Spatiotemporal reconstruction of 4D point clouds at different time scales through implicit neural representations for topographic monitoring applications

Mathilde Letard[1], Peter Naylor[2], Katharina Anders[1]

[1] TUM School of Engineering and Design; Technical University of Munich, Germany - (mathilde.letard, k.anders)@tum.de
[2] Φ-lab, ESRIN, ESA, Frascati, Italy - peter.naylor@esa.int

Keywords: Implicit neural representations, multitemporal point clouds, spatiotemporal reconstruction, environmental monitoring.

Abstract

Monitoring surface change in dynamic environments is essential to preserve the integrity of human infrastructure and livelihoods from natural hazard consequences. With the advent of 4D remote sensing, near-continuous monitoring of dynamic scenes is unlocked. However, the unordered and irregular nature of point clouds, compounded by temporally variable occlusions and diverse acquisition conditions, hinders the accurate analysis of highly information-rich 4D data. This work addresses the challenge of irregular spatiotemporal sampling in time series of 3D point clouds for the case study of a dynamic sandy beach at different time scales. We explore the use of implicit neural representations (INRs) to model 4D data as continuous spatiotemporal functions that are optimised to estimate the beach topography continuously through space and time. By comparing four model variants and assessing their performance to reconstruct spatially and temporally subsampled data, we evaluate the applicability of INRs to high-frequency topographic monitoring, especially in the context of 4D change analysis. Our results show the ability to reconstruct missing epochs from time series of 3D point clouds with centimetric to decimetric accuracy at time scales ranging from seasonal to daily observations. Our findings highlight the importance of hyperparameter tuning to enable the capture of local details in complex spatiotemporal datasets. Through this, our work lays the foundation for continuous spatiotemporal representation of dynamic scenes, supporting a potentially broad range of change analysis applications.

1. Introduction

In the context of climate change, Earth surface processes like landslides, coastal erosion, or glacier retreat are set to intensify, threatening the integrity of human lives and infrastructures (IPCC, 2023). Monitoring topographic dynamics is necessary to better quantify, understand, and mitigate related risks. Remotely sensed 4D data, obtained through repeated 3D surveys of scenes of interest, provide an opportunity to improve the observation of environments subject to topographic changes. They can be acquired using various platforms and sensors, ranging from terrestrial laser scanning (TLS) to aerial photogrammetry. Over the last decade, continued multi-temporal acquisition and lately the emergence of permanent observation setups have led to the availability of extremely dense observation of dynamic scenes in different natural environments (Lindenbergh et al., 2025; Eitel et al., 2016). However, the irregular nature of 3D point clouds (PCs) combined with the complexity added by the temporal dimension pose data processing challenges. Performing 3D change detection within a pair of PCs is indeed challenging, first because the points are unordered and irregularly spread (Qin et al., 2016). Matching 3D points between two acquisitions is secondly not straightforward since the location of the points changes significantly, even in the absence of surface change (Lague et al., 2013). These difficulties propagate through the many observation epochs of 4D datasets and compound those resulting from occlusions in the scenes. Whether caused by environmental conditions, temporary objects, or sensor-related factors, occlusions cause gaps in the spatial coverage, subsequently affecting the temporal sampling regularity. These factors introduce additional challenges to 4D change analysis based on elevation change time series processing, an already complex task due to the massive and heterogeneous aspects of 4D PCs (Lindenbergh et al., 2025).

This work addresses the issue of irregular spatiotemporal sampling in 4D datasets, which hinders accurate subsequent topographic time series analyses. We explore the applicability of implicit neural representations (INRs) for 4D data processing by analysing their performance for spatiotemporal reconstruction in a dense time series of a dynamic sandy beach environment. INRs are used to compress multitemporal 3D PCs of this changing scene into continuous functions that can be queried at spatiotemporal coordinates of interest to access the surface elevation across time. We evaluate in particular their ability to reconstruct missing epochs and to perform spatial super-resolution. Finally, we explore the accuracy of elevation time series generation through INRs and thereby assess their potential to advance spatiotemporal scene representation for change analysis tasks and environmental monitoring applications.

2. Related work

2.1 Monitoring environmental dynamics with topographic laser scanning

The application of TLS for monitoring dynamic natural scenes, such as beaches, presents unique challenges (Lindenbergh et al., 2025). The off-grid and noisy nature of 3D PCs, combined with the often subtle and distributed nature of surface changes (e.g., erosion, accretion, sediment redistribution), can introduce artifacts and uncertainty into change detection and quantification (Anders et al., 2019). A common alternative to processing full-3D data is to use 2D raster derivatives. In this paradigm, 3D PCs are interpolated onto a regular 2D grid to derive Digital Elevation Models (DEMs) (Stular et al., 2012). The regularity of these grids makes them ideally suited for convolutional neural networks (CNN), and several studies have applied 2D CNN architectures to DEM pairs for change detection (Shi et

al., 2020; Zhang et al., 2019). However, this projection and interpolation process inevitably leads to a loss of the original precision and detail captured by the LiDAR data, potentially smoothing out fine-scale features critical for understanding environmental dynamics (Naylor et al., 2025). Besides, DEMs are not able to preserve the 3D nature of surface changes that is inherent in many environments (Qin et al., 2016). Alternatives include cloud-to-cloud (Girardeau-Montaut et al., 2005) or M3C2 (Lague et al., 2013) distance computation, which takes into account the local surface orientation to return signed distances. These approaches have been extended to time series-based analysis of dense 4D PCs for near-continuous monitoring of natural environments (Anders et al., 2021; Winiwarter et al., 2023). However, the need to find point neighborhood correspondences within pairs of acquisitions to compute surface change remains a limitation. Indeed, when occlusions occur, even dedicated methods like M3C2 are not able to compute distances in these local areas. Another issue, especially in the context of near-continuous acquisition, is the total absence of one acquisition, i.e. temporal gap, disrupting the regularity of observation commonly required by algorithms. To account for this, time series-based 4D change analysis methods commonly rely on temporal smoothing to recover a regular temporal distance between data points (Anders et al., 2019; Eltner et al., 2017). While this leverages temporal sampling redundancy also to reduce noise in the data (Kromer et al., 2015), simple rolling median approaches introduce precision loss and may prevent making use of the complete dataset at its original spatiotemporal resolution. An alternative for advanced time series smoothing is to use Kalman filtering on time series of the M3C2 distances (Winiwarter et al., 2023). However, these strategies all consider the spatial and temporal dimensions separately, which may not exploit the spatiotemporal context to its full extent.

In contrast to direct interpolation, our method uses an INR to model the continuous surface in space and time, allowing for the generation of 3D scene representations at an arbitrary spatial and temporal resolution, thereby potentially increasing the accuracy and completeness of topographic information.

2.2 Implicit Neural Representations (INRs)

Dealing with unstructured PC data in space and time is not straightforward and requires flexibility in the model, which is possible with Neural Networks (NNs). INRs (Xie et al., 2022) are coordinate-based NNs that aim to encode a *field* implicitly and continuously (e.g., from pixel position to colour in an image). Unlike traditional deep learning, where a dataset is an ensemble of input data associated with annotations, INRs are usually trained to reconstruct a single instance over delimited space and time. INRs have had wide success in 1D with audio reconstruction (Kazerouni et al., 2024), in 2D with image representation (Molaei et al., 2023), in 3D object reconstruction (Huang et al., 2024; Mildenhall et al., 2021), and in 4D for bi-temporal surface reconstruction (Naylor et al., 2024). The strength of INR lies in the representational power of deep learning to map arbitrary inputs, here the coordinates (Tancik et al., 2020; Sitzmann et al., 2020), to the *field*. Such a link allows us then to penalize the model by quantities like the function derivative with respect to the input coordinates, methods that have been at the core of Physics-Informed Neural Networks (Raissi et al., 2019; Di Carlo et al., 2022; Karniadakis et al., 2021). The framework also allows for signal compression (Pistilli et al., 2022; Strümpler et al., 2022) where the model size scales with the complexity rather than the output dimension. Additionally, INR can natively interpolate the target *field*. Hence, INRs are trained by fitting observations to a *field* and inducing desired properties by application-based regularisation terms.

An important line of research is to tackle the limited representational power of INR to capture high-frequency patterns in the signal, which is referred to as the spectral bias. In this respect, modifying the input basis representation of coordinates with positional encoding or Random Fourier Features (RFF) (Rahimi and Recht, 2007; Tancik et al., 2020) has enabled more precise reconstruction of the *fields*. An alternative approach to mitigate the spectral bias is to modify the nonlinearity functions by periodic sine functions, provided with SIREN (Sitzmann et al., 2020). This architecture has shown a very low reconstruction error for the target *field* and its gradient. Despite their success, training SIRENs is difficult as they are prone to overfitting, and careful parameter tuning is required. These issues have been addressed in subsequent works (Fathony et al., 2021; Saragadam et al., 2023; Lindell et al., 2022; Kazerouni et al., 2024), where novel architectures have been proposed, achieving easier training and faster convergence. Nevertheless, the respective model choice depends on the application task. As reported in many works, neither SIREN nor more recent approaches consistently outperform RFF-based architecture (Naylor et al., 2024).

INRs have been applied to the field of Earth Observation in recent years (Cole et al., 2023; Rußwurm et al., 2023; Cai and Balestriero, 2025). However, their usage is still somehow limited due to the inherently complex nature of remote sensing data (Yu et al., 2025). For sparse or off-grid data, INRs have been used to improve 3D reconstruction, leading to sharper visual features (Lutio et al., 2019; Liu et al., 2025). They have been applied to monitor urban sprawling (Naylor et al., 2024) and for reconstructing the Greenland ice sheet (Naylor et al., 2025) from satellite altimetry PCs. A major difficulty in the application of INR to 4D reconstructions in the field of geosciences is to adequately capture both the spatial and temporal dynamics in the data (Krishnapriyan et al., 2021; Chen et al., 2023).

In this paper, we investigate the applicability of 4D INR methods for spatiotemporal reconstruction of 4D PCs at variable time scales with data acquired in the complex environment of a sandy beach through permanent TLS.

3. Materials and method

We employ INRs with the aim of modelling spatiotemporal surfaces from 4D laser scanning data. Our study focuses on reconstructing the surface and its dynamics of a sandy beach. The objective is hence to estimate the (changing) elevation z within the beach scene continuously across space and time. Once optimised, an INR model acts as a function $f(t, x, y) = z$, where t represents time and x, y are spatial coordinates. f can then be queried at any t, x, y within the extent covered by the training data to recover z. This enables the estimation of z for values of t, x, y not initially measured, i.e., interpolation, thus unlocking seamless spatiotemporal analysis of the scene.

3.1 Data and study area

We use openly available data published within the CoastScan project (Vos et al., 2023), consisting of hourly PCs of the beach-dune system at Noordwijk in the Netherlands. Data were acquired through permanent laser scanning between July 2019 and June 2022. This results in a near-continuous 4D dataset of

21,812 epochs encapsulating the daily, seasonal, and annual dynamics of this coastal area at high resolution (from 50 points/m² down to 1-5 points/m² in the outer parts of the scene).

In this work, we crop the spatial data extent to only consider the core area of the sandy beach (i.e., removing the dune) in order to assess the ability of INRs to capture gradual and temporary changes caused by wave action (including tides) and coastal meteorological conditions. The beach has a relatively planar topography at the macroscale, and its elevation varies more strongly in the cross-shore (here, longitudinal) direction than in the along-shore (here, latitudinal) direction. Several aspects add challenges to the analysis of the 4D dataset, as shown in Fig. 1. First, the tides and related varying water levels result in variable coverage, especially of the lower beach area. Second, due to sensor outages and adverse weather conditions, some epochs are missing entirely or feature particularly noisy or reduced PCs. Last, the presence of temporary objects on the beach causes sudden variations in elevation that do not represent the beach topography nor relevant surface processes thereon (e.g., persons, vehicles, seasonal buildings for tourism). By focusing on the use case of a sandy beach sampled at high frequency (sub-daily) over a long period (three years), we address the problem of reconstructing a highly dynamic spatiotemporal surface characterised by strong variations between spatial and temporal frequencies. Coupled with the many instances of spatiotemporal occlusions, these aspects render the dataset particularly challenging and thus interesting to the evaluation of INR performance in dynamic environments captured with 4D data.

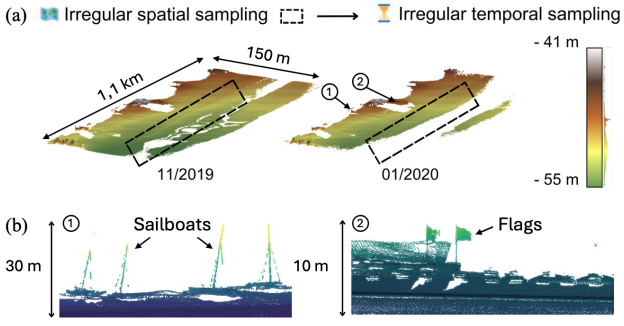

Figure 1. Example point clouds of the beach at Noordwijk, NL, illustrating the challenges of 4D data sampled irregularly in space and time: a) two epochs with differing surface coverage, b) temporary objects present only in some epochs.

3.2 Implicit neural representation architectures

We compare four INR models for 4D topographic reconstruction, which are all coordinate-based NNs for predicting the elevation z from spatiotemporal coordinates $\mathcal{X} = (t, x, y)$ with distinct mechanisms to capture spatial and temporal patterns.

RFF: The first model we evaluate combines a multilayer perceptron (MLP) with an RFF projection layer in order to map the input spatiotemporal coordinates $\mathcal{X}$ in a higher-dimensional space. This transformation γ aims to disentangle close coordinates and capture higher frequencies in the input data, ultimately improving reconstruction (Tancik et al., 2020). The RFF Gaussian projection is defined as:

$$\gamma(\nu) = [\cos(2\pi\nu),\ \sin(2\pi\nu)] \tag{1}$$

where $\nu = B\mathcal{X}$ and $B \in \mathbb{R}^{M\times 3}$, $\forall(i,j), B_{ij} \sim \mathcal{N}(0,\sigma)$. The size of B depends on the size of $\mathcal{X}$, here $\mathcal{X} \in \mathbb{R}^3$. The mapping size M and the scale σ impact the learned levels of detail and require fine-tuning. After transformation, an MLP is used to predict the elevations from the features. In this paper, we name this model the *original RFF*.

Spatiotemporal RFF: To accommodate strong differences between spatial and temporal frequencies in the data, the original RFF can be modified to encode the spatial and temporal dimensions separately. In this version, which we name *modified RFF*, two random matrices are drawn independently:

$$\nu = [B_s\mathcal{Y}, B_t\mathcal{T}] \tag{2}$$

where $\mathcal{T}$ denotes temporal coordinates, $\mathcal{Y}$ spatial coordinates, $B_s \sim \mathcal{N}(0,\sigma_s)$ and $B_t \sim \mathcal{N}(0,\sigma_t)$ with σ_s and σ_t as two scaling parameters. The obtained ν is then passed to γ as in Eq. 1. σ_s and σ_t control the range of spatial and temporal frequencies learned by the network, and require optimisation.

SIREN: The SIREN network (Sitzmann et al., 2020) replaces the commonly used ReLU or tanh activation functions in MLPs with sine activation. These functions enable the network to represent continuous complex signals without the need for a specific projection layer. It has been demonstrated that the RFF projection layer can be represented by a single SIREN layer, the only difference being that SIREN possesses learnable parameters (Benbarka et al., 2022). SIREN models are sensitive to weight initialisation due to the tendency of the sine activation to saturate. To counterbalance this, a parameter σ_{SIREN} is used to scale the weights of the activation layer, similar to RFF models. Mathematically, a SIREN network can be expressed as:

$$\begin{aligned} f(\mathcal{X}) &= W_n(g_{n-1} \circ g_{n-2} \circ \ldots \circ g_0(\mathcal{X})) + h_n, \\ \text{where} \quad g_i(\mathcal{X}) &= \sin(W_i\mathcal{X} + h_i), \end{aligned} \tag{3}$$

where h_i and W_i, for all i, are learnable parameters, $\circ$ denotes the function composition operator, and n the number of layers.

KAN: The Kolmogorov-Arnold Network (KAN) (Liu et al., 2024) follows a different architecture than MLPs. It is inspired by the Kolmogorov-Arnold representation theorem, which states that any multivariate continuous function can be approximated by superposing continuous univariate functions and additions. Contrary to traditional MLPs founded on the universal representation theorem, KANs employ learnable activation functions, which are most commonly splines. With this structure, they can capture more complex patterns while employing a smaller layer width and network depth than MLP-based INRs. Mathematically, a two-layer KAN of width q corresponds to:

$$f(\mathcal{X}) = \sum_{i=1}^{q} \Phi_i \left(\sum_{j=1}^{p} \phi_{ij}(\mathcal{X}_j) \right) \tag{4}$$

where Φ_i and ϕ_{ij} are learnable functions. Here, ϕ_{ij} transforms the input dimensions of the p samples $\mathcal{X}$, while Φ_i combines these transformations to produce the final output. The performance of KANs is highly dependent on the initialisation and optimisation of the univariate functions. While their flexibility is essential to fit complex signals, it can hinder training stability.

3.3 Experimental setup

We apply the different INR architectures to the modelling of surface dynamics in the sandy beach dataset. Since the beach environment involves different scales of spatial and temporal process patterns, we rely on different priors to optimise the resulting spatiotemporal surfaces.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-527-2026 | © Author(s) 2026. CC BY 4.0 License.

Surface priors: Globally, the beach topography features low-frequency variations in the spatial domain, but the surface roughness can be higher at the closer local scale in the presence of sand ripples (micro-scale) or sand bars (meso-scale). To enable fitting these stronger local variations without the effect of overall less accurate spatial gradients, we use spatial gradient-based regularisation. As our study area is also characterised by higher elevation variations in the cross-shore direction than along the shore, we apply different magnitudes of constraint for longitudinal and latitudinal surface gradients. Since coastal areas are highly dynamic due to tidal and meteorological forcing, the temporal patterns in the data feature high frequencies. To preserve these actual variations while avoiding overestimation of changes between observed epochs, we also apply regularisation to the temporal gradient. We use Total Variation Norm (TVN) regularisation to enforce these priors and extend it to the time domain to constrain the spatial and temporal gradients of the learned surface, as detailed in Eq. 5. This also counterbalances the typical overfitting of INRs.

$$\mathcal{R}_{\text{TVN}} = \alpha_x \left\| \frac{\partial f}{\partial x} \right\|_2 + \alpha_y \left\| \frac{\partial f}{\partial y} \right\|_2 + \alpha_t \left\| \frac{\partial f}{\partial t} \right\|_2 \quad (5)$$

where $\alpha_x, \alpha_y, \alpha_t$ are regularisation weights, f the spatiotemporal function, x, y spatial coordinates, and t time.

With this, we aim to learn finer details of the training dataset while limiting errors on validation and test data. As a result, our loss function $\mathcal{L}$ is a combination of Mean Absolute Error (MAE) and spatiotemporal TVN, as expressed in Eq. 6.

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} |z_i - \hat{z}_i| + \mathcal{R}_{\text{TVN}} \quad (6)$$

where z and $\hat{z}$ correspond to reference and reconstructed elevations, respectively, and N is the batch size.

Experiments: We conduct a set of experiments to evaluate our approach of spatiotemporal reconstruction on the Noordwijk beach dataset. Our goal is to investigate different application scenarios in which INRs could enhance the analysis of 4D data. The first scenario is the reconstruction of a full epoch, which can be relevant in cases of missed acquisition, e.g., due to adverse weather conditions or sensor outage. We implement this scenario by removing entire epochs within the 4D dataset and reconstructing them using the trained model. In practice, in a time series $\mathcal{S} = \{\mathcal{P}_0, ..., \mathcal{P}_N\}$, where all $\mathcal{P}_i$ are spaced regularly in time, we leave out one epoch $\mathcal{P}_t$ and compare its reconstruction to the original point cloud epoch.

The second application scenario consists of performing spatial densification within one epoch and reconstructing the full resolution of a low-resolution PC. It explores the feasibility of alternating acquisitions at full and lower resolution when setting up permanent observation systems or planning multitemporal acquisition campaigns. This reduces acquisition cost and effort, optimises data storage, and enables an optimised trade-off between acquisition time for a field of view and repetition. In practice, we compute a low-resolution version $\mathcal{P}_{t_{\text{LR}}}$ of a test acquisition $\mathcal{P}_{t_{\text{HR}}}$ of the time series $\mathcal{S}$ through voxel downsampling at 2.0 m. $\mathcal{P}_{t_{\text{LR}}}$ is then included in the training dataset, and the evaluation is performed by reconstructing $\mathcal{P}_{t_{\text{HR}}}$.

Data splits: We build four experimental datasets from the full 4D beach PCs with different temporal sampling to assess the impact of temporal resolution - reflecting variable scales of surface change dynamics and monitoring scenarios - on the ability to reconstruct the topography of the beach. The *seasonal* dataset includes one acquisition every three months for three years; *monthly* has one acquisition every month for one year; *weekly* one acquisition every week for three months; and *daily* one acquisition every day for one month. From each of these datasets, two epochs are used for validation, and one epoch for testing. We ensure that validation and test epochs are neither the first nor the last of the time series to avoid extrapolation, and constrain their selection so that validation and test epochs cannot be consecutive and are separated by at least one training epoch. The properties of each dataset are presented in Tab. 1.

Dataset	Seasonal	Monthly	Weekly	Daily
Points	5,070,366	5,236,257	5,420,952	8,920,862
Time range	01/09/19 - 01/06/22	01/09/19 - 01/09/20	04/12/19 - 01/04/20	01/01/20 - 31/01/20
Epochs	12	13	17	29

Table 1. Properties of each dataset. Points are the total number of points in the dataset, dates are given in format DD/MM/YY.

Optimisation: Each model is trained for 300 training epochs (iterations over the entire training dataset) with a learning rate of 10^{-3} and a batch size of $4,096$. For each dataset, we perform 200 trainings to optimise the model hyperparameters. This procedure is repeated for each INR model. The parameters σ_{RFF}, σ_s, σ_t, σ_{SIREN}, α_x, α_y, α_t, and the mapping size M of the RFF are optimised using the Optuna framework (Akiba et al., 2019). We use the MAE to monitor the performance of the models during training and hyperparameter optimisation. The model with the best (here, lowest) validation metrics is then retained.

Baseline: We compare our INR models to two baseline spatiotemporal interpolation methods. The first, *Baseline*^{T}, is a 2D bilinear interpolation based on spatial and temporal neighborhoods. For each test point, we add temporal neighbors from the previous and following epoch. Depending on the temporal scheme, these points can be distanced by a season, month, week, or day. We collapse the time axis and perform an interpolation weighted by the spatial point distances to obtain the target value. The second baseline method, *Baseline*^{S}, only applicable to the spatial densification experiments, is similar to the first, except that we do not consider neighboring temporal points and only consider points within the considered epoch.

4. Results

4.1 Spatiotemporal reconstruction metrics

The MAEs obtained in the spatial and temporal densification experiments for each dataset and each model are reported in Tab. 2 and 3, respectively. Our results show that INRs reconstruct elevations with lower error than the baseline bilinear interpolations when considering spatial densification, except in the *seasonal* case. However, temporal reconstruction is less accurate: for the *monthly* and *daily* datasets, the baseline MAE is lower than the best-performing INR model by 4 cm and 1 cm, respectively. In the temporal leave-out experiments, the performance of all methods tends to decrease with increasing temporal distance between observations, except for KAN in the *weekly* scenario. For example, the MAE of the modified RFF reconstruction is seven times lower on the *daily* dataset than on the *seasonal* dataset. Finally, reconstructions of spatially subsampled data tend to be more accurate than for temporal subsampling, where complete epochs are missing.

Method	Mean Absolute Error (m)			
	Seasonal	Monthly	Weekly	Daily
Baseline^{T}	**0.032**	0.108	0.032	0.040
Baseline^{S}	0.345	0.167	0.053	0.030
SIREN	0.176	0.188	0.057	0.051
KAN	0.196	0.238	0.079	0.083
Original RFF	0.051	**0.106**	**0.027**	0.029
Modified RFF	0.060	0.116	0.029	**0.027**

Table 2. Results for the spatial subsampling experiments. Best values shown in bold.

Method	Mean Absolute Error (m)			
	Seasonal	Monthly	Weekly	Daily
Baseline^{T}	0.353	**0.197**	0.066	**0.038**
SIREN	0.512	0.256	0.117	0.056
KAN	**0.299**	0.266	**0.057**	0.071
Original RFF	0.644	0.239	0.067	0.056
Modified RFF	0.342	0.253	0.064	0.048

Table 3. Results for the temporal subsampling experiments. Best values shown in bold.

4.2 Spatial distribution of errors

Fig. 2 shows the spatial distribution of estimation errors for the best INR model in each experiment on the *seasonal* and *weekly* datasets. The reconstruction of the left-out epoch in the *seasonal* data shows a tendency to underestimate the original elevation, but the other experiments do not share this bias. The distribution of spatial densification errors suggests that fine-grained details are not precisely captured by the models.

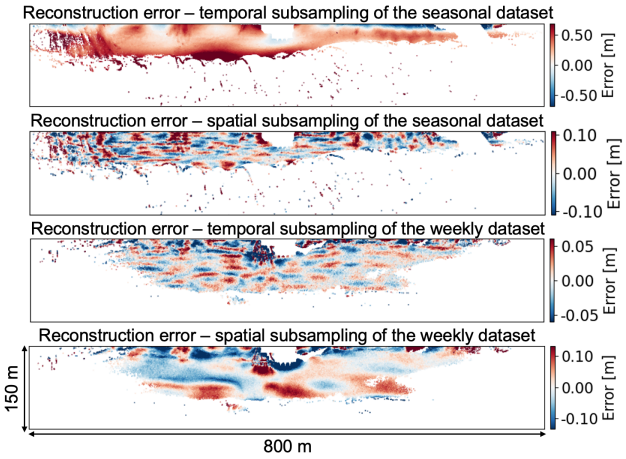

Figure 2. Reconstruction error for spatial and temporal subsampling experiments obtained with the best INR model for the seasonal and weekly datasets (reference-reconstructed).

4.3 Time series

By querying the trained models on each epoch covered by the datasets, it is possible to generate time series of elevations and assess the ability of INRs to reconstruct the details of the temporal patterns. Fig. 3a and 3b show examples of time series obtained from the same two datasets for the temporal gap and spatial subsampling experiment, respectively.

Overall, our results show that INRs are able to retrieve global temporal patterns in the variable time scale datasets. When performing spatial densification, detailed patterns in the time series are reconstructed more accurately than in experiments with missing epochs. In particular, abrupt transitions between

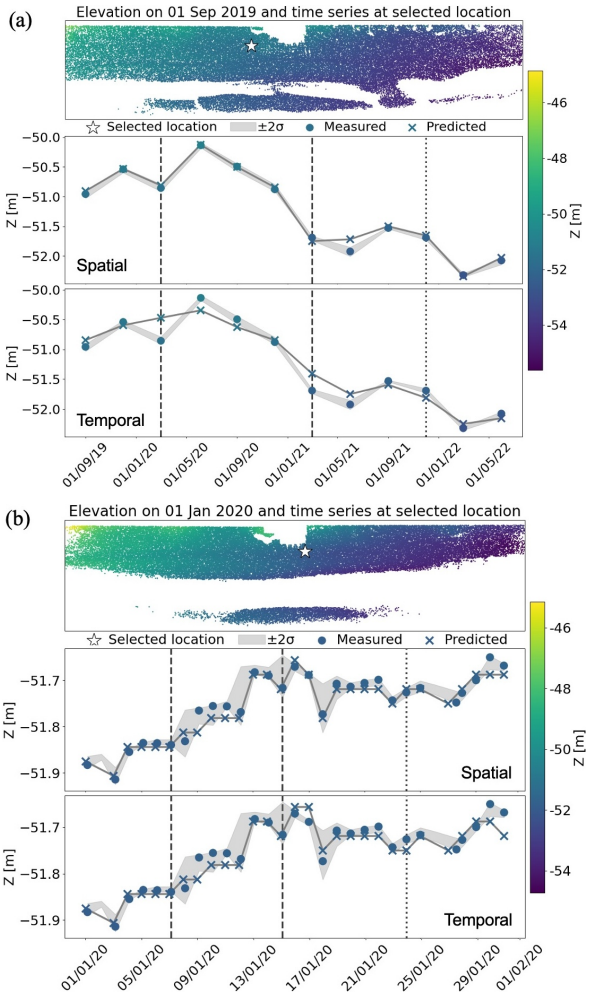

Figure 3. Comparison of original and reconstructed elevation time series obtained with the best-performing spatiotemporal reconstruction models on the a) seasonal and b) daily datasets for each experiment. Validation and test epochs are indicated with dashed and dotted lines on the time axis, respectively. The gray buffer represents two standard deviations of local elevation in the original data within four days (seasonal) or 12 hours (daily) of each selected epoch, thus illustrating the influence of single epoch selection on the quantification of reconstruction accuracy.

accretion and erosion (slope sign changes) are underestimated in validation or test epochs. We observe this across experiments of different time scales. However, the experiments on the daily data show a saturation of the (higher) frequencies captured by the INR. Small variations between points that are close in time are not as accurately represented as greater changes, leading to a step effect in the reconstructed time series.

5. Discussion

5.1 Spatiotemporal reconstruction performance

We evaluate spatiotemporal reconstruction in distinct experiments on datasets with spatial or temporal gaps. As reflected in Tab. 2, RFF-based INRs perform spatial super-resolution with errors ranging from 3 cm to 11 cm across time scales, and surpass other architectures in terms of accuracy. The seasonal dataset is the only experiment for which spatiotemporal bilinear interpolation (Baseline^{T}) outperforms the RFF reconstruction. In

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-527-2026 | © Author(s) 2026. CC BY 4.0 License.

that case, the complete change history might complicate epoch reconstruction as opposed to the integration of temporally adjacent acquisitions only, as done in Baseline^{T}. This might be related to the globally higher change magnitudes that occur across the larger time scales covered by this dataset. However, RFF models consistently outperform purely spatial bilinear interpolation (Baseline^{S}) even though they incorporate highly variable multi-temporal data, demonstrating their strength in operating on a full 4D dataset at once. Overall, temporal redundancy appears to be important even for spatial densification, as Baseline^{T} also outperforms Baseline^{S}, except for the daily case.
Regarding temporal reconstruction, Tab. 3 shows overall higher errors. This can be attributed to the difficulty of predicting temporary or sudden changes occurring between two epochs used for training in the absence of any additional information about the onset of such changes. This translates as a form of inertia in the time series, especially for temporal subsampling experiments. For validation and test epochs, the reconstructed elevation tends to follow the local temporal trend, leading to an underestimation of the change. The start of gradual erosion or accretion dynamics is typically reconstructed with a delay, showing the impact of the training and validation strategy on the results. Abrupt temporary changes are even harder to capture, as shown with the first validation epoch of Fig. 3a. This explains the higher MAE for experiments where more temporary objects are present (e.g., summer touristic infrastructure), as in the *seasonal* and *monthly* datasets, where the lowest errors are decimetric while *weekly* and *daily* data are reconstructed at centimetric accuracy. For the *monthly* and *daily* cases, Baseline^{T} outperforms all INRs, suggesting that the reconstruction of their test epochs is facilitated by the integration of adjacent epochs only. In the end, in all but one experiment (seasonal spatial densification), there is always an INR method performing worse than the baseline, which suggests that the choice of INR over bilinear interpolation depends on the relevance, for each application, of employing a method that accounts for the full history of topographic change.

Importantly, our results are in line with Naylor et al. (2024) and Naylor et al. (2025), showing that the best INR architecture depends on the application and spatiotemporal data situation. Indeed, contrary to earlier works on INRs, we find that SIREN does not consistently outperform other models, including the original RFF. Overall, RFF-based models largely outperform SIREN and KAN for spatial densification. When reconstructing left-out epochs, KAN obtains the lowest MAE of all INR models for half the datasets, while RFF models rank second. For environmental applications of topographic monitoring, RFF models can thus be recommended, resulting in low reconstruction error while maintaining a lower computational cost (180 iterations/sec) than KANs (38.5 iterations/sec).

5.2 Optimisation of INR models

Our findings emphasize the importance of the used validation scheme and the model hyperparameters, in agreement with other studies (Essakine et al., 2024). INR parametrization, therein, particularly impacts the range of frequencies that can be captured both in time and space, and subsequently the level of detail and accuracy in the reconstruction (Essakine et al., 2024). The ability to capture local spatial texture and roughness largely depends on parameters controlling the spectral bias, such as σ, σ_s, σ_t, or M. Identifying suitable ranges for these parameters is a laborious tuning process that is largely application- and data-dependent, and automatic optimisation is a practical solution that can be recommended for deployment to new data. To optimise the reconstruction of both spatial and temporal fine-grained details, a suitable balance must be found between high-frequency learning and regularisation. Without regularisation, the models tend to overfit on the training data and underperform on validation and test epochs. Another pitfall is fitting only low frequencies, resulting in similar MAEs but producing largely smoothed out reconstructions. To tackle this, the use of separate parameters for spatial and temporal domains, as in our modified RFF, can be useful. Its higher performance on both *daily* experiments reflects the larger difference between spatial and temporal frequencies in this dataset, where the beach topography undergoes high-frequency variations on a daily basis due to tides and short-term forcing, as well as typical high-frequency measurement noise in near-continuous laser scanning data (Lindenbergh et al., 2025). The difficulty of finding adapted parameters, however, cannot be fully remedied when treating spatial and temporal dimensions separately, as indicated by the step-like effect in the *daily* time series (Fig. 3b), where the RFF frequency range likely does not enable disentangling small-magnitude, short-term temporal patterns.

Our experiments and results also highlight the challenge of designing an appropriate validation procedure for reconstruction tasks. Automatic hyperparameter optimisation relies on a validation criterion that is challenging to formulate for both spatial and temporal reconstruction. For example, with the current data split, it is impossible to quantify time series reconstruction metrics without implicitly leaking validation information into training, since constructing each time series necessarily mixes epochs from both splits. Enforcing accuracy in single epoch and global temporal pattern reconstruction simultaneously is thus difficult. Epoch selection also impacts reconstruction: since surface dynamics are not homogeneous in space and time, generalisation to validation and test epochs is harder when they include abrupt or temporary changes unseen in training. The elevation variance around each selected epoch in Fig. 3a and 3b illustrates the difficulty of determining a precise reference to use for validation amid high-frequency surface dynamics.

5.3 Potential for environmental monitoring

In the context of environmental topographic monitoring, INRs show promising potential to address irregular spatiotemporal sampling that may arise due to acquisition conditions (weather, infrastructure-related) or variable survey situations (temporary occlusion). The ability to reconstruct temporal variations, even with a degree of underestimation, should be further explored. Further research should aim to reduce errors in missing epoch reconstruction for deployment in operational monitoring tasks. Indeed, the capacity of INRs to perform spatiotemporal reconstruction from a time series of PCs could be particularly interesting for gap filling, a form of spatial interpolation where information from temporally close epochs can help constrain the reconstruction. The capacity to perform spatial super-resolution could also allow setting up permanent laser scanning systems to acquire lower resolution PCs in order to save on acquisition time, thus either increasing possible acquisition frequency for capturing rapid dynamics or reducing data storage. Temporal super-resolution is also feasible, though it requires further evaluation and extended experiments on different datasets, along with gap filling in future research.

Some limitations remain, in particular linked to the spectral bias of INRs. In this work, we applied INRs only to the use case of a sandy beach, but the ability to reconstruct scenes with

a higher range of spatial frequencies, i.e., more complex geometry and noise, needs to be evaluated. Applicability to scenarios covering mainly rapid temporal changes should also be further assessed, as current results suggest limited performance due to the observed inability to reconstruct full epochs with abrupt changes (see Fig. 3a). Beyond this, INRs are not suitable for extrapolation without a priori knowledge of the dynamics, and their generalisation ability is limited to the spatial and temporal extent covered by the training data. When enlarging the input data domain in time or space, or when considering a new study area, the models need to be retrained or fine-tuned. Though simple models based on RFF take only a few minutes to train (approx. 4 mins for 300 training epochs), hyperparameters likely need to be re-optimised, especially for new locations.

Despite these limitations, INRs open promising perspectives for change analysis methods. Under irregular sampling, spatiotemporal reconstruction allows exploiting the 4D PCs at their full native resolution. The continuous model obtained can then be queried to generate time series and apply time series-based 4D change analysis methods that commonly rely on regular, gap-free time series (Anders et al., 2021; Kuschnerus et al., 2021).

6. Conclusion

In this paper, we assess the performance of INR in spatiotemporal reconstruction of 4D PCs in the context of topographic monitoring at variable spatial and temporal scales. Our findings highlight promising avenues for INR in different scenarios, including the reconstruction of missing epochs or the densification of low-resolution PCs. While hyperparameter optimisation remains a critical point, the capacity of INR to provide continuous representations of spatiotemporal data has the potential to enhance accurate 4D change analysis from topographic time series. Future work should explore their extension to more complex areas where different geometry and levels of noise coexist, to scenarios of temporal super-resolution and gap filling, and their integration with change analysis methods.

Acknowledgements

This research was funded by the Bavarian State Ministry of Science and the Arts (bidt Graduate Center for Postdocs), enabled by the SpaceHPC computing facility at ESRIN, and partly supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project 535733258 (Extract4D).

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*.

Anders, K., Lindenbergh, R., Vos, S., Mara, H., de Vries, S., Höfle, B., 2019. High-Frequency 3D Geomorphic Observation Using Hourly Terrestrial Laser Scanning Data Of A Sandy Beach. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W5, 317-324. ISPRS-Annals.

Anders, K., Winiwarter, L., Mara, H., Lindenbergh, R., Vos, S. E., Höfle, B., 2021. Fully automatic spatiotemporal segmentation of 3D LiDAR time series for the extraction of natural surface changes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 173, 297–308.

Benbarka, N., Höfer, T., Zell, A. et al., 2022. Seeing implicit neural representations as fourier series. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2041–2050.

Cai, D., Balestriero, R., 2025. No Location Left Behind: Measuring and Improving the Fairness of Implicit Representations for Earth Data. *arXiv preprint arXiv:2502.06831*.

Chen, H., Wu, R., Grinspun, E., Zheng, C., Chen, P. Y., 2023. Implicit neural spatial representations for time-dependent pdes. *International Conference on Machine Learning*, PMLR.

Cole, E., Van Horn, G., Lange, C., Shepard, A., Leary, P., Perona, P., Loarie, S., Mac Aodha, O., 2023. Spatial implicit neural representations for global-scale species mapping. *International conference on machine learning*, PMLR, 6320–6342.

Di Carlo, D., Heitz, D., Corpetti, T., 2022. Post processing sparse and instantaneous 2d velocity fields using physics-informed neural networks. *20th International Symposium On Application Of Laser And Imaging Techniques To Fluid Mechanics*.

Eitel, J. U. H., Höfle, B., Vierling, L. A., Abellán, A., Asner, G. P., Deems, J. S., Glennie, C. L., Joerg, P. C., LeWinter, A. L., Magney, T. S., Mandlburger, G., Morton, D. C., Müller, J., Vierling, K. T., 2016. Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences. *Remote Sensing of Environment*, 186, 372-392.

Eltner, A., Kaiser, A., Abellan, A., Schindewolf, M., 2017. Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring. *Earth Surface Processes and Landforms*, 42, 2240-2253.

Essakine, A., Cheng, Y., Cheng, C.-W., Zhang, L., Deng, Z., Zhu, L., Schönlieb, C.-B., Aviles-Rivero, A. I., 2024. Where Do We Stand with Implicit Neural Representations? A Technical and Performance Survey. *Transactions on Machine Learning Research*, 105–110.

Fathony, R., Sahu, A. K., Willmott, D., Kolter, J. Z., 2021. Multiplicative filter networks. *International Conference on Learning Representations*.

Girardeau-Montaut, D., Roux, M., Marc, R., Thibault, G., 2005. Change detection on points cloud data acquired with a ground laser scanner. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(part 3), W19.

Huang, Z., Wen, Y., Wang, Z., Ren, J., Jia, K., 2024. Surface reconstruction from point clouds: A survey and a benchmark. *IEEE transactions on pattern analysis and machine intelligence*, 46(12), 9727–9748.

IPCC, 2023. *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440.

Kazerouni, A., Azad, R., Hosseini, A., Merhof, D., Bagci, U., 2024. Incode: Implicit neural conditioning with prior knowledge embeddings. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 1298–1307.

Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R., Mahoney, M. W., 2021. Characterizing possible failure modes in physics-informed neural networks. *Adv. Neural Inf. Proc. Syst.*, 34.

Kromer, R., Abellán, A., Hutchinson, D., Lato, M., Edwards, T., Jaboyedoff, M., 2015. A 4D Filtering and Calibration Technique for Small-Scale Point Cloud Change Detection with a Terrestrial Laser Scanner. *Remote Sensing*, 7, 13029-13052.

Kuschnerus, M., Lindenbergh, R., Vos, S., 2021. Coastal change patterns from time series clustering of permanent laser scan data. *Earth Surf. Dynam.*, 9(1), 89-103. ESurf.

Lague, D., Brodu, N., Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ). *ISPRS journal of photogrammetry and remote sensing*, 82, 10–26.

Lindell, D. B., Van Veen, D., Park, J. J., Wetzstein, G., 2022. Bacon: Band-limited coordinate networks for multiscale scene representation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16252–16262.

Lindenbergh, R., Anders, K., Campos, M., Czerwonka-Schröder, D., Höfle, B., Kuschnerus, M., Puttonen, E., Prinz, R., Rutzinger, M., Voordendag, A., Vos, S., 2025. Permanent terrestrial laser scanning for near-continuous environmental observations: Systems, methods, challenges and applications. *ISPRS Open Journal of Photogrammetry and Remote Sensing*.

Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T. Y., Tegmark, M., 2024. Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*.

Liu, Z., Zhang, F., Jiao, J., Lao, N., Mai, G., 2025. Gair: Improving multimodal geo-foundation model with geo-aligned implicit representations. *arXiv preprint arXiv:2503.16683*.

Lutio, R. d., D’aronco, S., Wegner, J. D., Schindler, K., 2019. Guided super-resolution as pixel-to-pixel transformation. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 8829–8837.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., Ng, R., 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1), 99–106.

Molaei, A., Aminimehr, A., Tavakoli, A., Kazerouni, A., Azad, B., Azad, R., Merhof, D., 2023. Implicit neural representation in medical imaging: A comparative survey. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2381–2391.

Naylor, P., Di Carlo, D., Traviglia, A., Yamada, M., Fiorucci, M., 2024. Implicit neural representation for change detection. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 935–945.

Naylor, P., Stokholm, A., Andersen, N. H., Dionelis, N., Paletta, Q., Simonsen, S. B., 2025. Implicit Neural Representation for Ice Sheet Surface Elevation Reconstruction to Assess Elevation Change in High-Spatiotemporal Resolution. *Available at SSRN 5447938*.

Pistilli, F., Valsesia, D., Fracastoro, G., Magli, E., 2022. Signal compression via neural implicit representations. *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 3733–3737.

Qin, R., Tian, J., Reinartz, P., 2016. 3D change detection – Approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 41-56.

Rahimi, A., Recht, B., 2007. Random features for large-scale kernel machines. *Adv. Neural Inf. Proc. Syst.*, 20.

Raissi, M., Perdikaris, P., Karniadakis, G. E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707.

Rußwurm, M., Klemmer, K., Rolf, E., Zbinden, R., Tuia, D., 2023. Geographic location encoding with spherical harmonics and sinusoidal representation networks. *arXiv preprint arXiv:2310.06743*.

Saragadam, V., LeJeune, D., Tan, J., Balakrishnan, G., Veeraraghavan, A., Baraniuk, R. G., 2023. Wire: Wavelet implicit neural representations. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Shi, W., Zhang, M., Zhang, R., Chen, S., Zhan, Z., 2020. Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges. *Remote Sensing*, 12(10).

Sitzmann, V., Martel, J. N., Bergman, A. W., Lindell, D. B., Wetzstein, G., 2020. Implicit neural representations with periodic activation functions. *Proc. NeurIPS*.

Strümpler, Y., Postels, J., Yang, R., Gool, L. V., Tombari, F., 2022. Implicit neural representations for image compression. *European Conference on Computer Vision*, Springer, 74–91.

Stular, B., Žiga Kokalj, Oštir, K., Nuninger, L., 2012. Visualization of lidar-derived relief models for detection of archaeological features. *Journal of Archaeological Science*, 39(11), 3354-3360.

Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J., Ng, R., 2020. Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems*, 33, 7537–7547.

Vos, S., Kuschnerus, M., Lindenbergh, R., de Vries, S., 2023. 4D spatio-temporal laser scan dataset of the beach-dune system in Noordwijk, NL (dataset).

Winiwarter, L., Anders, K., Czerwonka-Schröder, D., Höfle, B., 2023. Full four-dimensional change analysis of topographic point cloud time series using Kalman filtering. *Earth Surf. Dynam.*, 11(4), 593-613. ESurf.

Xie, Y., Takikawa, T., Saito, S., Litany, O., Yan, S., Khan, N., Tombari, F., Tompkin, J., Sitzmann, V., Sridhar, S., 2022. Neural fields in visual computing and beyond. *Computer Graphics Forum*, 41, Wiley Online Library, 641–676.

Yu, Z., Idris, M. Y. I., Wang, H., Wang, P., Chen, J., Wang, K., 2025. From Physics to Foundation Models: A Review of AI-Driven Quantitative Remote Sensing Inversion. *arXiv preprint arXiv:2507.09081*.

Zhang, Z., Vosselman, G., Gerke, M., Persello, C., Tuia, D., Yang, M. Y., 2019. Detecting Building Changes between Airborne Laser Scanning and Photogrammetric Data. *Remote Sensing*, 11(20).