IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10642253

DATA-DRIVEN PREDICTION OF LARGE INFRASTRUCTURE MOVEMENTS THROUGH PERSISTENT SCATTERER TIME SERIES MODELING

Gideon Stein[1], Jonas Ziemer[2], Carolin Wicker[3], Jannik Jänichen[2], Gabriele Demisch[3], Daniel Klöpper[3], Katja Last[3], Joachim Denzler[1], Christiane Schmullius[2], Maha Shadaydeh[1] and Clémence Dubois[2,4]

[1] Computer Vision Group, Friedrich Schiller University Jena,
Ernst-Abbe-Platz 2, 07743 Jena, Germany
[2] Department for Earth Observation, Friedrich Schiller University Jena,
Leutragraben 1, 07743 Jena, Germany
[3] Ruhrverband, Kronprinzenstraße 37, 45128 Essen, Germany
[4] Institute of Data Science, German Aerospace Center, Maelzerstraße 3-5, 07745 Jena, Germany

ABSTRACT

Deformation monitoring is a crucial task for dam operators, particularly given the rise in extreme weather events associated with climate change. Further, quantifying the expected deformations of a dam is a central part of this endeavor. Current methods rely on in situ data (i.e., water level and temperature) to predict the expected deformations of a dam (typically represented by plumb or trigonometric measurements). However, not all dams are equipped with extensive measurement techniques, resulting in infrequent monitoring. Persistent Scatterer Interferometry (PSI) can overcome this limitation, enabling an alternative monitoring scheme for such infrastructures. This study introduces a novel monitoring approach to quantify expected deformations of gravity dams in Germany by integrating the PSI technique with in situ data. Further, it proposes a methodology to find proper statistical representations in a data-driven manner, which extends established statistical approaches. The approach demonstrates plausible deformation patterns as well as accurate predictions for validation data (mean absolute error=1.81 mm), confirming the benefits of the proposed method.

Index Terms— PSI, Sentinel-1, Deformation prediction, Dam monitoring

1. INTRODUCTION

The monitoring program for gravity dams comprises a variety of methods to ensure their safety [1]. These methods typically include plumb and trigonometric measurements. To assess whether measured movements correspond to expected deformations, statistical predictions are made about the expected deformation based on relevant exogenous factors such as temperature and water level. This is ideally done on a daily or weekly basis. However, not all dams have extensive measurement techniques (plumbs), allowing only for bi-annually and expensive monitoring strategies due to time-consuming field campaigns. This is especially problematic, given that climate change increases the probability of extreme weather events, which, in turn, potentionally increase environmental stress on infrastructure.

In this context, Persistent Scatterer Interferometry (PSI, [2]) has proven to be a valuable alternative, enabling high-resolution spatial monitoring of dams with revisit cycles of a few days [3, 4]. The implementation of this technique into nationwide ground motion services has opened up the possibility of monitoring dam deformations on a large scale. The German Ground Motion Service (BBD) provides freely available Sentinel-1 PS time series starting in 2015 to detect nationwide ground motions and deformations [5]. Further, current statistical models of dam deformation typically utilize multiple linear regression models to predict the expected deformation of a dam based on only a few exogenous regressors (water level and temperature, [6]). While these models provide robust deformation forecasts, they are unlikely to describe the dynamics of gravity dams exhaustively.

With this background, this study aims to provide the first complete example of a monitoring strategy, including data-driven dam deformation forecasting based on Sentinel-1 PS time series, complemented with in situ data and appropriate time series modeling.

2. MATERIALS AND METHODS

2.1. Study Area

The study was conducted on a gravity dam in North Rhine-Westphalia, Germany. The Lister Dam is located in the Sauerland region and is part of the Bigge reservoir with a total capacity of 171.6 million m³. It primarily serves for water release into the Bigge reservoir. The arched gravity dam is 42 m high, 264 m long, and collects water from a catchment

Fig. 1. The Lister gravity dam, as seen from the impounded Bigge reservoir on the downstream side.

area of 68.2 km², creating a reservoir with a capacity of 21.6 million m³ [7]. It is operated by the Ruhrverband, a non-profit-oriented water management company based on German public law. The dam is equipped with several measurement techniques, including trigonometric measurement points, water level, and temperature monitoring. However, no plump system has been installed on the Lister gravity dam, reducing the deformation monitoring to a few samples yearly. Figure 1 shows the Lister Dam from the downstream side.

2.2. Data

The study period covers almost six years, from April 2015 to December 2020, with the last 12 months of the time series used as test data to validate predictions. The PS time series used in this study were provided by the Federal Institute for Geosciences and Natural Resources (BGR) and complemented with freely available data from BBD's WebGIS platform [8]. Time series were acquired by the Copernicus Sentinel-1 satellites, utilizing the PSI technique. Data originating from the ascending and descending directions in the sensor's line of sight (LOS) were used, starting in 2015. Furthermore, vertical and east-west deformation components were analyzed, which are also provided by BBD. All time series were downloaded at a temporal resolution of six to 12 days. Finally, in this study, the PS time series were complemented by in situ measurements provided by the Ruhrverband. Variables such as water level and air temperature influence the seasonal deformation pattern of a gravity dam. They are thus essential for accurate predictions of dam movements. Both variables were provided at a daily resolution.

3. METHODS

3.1. Exogenous variables and data preprocessing

As a gravity dam is mainly influenced by the seasonal fluctuations in water level (W) and temperature (T), these variables on the current day are considered as potential exogenous influences on the dam deformation. Further, it was estimated that the immediately preceding values of these exogenous variables (e.g., the day before) might be beneficial for predicting movements. The recent past of W and T as well as the mean of the last few days (denoted as W_n^M, T_n^M, where n specifies the previous days considered) are additionally considered as additional exogenous regressors. Next, multiple equally valid options are considered to represent seasonal cycles beyond what can be inferred from the exogenous variables. First, Fourier terms (F), which are defined as follows, are tested as additional exogenous variables:

$$F_{k,t} = \left[sin\left(\frac{2\pi kt}{m}\right), cos\left(\frac{2\pi kt}{m}\right)\right] \quad (1)$$

Here, m specifies the length of the seasonal cycle, and k is the order of the term. Second, the annual cycle (seasonal mean) subtraction from all time series to only model the differences to the annual cycle is considered. Finally, the differencing order (d) of all time series (subtracting previous values of itself) is optionally evaluated for models that do not feature spatial interaction (see below).

To further prepare these regressors as well as the PSI-time series for modeling, the following preprocessing steps are performed on every time series that is used during training: 1. Outliers (measurement errors) are filtered via boxplot filtering (IQR=2.25) by removing measurements outside the whiskers. 2. Missing values are linearly interpolated (including measurements filtered out before). 3. Every time series is normalized to lie in the range $[0, 1]$.

3.2. Model classes

To link to previous work on modeling dam deformations (typically via plumb bobs, [6]), well-established linear regression models are considered. As there are typically multiple PS-point time series available and since it is unclear whether utilizing the spatial interaction between the different PS point deformations can improve prediction, these dynamics are additionally considered. This means it is explored whether there are benefits of considering other PS points as potential regressors. Therefore, the following two model classes are investigated: The Autoregressive Integrated Moving Average model with exogenous regressors (ARIMAX) and the Vectorized Autoregressive Moving Average model with exogenous regressors (VARMAX):

$$Y_t = \sum_{i=1}^{p} \Phi_i Y_{t-i} + \sum_{j=1}^{q} \Theta_j \mathcal{E}_{t-j} + \mathcal{A} X_t + \mu_t + c + \mathcal{E}_t \quad (2)$$

Here, Y specifies a multivariate time series. Correspondingly, $\mathcal{E}$ specifies a vector of error terms. In the case of ARIMAX, Y and $\mathcal{E}$ are reduced to scalars. p and q denote the AR and MA model orders, respectively. Φ_i, Θ_i and $\mathcal{A}$ are vectors that hold the parameters of the model. X details a vector that holds all exogenous variables that are considered. Finally, μ specifies a linear trend and c a constant. While the ARIMAX model considers only a single-point PS time series to account for individual dynamics, the VARMAX model utilizes multiple PS point time series to explore the benefit of using spatial dynamics. For both model classes, the implementation provided by Statsmodels is used.

3.3. Model search

Since there are different combinations of exogenous variables, model classes, and seasonal representations that might be relevant for a specific PS point, an evaluation of various combinations is performed, which also includes the evaluation of different model orders (p, q, and d). As a model selection criteria, a prediction of unseen data (validation data) and a calculation of the Mean Average Percentage Error (MAPE) between the prediction and the actual values is performed. Since each PS time series represents a different position on the gravity dam, it is estimated that the deformation patterns might vary greatly. An individual model for each PS point is found to account for this. The specification with the lowest error for each specific PS point is selected as the final model. Notably, this model search procedure extends previous standards such as [6] by selecting the best statistical representation from many possible candidates in a data-driven manner instead of formulating a single statistical model based on expert knowledge.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The following search space was exhaustively evaluated: $p \leq 2$, $q \leq 2$, $d \leq 1$, the order k of $F_k \leq 1$. T and W are inclusively evaluated until a lag of two. For T^M and W^M, the mean of the last three and seven days before the PS point measurement was evaluated. All of the named Hyperparameters are optional, meaning models that exclude these terms completely were also evaluated down to a null model. All combinations are evaluated with and without the subtraction of seasonal mean. Each combination was scored by predicting the year 2019, which was excluded during model fitting. In total, more than 14,000 combinations were evaluated for each PS point to select the final models. These final models were then retrained on the complete time series until 2020 and used to predict the complete year 2020. All predictions are performed on a six-day resolution. Finally, all PS points of the ascending and descending direction and the east-west and the vertical components were predicted, performing no sub-selection of the available PS points. To visualize the prediction quality

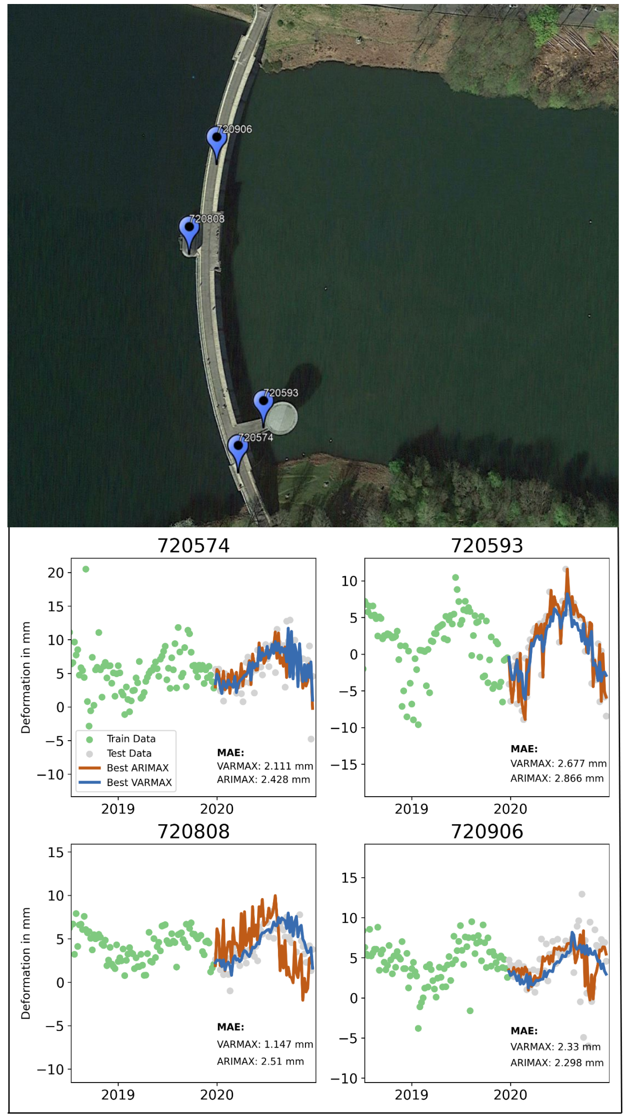

Fig. 2. Location of all PS points in ascending direction on the Lister Dam (top) and their corresponding deformation predictions for 2020 (bottom). For comparison, the best individual ARIMAX and VARMAX models are shown.

exemplarily, all PS points of the ascending direction and the corresponding prediction for 2020 are shown for both model classes in Figure 2. In total, four PS points were detected on the Lister Dam in the ascending direction. In Figure 3, the residual distribution for the ascending PS points is displayed. Additionally, a comparison of exogenous variable usage in the final models is included in Table 1. Error statistics for all LOS directions and additional components (east-west and vertical) are provided in Table 2.

The final selection of models generally provides reasonable results for all PS-points, with a small mean absolute error of, in the case of the ascending direction, 2.06 mm. Most importantly, the resulting forecasts are more accurate than what

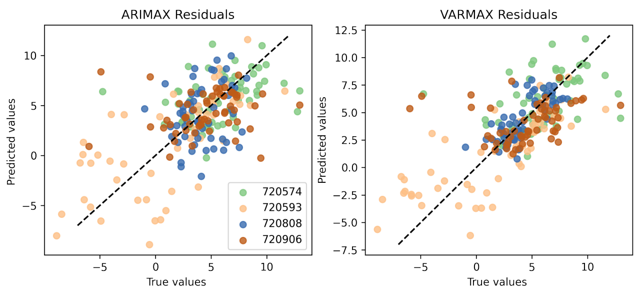

Fig. 3. Residual plot of the best-performing ARIMAX (left) and VARMAX (right) model for the Lister Dam for all PS points in ascending direction.

could be predicted by only using W and T of the current day (MAE of these models are included in Table 2 for comparison). This confirms the benefits of selecting models for specific PS-point time series in a data-driven manner. Further, the final specifications of model order, exogenous variables, and preprocessing differ completely for each PS point, validating the notion that an individual modeling approach should be deployed. While most models include the influence of water level and air temperature in some form, additional influences differ from model to model (Table 1). Intuitively, since the position of the PS points (Figure 2) lies in different dam sections, this is deemed a valid phenomenon. As a slight tendency, it was observed that the water level seems to have a more pronounced impact in the final models (Table 1). Further studies are, however, required to test whether this observation also corresponds to actual physical phenomena.

Finally, for many PS points, the performance of the best VARMAX model is slightly more robust for the year 2020, suggesting that the PS points hold at least some relevant information for each other. Especially for PS-point 720808, the ARIMAX forecasts fail to predict the second half of 2020 properly. While investigating the reason for this bad performance, it was found that the steep drop in predicted deformation can most likely be attributed to a change in water levels at the Lister gravity dam, which drops strongly simultaneously. This, however, also means that the relationship between the PS-point time series and water level was not properly learned. While this could suggest that the model selection is not optimal in this case, it could also mean that additional PS-point preselection is necessary for a robust monitoring strategy to filter out time series that show no relationship to known deformation drivers. As the quality of the PS-points time series can vary with how well they are aligned with the line-of-sight of Sentinel-1, and there are typically many points available, this seems beneficial for deploying such a monitoring strategy in practice.

ID	720574	720593	720808	720906
W_t	✓	✓	✓	✓
T_t	✓	✗	✓	✗
W_{t-n}	✗	✓	✓	✓
T_{t-n}	✗	✗	✓	✗
$W_{3/7}^M$	✗	✗	✓	✓
$T_{3/7}^M$	✓	✗	✓	✓

Table 1. Exogenous variables related to either water level or air temperature that are represented in the final models. For PS points 720574, 720593, and 720808, results for VARMAX models are shown, while for PS point 720906, the ARIMAX model results are shown, as they produce a lower MAE. Interestingly, the specific variable selection is PS point specific, emphasizing the need for an individual model specification.

MAE (mm)	Asc.	Desc.	EW	Vert.	Total
ARIMAX	2.53	2.32	2.02	1.73	2.15
VARMAX	2.07	1.86	1.80	1.51	1.81
Baseline	4.10	4.21	4.10	2.92	3.83

Table 2. Mean absolute error (MAE) in millimeter over all points of a specific LOS (ascending and descending) or additional components (east-west and vertical) for the Lister Dam.

5. CONCLUSION AND OUTLOOK

This work investigated a dam monitoring strategy by predicting deformation patterns based on freely available PS time series. Combined with a proper representation and selection of exogenous regressors and preprocessing, it was found that well-established statistical forecasting models can achieve sufficient accuracy for this endeavor, especially when deploying a data-driven model selection. It can be concluded that the approach provides a potential low-cost alternative to established monitoring strategies. Future work could address the following three topics: Firstly, while the two main drivers of dam deformation are clear (water level and air temperature), other exogenous variables, such as frost, groundwater, or tectonic movement, might additionally influence the deformation pattern of the dam. It is planned to test these additional influences to potentially increase prediction accuracy. Secondly, while this work currently only considers linear relationships between all regressors and the target, other alternative modeling approaches should be tested, which might allow for even more precise predictions. Specifically, nonlinear relationships and interaction effects might be of great use. Finally, to determine the flexibility of our approach, we deem it necessary to extend our analysis to other dams and include some form of PS-point preselection. With this, we hope that it will be possible to integrate the approach into real-world monitoring strategies.

6. ACKNOWLEDGEMENTS

The authors would like to thank the Federal Institute for Geosciences and Natural Resources for providing a PSI dataset from 2015-2020 encompassing additional information to the one freely available online and for their support in the data interpretation.

7. REFERENCES

[1] German Institute for Standardization (DIN), "DIN 19700-10:2004-07, Stauanlagen- Teil 10: Gemeinsame Festlegungen," 2019.

[2] Michele Crosetto, Oriol Monserrat, María Cuevas-González, Núria Devanthéry, and Bruno Crippa, "Persistent scatterer interferometry: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 78–89, 2016.

[3] Jänichen, J., Schmullius, C., Baade, J., Last, K., Bettzieche, V., Dubois, C., "Monitoring of Radial Deformations of a Gravity Dam using Sentinel-1 Persistent Scatterer Interferometry," *Remote Sensing*, vol. 14, no. 5, pp. 1112, 2022.

[4] Ziemer, J., Jänichen, J., Stumpf, N., Wicker, C., Klöpper, D., Last, K., Demisch, G., Schmullius, C., Dubois, C., "Sensor Fusion for detecting Deformation Patterns of Gravity Dams in Germany using SAR Interferometry," *(submitted)*, 2024.

[5] Kalia, A., Frei, M., Lege, T., "BodenBewegungsdienst Deutschland (BBD): Konzept, Umsetzung und Service-Plattform," *ZfV-Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 2021.

[6] Volker Bettzieche, "Mathematisch-statistische analyse von messwerten der talsperrenüberwachung," *Wasserwirtschaft*, vol. 94, no. 01-02, pp. 1–5, 2004.

[7] Ruhrverband, "Biggetalsperre. Sicherheitsbericht, Teil a. Unpublished security report.," 2020.

[8] Federal Institute for Geosciences and Natural Resources (BGR), "BodenBewegungsdienst Deutschland (BBD)," 2023.