are analyzed through a comparison with the original flow field, and the metrics are determined in order to measure the accuracy of the machine learning (ML) models.

The tailored U matrix is then used to determine the best sensor positions by performing a QR decomposition on it. The reconstruction by the estimated mode amplitudes based on the few sensor positions is compared to the original flow field and analysed with the ML metrics given by equation **(3)**, **(4)**, **(5)** and **(6)**. An analysis of a further decrease in number of sensors were performed.

4. Results

This section presents the detected modes, their dynamic, the best sensor placements, and the reconstructed flow fields. A detailed discussion of the here shown results are given in chapter 5.

4.1 Detected modes

The by the POD detected modes are shown in Figure 6 and arranged by their energies. In conjunction with the mode amplitudes over time, as visualized in Figure 7, the modes can be interpreted. It can be seen that the first mode has an equal impact over the whole manifold flow field and a linear increase over time, which will represent the increase of the overall pressure ratio. Mode 2 caught the dynamics of the region where the flow separation occurs.

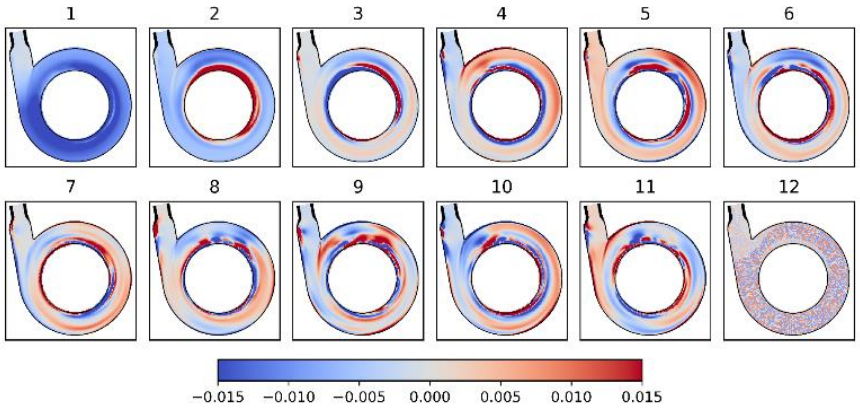

Figure 6: determined modes

Mode 3 describes the circumferential variance of the thickest part of the separation along the inner radius. It can be recognized that the details of the modal patterns increases, and their periods getting shorter over a decreasing energy level. This excludes mode 7, which only shows significant dynamics between operational point 7 and 10, which might be correlated to the disappearance of the flow separation during this operational range. Mode 12 represents some weak noise with an energy level and an amplitude next to zero.

DOI: 10.13009/EUCASS2025-646

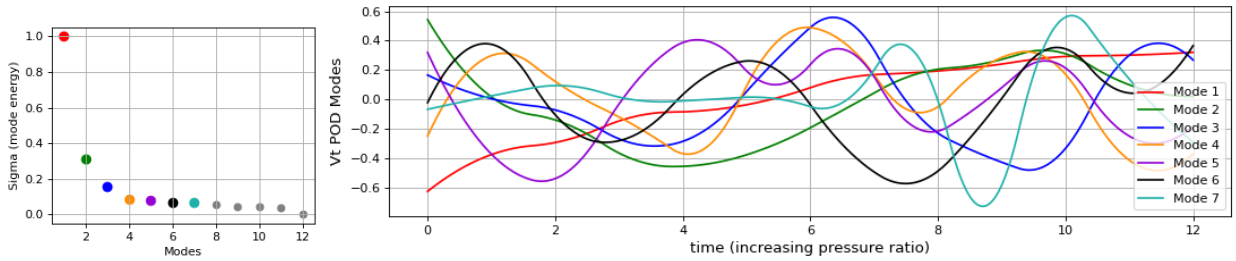

Figure 7: left: mode energy, right: mode amplitude

In the Σ-plot on the left side of Figure 7 a larger energy difference between the first mode and the second mode compared to the distances between the other modes can be recognized. As the first three modes containing the most energies, these modes will be chosen to generate the ROM model by setting the rank to three.

4.2 Latent space

The compressed dynamic behaviour of the ROM can be visualized by plotting the latent space with the amplitudes of the first three modes, as shown in Figure 8. As there is no convergence towards a limit cycle visible in the latent space, the system does not contain an oscillating dynamic.

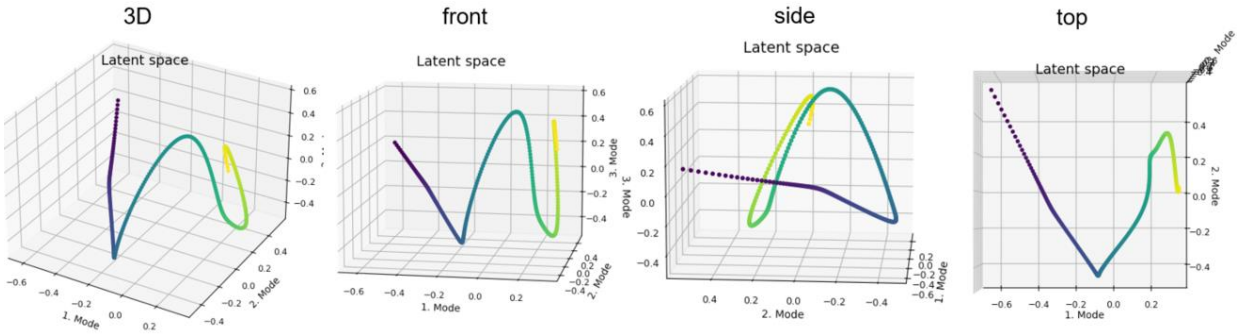

Figure 8: latent space, first three mode amplitudes

The latent space can be also used to compare captured system dynamics of different ROMs.

4.3 Best sensor positions

Performing the QR decomposition on the U matrix, gives the most sensitive element for each mode, which can be interpret as best sensor position to detect the corresponding mode. Figure 9 shows the best sensor positions to detect all twelve modes, the three most dominant modes and the most important position related to the first mode. The so determined sensor positions identify also the positions of highest variance during operation, which is here along the inner radius.

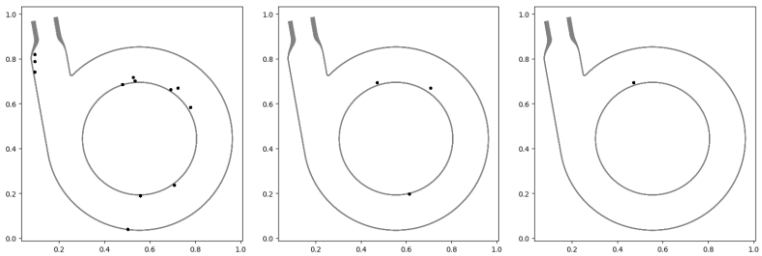

Figure 9: best sensor placement: left: 12 modes, middle: 3 modes, right: 1 mode

DOI: 10.13009/EUCASS2025-646

4.4 Flow field reconstruction

As mentioned in section 4.1 the first three modes containing the most energies and will be taken to generate the ROM model. The reconstruction of the flow field by the ranked matrices U_r, Σ_r and V_r^T with a rank of three is shown in Figure 10. An analysation of the reconstruction accuracy is observed by determining the difference to the original snapshots and the ML metrics.

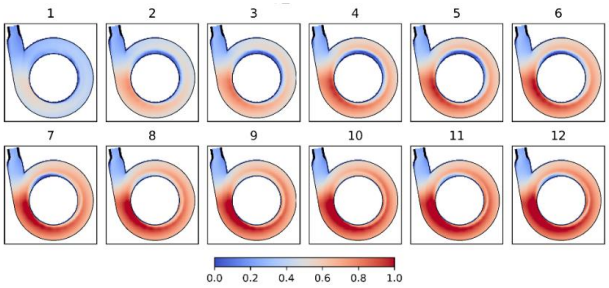

Figure 10: reconstructed flow field with first three modes

*Figure 11*Figure 10 shows the reconstructed flow field of the ROM with the first three modes. The difference between the reconstructed flow field and the original one is given in Figure 11.

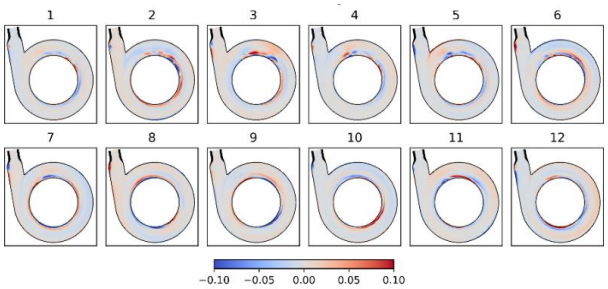

Figure 11: difference of reconstructed flow field with first three modes to original snapshots

It has to be mentioned that the colour scale here is zoomed in for a better contrast of the small differences.
Figure 12 shows the reconstruction of the flow field by tailored POD with rank three and a determined mode amplitude on three sensors, which positions were determined with the QR decomposition method.

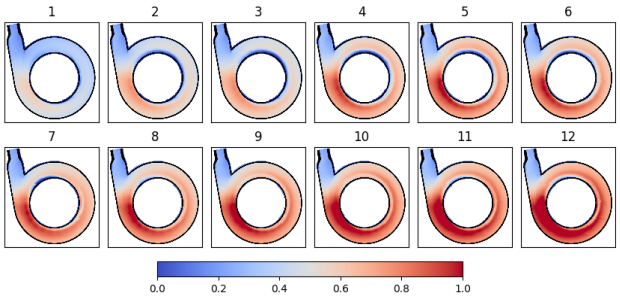

Figure 12: reconstructed flow field with first three modes and three sensors

The pure reconstructed fluctuation is shown in Figure 13 and the difference to the original snapshots are shown in Figure 14. The bigger visible difference to the original snapshots can be also realized in the values of the metrics.

DOI: 10.13009/EUCASS2025-646

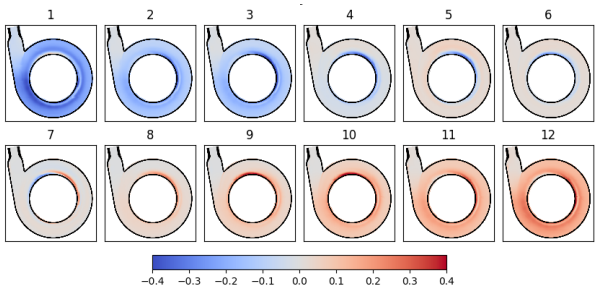

Figure 13: fluctuation of reconstructed flow field with first three modes and three sensors

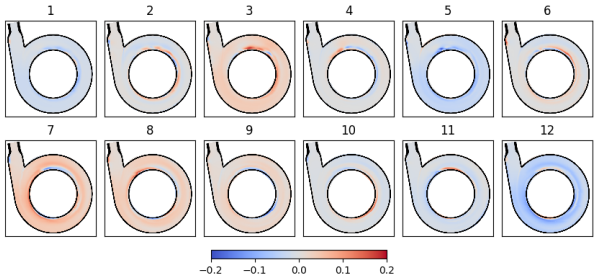

Figure 14: difference of reconstructed flow field with first three modes and three sensors to original snapshots

MSE	0.001	R^2	0.978
MAPE	0.183	PSNR	40.969

The following figures shows the reconstruction by only one sensor and tailored POD matrices of rank 3. Figure 15 shows the mean flow field with small variances over the operational points, which is also recognizable in the caught fluctuation in Figure 16 and leads to a huge difference to the original as shown in Figure 17.

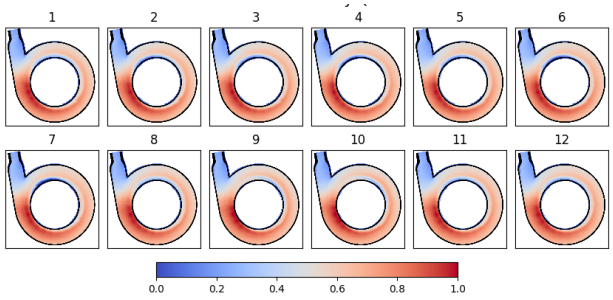

Figure 15: reconstructed flow field with first three modes and one sensor

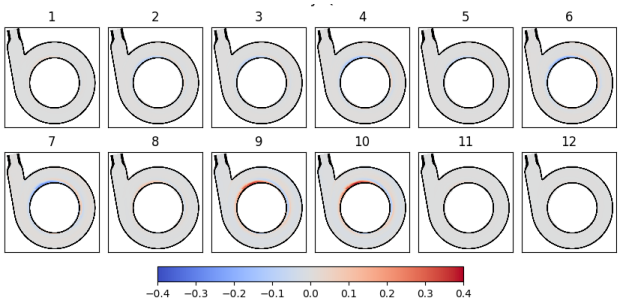

Figure 16: fluctuation of reconstructed flow field with first three modes and one sensor

DOI: 10.13009/EUCASS2025-646

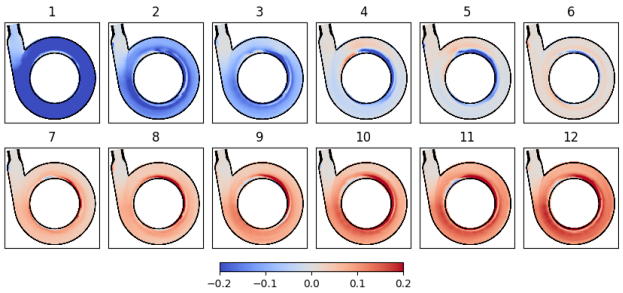

Figure 17: difference of reconstructed flow field with first three modes and one sensor to original snapshots

The poor reconstruction accuracy is also represented in the following metrics.

MSE	0.013	R^2	0.527
MAPE	0.689	PSNR	29.142

The here shown first results of the dimension reduced reconstruction model based on POD are promising and will be further improved by variating the number of sensors and tailored modes. Further, a comparison to experimental data will be performed by reconstruct the flow field with the same sensor positions than in the test. As already seen here a minimum number of three sensors and modes is necessary.

5. Conclusion & Outlook

With the here generated reduced order model a reconstruction of the flow field during the test can be realized to recognize flow separation in an early stage. These first results offers a good foundation to optimize the POD model by variating the number of sensors and ranked modes. With a validation on experimental data the accuracy of the model can be further analysed and improved. Further, the model will be trained on more simulations and tested on operational points outside of the training data set. As the POD provides a linear solution, a comparison to a Variational Autoencoder, which also generates nonlinear solutions is planned. The presented approach of best sensor placement offers promising opportunities for experiment design.

6. Acknowledgements

Part of the work was supported by the DLR under the federal grant project ANNA with the contract number 50RL2210.

References

[1] D NASA SP-8110; Liquid rocket engine turbines; NASA; 1974 ;
[2] Goodfellow, I., Bengio, Y., Courville, A.; Deep Learning; The MIT Press; 2016;
[3] Brunton, S.L., Kutz, J.N.; Data-Driven Science and Engineering; Cambridge University Press; 2021;
[4] Celik Celik, I.B., (2008). Procedure for Estimation and Reporting of Uncertainty Due to Discretization in CFD Applications. ASME, 078001-4, Vol.130