

Technische Universität Ilmenau  
Fakultät für Elektrotechnik und Informationstechnik



# Master Thesis

Data Driven Hybrid Algorithms for Preprocessing of  
Manually Acquired Ultrasound NDT Data

<b>Submitted by</b>	Sayako Kodera
<b>Submission Date</b>	20.07.2021
<b>Faculty</b>	Electrical Engineering and Information Technology
<b>Major</b>	Communications and Signal Processing
<b>Enrolment Number</b>	53168
<b>Supervisor</b>	Dr.-Ing Florian Römer, Fraunhofer IZFP
<b>Responsible Professor</b>	Univ.-Prof. Dr.-Ing. Giovanni Del Galdo
<b>Department</b>	Electronic Measurements and Signal Processing Group EMS

# Abstract

Ultrasonic testing (UT) is a nondestructive inspection method to localize the flaws in test objects. Although automation in UT is progressing, there are still great needs for manual operations such as inspection of objects with complex geometries. To ease the operation and improve its reliability, an assistance system can be integrated into manual UT. Such a system records the measurements along with their scan positions and provides visual feedback which can be improved by applying postprocessing. Yet, conventional postprocessing methods, such as synthetic aperture focusing technique (SAFT), may not lead to a satisfactory resolution when applying to manual UT data. One possible reason is spatial undersampling, as the distribution of manual sampling tends to be sparse and irregular compared to densely scanned automatic UT. As a result, spatial aliasing in SAFT reconstructions occurs, leading to the appearance of artefacts.

In an attempt to prevent spatial aliasing in SAFT reconstructions, the main goal of this work is to predict and interpolate the missing UT data of the unobserved positions as preprocessing of the measurement system. To develop a computationally efficient method, we propose a data driven hybrid approach for interpolating spatio-temporal lattice data, where the frequency responses of multiple time series data are predicted via a minimum mean square estimator. This requires the estimate of the second order statistical moments in space-frequency domain, which are computed via deep neural network. This method can compute the variance of the predictions, which can be utilized to suggest sampling positions strategically. In theory, the proposed method is generic and not confined to UT data, which is one of its main advantages.

To evaluate the validity of the proposed method, simulations are conducted, where automatic UT data are irregularly subsampled. Overall, the proposed interpolation method yields better prediction results than an existing generic method. Moreover, the proposed preprocessing scheme is shown to contribute to an enhanced reconstruction resolution, if there are adequate datasets available, which is in our case 0.61 samples per squared wavelength. Furthermore, sampling in high variance area appears to improve the reconstruction resolution than random sampling, indicating the usefulness of suggesting the sampling positions.

---

# Zusammenfassung

Ultraschallprüfung (UP) ist eine Zerstörungsfreie Prüfmethode, die zur Lokalisation von Fehlern in Bauteilen genutzt wird. Obwohl in einigen Fällen automatisiert werden kann, ist händische Arbeit oft eine Notwendigkeit. Um die Zuverlässigkeit der Hand-aufgenommenen UP (HaUP) zu erhöhen, können Assistenzsysteme integriert werden. Ein solches System speichert die Messdaten und die Messposition und nutzt sie, um ein visuelles Feedback zu erzeugen, welches durch Post-processing Methoden verbessert werden kann. Konventionelle Methoden zur Nachbearbeitung, wie zum Beispiel die synthetic aperture focusing technique (SAFT), verbessern häufig die HaUP Daten nicht. Ein möglicher Grund ist, dass die nötige räumliche Abtastung durch das HaUP Verfahren nicht erreicht werden kann. Bei Anwendung des SAFT Algorithmus auf solche HaUP Daten kann es deshalb, bei der Rekonstruktion, zu räumlichen Aliasing kommen.

Mit dem Ziel räumliches Aliasing zu verhindern, werden in dieser Arbeit, als Pre-processing Schritt, UP Daten von fehlenden Positionen prognostiziert und interpoliert. Wir schlagen eine an Daten orientierte hybride Vorgehensweise vor, um die räumlich-zeitlich diskreten Daten zu interpolieren. Dabei werden die Frequenzantworten mehrerer Zeitreihen mit einer minimalen mittleren quadratischen Fehlereinschätzung prognostiziert. Dafür wird, im Raum-Frequenz Bereich, das statistische Moment zweiter Ordnung eingeschätzt, welches durch ein deep neural network berechnet wird. Diese Methode kann die Varianz der Vorhersagen berechnen, welche genutzt werden können, um wichtige Messpositionen vorzuschlagen. Theoretisch ist die vorgestellte Methode robust genug, um nicht nur auf UP Daten, sondern auf eine Vielzahl von Datensätzen angewendet werden zu können.

Um die Validität der vorgeschlagenen Methode zu überprüfen wurden Simulationen durchgeführt, bei denen die automatisierten UP Daten unregelmäßig abgetastet waren. Insgesamt berechnet die vorgestellte Interpolation eine bessere Vorhersage, als eine andere gängige Methode. Zusätzlich erzeugt der vorgestellte Pre-processing Algorithmus, bei geeigneten Datensätzen ( $0,61 \text{ Sample/Wellenlänge}^2$ ) eine verbesserte Rekonstruktions-Auflösung. Weiterhin hat sich gezeigt, dass das Messen, in einem Gebieten mit hoher Varianz, zu einer höheren Rekonstruktions-Auflösung führt, was auf die Nützlichkeit des Messpositionsvorschlags hinweist.

---

# Contents

**Abstract**

**Zusammenfassung**

**Contents**

**List of Figures**

**List of Tables**

<b>1. Introduction</b>	<b>1</b>
1.1. Background and Motivation . . . . .	1
1.2. Objective and Contributions of the Thesis . . . . .	3
<b>I. Data Model and SoA</b>	<b>4</b>
<b>2. UT Data Model</b>	<b>5</b>
2.1. Measurement System: 3D SmartInspect . . . . .	5
2.2. Data Model . . . . .	6
2.3. Problem Formulation . . . . .	8
<b>3. SoA: DL</b>	<b>10</b>
3.1. Conventional Neural Networks . . . . .	11
3.2. Data Driven Interpolation Methods . . . . .	11
<b>4. SoA: ST Interpolation</b>	<b>13</b>
4.1. IDW . . . . .	14
4.2. SRF Model . . . . .	14
4.3. Geostatistical Inference Technique: Kriging . . . . .	17
4.4. Variogram Estimation . . . . .	20
4.5. STRF Model . . . . .	21
4.6. Prediction in Space-Time Domain . . . . .	23
4.7. Spectral Representation of STRF . . . . .	24
4.8. Prediction in Space-Frequency Domain: Space-Frequency Kriging . . . . .	25
4.9. Non-Parametric Estimation of Frequency Variogram . . . . .	27
4.10. Parametric Estimation of Frequency Variogram . . . . .	28
4.11. Extension to Multiple Locations . . . . .	29

---

<b>5. Summary of Part I</b>	<b>31</b>
<b>II. Data Driven Space-Frequency Domain Interpolation</b>	<b>32</b>
<b>6. Space-Frequency Domain Block Interpolation of UT Data</b>	<b>34</b>
6.1. Complex Spatial Random Field Modeling of UT Data . . . . .	34
6.2. Space-Frequency Block Interpolation . . . . .	35
<b>7. Data Driven FV Estimation</b>	<b>38</b>
7.1. Network Inputs and Output . . . . .	38
7.2. Network Architecture . . . . .	39
7.3. Training Scheme . . . . .	41
<b>8. Summary of Part II</b>	<b>43</b>
<b>III. Results</b>	<b>44</b>
<b>9. Simulation Setting</b>	<b>46</b>
9.1. Specimen and Its Raw Measurement Data . . . . .	46
9.2. Interpolation Parameters . . . . .	46
<b>10. Simulation I: Batch Interpolation</b>	<b>48</b>
10.1. Simulation Parameters and Scenario . . . . .	48
10.2. Interpolation Results and their Evaluation . . . . .	49
<b>11. Simulation II: Reconstruction of Subsampled UT Data</b>	<b>54</b>
11.1. Simulation Parameters and Scenario . . . . .	54
11.2. Evaluation Metrics . . . . .	55
11.3. Interpolation and Reconstruction Results . . . . .	57
11.4. Evaluation of the Results . . . . .	63
<b>12. Summary of Part III</b>	<b>65</b>
<b>Conclusion</b>	<b>67</b>
<b>Bibliography</b>	<b>73</b>
<b>Notation and Symbols</b>	<b>74</b>
<b>Glossary</b>	<b>76</b>
<b>Acronyms</b>	<b>77</b>

---

# List of Figures

2.1. Manual measurement setup . . . . .	6
2.2. 3D SmartInspect . . . . .	6
2.3. System workflow for the research problem of this work . . . . .	9
4.1. Covariance vs Semivariogram . . . . .	17
4.2. Variogram types . . . . .	22
4.3. Destructive behavior of samples farther away . . . . .	30
5.1. Entire architecture of the proposed scheme . . . . .	33
7.1. Frequency variogram of neighboring frequencies . . . . .	40
7.2. Network architecture of FVnet . . . . .	41
9.1. Illustration of the geometry of MUSE specimen . . . . .	46
10.1. Simulation I: sampling and prediction positions . . . . .	48
10.2. Simulation I: prediction accuracy of A-Scans . . . . .	50
10.3. Simulation I: batch-wise interpolation results (C-Scan) . . . . .	51
10.4. Simulation I: SF-Kriging prediction error vs estimated variance . . . . .	52
10.5. Simulation I: comparison of frequency variogram estimations . . . . .	53
11.1. Simulation II: Illustration of the ROI . . . . .	55
11.2. Simulation II: C-Scan of the actual data in the ROI . . . . .	56
11.3. Simulation II: sampling and resampling positions with 15% coverage . . . . .	58
11.4. Simulation II: reconstruction results for 15 % coverage . . . . .	58
11.5. Simulation II: sampling and resampling positions with 12% coverage . . . . .	59
11.6. Simulation II: reconstruction results for 12 % coverage . . . . .	59
11.7. Simulation II: sampling and resampling positions with 10% coverage . . . . .	60
11.8. Simulation II: reconstruction results for 10 % coverage . . . . .	60
11.9. Simulation II: sampling and resampling positions with 7% coverage . . . . .	61
11.10Simulation II: reconstruction results for 7 % coverage . . . . .	61
11.11Simulation II: sampling and resampling positions with 5% coverage . . . . .	62
11.12Simulation II: reconstruction results for 5% coverage . . . . .	62
11.13Simulation II: results evaluation . . . . .	64

# List of Tables

7.1. Parameter values for forward modeling, fixed for all datasets . . . . .	41
7.2. Parameter values for forward modeling, varied with each dataset . . . . .	42
7.3. Parameter values related to training of FVnet . . . . .	42
9.1. Parameter values for MUSE dataset . . . . .	47
9.2. Parameter values for interpolation . . . . .	47

# 1. Introduction

## 1.1. Background and Motivation

Nondestructive Testing (NDT) is a method to inspect the structure of test objects without inducing damage in order to identify whether a test object contains any flaws, or to detect their locations. One of the prevalent NDT methods is Ultrasonic Testing (UT). Conventionally, an UT inspection requires either manual operation by a human operator or automated measurement systems. In manual UT, where a human operator observes the change in the echoed signals, its inspection quality is highly dependent on the expertise of the operator [1]. In automatic UT, on the other hand, measurement data and the corresponding scan positions are recorded, which enables to visualize the inner structure of the test object and further process the obtained data to improve the imaging quality, leading to more reliable inspection quality than its manual counterpart.

Nevertheless, there are still needs for manual UT, for instance when a complex structure is inspected, and its inspection reliability has been of great concern. In order to improve the inspection reliability of manual UT, an assistance system can be employed, which records measurement data while logging the scan positions through an optical tracking system. One of such systems is 3D SmartInspect, which is being developed by the Fraunhofer Institute for Nondestructive Testing IZFP. This system allows us not only to visualize the measurement data but also to process it further to increase its imaging quality [2, 3].

Although their application to manual UT data has been typically excluded, several post-processing techniques have been developed and extensively employed for automatic UT data [4–6]. One of the commonly employed post-processing methods is the Synthetic Aperture Focusing Technique (SAFT) [5, 6]. The aim of SAFT is to improve the spatial resolution through performing a superposition of the delayed versions of the observed signals [7]. In other words, SAFT regards the measurement Region of Interest (ROI) as a synthetic aperture and each measurement as its spatial sampling, indicating that regular and fine spatial sampling is required for successful SAFT reconstruction.

On the contrary, the application of such techniques to manual UT data has not been investigated extensively. Previously, we revealed that several stochastic observational errors, such as irregular and sparse spatial sampling, varying contact pressure or inaccurate positional information, may result in strong artefacts formation in SAFT reconstructions [2]. Since such errors are inevitable in manual measurements, incorporating the strategy to reduce those artefacts could improve the reconstruction quality.

One approach is preventing spatial aliasing in SAFT reconstructions by decreasing the sampling interval, such as through interpolating missing measurement data. It has



been shown that the artefacts in reconstructed images are significantly reduced when the signal responses at the unobserved positions are predicted as a part of preprocessing and fed into postprocessing together with the available measurement data [8].

Prediction of UT data is, however, not necessarily straightforward. First, we should predict several values in spatial and temporal domains. The difficulty here is that the missing signal responses behave nonlinearly in both domains, which requires the understanding of the behavior within each domain and the interaction between these two. Second, the size of the available datasets may be very small due to the nature of manual inspection. With such inadequate datasets, general interpolation schemes are likely to fail in properly modeling the missing data. Furthermore, such preprocessing scheme is ultimately to be incorporated into the whole measurement system of 3D SmartInspect, requiring an on-site execution.

Spatio-temporal interpolation has, in fact, gained growing interests from many different fields and applications, such as climate science, epidemiology, remote sensing as well as neuroimaging [9–12]. As a generalization of purely spatial data, analysis and processing techniques of spatio-temporal data have been mostly studied in the geoscientific community. Geoscience seeks to understand the underlying natural phenomena through careful interpretation of spatial or spatio-temporal measurements. Those measurements are in most cases taken with a much lower sampling rate than desired, necessitating statistical approaches to accommodate the uncertainties. As one of the well-established fields in statistics, geostatistical approaches are utilized in many applications, such as prediction of gold mining or weather forecast, which have long been proven to be very effective.

Nevertheless, handling spatio-temporal data has remained a very challenging task [13–16]. Considering that spatio-temporal data not only exhibit the dependencies in space and time but also there is an interaction between these two domains, it is of no doubt a nontrivial task to properly capture the space-time variability in data. Moreover, as the temporal dimension is typically larger than the spatial one, inclusion of temporal domain can easily lead to a large dimensional computation. Without any measures to reduce the computation, conventional geostatistical techniques are thus not appropriate for applications requiring a fast computation, such as on-site spatio-temporal interpolation.

To overcome the aforementioned issues regarding nonlinear spatio-temporal interpolation, the author of [8] developed an autoencoder Deep Neural Networks (DNN) which is trained to extract the data structure and model the missing data based on this knowledge. Since a DNN is capable of computing complex mapping and fast in execution, this deep learning (DL) approach is suitable for our manual UT scenario. Yet, this DL approach is rather of black-box nature and lacks the interpretability. The interpretability of a network is well known to play a crucial role in identifying its limitation and further advancing the system, making it an important concern when employing DL in many application areas [17–19].

## 1.2. Objective and Contributions of the Thesis

In an attempt to improve the reconstruction resolution of manual UT data, in this thesis an interpretable DL scheme is proposed for interpolating missing UT data, which is in essence nonlinear spatio-temporal interpolation. In particular, the primary goal of this thesis is to formulate our research problem in a way that we can understand the role DNN plays during interpolation.

In pursuit of this goal, our work is inspired by geostatistics since the aforementioned nature of data collection in manual UT inspection, such as sparse and irregular sampling, is very similar to that of geoscience. Prioritizing the computational feasibility for an online application, we decided to process our UT data in space-frequency domain, where the dimensionality can be significantly reduced due to orthogonal Fourier bases [16]. After having addressed the limitations of previous works [16, 20, 21] and examined the property of our data, we found a possibility to overcome those limitations by utilizing DNN. Based on these findings, we decompose the interpolation task into two subtasks, one of which is solved by theory-based geostatistical approach and the other by DNN. In this way, we can perform predictions with a reliable theory-based technique, but at the same time can make use of DNN, which are capable of complex mapping and fast in execution.

To this end, the contributions of this thesis are summarized as follows. First, an extensive overview of geostatistics is provided to show its application versatility and limitations for interpolating spatio-temporal data. Second, we show how UT data can be modeled in geostatistical manner in space-frequency domain. By doing so, we verify that the assumptions often made in geostatistics hold also true for UT data. As the third contribution, we extend the space-frequency domain spatio-temporal interpolation method proposed in [16] to the multi-points interpolation, demonstrating its validity for UT data. This is, to our knowledge, its first practical application. The last contribution of this thesis is the development of a DNN to estimate spatio-temporal statistics in space-frequency domain. This network, called FVnet in this work, enables online interpolation of spatio-temporal lattice data and, in theory, is not confined to data types or applications. Furthermore, the proposed preprocessing scheme is equipped to provide the estimates of the prediction error, which can be further utilized as feedback to the inspection operator to improve the measurement quality.

The rest of this work is divided into three parts. In Part I, mathematical modeling of UT data and its SAFT reconstruction is introduced, followed by an overview of conventional DL schemes and spatio-temporal interpolation techniques. Based on these concepts, the aforementioned research problem is mathematically formulated, and the proposed method is explained in Part II. To evaluate the validity of the proposed scheme, simulations are conducted, whose results are presented in Part III. The important findings from this thesis are then summarized in Conclusion.

## Part I.

# Manual UT and Conventional Techniques for Spatiotemporal Interpolation and Deep Learning

## 2. Manual UT Measurements and Assistance System

### 2.1. Measurement System: 3D SmartInspect

Object inspection through UT requires to stimulate the test object by inserting a short duration ultrasonic pulse. This inserted pulse reflects at any place where the acoustic impedance changes, such as a back wall, flaws or defects in the object. By analyzing the reflected signals, the location of those flaws can be identified.

Typically, manual UT inspections are performed with a contact transducer which is directly placed on the object surface. With a contact transducer, longitudinal ultrasonic waves are transmitted, and the echoed signals are received at the same position. The inspection operator can move the transducer freely on the object surface and takes measurements at each scan position (Figure 2.1a). In this simple setup, the received signals are displayed only in terms of the change in amplitude over time as it lacks the spatial information of the data. Such way of data presentation is called A-Scan (Figure 2.1b) and requires expertise of the operator to properly evaluate the obtained results [1].

However, when an assistance system, such as 3D SmartInspect, is deployed, the measurement positions can be traced and recorded, making the spatial information of the data available. With 3D SmartInspect, a tracking camera recognizes the position of the measurement probe, and the received ultrasonic signals are stored in the system together with their scan positions (Figure 2.2a). This enables the measurement results to be displayed in terms of space as well (Figure 2.2b) and even to be processed further to enhance the imaging quality [2]. This way of visualization serves as a direct feedback to the operator, which can help to improve scanning, assist the documentation and/or can be used for training of the operator.

One possible system workflow is divided into three steps: (1) data acquisition, (2) data pre-processing, (3) data post-processing. In the first step, raw measurement data are recorded and assigned to the nearest measurement grid points of the system. This raw dataset is then fed into the pre-processing step, where, for instance, the data can be denoised or the signals at the missing grid points are predicted from the available dataset. In the last step, the pre-processed data are reconstructed using any reconstruction method, such as SAFT.

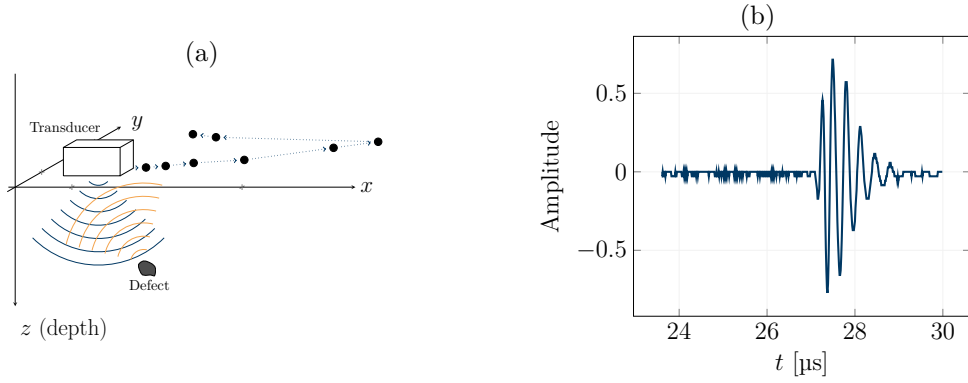


Figure 2.1.: Manual measurement setup: (2.1a) Simplified illustration of manual measurement setup using a contact transducer. The transducer is moved freely on the object surface and takes measurements at each scan position. The inserted ultrasonic pulse (blue wave) is reflected back at the defect to the scan position (orange wave). (2.1b) An example raw measurement result (A-Scan).

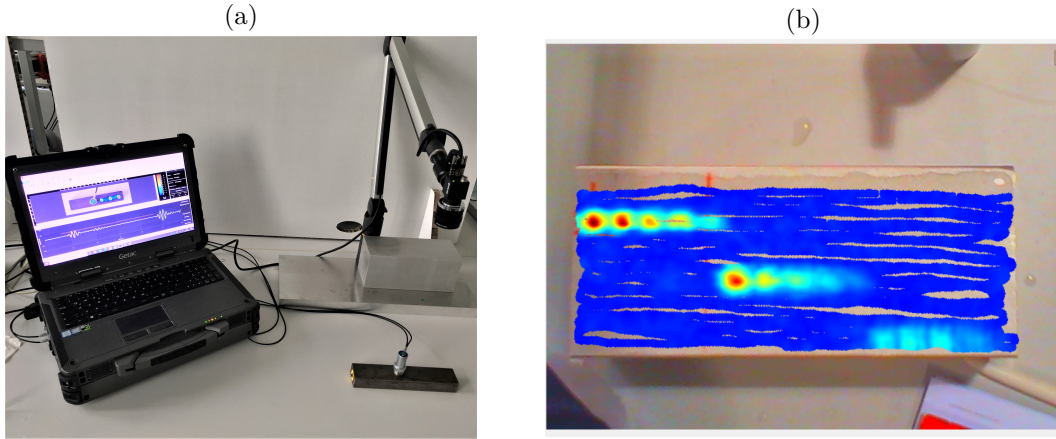


Figure 2.2.: Measurement system: 3D SmartInspect (2.2a) primary components are a measurement computer, an optical position tracker and a contact transducer (adapted from Figure 4.2 in [22]), (2.2b) C-Scan presentation of an example dataset. In this presentation, the maximum signal amplitude at each measurement position is displayed as a heat map.

## 2.2. Data Model

For proper modeling of the measurement data, so that it can be appropriately processed in both pre- and post-processing phases, the following assumptions are made for the measurement setup. The object under test is assumed to be homogeneous and isotropic

with the constant speed of sound  $c_0$  and to have a flat surface. The contact pressure between the transducer and the test object is considered to be constant, such that in the measurement data there is no temporal shift or amplitude change caused by improper coupling. We suppose our ROI is a part of the test object, recognized in the measurement system as the region with the size of  $N_x \times N_y \times M$ . Here,  $N_x$ ,  $N_y$  and  $M$  are number of grid points within the ROI along  $x$ ,  $y$  and  $z$  axis, respectively. The spacing between two grid points along each axis is kept constant and denoted as  $dx$ ,  $dy$  and  $dz$ . The ROI is assumed to contain  $J$  point scatterers inside, where  $J \in [1, MN_xN_y] \subset \mathbb{N}$ .

Suppose we take a measurement at a position which is recognized by the system as the  $i$ -th grid point,  $\mathbf{s}_i = [x_i, y_i, 0]^T$ . At  $\mathbf{s}_i$ , an ultrasonic pulse  $h(t)$  is inserted to stimulate the object, and the echoed signal,  $a(\mathbf{s}_i, t)$ , is received as a measurement data. The inserted pulse can be modeled as a real-valued Gabor function [23], as

$$h(t) = e^{-\alpha t^2} \cdot \cos(2\pi f_C t + \phi), \quad (2.1)$$

where  $f_C$ ,  $\alpha$  and  $\phi$  are the carrier frequency, the window width factor and the phase, respectively. In this way of pulse modeling, its envelope is assumed to be symmetric, which yields a satisfactory approximation of UT pulse with less model parameters compared to the one with an asymmetric envelope. The time which the sound travels from  $\mathbf{s}_i$  to a scatterer  $\mathbf{r}_j$  and reflects back to  $\mathbf{s}_i$  is so called Time-of-Flight (ToF). ToF can be obtained through

$$\tau(\mathbf{s}_i, \mathbf{r}_j) = \frac{2}{c_0} \cdot \|\mathbf{s}_i - \mathbf{r}_j\|_2, \quad (2.2)$$

with  $\|\mathbf{s}_i - \mathbf{r}_j\|_2$  being the Euclidean distance between  $\mathbf{s}_i$  and  $\mathbf{r}_j$ , suggesting that the ToF is determined by the position of both the measurement and the scatterer.

The obtained A-Scan is the sum of all  $J$  reflected echoes, which are delayed version of the inserted pulse  $h(t)$

$$a(\mathbf{s}_i, t) = \sum_{j=1}^J b_{i,j} \cdot h(t - \tau(\mathbf{s}_i, \mathbf{r}_j)) + n(\mathbf{s}_i, t), \quad (2.3)$$

where  $b_{i,j}$  is the reflection coefficient for the position  $\mathbf{s}_i$  and a scatterer  $\mathbf{r}_j$ , and  $n(t)$  is the measurement noise considered as Additive White Gaussian Noise (AWGN). Since we process the data digitally with the sampling interval of  $dt = \frac{1}{f_s}$ , (2.3) can be formulated as a vector  $\mathbf{a}_i \in \mathbb{R}^M$  with  $M$  being the number of temporal samples, which is the same dimension as the ROI along  $z$ -axis.

Although the scatterers can be in theory located anywhere in the ROI, their positions can be assumed to be on the measurement grids, as the manifold changes smoothly in all directions. This enables us to form the measurement dictionary  $\mathbf{H}_i$  by collecting the impulse response at  $\mathbf{s}_i$  for all possible scatterer positions  $\mathbf{r}_j \forall j = 1 \dots ML$  with  $L = N_x N_y$

$$\mathbf{H}_i = [\mathbf{h}_i^{(1)} \quad \mathbf{h}_i^{(2)} \quad \dots \quad \mathbf{h}_i^{(ML)}] \in \mathbb{R}^{M \times ML}. \quad (2.4)$$

Here,  $\mathbf{h}_i^{(j)} \in \mathbb{R}^M$  is the  $j$ -th column vector of  $\mathbf{H}_i$ , corresponding to the  $j$ -th scatterer position in the ROI. Its  $m$ -th element  $\mathbf{h}_i^{(j)}[m]$  can be obtained as

$$\mathbf{h}_i^{(j)}[m] = b_{i,j} \cdot h(m \, \text{dt} - \tau(\mathbf{s}_i, \mathbf{r}_j)). \quad (2.5)$$

This enables us to model the A-Scan for a single measurement position,  $\mathbf{a}_i$ , as a forward model

$$\mathbf{a}_i = \mathbf{H}_i \mathbf{b} + \mathbf{n}_i, \quad (2.6)$$

where  $\mathbf{n}_i \in \mathbb{R}^M$  is a noise vector, and  $\mathbf{b} \in \mathbb{R}^{ML}$  is the vectorized *defect map* which represents the scatterer positions and their amplitudes [24].

After taking an A-Scan at every grid point on the surface of the ROI and assigning each one to the corresponding scan position, a data tensor  $\mathcal{A} \in \mathbb{R}^{M \times N_x \times N_y}$  can be built. In order to formulate this data tensor as a forward model analogous to (2.6),  $\mathcal{A}$  is to be vectorized as

$$\boldsymbol{\alpha} = \text{vec}\{\mathcal{A}\} \in \mathbb{R}^{ML}, \quad (2.7)$$

where  $\text{vec}\{\cdot\}$  denotes the vectorization operation. In this work vectorization of a multidimensional array is applied to 1-mode unfolding of the array [25]. Having the data tensor vectorized, the forward modeling of the dataset is represented as

$$\boldsymbol{\alpha} = \mathbf{H} \mathbf{b} + \mathbf{n}. \quad (2.8)$$

Here  $\mathbf{n} \in \mathbb{R}^{ML}$  denotes all noise vectors, and  $\mathbf{H} \in \mathbb{R}^{ML \times ML}$  is the concatenation of each dictionary matrix in the ROI as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_L \end{bmatrix}. \quad (2.9)$$

The ultimate goal of UT inspections is to locate the scatterer positions, which is to recover  $\mathbf{b}$  from  $\boldsymbol{\alpha}$ . Since SAFT is based on the idea of correlator, (2.8) yields a SAFT reconstruction as [24]

$$\hat{\mathbf{b}} = \mathbf{H}^T \boldsymbol{\alpha} \in \mathbb{R}^{ML}. \quad (2.10)$$

### 2.3. Problem Formulation

Although the proposed method is not confined to UT data but general for any spatio-temporal data, in this work the following scenario is considered (Figure ??).

For a manual UT inspection, 3D SmartInspect system is thought to be employed. We made here the same assumptions with regard to the test object and the ROI as in Section 2.2. During the course of inspection, total of  $N_S$  distinctive A-Scans are to be registered in the system for the set of scan positions  $S_n = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{N_S}\}$ . The obtained A-Scans are collected in a single measurement matrix  $\mathbf{A}_{N_S} \in \mathbb{R}^{M \times N_S}$  as

$$\mathbf{A}_{N_S} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_{N_S} \end{bmatrix}. \quad (2.11)$$

This dataset is then fed into pre-processing, where the missing  $(L - N_S)$  A-Scans are predicted based on  $\mathbf{A}_{N_S}$  and  $S$ . The predicted dataset,  $\mathbf{A}_{N_S}^C \in \mathbb{R}^{M \times (L - N_S)}$ , is then combined with the available measurement data  $\mathbf{A}_{N_S}$  to form an interpolated data tensor  $\hat{\mathbf{A}}_{\text{itp}} \in \mathbb{R}^{M \times N_x \times N_y}$ . This interpolated data tensor is fed into post-processing, which computes SAFT reconstruction  $\hat{\mathbf{b}}_{\text{itp}} \in \mathbb{R}^{ML}$  as described in (2.10).

Seeking to reduce the artefacts in the SAFT reconstruction  $\hat{\mathbf{b}}_{\text{itp}}$ , this thesis aims to develop an algorithm to predict the missing A-Scans  $\mathbf{A}_{N_S}^C$  in the pre-processing step. Since the entire steps of the measurement system are ultimately performed in real-time, there are few more desirable attributes to be considered. First, the algorithm should be online capable, which includes parallel computing and processing the samples batch-wise instead of over the entire ROI. Second, it is desirable to be able to estimate the prediction accuracy of each predicted position, so that the system can provide feedback to encourage the inspection operator to increase sampling in the uncertain regions.

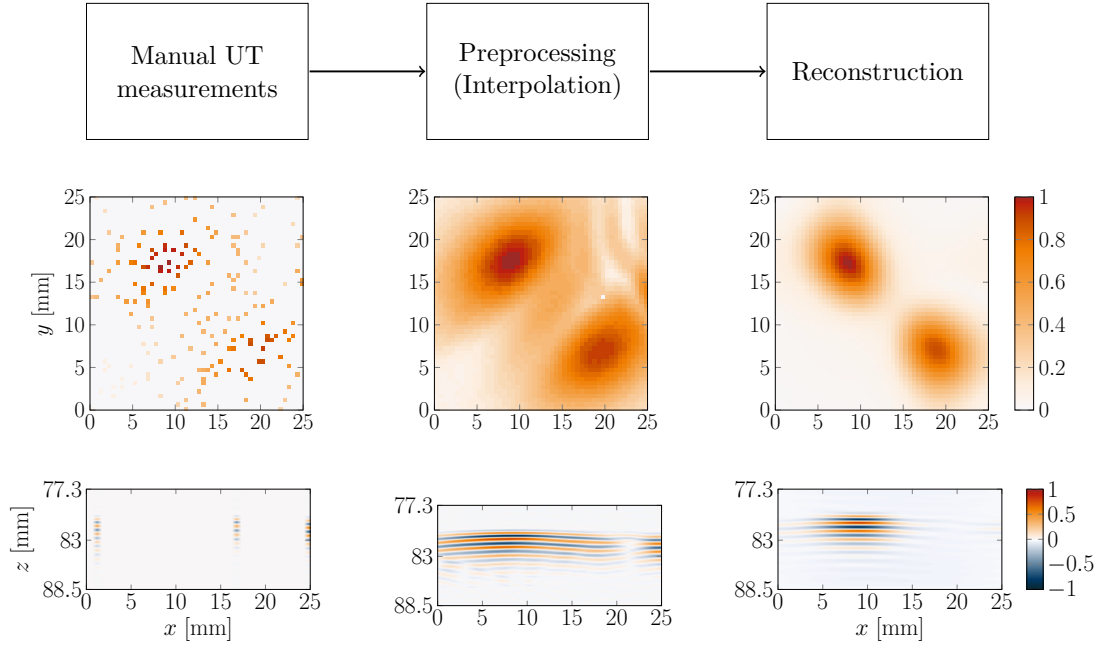


Figure 2.3.: Illustration of 3D SmartInspect workflow for the research problem of this work. The left figures are an example set of irregularly subsampled measurements, whereas the figures in middle are fully sampled measurements, representing the results after interpolation. The right figures are reconstruction of the fully sampled data. The figures in top are top view images (C-Scan), while those in bottom are from side with  $y = 15$  mm.



### 3. Deep Learning and Neural Networks

Recent years we have witnessed tremendous success of artificial learning, such as Deep Neural Networks (DNN), in many fields. In the context of scientific applications, there are not a few scenarios where the use of DNN may provide significant benefits which traditional theory-based approaches cannot offer. For instance, underlying processes may not be fully describable or require complicated operations which makes the computation infeasible. In such cases, their powerful capability of nonlinear modeling may be appealing, and it has been shown to even outperform theory-based methods [9, 26–28]. Moreover, once a network is trained, it can perform the given tasks much faster than traditional numerical methods, as it shifts the expensive computation from on-site iterative processing to prior training. This allows DNN to be deployed for online applications, which attracts further attentions from many scientific fields.

Yet, there are concerns to depend solely on DNN to solve complex tasks in scientific domain. One of the prominent problems is its black-box nature. If a conventional DNN is employed in end-to-end manner, it is hard to understand what the network actually learns. This is due to high dimensionality of the network parameters, which prevents us to identify the roles of individual parameters [19]. As a result, it is not possible to specifically train the parameters of such black-box DNN, requiring large amount of data to compensate [17, 18]. This strong dependency on the quantity of training data is another issue, especially for the applications where abundant amount of training samples are not available, such as medical imaging and 3D reconstruction [19]. If we fail to provide sufficient training data, conventional DNN are prone to overfit and thus cannot be generalized [18], which compounds a problem.

To address these limitations, there are emerging attempts to combine DNN and traditional theory-based approaches [17–19]. One such approach is called hybrid DNN, which seek to find physically aware models by integrating domain knowledge. Hybrid DNN can be established by decomposing the problem into several aspects, some of which are modeled and/or solved via theory-based approaches and the other by the network [17]. Consequently, a network can focus on physically consistent learning space, resulting in generalizable and scientifically interpretable models [17]. In this way, we can benefit from both the domain knowledge rich theory-based approach and complex mapping capable DNN.

In this chapter, our goal is to illustrate the possibilities and the limitations of DNN to perform spatio-temporal interpolation. For this purpose, a short description on conventional DNN is first provided, before we discuss their state of the art applications.

### 3.1. Conventional Neural Networks

For any art of artificial learning, a network of neurons, i.e. Neural Network (NN), makes up the backbone. These neurons are designed to mimic the human nerve systems through a set of algorithms. At a basic level, a neuron processes its inputs by individually weighting them and adding some biases. The weighted sum of inputs are then evaluated using activation functions and generates its outputs based on this evaluation. The generated outputs are then passed to the next neurons. Despite the simplicity of each operation, such network of neurons can infer nonlinear functions, especially well for classifications and nonlinear regressions [29], if their weights and biases are properly selected. Hence, learning process of NN means to tuning these parameters based on the given training datasets.

Such neurons can be stacked one another to build multiple layers of networks, which is called Multilayer Perceptrons (MLP). Training of MLP is accomplished through back propagation, which incorporates the feedback from the successive layers [30]. With back propagation, the weights of each layer are adapted to minimize the loss function of their next layer via gradient descent. The problem with MLP is, however, its fully-connected nature. Since each neuron is connected to every neuron in the previous and succeeding layers, their trainable parameters can increase rapidly, which makes its training intractable. As a trade-off, the number of layers is to be reduced.

To overcome these limitations, Convolutional Neural Networks (CNN) are introduced to reduce the connectivity of neurons. Based on the assumption that there are some hierarchical patterns in data, connections of neurons in CNN are restricted to local neighbors, and the same weights are shared across different locations [19]. This contributes to significant reduction in parameter dimensionality, enabling to train deeper networks. CNN have been proven to be very effective in many application fields, especially in imaging and computer vision.

For applications with sequential data, Recurrent Neural Networks (RNN) are often preferred over other network architectures. RNN contain global parameters which are shared across all the time steps, allowing it to include the temporal dependencies [19]. In this way, RNN are applicable to tasks such as speech recognition. However, having global parameter for all time step may not be useful, if the data dynamically change over time. For this reason, a more flexible architecture is introduced, called Long Short Term Memory (LSTM) [31]. LSTM allows those global parameters to "vanish" if necessary, enabling it to adjust the temporal dynamics.

### 3.2. Data Driven Interpolation Methods

In this section, a brief overview of state of the art DNN approaches is provided. Since application fields of DNN are vast and growing further due to its popularity in recent years, in this section we restrict ourselves to the topic of spatial and spatio-temporal interpolation.

By virtue of its capability to learn nonlinear functions, DNN have been extensively

applied to interpolation tasks in different fields. Especially, some of recent works have demonstrated remarkable performance of their DNN being able to capture the spatial structures from extremely sparsely subsampled data [32, 33]. Yet, they are rather of *black box* nature and their generality have not been discussed.

Another popular approach is to extend RNN or LSTM to include spatial dimension [34–38]. In general, what is sought are values of a single time step in the future and/or at unknown positions. It is normally assumed that there is no past values available for the prediction positions, thus predictions should be made given the past values of the neighboring regions. In view of our UT scenarios described in Section 2.3, this approach is less appealing, as our data is, in most cases, static. This means that *when* a single A-Scan is collected does not affect other A-Scans. It is *where* this A-Scan is collected that matters. In such cases, assuming temporal dependencies may not be helpful, because some areas might be revisited after a farther area is inspected, where there is no correlation to the current measurements.

The effect of DNN trends are also observable in geostatistics [9, 39–41]. Yet, the majority of them are rather based on black-box approaches, which is due to the complexity in dealing with spatio-temporal datasets [9]. In an attempt to ease the computational burden of spatiotemporal Kriging, which is a geostatistical interpolation scheme discussed later in Section 4.3, the authors of [41] incorporate CNN to predict a set of vectors at unsampled positions. The significance of their approach is its interpretability. In their approach, a small set of unknown vectors are predicted using Kriging, before they are fed into the network along with the available samples. To include the spatial statistics, spatial covariance values between samples are also fed into the network. What the network should learn is to estimate the parameter values of the spatio-temporal covariance function, with which predictions at unsampled positions can be made based on the available data. However, the spatio-temporal model is fixed to a particular model, which may not be applicable to some cases. Indeed, selecting a proper model for spatio-temporal covariance itself is a nontrivial task and should be done with caution [13, 14].

## 4. Spatio-Temporal Data Analysis and Inference

In this chapter analysis and inference schemes for spatio-temporal data are presented. Spatio-temporal datasets result from processes that evolve simultaneously in space and time, which can be observed in nearly all the fields of applied sciences [42]. Indeed, increasing number of such data is collected in diverse fields such as geoscience, social sciences, neuroscience and transportation [11]. As a result, there have been increasing needs in analysis and data mining for spatio-temporal data [11, 43].

In general, the values of such data vary in an inhomogeneous manner over space and/or are nonstationary in time. Thus, the objective of spatio-temporal data analysis is to address the spatio-temporal variability of the underlying process, such as degree of continuity and nonstationary characteristics [42]. This enables us to infer the unknown values of points in space and/or instances in time from the available samples.

However, handling spatio-temporal data is not a trivial task. First, space is different from time. Space is not confined to directions, whereas time only forwards from the past to the present into the future [13]. This obvious but crucial difference implies that temporal dimension cannot be treated identically to its spatial counterpart as 'just another spatial dimension' [16]. Second, as such processes exhibit different behavior in space while evolving through time, it is not sufficient to account for only its dependency on space and time. Any interaction between these two should also be taken into account, which inevitably increases the computational cost [13, 16]. As a consequence, while the analysis and inference schemes of purely spatial data are well-investigated, there are relatively few studies which deal with spatio-temporal processes [10, 16, 42].

Although it is initially derived by the needs to analyze purely spatial data, statistical approaches in geoscience, *geostatistics*, provide fundamental insight as to handling of spatio-temporal data. As opposed to general inference approaches, geostatistics acknowledge inherent uncertainties in the available dataset. For instance, the available dataset may be inadequate in size, or the measurements can be taken only in a certain part of the ROI. This suggests that the dataset is merely a single possibility of the underlying phenomena and by no means its representation [42]. As one of the well-established branches of statistics, analysis and inference techniques of geostatistics are beneficial for many spatial and spatio-temporal data.

Aligned to our research problem described in Section 2.3, our primal goal is to make inference for unsampled positions. For this purpose, the following scenario is considered in this chapter. Within our ROI  $D \subset \mathbb{R}^d$  in  $d$ -dimensional Euclidean space, a set of time series data is taken at  $N$  different positions  $S_N = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N\} \subset D$ . Based on this dataset, another set of unknown time series data is to be predicted for the unsampled

positions in  $\{D \setminus S_N\}$ . For the sake of simplicity, we first focus on predicting a single time series at a position  $\mathbf{s}_0 \in \{D \setminus S_N\}$ . This notion can be easily extended to infer multiple time series at different locations, which is described later in this chapter.

This chapter is structured as follows. First an overview of a general spatial interpolation method is provided, which in the later part serves as a comparison to the proposed method. Then, the fundamental ideas of geostatistics are introduced first in purely spatial domain. The same notion is then extended to spatio-temporal domain, which is generalized form of the purely spatial ones. Here, it is highlighted that handling spatio-temporal data in space-time domain can easily become computationally intractable. To overcome this limitation, spectral domain analysis is explained, and a more computationally efficient handling of such data is presented. Yet, there still remain some difficulties, which are discussed in the later section.

## 4.1. Inverse Distance Weighting (IDW)

Inverse Distance Weighting (IDW) is one of the most widely used interpolation methods for spatial data. It calculates the values at the position of interest  $\mathbf{s}_0$  as a linear combination of the available samples within its neighboring region. The prediction using IDW is based on the assumption that the values of  $\mathbf{s}_0$  is more similar to that of its neighboring samples than those locating further away. Hence, the weight for each sample is set inversely proportional to the distance between  $\mathbf{s}_0$  and the sampled position as [44].

Although its simplicity is very attractive for many applications, IDW may perform poorly if the samples within its neighboring region are inadequate or clustered [44]. This issue can be solved by adjusting the size of neighborhood for each prediction position, nonetheless determining the proper neighbor size requires  $k$ -nearest neighbors search which may result in an expensive computation [45]. Moreover, IDW is incapable of providing its prediction accuracy.

Its application to spatio-temporal data is not as effective as to purely spatial ones either. A prediction can be made based only on the spatial dependency of the available dataset by weighting the samples uniformly over the temporal domain. Yet, in this way the obtained prediction is likely to fail in capturing important structures in the temporal domain. On the other hand, if the weights are to be adapted for each time sample, the temporal domain should be included for distance calculation. This necessitates an appropriate conversion of the time unit with regard to the spatial one, which itself requires the knowledge regarding the spatio-temporal variability of the underlying process [10].

## 4.2. Spatial Random Field (SRF) Model

Conventionally, data modeling in geostatistics is based on the idea of Spatial Random Field (SRF). With SRF modeling, all observed values are considered to be the outcome of random field, or, in other words, random process. The SRF modeling of our process of interest can be described as

$$\{Y(\mathbf{s}) : \mathbf{s} \in D\}, \quad (4.1)$$

where  $Y$  denotes the random quantity of interest observed at a location  $\mathbf{s}$ . Based on this notion, a dataset obtained at the positions  $S_N = \{\mathbf{s}_1, \mathbf{s}_2 \dots \mathbf{s}_N\}$  within our ROI  $D$  is regarded as a set of its *single* realization,  $\{y(\mathbf{s}_i), \forall i = 1 \dots N\}$  [14, 42, 46]. This dataset can be described as a vector form as

$$\mathbf{y}_N = \begin{bmatrix} y(\mathbf{s}_0) \\ y(\mathbf{s}_1) \\ \vdots \\ y(\mathbf{s}_N) \end{bmatrix} \in \mathbb{R}^N. \quad (4.2)$$

In the same manner, the aforementioned inference problem can also be formulated as a SRF problem. Our goal here is to predict the unknown value of the same SRF for an unsampled position  $\mathbf{s}_0$  given  $\mathbf{y}_N$

$$\mathbb{E}\{Y(\mathbf{s}_0) \mid \mathbf{y}_N\}, \quad (4.3)$$

where  $\mathbb{E}\{\cdot\}$  represents the expected value.

In order to obtain an optimal prediction, it is essential to gain proper understanding of the SRF of interest. For this purpose, let us consider its statistical aspects. Without any knowledge regarding the underlying process, the distribution function of the SRF is normally not attainable due to the limited size of the available dataset, which only represents an incomplete sampling of a single realization [42, 46]. Its statistical moments, on the other hand, can be obtained. Typically, the first and the second moments are used for inference as follows:

$$\mathbb{E}\{Y(\mathbf{s}_i)\} := \mu(\mathbf{s}_i) \quad (4.4)$$

$$\begin{aligned} \text{Corr}\{Y(\mathbf{s}_i), Y(\mathbf{s}_j)\} &= \mathbb{E}\{Y(\mathbf{s}_i)Y(\mathbf{s}_j)\} \\ &:= \varphi_Y(\mathbf{s}_i, \mathbf{h}_{ij}) \end{aligned} \quad (4.5)$$

$$\begin{aligned} \text{Cov}\{Y(\mathbf{s}_i), Y(\mathbf{s}_j)\} &= \mathbb{E}\{(Y(\mathbf{s}_i) - \mu(\mathbf{s}_i))(Y(\mathbf{s}_j) - \mu(\mathbf{s}_j))\} \\ &:= \mathbf{C}_Y(\mathbf{s}_i, \mathbf{h}_{ij}) \\ &= \varphi_Y(\mathbf{s}_i, \mathbf{h}_{ij}) - \mu(\mathbf{s}_i)\mu(\mathbf{s}_j) \end{aligned} \quad (4.6)$$

$$\begin{aligned} \text{Var}\{Y(\mathbf{s}_i)\} &= \mathbb{E}\{(Y(\mathbf{s}_i) - \mu(\mathbf{s}_i))^2\} \\ &= \text{Cov}\{Y(\mathbf{s}_i), Y(\mathbf{s}_i)\} \\ &= \mathbf{C}_Y(\mathbf{s}_i, \mathbf{0}) \\ &= \varphi_Y(\mathbf{s}_i, \mathbf{0}) - \mu^2(\mathbf{s}_i). \end{aligned} \quad (4.7)$$

Here,  $\mu(\mathbf{s}_i)$ ,  $\text{Corr}\{Y(\mathbf{s}_i), Y(\mathbf{s}_j)\}$ ,  $\text{Cov}\{Y(\mathbf{s}_i), Y(\mathbf{s}_j)\}$  and  $\text{Var}\{Y(\mathbf{s}_i)\}$  are the mean, the correlation, the covariance and the variance of the SRF for  $\mathbf{s}_i$ , respectively. The vector  $\mathbf{h}_{ij} = \mathbf{s}_i - \mathbf{s}_j \in \mathbb{R}^d$  represents the spatial lag vector between  $\mathbf{s}_i$  and  $\mathbf{s}_j$ . The correlation and the covariance can be expressed as a function of the position and the lag vector, denoted as  $\varphi_Y(\cdot)$  and  $\mathbf{C}_Y(\cdot)$  in this work. The covariance function for a SRF should be nonnegative definite as the variances of linear combination of  $Y(\mathbf{s})$  are non-negative [13].

One of the common inference approaches utilizes the covariance of the SRF, which needs to be estimated from the available samples. However, without any assumptions

its estimate cannot be obtained, as the available dataset is inadequate to represent the SRF. Hence, for deriving a covariance-based inference, it is assumed that second-order stationarity holds true for the SRF. Second-order stationarity is the equivalent notion of spatial Wide Sense Stationarity (WSS) and satisfies

$$\begin{aligned} \text{(i)} \quad & \mathbb{E}\{Y(\mathbf{s})\} = \mu_Y \\ \text{(ii)} \quad & \text{Corr}\{Y(\mathbf{s}), Y(\mathbf{s} + \mathbf{h})\} = \varphi_Y(\mathbf{h}) \quad \forall \mathbf{s} \\ \text{(iii)} \quad & \text{Cov}\{Y(\mathbf{s}), Y(\mathbf{s} + \mathbf{h})\} = c_Y(\mathbf{h}) \quad \forall \mathbf{s} \end{aligned} \quad (4.8)$$

This indicates that (i) the mean of the SRF is constant and (ii) and (iii) its second-order moments are shift invariant and depend only on the lag vector  $\mathbf{h} \in \mathbb{R}^d$  between two points in  $D$  and not on the two locations themselves [14, 42, 46]. The variance can be described with the covariance function as  $c_Y(\mathbf{0})$ , which is called *sill* in geostatistics.

However, this assumption of second-order stationarity can be sometimes unrealistic, making covariance-based inference inappropriate in many scenarios. To accommodate this problem, we can weaken the assumption, such that an inference can be made even for a potentially nonstationary process. This assumption is called intrinsic stationarity which applies spatial WSS only to the *increments* of an SRF. A spatial increments of our SRF for a lag  $\mathbf{h}$  can be described as

$$X^{[\mathbf{h}]}(\mathbf{s}) = Y(\mathbf{s}) - Y(\mathbf{s} + \mathbf{h}), \quad (4.9)$$

which depends on the position  $\mathbf{s}$  for a fixed  $\mathbf{h}$ . Under the assumption of intrinsic stationarity, this incremental process of a fixed lag  $\mathbf{h}$  is regarded as zero-mean second-order stationary, meaning it satisfies followings

$$\begin{aligned} \text{(i)} \quad & \mathbb{E}\{X^{[\mathbf{h}]}(\mathbf{s})\} = \mathbb{E}\{Y(\mathbf{s}) - Y(\mathbf{s} + \mathbf{h})\} \\ & = 0 \\ \text{(ii)} \quad & \text{Corr}\{X^{[\mathbf{h}]}(\mathbf{s}), X^{[\mathbf{h}]}(\mathbf{s} + \boldsymbol{\vartheta})\} = \mathbb{E}\{X^{[\mathbf{h}]}(\mathbf{s}) \cdot X^{[\mathbf{h}]}(\mathbf{s} + \boldsymbol{\vartheta})\} \\ & := \varphi_X^{[\mathbf{h}]}(\boldsymbol{\vartheta}) \\ \text{(iii)} \quad & \text{Var}\{X^{[\mathbf{h}]}(\mathbf{s})\} = \mathbb{E}\{(X^{[\mathbf{h}]}(\mathbf{s}))^2\} \\ & = \varphi_X^{[\mathbf{h}]}(\mathbf{0}) \\ & = \text{Var}\{Y(\mathbf{s}) - Y(\mathbf{s} + \mathbf{h})\} \\ & := 2\gamma_Y(\mathbf{h}). \end{aligned} \quad (4.10)$$

As is the case with a second-order stationary SRF, the spatial correlation of two incremental processes with the same lag,  $\mathbf{h}$ , depends solely on the distance between two initial points,  $\boldsymbol{\vartheta}$ , regardless of these initial points themselves, denoted as  $\varphi_X^{[\mathbf{h}]}(\boldsymbol{\vartheta})$ . The quantity  $2\gamma_Y(\mathbf{h})$  and  $\gamma_Y(\mathbf{h})$  are called variogram and semivariogram of the SRF and can be expressed as a function of the lag for the considered incremental process,  $\mathbf{h}$ . The definition in (4.10) implies that the smaller variogram value means the stronger spatial correlation. Hence, it is of our interest to determine how *fast* the variogram value increases over the lags.

It is relatively easy to see the relationship between the autocovariance function and the semivariogram of a SRF. Under the assumption of second-order stationarity, the semivariogram can be expressed based on (4.10) as [42]

$$\begin{aligned}\gamma_Y(\mathbf{h}) &= \frac{1}{2} [\text{Var} \{Y(\mathbf{s})\} + \text{Var} \{Y(\mathbf{s} + \mathbf{h})\}] - \text{Cov} \{Y(\mathbf{s}), Y(\mathbf{s} + \mathbf{h})\} \\ &= c_Y(\mathbf{0}) - c_Y(\mathbf{h}).\end{aligned}\quad (4.11)$$

This relationship is illustrated in Figure 4.1. Moreover, as (4.11) suggests, the semi-variogram is nonnegative, yet restricted to be nonpositive definite, since the covariance function  $c_Y(\mathbf{h})$  is nonnegative definite [14, 46].

As opposed to covariance-based inferences, there are several advantages to make inference using a variogram. For determining an autocovariance function, second-order stationarity should hold true for the SRF of interest. On the contrary, computing a variogram does not require the stationarity of the SRF since it is based on a weaker assumption of intrinsic stationarity. As a consequence, an inference using variogram is more robust against nonstationary processes than its covariance-based counterpart [14, 21]. In addition to that, it has been shown that a variogram is easier to estimate than an autocovariance function [46].

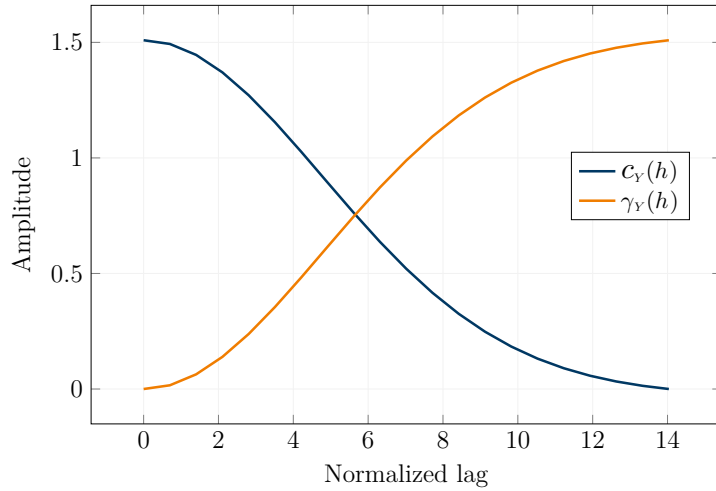


Figure 4.1.: Illustration of the relationship between the covariance (blue) and the semi-variogram (orange) of a SRF if second-order stationarity holds true, which is shown in (4.11).

### 4.3. Geostatistical Inference Technique: Kriging

Our objective of the SRF analysis is to predict an unknown value of the SRF from the given dataset as described in (4.3). One of the well-established methods for spatial prediction is a variogram-based inference scheme called Kriging. Kriging can be characterized as a linear Minimum Mean Squared Error (MMSE) predictor, which, under



proper assumptions, yields the best linear unbiased prediction. Indeed, any unbiased linear predictor has the property that its mean squared prediction error can be expressed in terms of the variogram [13], which is presented in this section. Although Kriging is robust again nonstationarity as it is based on variogram, in this section our SRF of interest  $Y(\mathbf{s})$  is considered to be second-order stationary for the sake of simplicity. This means that  $Y(\mathbf{s})$  satisfies (4.8) and (4.11). Derivations of this section can be found in [13, 14, 42, 47].

A linear prediction of  $\hat{Y}(\mathbf{s}_0)$  can be expressed as

$$\hat{Y}(\mathbf{s}_0) = \mathbf{w}^T \mathbf{y}_N, \quad (4.12)$$

with  $\mathbf{w} \in \mathbb{R}^N$  being weights. Based on this notion, a linear MMSE predictor seeks to solve the following problem

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad f(\mathbf{w}) &= \mathbb{E} \left\{ \left( Y(\mathbf{s}_0) - \hat{Y}(\mathbf{s}_0) \right)^2 \right\} \\ &= \mathbb{E} \{ Y(\mathbf{s}_0)^2 \} - 2\mathbf{w}^T \mathbf{u}_{0y_N} + \mathbf{w}^T \mathbf{R}_{y_N y_N} \mathbf{w}, \end{aligned} \quad (4.13)$$

where

$$\begin{aligned} \mathbf{u}_{0y_N} &= \mathbb{E} \{ Y(\mathbf{s}_0) \mathbf{y}_N \} \in \mathbb{R}^N \\ \mathbf{R}_{y_N y_N} &= \mathbb{E} \{ \mathbf{y}_N \mathbf{y}_N^T \} \in \mathbb{R}^{N \times N}. \end{aligned} \quad (4.14)$$

The optimal solution of (4.13) is well known as

$$\mathbf{w}_{\text{opt}} = 2\mathbf{R}_{y_N y_N}^{-1} \mathbf{u}_{0y_N}. \quad (4.15)$$

In a very similar manner, Kriging aims to solve the same problem with a constraint to reinforce unbiasedness as

$$\underset{\mathbf{w}}{\text{minimize}} \quad f(\mathbf{w}) \quad \text{s. t.} \quad \mathbb{E} \{ Y(\mathbf{s}_0) - \hat{Y}(\mathbf{s}_0) \} = 0. \quad (4.16)$$

From (4.8) we can see that the expected value of the available dataset is constant  $\mu_Y$ , suggesting that the constraint in (4.16) is satisfied if

$$\sum_{i=1}^n w_i = 1. \quad (4.17)$$

Furthermore, considering that its bias should be equal to 0, the Mean Squared Error (MSE) of  $\hat{Y}(\mathbf{s}_0)$  becomes identical to its variance, letting us to reformulate (4.16) as

$$\begin{aligned} f(\mathbf{w}) &= \mathbb{E} \left\{ \left( Y(\mathbf{s}_0) - \hat{Y}(\mathbf{s}_0) \right)^2 \right\} \\ &= \text{Var} \left\{ Y(\mathbf{s}_0) - \hat{Y}(\mathbf{s}_0) \right\} \\ &= \text{Var} \{ Y(\mathbf{s}_0) \} + \text{Var} \{ \hat{Y}(\mathbf{s}_0) \} - 2 \text{Cov} \{ Y(\mathbf{s}_0), \hat{Y}(\mathbf{s}_0) \}. \end{aligned} \quad (4.18)$$

Since (4.11) holds true for  $Y(\mathbf{s})$ , and the optimal weights should satisfy (4.17), each component of (4.18) can be expressed in terms of the covariance and the semivariogram of  $Y(\mathbf{s})$  as

$$\text{Var} \{ Y(\mathbf{s}_0) \} = c_Y(\mathbf{0}) \quad (4.19)$$

$$\begin{aligned}
\text{Var} \left\{ \hat{Y}(\mathbf{s}_0) \right\} &= \text{Var} \left\{ \sum_{i=1}^n w_i y(\mathbf{s}_i) \right\} \\
&= \sum_{i=1}^N \sum_{j=1}^N w_i w_j \text{Cov} \{ y(\mathbf{s}_i), y(\mathbf{s}_j) \} \\
&= \sum_{i=1}^N \sum_{j=1}^N w_i w_j [c_Y(\mathbf{0}) - \gamma_Y(\mathbf{h}_{ij})] \\
&= c_Y(\mathbf{0}) - \sum_{i=1}^N \sum_{j=1}^N w_i w_j \gamma_Y(\mathbf{h}_{ij})
\end{aligned} \tag{4.20}$$

$$\begin{aligned}
\text{Cov} \left\{ Y(\mathbf{s}_0), \hat{Y}(\mathbf{s}_0) \right\} &= \text{E} \left\{ (Y(\mathbf{s}_0) - \mu_Y) (\hat{Y}(\mathbf{s}_0) - \mu_Y) \right\} \\
&= \sum_{i=1}^N (w_i \text{E} \{ Y(\mathbf{s}_0) y(\mathbf{s}_i) \}) - \mu_Y^2 \\
&= \sum_{i=1}^N w_i (c_Y(\mathbf{0}) - \gamma_Y(\mathbf{h}_{0i}) + \mu_Y^2) - \mu_Y^2 \\
&= c_Y(\mathbf{0}) - \sum_{i=1}^N w_i \gamma_Y(\mathbf{h}_{0i}),
\end{aligned} \tag{4.21}$$

with  $\mathbf{h}_{ij} = \mathbf{s}_i - \mathbf{s}_j \in \mathbb{R}^d$  being the spatial lag vector.

Consequently, (4.16) can be simplified as

$$\begin{aligned}
\underset{\mathbf{w}}{\text{minimize}} \quad & f(\mathbf{w}) = -\mathbf{w}^T \mathbf{\Gamma}_{\mathbf{y}_N \mathbf{y}_N} \mathbf{w} + 2\mathbf{w}^T \mathbf{v}_{0y_N} \\
\text{s. t.} \quad & \sum_{i=1}^n w_i = 1.
\end{aligned} \tag{4.22}$$

The matrix  $\mathbf{\Gamma}_{\mathbf{y}_N \mathbf{y}_N} \in \mathbb{R}^{N \times N}$  and the vector  $\mathbf{v}_{0y_N} \in \mathbb{R}^N$  account for the SRF statistics defined as

$$\mathbf{\Gamma}_{\mathbf{y}_N \mathbf{y}_N} = \begin{bmatrix} \gamma_Y(\mathbf{0}) & \gamma_Y(\mathbf{h}_{21}) & \dots & \gamma_Y(\mathbf{h}_{N1}) \\ \gamma_Y(\mathbf{h}_{12}) & \gamma_Y(\mathbf{0}) & \dots & \gamma_Y(\mathbf{h}_{N2}) \\ \vdots & & & \vdots \\ \gamma_Y(\mathbf{h}_{1N}) & \gamma_Y(\mathbf{h}_{2N}) & \dots & \gamma_Y(\mathbf{0}) \end{bmatrix} \tag{4.23}$$

$$\mathbf{v}_{0y_N} = \begin{bmatrix} \gamma_Y(\mathbf{h}_{01}) \\ \gamma_Y(\mathbf{h}_{02}) \\ \vdots \\ \gamma_Y(\mathbf{h}_{0N}) \end{bmatrix}. \tag{4.24}$$

The optimal solution for (4.22) can be obtained via the method of Lagrange multiplier as

$$\begin{bmatrix} \mathbf{w}_{\text{opt}} \\ \frac{\lambda}{2} \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}_{\mathbf{y}_N \mathbf{y}_N} & \mathbf{1}_N \\ \mathbf{1}_N^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{v}_{0y_N} \\ 1 \end{bmatrix} \tag{4.25}$$

where  $\lambda \in \mathbb{R}$  is the Lagrange multiplier and  $\mathbf{1}_N \in \mathbb{R}^N$  represents a vector containing only 1.

As opposed to other general interpolation schemes, Kriging allows us to estimate its prediction accuracy since it is based on the statistical moments of the SRF. This is very useful, especially when used for multiple points inference, to highlight inadequate sampling in regions with lower accuracy. This accuracy estimate  $\sigma_{s_0}$  is the sum of the weighted variances and the Lagrange multiplier as

$$\sigma^2(s_0) = \mathbf{w}_{\text{opt}}^T \mathbf{v}_{0_{y_s}} + \frac{\lambda}{2}. \quad (4.26)$$

#### 4.4. Variogram Estimation

In order to perform Kriging, the variogram of a SRF is to be estimated. Typically, there are two approaches to obtain an estimate: empirical estimate and parametric modeling.

An empirical approach is based solely on the available data  $\mathbf{y}_N$ , by computing the average of semivariances of all possible pairs in  $S_N$ . Such estimate is called experimental variogram or sample variogram and expressed as [14]

$$\hat{\gamma}_Y(\mathbf{h}) = \frac{1}{2|N(\mathbf{h})|} \sum_{N(\mathbf{h})} [Y(\mathbf{s}_i) - Y(\mathbf{s}_j)]^2. \quad (4.27)$$

Here,  $\hat{\gamma}_Y(\mathbf{h})$  represents the estimated semivariogram, while  $N(\mathbf{h}) = \{(\mathbf{s}_i, \mathbf{s}_j) : \mathbf{s}_i - \mathbf{s}_j = \mathbf{h}\}$  denotes the set of all pairs of locations separated by the lag vector  $\mathbf{h}$ .

Although it is intuitive and unbiased, experimental variograms fails to provide a valid estimate if there are not sufficient samples available. This is due to the following reasons. First, it can only provide estimates at particular lags available within the sampling positions  $S_N$ . This implies that some lag vectors may be missing if there are not sufficient samples available, making some inferences impossible. Second, it is not assured to be nonpositive definite, which is required as a valid model. Furthermore, it may end up a non-smooth modeling with respect to increasing lag due to random perturbations. Hence, it is desirable to find a continuous function of lags, which fits well with the experimental variogram and at the same time satisfies the nonpositive definiteness.

Upon modeling, an assumption of *isotropy* is commonly made [14, 42, 46]. If the SRF is isotropic, then its variogram depends solely on distance between two locations  $h = \|\mathbf{h}\|_2$ , regardless of direction or orientation between them, as

$$\gamma_Y(\mathbf{h}) = \gamma_Y(h). \quad (4.28)$$

This holds also true for the autocovariance function of  $Y(\mathbf{s})$ , if it is attainable.

Suppose that the underlying process is a stable process and yields finite-valued results within the ROI, then its variogram converges for  $h \rightarrow \infty$ . It is expected that variogram increases monotonically up to a certain  $h$ , since the values of nearby locations are similar than those further away (Figure 4.2). In this case, an estimate can be made based on (4.11) [14]

$$\hat{\gamma}_Y(h) = \hat{c}_Y(0) - \hat{c}_Y(h), \quad (4.29)$$

indicating that we essentially need to model the autocovariance function.

One of the common covariance models is exponential model, which can be expressed as [14]

$$\hat{c}_Y(h|\boldsymbol{\theta}) = \sigma^2 e^{-\nu h}, \quad (4.30)$$

where  $\boldsymbol{\theta} = (\sigma, \nu)$  is the parameters to estimate. The parameter  $\sigma$  determines the overall variability of the process, while the parameter  $\nu$  controls the rate of decay of correlation with regard to the distance.

Another common model is spherical model, which is modeled as [14]

$$\hat{c}_Y(h|\boldsymbol{\theta}) = \sigma^2 \left[ 1 - \frac{3h}{2r} + \frac{h^3}{2r^3} \right] \quad (4.31)$$

Here again, the parameter  $\sigma$  has the same effects on the autocovariance function, while the range of the correlation is now controlled by  $r$ . With the spherical model, the variables separated by more than the range  $r$  are considered to be uncorrelated, imposing their correlation to be zero.

A more flexible covariance model is Matérn model defined as [48]

$$\hat{c}_Y(h|\boldsymbol{\theta}) = \frac{\sigma^2}{2^{\eta-1}\Gamma(\eta)} \left( \frac{\|\mathbf{h}\|_2}{\frac{r}{2}} \right)^\eta K_\eta \left( \frac{\|\mathbf{h}\|_2}{\frac{r}{2}} \right), \quad (4.32)$$

which includes the parameter  $\eta > 0$  to control the smoothness of the covariance function. The parameters  $\sigma$  and  $r$  have the same implications as the aforementioned two models, while  $K_\eta(\cdot)$  denotes the modified Bessel function of second kind of order  $\eta$ .

However, it is not always possible to attain an autocovariance function, since  $Y(\mathbf{s})$  may not be second-order stationary. For instance, the expected value of  $Y(\mathbf{s})$  may vary with locations, resulting in a nonconvergent incremental process. In such cases,  $Y(\mathbf{s})$  can still be considered as intrinsic stationary, allowing its variogram to be modeled directly. A common nonconvergent model is the power model defined as [46]

$$\hat{\gamma}_Y(\mathbf{h}|\boldsymbol{\theta}) = a + b \|\mathbf{h}\|^v. \quad (4.33)$$

Here, the parameters  $a \geq 0$ ,  $b \geq 0$  and  $0 \leq v < 2$  are the offset at  $h = 0$  called *nugget*, the scaling factor and the range to control the correlation level.

Choosing the right values for the model parameters also plays crucial role in variogram estimation. Generally, the parameter values are selected by finding the best fit to the experimental variogram of the available dataset. Typical approaches are Maximum Likelihood (ML) (and its variant) and least squares based methods [46].

## 4.5. Spatio-Temporal Random Field (STRF) Model

Let us now consider spatio-temporal processes and their data modeling. As a generalized form of purely spatial setting, spatio-temporal datasets can be formulated as Spatio-Temporal Random Field (STRF). For  $\mathbf{s} \in D$  and  $t \in T \subset \mathbb{R}$ , a STRF can be formulated as

$$\{Y(\mathbf{s}, t) : \mathbf{s} \in D, t \in T\}, \quad (4.34)$$

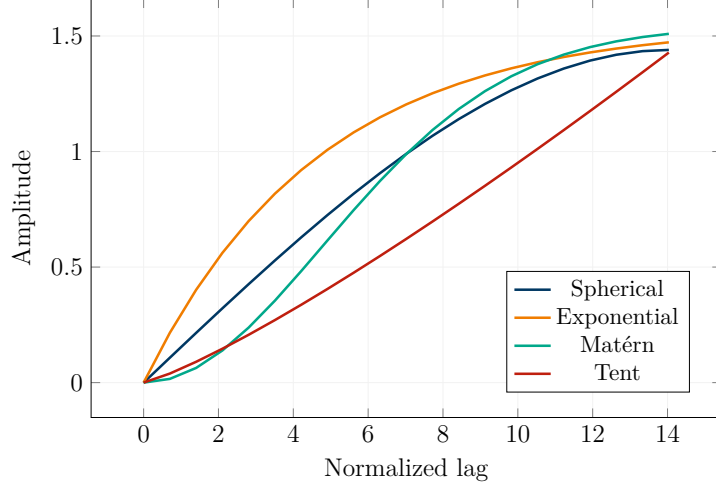


Figure 4.2.: Illustration of different types of variogram model

where  $T$  is the duration of each observation. In the same manner as (4.2), the available dataset sampled at the positions  $S_N = \{\mathbf{s}_1, \mathbf{s}_2 \dots \mathbf{s}_N\}$  is considered to be a single realization of the STRF. Difference is that each data is now a time series instead of a scalar value, which is then discretized with a temporal sampling interval  $t_s$  to be stored in a measurement system. This enables us to formulate each time series as a vector

$$\mathbf{y}(\mathbf{s}_i) = \begin{bmatrix} y(\mathbf{s}_i, 0) \\ y(\mathbf{s}_i, t_s) \\ \vdots \\ y(\mathbf{s}_i, (M-1)t_s) \end{bmatrix} \in \mathbb{R}^M, \quad (4.35)$$

with  $M$  being the temporal dimension of each data. This vector-valued time series can be concatenated into a single matrix to form an observation matrix  $\mathbf{Y}_N$  as

$$\mathbf{Y}_N = [\mathbf{y}(\mathbf{s}_1) \quad \mathbf{y}(\mathbf{s}_2) \quad \dots \quad \mathbf{y}(\mathbf{s}_N)] \in \mathbb{R}^{M \times N}. \quad (4.36)$$

Analogous to (4.3), our inference problem now can be formulated as

$$\mathbb{E}\{Y(\mathbf{s}_0, t) \mid \mathbf{Y}_N\}. \quad (4.37)$$

Stationarity assumptions are useful to estimate the statistical moments. Equivalent to the purely spatial processes, spatio-temporal stationarities are defined as follows. Spatio-temporal second-order stationarity (equivalent to (4.8)):

$$\begin{aligned} \text{(i)} \quad & \mathbb{E}\{Y(\mathbf{s}, t)\} = \mu_Y \\ \text{(ii)} \quad & \text{Cov}\{Y(\mathbf{s}, t), Y(\mathbf{s} + \mathbf{h}, t + \tau)\} = c_Y(\mathbf{h}, \tau) \\ & = \varphi_Y(\mathbf{h}, \tau) - \mu_Y \end{aligned} \quad (4.38)$$

Spatio-temporal intrinsic stationarity (equivalent to (4.10)):

$$\begin{aligned} \text{(i)} \quad & \mathbb{E}\{Y(\mathbf{s}, t) - Y(\mathbf{s} + \mathbf{h}, t + \tau)\} = 0 \\ \text{(ii)} \quad & \text{Var}\{Y(\mathbf{s}, t) - Y(\mathbf{s} + \mathbf{h}, t + \tau)\} := 2\gamma_Y(\mathbf{h}, \tau), \end{aligned} \quad (4.39)$$

where  $\tau \in \mathbb{R}$  is the temporal lag, while  $c_Y(\cdot)$ ,  $\varphi_Y(\cdot)$ ,  $2\gamma_Y(\cdot)$  denote covariance, correlation and variogram of  $Y(\mathbf{s}, t)$ , respectively. Spatio-temporal semivariogram can be also expressed in terms of spatio-temporal covariance function similar to (4.11) as

$$\gamma_Y(\mathbf{h}, \tau) = c_Y(\mathbf{0}, 0) - c_Y(\mathbf{h}, \tau). \quad (4.40)$$

## 4.6. Prediction in Space-Time Domain

Although in theory space and time should not be treated equally, from the point of view of linear predictors, time is just another dimension [13]. It is how we model the second-order statistics of a Spatio-Temporal Random Field (STRF) that reflects the different behaviors of the data with respect to space and time. This suggests that a prediction can be obtained in a very similar way as with the purely spatial scenario.

Equivalent to (4.12), a linear predictor of the value at the location  $\mathbf{s}_0$  of the time point  $t_0$  is expressed as

$$\hat{Y}(\mathbf{s}_0, t_0) = \mathbf{w}^T \cdot \text{vec}\{\mathbf{Y}_N\}. \quad (4.41)$$

This enables us to formulate the optimization problem for spatio-temporal Kriging as

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad & f(\mathbf{w}) = -\mathbf{w}^T \cdot \mathbf{\Gamma}_{Y_N Y_N} \cdot \mathbf{w} + 2\mathbf{w}^T \cdot \mathbf{v}_{0Y_N} \\ \text{s. t.} \quad & \sum_{i=1}^{nm} w_i = 1. \end{aligned} \quad (4.42)$$

The matrix  $\mathbf{\Gamma}_{Y_N Y_N} \in \mathbb{R}^{NM \times NM}$  and the vector  $\mathbf{v}_{0Y_N} \in \mathbb{R}^{NM}$  account for the second-order statistics, which are defined as

$$\mathbf{\Gamma}_{Y_N Y_N} = \begin{bmatrix} \gamma_Y(\mathbf{0}, 0) & \gamma_Y(\mathbf{0}, -t_S) & \dots & \gamma_Y(\mathbf{h}_{N1}, -(M-1)t_S) \\ \gamma_Y(\mathbf{0}, t_S) & \gamma_Y(\mathbf{0}, 0) & \dots & \gamma_Y(\mathbf{h}_{N1}, -(M-2)t_S) \\ \vdots & & & \vdots \\ \gamma_Y(\mathbf{h}_{1N}, (M-1)t_S) & \gamma_Y(\mathbf{h}_{1N}, (M-2)t_S) & \dots & \gamma_Y(\mathbf{0}, 0) \end{bmatrix} \quad (4.43)$$

$$\mathbf{v}_{0Y_N} = \begin{bmatrix} \gamma_Y(\mathbf{h}_{01}, t_0) \\ \gamma_Y(\mathbf{h}_{01}, t_0 - t_S) \\ \vdots \\ \gamma_Y(\mathbf{h}_{0n}, t_0 - (m-1)t_S) \end{bmatrix}. \quad (4.44)$$

The optimal weights can be obtained in the exact same manner as (4.25), which requires the inverse of  $\mathbf{\Gamma}_{Y_N Y_N}$ .

There is, however, one significant difference in spatio-temporal Kriging compared to its spatial counterpart: the dimension of the variogram matrix  $\mathbf{\Gamma}_{Y_N Y_N}$ . This matrix now

contains the temporal dimension as well, which can easily make its inversion computationally expensive. In general, temporal dimension tends to be much larger than the spatial one, necessitating some kinds of dimensional reduction.

## 4.7. Spectral Representation of STRF

A well known approach for dimension reduction is to represent time series data as their spectra in frequency domain. With spectral representation, the same information is attainable while the underlying correlation structure can be simplified without much loss of information. This is due to the orthogonal nature of Fourier basis functions, making correlation between spectra of two distinctive frequencies negligible. This notion can be extended to the spatial domain as well. In this section, spectral representation of STRFs is introduced, highlighting its usefulness compared to space-time domain.

For a STRF data, its spectrum in space-frequency domain is defined as

$$\begin{aligned} Z_Y(\mathbf{s}, \omega) &:= \mathcal{F}_t \{Y(\mathbf{s}, t)\} \\ &= \int_{-\infty}^{\infty} Y(\mathbf{s}, t) e^{-j\omega t} dt \end{aligned} \quad (4.45)$$

with  $\omega = 2\pi f$  being the angular frequency. This frequency response (4.45) can be further transformed into wave number domain as

$$\begin{aligned} Z_Y(\mathbf{k}, \omega) &:= \mathcal{F}_s \{ \mathcal{F}_t \{Y(\mathbf{s}, t)\} \} \\ &= \int_{\mathbb{R}^d} Z_Y(\mathbf{s}, \omega) e^{-j\mathbf{k}^T \mathbf{s}} d\mathbf{s} \\ &= \int_{\mathbb{R}^d} \int_{-\infty}^{\infty} Y(\mathbf{s}, t) e^{-j\omega t} e^{-j\mathbf{k}^T \mathbf{s}} dt d\mathbf{s} \end{aligned} \quad (4.46)$$

where  $\mathbf{k} \in \mathbb{R}^d$  denotes the vector of wave numbers, which is the Fourier pair of  $\mathbf{s}$ . From the spectrum (4.46), the spectral density of  $Y(\mathbf{s}, t)$  can be obtained

$$\Phi_{Z_Y}(\mathbf{k}, \omega) := \mathbb{E} \left\{ |Z_Y(\mathbf{k}, \omega)|^2 \right\}, \quad (4.47)$$

where  $\Phi_{Z_Y}(\mathbf{k}, \omega)$  represents the spectral density function.

### Statistical Moments of Space-Frequency STRF Spectra

When it comes to its statistical moments, which we need for making an inference, we can see the benefits of transforming a STRF into space-frequency domain. For a fixed position, correlation of space-frequency spectra of two distinctive frequencies is almost negligible. This suggests that we can separate the spectra into each frequency bin and compute their statistical moments individually. In this way, we need to consider only the spatial correlation, which is equivalent to the purely spatial scenarios discussed in 4.2.

Under the assumption of spatial second-order stationarity, the second-order moments of  $Z_Y(\mathbf{s}, \omega_m)$  can be described only in terms of the lag vector  $\mathbf{h}$ , which is analogous to (4.8),

$$\begin{aligned}\varphi_{Z_Y}^{[\omega_m]}(\mathbf{h}) &:= \text{Corr} \{Z_Y(\mathbf{s}, \omega_m), Z_Y(\mathbf{s} + \mathbf{h}, \omega_m)\} \\ &= \text{E} \{Z_Y(\mathbf{s}, \omega_m)^* \cdot Z_Y(\mathbf{s} + \mathbf{h}, \omega_m)\} \\ &= \mathcal{F}_{\mathbf{h}}^{-1} \left\{ \Phi_{Z_Y}^{[\omega_m]}(\mathbf{k}) \right\}\end{aligned}\tag{4.48}$$

$$\begin{aligned}\mathcal{C}_{Z_Y}^{[\omega_m]}(\mathbf{h}) &:= \text{Cov} \{Z_Y(\mathbf{s}, \omega_m), Z_Y(\mathbf{s} + \mathbf{h}, \omega_m)\} \\ &= \varphi_{Z_Y}^{[\omega_m]}(\mathbf{h}) - \left| \mu_{Z_Y}^{[\omega_m]} \right|^2,\end{aligned}\tag{4.49}$$

with  $\mu_{Z_Y}^{[\omega_m]} \in \mathbb{C}$  being its constant mean.

If the STRF is intrinsic stationary in space, we are interested in its spatial incremental process. Let the spatial incremental process of  $Y(\mathbf{s}, t)$  be

$$X^{[h]}(\mathbf{s}, t) = Y(\mathbf{s}, t) - Y(\mathbf{s} + \mathbf{h}, t),\tag{4.50}$$

whose spectral representation in space-frequency domain is denoted as

$$Z_X^{[h]}(\mathbf{s}, \omega) := \mathcal{F}_t \left\{ X^{[h]}(\mathbf{s}, t) \right\}.\tag{4.51}$$

In the same way as (4.10), under intrinsic stationarity, variance of the incremental process becomes identical to the expected value of its periodogram as

$$\begin{aligned}2\gamma_{Z_Y}^{[\omega_m]}(\mathbf{h}) &:= \text{Var} \{Z_Y(\mathbf{s}, \omega_m) - Z_Y(\mathbf{s} + \mathbf{h}, \omega_m)\} \\ &= \text{Var} \left\{ Z_X^{[h]}(\mathbf{s}, \omega_m) \right\}. \\ &= \text{Corr} \left\{ Z_X^{[h]}(\mathbf{s}, \omega_m), Z_X^{[h]}(\mathbf{s}, \omega_m) \right\} \\ &= \text{E} \left\{ \left| Z_X^{[h]}(\mathbf{s}, \omega_m) \right|^2 \right\} \\ &= \varphi_{Z_X}^{[\mathbf{h}, \omega_m]}(\mathbf{0}).\end{aligned}\tag{4.52}$$

Here,  $\varphi_{Z_X}^{[\mathbf{h}, \omega_m]}(\mathbf{v})$  represents the spatial correlation function of  $Z_X^{[h]}(\mathbf{s}, \omega_m)$ . This real-valued nonnegative function (4.52) is called frequency variogram of the STRF  $Y(\mathbf{s}, t)$ , which is first introduced in [21]. From (4.52), we can see that estimating the frequency variogram of a STRF is equivalent to estimating the periodogram of its spatial incremental process. This implies that second-order stationarity assumption is not necessary to compute predictions in space-frequency domain, which makes the prediction more robust against nonstationarity of the underlying process.

## 4.8. Prediction in Space-Frequency Domain: Space-Frequency Kriging

In this section, we seek to solve the same prediction problem described in Section 4.6. The difference is here we that perform prediction in space-frequency domain instead of



space-time domain. This means that we aim to obtain a prediction for the frequency response of  $\mathbf{y}(\mathbf{s}_0)$  from the available observation data and then back-transform into time domain [16]. Nevertheless, the method is very similar to the ones for purely spatial SRF described in Section 4.3. We call this approach Space-Frequency Kriging (SF-Kriging) in this work.

The ultimate goal is to predict the time series at the position  $\mathbf{s}_0$ , which can be obtained through inverse transform of its frequency response

$$\hat{\mathbf{y}}(\mathbf{s}_0) = \mathbf{F}_M^{-1} \hat{\mathbf{z}}_y(\mathbf{s}_0). \quad (4.53)$$

Here,  $\hat{\mathbf{y}}(\mathbf{s}_0) \in \mathbb{R}^M$  is the estimate of the time series, i.e. in time domain, whereas  $\hat{\mathbf{z}}_y(\mathbf{s}_0) \in \mathbb{C}^M$  is the estimate of the frequency response of the desired time series, i.e. in frequency domain. The matrix  $\mathbf{F}_M^{-1} \in \mathbb{C}^{M \times M}$  are the inverse Discrete Fourier Transform (DFT) matrix of size  $M$ . The available data set  $\mathbf{Y}_N \in \mathbb{R}^{M \times N}$  is defined in (4.36), whose frequency responses can be concatenated into a matrix as

$$\mathbf{Z}_{\mathbf{Y}_N} = \begin{bmatrix} \mathbf{z}_y(\mathbf{s}_1) & \mathbf{z}_y(\mathbf{s}_2) & \dots & \mathbf{z}_y(\mathbf{s}_N) \end{bmatrix} \in \mathbb{C}^{M \times N}. \quad (4.54)$$

This matrix can be also expressed in terms of its Fourier coefficients for each component as

$$\mathbf{Z}_{\mathbf{Y}_N} = \begin{bmatrix} \boldsymbol{\zeta}_y(\omega_0)^T \\ \boldsymbol{\zeta}_y(\omega_1)^T \\ \vdots \\ \boldsymbol{\zeta}_y(\omega_{M-1})^T \end{bmatrix} \quad (4.55)$$

where the Fourier coefficients for the frequency bin  $\omega_m$  is denoted as a vector

$$\boldsymbol{\zeta}_y(\omega_m) = \begin{bmatrix} z_y(\mathbf{s}_1, \omega_m) \\ z_y(\mathbf{s}_2, \omega_m) \\ \vdots \\ z_y(\mathbf{s}_N, \omega_m) \end{bmatrix} \in \mathbb{C}^N. \quad (4.56)$$

A prediction of the desired frequency response is then

$$\hat{\mathbf{z}}_y(\mathbf{s}_0) = \mathbb{E} \{ \mathbf{z}_y(\mathbf{s}_0) | \mathbf{Z}_{\mathbf{Y}_N} \}. \quad (4.57)$$

Because of the orthogonal Fourier basis,  $\hat{\mathbf{z}}_y(\mathbf{s}_0)$  can be estimated individually for each frequency bin, which significantly reduces the computational time. For the frequency component  $\omega_m$ , a linear estimator is expressed as

$$\hat{z}_y(\omega_m) = \mathbf{w}^H(\omega_m) \cdot \boldsymbol{\zeta}_y(\omega_m), \quad (4.58)$$

with weights  $\mathbf{w}(\omega_m) \in \mathbb{C}^N$  and  $\mathbf{w}^H$  representing its Hermitian transpose.

We seek to find a prediction which minimizes its Mean Squared Error (MSE) as

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad f(\mathbf{w}(\omega_m)) &= \mathbb{E} \left\{ |z_y(\mathbf{s}_0, \omega_m) - \hat{z}_y(\mathbf{s}_0, \omega_m)|^2 \right\} \\ &= \mathbb{E} \left\{ \left| z_y(\mathbf{s}_0, \omega_m) - \mathbf{w}^H(\omega_m) \cdot \boldsymbol{\zeta}_y(\omega_m) \right|^2 \right\}. \end{aligned} \quad (4.59)$$

Analogous to Kriging in other domains, the MSE of an unbiased linear predictor can be expressed as

$$\begin{aligned}
\mathbb{E} \left\{ |z_y(\mathbf{s}_0, \omega_m) - \hat{z}_y(\mathbf{s}_0, \omega_m)|^2 \right\} &= \text{Var} \{ z_y(\mathbf{s}_0, \omega_m) - \hat{z}_y(\mathbf{s}_0, \omega_m) \} \\
&= \text{Var} \left\{ z_y(\mathbf{s}_0, \omega_m) - \mathbf{w}^H(\omega_m) \cdot \boldsymbol{\zeta}_y(\omega_m) \right\} \\
&= -\mathbf{w}^H(\omega_m) \cdot \boldsymbol{\Gamma}_{\zeta\zeta}(\omega_m) \mathbf{w}(\omega_m) + 2\mathbf{w}^H(\omega_m) \cdot \mathbf{v}_{\zeta\zeta}(\omega_m).
\end{aligned} \tag{4.60}$$

The matrix  $\boldsymbol{\Gamma}_{\zeta\zeta}(\omega_m) \in \mathbb{R}^{N \times N}$  and the vector  $\mathbf{v}_{\zeta\zeta}(\omega_m) \in \mathbb{R}^N$  again account for the spatial statistics of the ROI and defined based on its frequency variogram as

$$\boldsymbol{\Gamma}_{\zeta\zeta}(\omega_m) = \begin{bmatrix} \gamma_{z_y}^{[\omega_m]}(\mathbf{0}) & \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{21}) & \dots & \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{N1}) \\ \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{12}) & \gamma_{z_y}^{[\omega_m]}(\mathbf{0}) & \dots & \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{N2}) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{1N}) & \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{2N}) & \dots & \gamma_{z_y}^{[\omega_m]}(\mathbf{0}) \end{bmatrix} \tag{4.61}$$

$$\mathbf{v}_{\zeta\zeta}(\omega_m) = \begin{bmatrix} \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{01}) \\ \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{02}) \\ \vdots \\ \gamma_{z_y}^{[\omega_m]}(\mathbf{h}_{0N}) \end{bmatrix}. \tag{4.62}$$

Taking the partial derivatives of (4.59) and solving it yields the optimal weights as [16]

$$\mathbf{w}_{\text{FK}}(\omega_m) = \boldsymbol{\Gamma}_{\zeta\zeta}^{-1}(\omega_m) \cdot \mathbf{v}_{\zeta\zeta}(\omega_m) \in \mathbb{R}^N. \tag{4.63}$$

Furthermore, as the obtained prediction is based on the field statistics, the possible prediction error can also be estimated as

$$\sigma^2(\omega_m) = \mathbf{w}_{\text{FK}}^T(\omega_m) \cdot \mathbf{v}_{\zeta\zeta}(\omega_m). \tag{4.64}$$

From (4.63), we can see the benefits of performing the prediction in space-frequency domain. Compared to the space-time domain approach, the dimension of the matrix  $\boldsymbol{\Gamma}_{\zeta\zeta}(\omega_m)$  is reduced from  $NM \times NM$  to  $N \times N$ , which is equivalent to the computational time reduction from  $O(NM \times NM)$  to  $O(M \cdot N \times N)$ . Considering that the temporal dimension is often much larger than its spatial counterpart, this gain can contribute significantly to online scenarios in many fields.

## 4.9. Non-Parametric Estimation of Frequency Variogram

As discussed in Section 4.7, estimating the frequency variogram of a STRF is equivalent to estimating the periodogram of its spatial incremental process for each frequency. To estimate a frequency variogram non-parametrically, method-of-moments estimator can be employed in a very similar manner as in the purely spatial domain. In [21], it has been

shown that a consistent estimator is attainable by computing weighted window average of the periodogram of the incremental process over neighboring frequency components.

For a fixed lag vector  $\mathbf{h}$ , an empirical estimate of the frequency variogram can be obtained as

$$\begin{aligned} 2\gamma_{z_y}^{[\omega_m]}(\mathbf{h}) &= \varphi_{z_y}^{[\mathbf{h}, \omega_m]}(\mathbf{0}) \\ &= \int_{-\infty}^{\infty} W_L(\omega_m - \lambda) \cdot \left[ \frac{1}{N(\mathbf{h})} \sum_{\forall i} \left| Z_x^{[\mathbf{h}]}(\mathbf{s}_i, \lambda) \right|^2 \right] d\lambda. \end{aligned} \quad (4.65)$$

Here,  $W_L(x)$  is a real-valued even window function with the window length  $L$  whereas  $N(\mathbf{h}) = \{(\mathbf{s}_i, \mathbf{s}_j) : \mathbf{s}_i - \mathbf{s}_j = \mathbf{h}\}$ . The window function should satisfy the following conditions:

$$\begin{aligned} W_L(x) &\geq 0 & \forall x \forall L \\ \int_{-\infty}^{\infty} W_L(x) dx &= 1 & \forall L \\ \int_{-\infty}^{\infty} W_L^2(x) dx &< \infty & \forall L \\ \lim_{M \rightarrow \infty} W_L(x) &= 0 & \forall x. \end{aligned} \quad (4.66)$$

This implies that the weights are concentrated in the neighborhood of  $\omega_m$  with relative weights determined by the tapers [49]. An example of such window function is a rectangular function of the length  $L$  with the amplitude of  $\frac{1}{L}$ .

However, the same problems with non-parametrical estimation remain as discussed in Section 4.4. This estimator is not continuous with respect to the lags and cannot be guaranteed to be nonpositive definite. These problems can be alleviated only through valid parametrical modeling. On the other hand, non-smoothness with respect to the lags can be handled to some extent. Some form of smoothing is often proven to be effective against any form of rough surfaces [49], though undoubtedly the intervals should be sufficiently small.

## 4.10. Parametric Estimation of Frequency Variogram

To parametrically model the frequency variogram of a STRF, a Stochastic Partial Differential Equation (SPDE) based approach is proposed in [21]. SPDE includes, as its name suggests, a stochastic term in Partial Differential Equation (PDE), such that the resulting model accounts for dynamic behavior of both macro structures and micro fluctuations [13]. In general, it is difficult to deal with SPDE of dynamic processes analytically [13]. Although their spectral density functions can be formulated analytically, it does not necessarily mean that the spatial correlation functions of such processes are attainable. This is due to the inclusion of the temporal derivatives, which hinders their spectral density functions to be transformed into spatial domain. It has been demonstrated that there is no closed form expression for the spatial correlation function of SPDE for the heat equation [15, 16].

To circumvent this problem, [16] proposes a generalized form of the SPDE modeling in space-frequency domain, called Complex Stochastic Partial Differential Equation

(CSPDE). According to [16], the implication of this generalized model is following. While the SPDE in space-time domain describes the local behavior of the underlying process during a small time interval, the SPDE in space-frequency domain accounts for the local behavior of the frequency response of the process within a small frequency window. The significance of this modeling is that it allows us to back transform its spectral density function into spatial domain, which can be described as a function of polynomial in frequency  $\omega$ . This polynomial of  $\omega$  corresponds to the temporal derivatives, which assures the inclusion of temporal dynamics. The resulting frequency variogram then becomes very similar to Matérn type of variogram introduced in (4.32).

For CSPDE based modeling of frequency variogram, two parameters are to be estimated, which are the spectral density of the random process and the polynomial of  $\omega$ . Imposing the assumption that the random process is white Gaussian noise both in space and time, [21] proves the possibility to estimate those parameters via Maximum Likelihood (ML) based approach. Yet, considering such estimation should be performed all possible lags, it seems still daunting for online applications. Moreover, the assumption for random process, especially white in space, is rather unrealistic. It may also be worth mentioning that, despite its significant reduction in dimensionality, SF-Kriging has not been, to our knowledge, applied in practical sense. This is, we suppose, partly due to this complex mathematical expression for computing frequency variograms.

#### 4.11. Extension to Multiple Locations

So far we consider the scenarios of point-wise prediction where the value or the time series of an unobserved position  $\mathbf{s}_0$  is to be inferred. This notion can be extended to prediction of multiple points in the ROI. As we have seen in the previous sections, there is essentially no difference in performing linear prediction of purely spatial data or spatio-temporal data. For Kriging or any linear prediction scheme, temporal dimension is just another dimension [13]. Hence, in this section we consider a purely spatial case for the sake of simplicity, though the method described here can be used for STRF in the same manner as well.

Let  $\{Y(\mathbf{s}) \mid \mathbf{s} \in D\}$  be the SRF representation of the process under test. Aligned to the scenario in Section 4.3, a dataset is obtained at  $N$  distinctive positions  $S_N = \{\mathbf{s}_1, \mathbf{s}_2 \dots \mathbf{s}_N\}$ , which is represented as a vector  $\mathbf{y}_N \in \mathbb{R}^N$ . The goal here is to predict the values of  $K$  different positions in the same ROI denoted as  $S_K = \{\mathbf{s}_1^{\sim}, \mathbf{s}_2^{\sim} \dots \mathbf{s}_{\tilde{K}}^{\sim}\} \in \{D \setminus S_N\}$ .

Based on the linear predictor presented in (4.12), we can obtain a prediction for  $S_K$  as

$$\hat{\mathbf{y}}_K = \begin{bmatrix} \hat{y}(\mathbf{s}_1^{\sim}) \\ \hat{y}(\mathbf{s}_2^{\sim}) \\ \vdots \\ \hat{y}(\mathbf{s}_{\tilde{K}}^{\sim}) \end{bmatrix} = \mathbf{W}^T \mathbf{y}_N. \quad (4.67)$$

The matrix  $\mathbf{W} \in \mathbb{R}^{N \times K}$  is a weight matrix containing the weights for each prediction

position as

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}(s_1) & \mathbf{w}(s_2) & \dots & \mathbf{w}(s_{\tilde{K}}) \end{bmatrix}. \quad (4.68)$$

There is one concern in (4.67), which is the inclusion of all the sampled positions for every prediction point. Undoubtedly there is stronger correlation between two closer points than the ones separated away, implying that the samples located farther away may be destructive for the prediction (Figure 4.3). One approach for compensating this issue is to impose the maximal number of the samples to be used for prediction [48]. Alternatively, the neighboring region of each prediction position can be limited by setting the maximal lag and including only the samples within this lag.

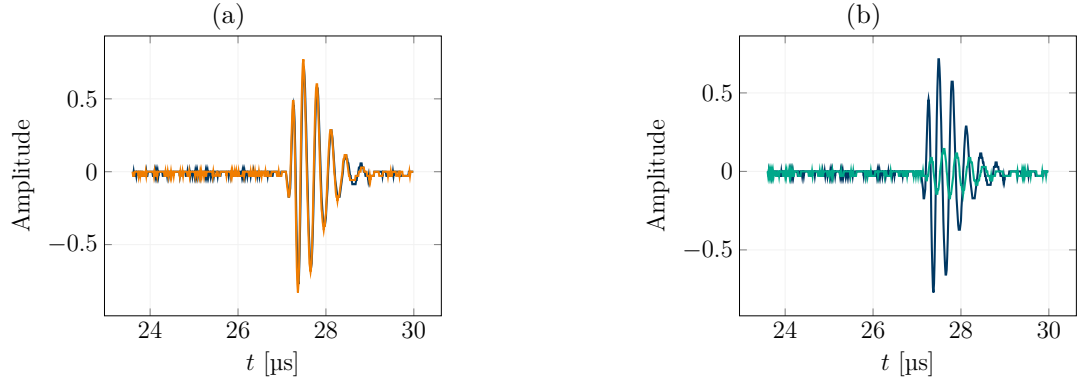


Figure 4.3.: Example A-Scans of the samples in the neighborhood (left) and the ones located farther away (right).

## 5. Summary of Part I

In Part I, the considered scenario and motivation for this thesis is first described, where measurement systems and the data modeling of manual UT are presented. This follows the problem formulation of our research problem, which is a development of the interpolation method as a preprocessing, where the missing A-Scans are to be predicted.

In the following chapter, an overview and the state of the art approaches of DNN are provided. Here, we highlight the interpretability of DL applications play a crucial role to ensure the generalization of the results. To overcome this limitation, DNN can be incorporated into a conventional theory-based approaches, which are interpretable yet may be slow in execution, if not computationally intractable. Such hybrid approaches are gaining interests from many fields, casting a light onto new approach in scientific computing.

In the last chapter of this part, fundamental idea of spatio-temporal interpolation is extensively studied. Through literature research, we found that there have been growing interests in spatio-temporal interpolation in many different applications. As one of the well-established scientific field, geostatistics can provide useful insights into such data, implying the possibility of applications in different fields. Nonetheless, we also pointed out that analysis and processing of spatio-temporal data are remained challenging. One reason for that is the difficulty in finding proper statistical models which accounts for not only the dependency in each domain but also the interactions between these two. Considering that the different characteristics of time (unidirectional) and space (non directional), spatio-temporal data should be carefully handled. Another reason is the large dimensionality, which can easily become computationally intractable, requiring any kind of dimension reduction.

One possible way to reduce the dimension is transforming the data into frequency domain. Because of the orthogonal Fourier bases, the correlation between two frequency bins is in theory negligible, and thus the Fourier coefficients of each frequency can be handled individually. This also allows parallel computing, which can further reduce the computational time. The problem here is, yet, the spatial statistics of the frequency responses, frequency variograms, are to be estimated, whose parameters vary for each frequency and each lag. In view of manual UT applications, there are still needs to find a measure to reduce the computation.

Part II.

# Data Driven Space-Frequency Domain Interpolation

In this chapter, a data-driven hybrid approach is presented for interpolating spatio-temporal data. The proposed method is generic and in theory not limited to UT data. Yet, in this chapter we restrict ourselves to the scenario described in Section 2.3 and aim to predict missing A-Scans at unobserved measurement grid points in the 3D SmartInspect system.

The primary work of this thesis is to formulate our research problem in a computational tractable manner while keeping the interpretability of the entire process. For this purpose, we decided to incorporate DNN into conventional theory-based interpolation scheme, such that we do not have to rely solely on DNN, yet can benefit from its fast computing in execution and capability of learning complex mappings. With that said, it is of our prime interest to understand the role the network plays in our preprocessing scheme. However, how the network accomplishes the given task is not.

Now the question is into which theory-based approach the network should be integrated. When it comes to spatial interpolation, geostatistical approaches are well-established and thus a good candidate for our method. Especially the possibility to estimate prediction error is a very useful feature, making it more appealing. Since we ultimately want to compute spatio-temporal interpolation on-site, Space-Frequency Kriging (SF-Kriging) appears to be an appropriate choice due to its reduced dimensionality. As we discussed in Section 4.10, difficulty in conducting SF-Kriging is, despite it indeed can be expressed in closed form, mathematically complex modeling for frequency variograms. Considering that frequency variograms can be to some extent obtained non-parametrically, there may be possibilities to utilize the powerful non-linear mapping of DNN.

Based on this idea and aligned to our scenario, we decided the structure of our preprocessing scheme as depicted in Figure 5.1. In this proposed approach, spatio-temporal interpolation is performed via batch-wise SF-Kriging, i.e. in space-frequency domain. As a part of our preprocessing, DNN plays a role to estimate frequency variograms, which we denote FVnet in this work. In the following chapters, the employed algorithm for batch-wise SF-Kriging is first described. This follows the discussion regarding the configuration and training details of FVnet.

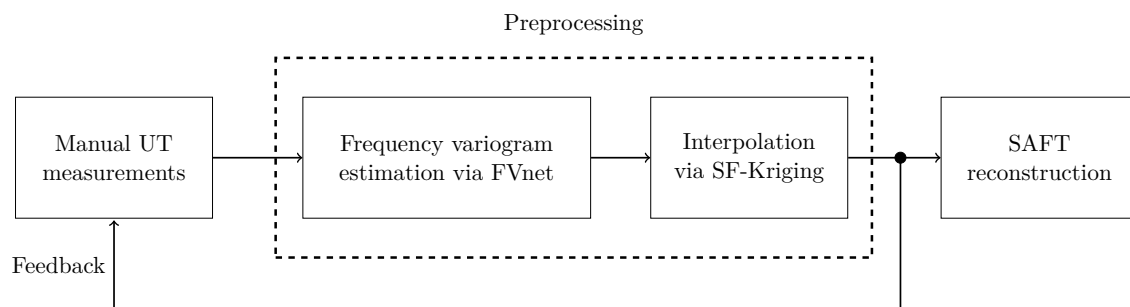


Figure 5.1.: Illustration of the entire system architecture including the proposed preprocessing scheme



## 6. Space-Frequency Domain Block Interpolation of UT Data

### 6.1. Complex Spatial Random Field Modeling of UT Data

For making predictions in a geostatistical manner, SRF based data modeling is desirable. As we aim to perform prediction in space-frequency domain, we consider the Fourier coefficients of UT data as a set of observations obtained in a Complex Spatial Random Field (CSRF).

Let us first consider the frequency response of a single A-Scan. We regard an A-Scan at the position  $\mathbf{s}_i$  as a superposition of the delayed version of the inserted pulse. In accordance with pulse modeling in (2.1), the frequency response of a Gabor model is then expressed as

$$\mathcal{F}_t\{h(t)\} := H(\omega) = \frac{1}{4\sqrt{\pi\alpha}} \cdot e^{\left(-\frac{(\omega-\omega_C)^2}{4\alpha} + j\phi\right)}, \quad (6.1)$$

with  $\omega_C = 2\pi f_C$ . Since Fourier transform is a linear operator, the frequency response of (2.3) results in

$$\mathcal{F}_t\{a(\mathbf{s}_i, t)\} := P(\mathbf{s}_i, \omega) = H(\omega) T(\mathbf{s}_i, \omega) + N(\mathbf{s}_i, \omega), \quad (6.2)$$

where

$$T(\mathbf{s}_i, \omega) = \sum_{j=1}^J e^{-j\omega\tau(\mathbf{s}_i, \mathbf{r}_j)}. \quad (6.3)$$

The quantity  $N(\mathbf{s}_i, \omega)$  represents the frequency response of the measurement noise modeled as a zero-mean complex Gaussian process over space, which may or may not be spatially correlated with finite variance  $\sigma_n^2$ .

The implication of (6.2) is following. For a fixed frequency  $\omega_m$ ,  $H(\omega_m)$  is a known complex constant, whereas  $T(\mathbf{s}_i, \omega_m)$  and  $N(\mathbf{s}_i, \omega_m)$  are unknown and vary with scan positions. Spatial variation of these two differs too. While  $T(\mathbf{s}_i, \omega_m)$  is deterministic and exhibits a distinctive yet unknown structure over space,  $N(\mathbf{s}_i, \omega_m)$  is stochastic but easy to be modeled statistically.

Let us now formulate the Fourier coefficients of  $\omega_m$  as a CSRF as

$$\{P_m(\mathbf{s}) | \mathbf{s} \in D\}, \quad (6.4)$$

with  $D$  representing our ROI. Based on the aforementioned domain knowledge, this CSRF can be decomposed into

$$P_m(\mathbf{s}) = \mu_m + T_m(\mathbf{s}) + \varepsilon_m(\mathbf{s}). \quad (6.5)$$

Here,  $\mu_m \in \mathbb{C}$  is a constant mean, and  $T_m(\mathbf{s})$  and  $\varepsilon_m(\mathbf{s})$  are both zero-mean CSRFs representing the macro structure and the micro fluctuation over  $D$ , respectively.

Interesting to consider is its spatial incremental process, which is defined as

$$\begin{aligned} Q_m^{[h]}(\mathbf{s}) &= P_m(\mathbf{s}) - P_m(\mathbf{s} + \mathbf{h}) \\ &= T_m(\mathbf{s}) - T_m(\mathbf{s} + \mathbf{h}) + \delta_m(\mathbf{s}), \end{aligned} \quad (6.6)$$

where  $\delta_m(\mathbf{s})$  represents the Gaussian noise component, which has much smaller variance than  $\varepsilon_m(\mathbf{s})$ . Since  $T_m(\mathbf{s}_i)$  has a finite range in magnitude, its increment can also be regarded as stationary. Hence, without loss of generality,  $P_m(\mathbf{s})$  can be considered to be intrinsic stationary.

## 6.2. Space-Frequency Block Interpolation

Analogous to the notations in Section 2.3 and 4.8, the frequency response of the A-Scan at the position  $\mathbf{s}_i$  is denoted as

$$\mathbf{p}_i = \mathbf{F}_M \mathbf{a}_i \in \mathbb{C}^M. \quad (6.7)$$

The frequency responses of the observed A-Scans can be concatenated into a single matrix  $\mathbf{P}_{N_S} \in \mathbb{C}^{M \times N_S}$  as

$$\mathbf{P}_{N_S} = [\mathbf{p}_0 \quad \mathbf{p}_1 \quad \dots \quad \mathbf{p}_{N_S}], \quad (6.8)$$

which is a representation of A-Scans in space-frequency domain. Alternatively,  $\mathbf{P}_{N_S}$  can be described in terms of its frequency components as

$$\mathbf{P}_{N_S} = \begin{bmatrix} \boldsymbol{\pi}_0^T \\ \boldsymbol{\pi}_1^T \\ \vdots \\ \boldsymbol{\pi}_{M-1}^T \end{bmatrix}, \quad (6.9)$$

where  $\boldsymbol{\pi}_m^T \in \mathbb{C}^{N_S}$  is collection of Fourier coefficients of the sampling positions for the frequency  $\omega_m$ .

Our objective here is to make a prediction of the missing A-Scans within the same ROI in space-frequency domain. This means that their frequency responses, instead of the A-Scans themselves, are to be estimated, which are then back transformed into time domain afterwards. The prediction positions are represented as a set  $S^C = \{\mathbf{s}_1^{\sim}, \mathbf{s}_2^{\sim}, \dots, \mathbf{s}_K^{\sim}\}$  with  $K = L - N_S$  and their space-frequency spectral are

$$\mathbf{P}_{N_S}^C = [\mathbf{p}_1^{\sim} \quad \mathbf{p}_2^{\sim} \quad \dots \quad \mathbf{p}_K^{\sim}] = \begin{bmatrix} (\boldsymbol{\pi}_0^C)^T \\ (\boldsymbol{\pi}_1^C)^T \\ \vdots \\ (\boldsymbol{\pi}_{M-1}^C)^T \end{bmatrix} \in \mathbb{C}^{M \times K}. \quad (6.10)$$

Furthermore, as discussed in Section 4.11, the effect of the farther samples should be suppressed to assure the prediction accuracy. For this purpose, the maximal lag  $h_{\max}$  is set to limit the neighboring region to be considered for each prediction.

As described in Section 4.8, the linear Minimum Mean Squared Error (MMSE) prediction of the frequency responses for  $\omega_m$  can be obtained through

$$\hat{\boldsymbol{\pi}}_m^C = \mathbf{W}_m^T \boldsymbol{\pi}_m, \quad (6.11)$$

where the weights  $\mathbf{W}_m \in \mathbb{R}^{N_S \times K}$  are calculated as

$$\mathbf{W}(\omega_m) = \boldsymbol{\Gamma}_m^{-1} \boldsymbol{\Upsilon}_m. \quad (6.12)$$

The matrix  $\boldsymbol{\Gamma}_m \in \mathbb{R}^{N_S \times N_S}$  is the frequency variogram matrix of the samples defined in (4.61), whereas  $\boldsymbol{\Upsilon}_m \in \mathbb{R}^{N_S \times K}$  is the frequency variogram matrix of the prediction positions which is the concatenation of (4.62) as

$$\boldsymbol{\Upsilon}_m = \begin{bmatrix} \mathbf{v}_m(\mathbf{s}_1) & \mathbf{v}_m(\mathbf{s}_2) & \dots & \mathbf{v}_m(\mathbf{s}_{\tilde{K}}) \end{bmatrix}. \quad (6.13)$$

As discussed in Section 4.11, including all sampling positions may become problematic and lead to inaccurate predictions. For this reason, the neighboring range for each prediction point should be limited by a certain maximal lag  $h_{\max}$ . This means that the weights for the samples outside the neighboring range should be set to 0. Yet, the neighboring range varies for each prediction, requiring an individual adjustment. This can be done using the following functions to determine each element of  $\boldsymbol{\Gamma}_m$  and  $\boldsymbol{\Upsilon}_m$

$$g_{\mathbf{r}}(h_i, h_{\max}) = \begin{cases} \gamma_{\mathbf{P}}^{[\omega_m]}(h_i) & \text{for } \|\mathbf{h}_{ij}\|_2 \leq h_{\max} \\ \max(\gamma_{\mathbf{P}}^{[\omega_m]}(h)) & \text{otherwise,} \end{cases} \quad (6.14)$$

$$g_{\mathbf{v}}(h_i, h_{\max}) = \begin{cases} \gamma_{\mathbf{P}}^{[\omega_m]}(h_i) & \text{for } \|\mathbf{h}_{0i}\|_2 \leq h_{\max} \\ 0 & \text{otherwise.} \end{cases} \quad (6.15)$$

Here,  $\gamma_{\mathbf{P}}^{[\omega_m]}(h)$  denotes the frequency variogram of the considered CSRF  $P_m(\mathbf{s})$ . Consequently,  $\boldsymbol{\Gamma}_m$  and each column of  $\boldsymbol{\Upsilon}_m$  become

$$\boldsymbol{\Gamma}_m = \begin{bmatrix} g_{\mathbf{r}}(0, h_{\max}) & g_{\mathbf{r}}(h_{21}, h_{\max}) & \dots & g_{\mathbf{r}}(h_{N_S 1}, h_{\max}) \\ g_{\mathbf{r}}(h_{12}, h_{\max}) & g_{\mathbf{r}}(0, h_{\max}) & \dots & g_{\mathbf{r}}(h_{N_S 2}, h_{\max}) \\ \vdots & \vdots & \ddots & \vdots \\ g_{\mathbf{r}}(h_{1 N_S}, h_{\max}) & g_{\mathbf{r}}(h_{2 N_S}, h_{\max}) & \dots & g_{\mathbf{r}}(0, h_{\max}) \end{bmatrix} \quad (6.16)$$

$$\mathbf{v}_m(\mathbf{s}_i) = \begin{bmatrix} g_{\mathbf{v}}(h_{i1}, h_{\max}) \\ g_{\mathbf{v}}(h_{i2}, h_{\max}) \\ \vdots \\ g_{\mathbf{v}}(h_{i N_S}, h_{\max}) \end{bmatrix}. \quad (6.17)$$

To elaborate the implication of the functions (6.14) and (6.15), let us consider the following scenario. Two of the available samples are obtained at  $\mathbf{s}_i$  and  $\mathbf{s}_j$  which are distanced away with  $h_{ij} > h_{\max}$ . A prediction should be made for the position  $\mathbf{s}_k$ , yet  $\mathbf{s}_i$  is outside its neighboring range. To ensure the weights for  $\mathbf{s}_i$  to be 0, the corresponding element in the matrix  $\mathbf{\Upsilon}_m$  should be set to 0, which is described in (6.15). On the other hand, since the matrix  $\mathbf{\Gamma}_m$  should remain the same dimension and is inverted for weights calculation, it should be handled more carefully. From the fundamental assumption of spatial statistics, the values nearby should be similar than the values farther apart, indicating the spatial correlation of  $\mathbf{s}_i$  and  $\mathbf{s}_j$  is to be weaker than a point which is within the radius of  $h_{\max}$  from  $\mathbf{s}_i$ . Since a larger frequency variogram values means a weaker spatial correlation, the corresponding elements in  $\mathbf{\Gamma}_m$  can be set to the maximal value of the given frequency variogram to minimize their effects on the predictions in (6.14).

Yet, there is another concern in this way of weight calculation. While the number of the possible lags is limited as our observations are lattice data, the size of  $\mathbf{\Gamma}_m$  in (6.12) increases with  $N_S$ . This may result in ill-determined rank in  $\mathbf{\Gamma}_m$ , i.e., its singular values decay gradually to zero [50]. As a result, its condition number increases significantly with number of samples  $N_S$ , yielding an unstable solution for (6.12). In general, rank reduction alone is not sufficient for an inverse problem with ill-determined rank, since there does not exist an obvious gap in the singular values which could be used to determine the threshold for rank reduction [50]. For this reason, some kind of regularization is often required to attain a stable solution.

One approach for such problem is Modified Truncated Singular Value Decomposition (MTSVD), which seeks to solve a Tikhonov regularized least squares problem with the low-rank approximated system matrix [51]. Consequently, a stable solution for (6.12) can be obtained by replacing  $\mathbf{\Gamma}_m^{-1}$  with its  $k$ -rank MTSVD inverse  $(\mathbf{\Gamma}_m^{[k]})^\# \in \mathbb{R}^{N_S \times N_S}$  as

$$(\mathbf{\Gamma}_m^{[k]})^\# = \left[ (\mathbf{\Gamma}_m^{[k]})^T \mathbf{\Gamma}_m^{[k]} + \lambda \mathbf{I}_{N_S} \right]^{-1} (\mathbf{\Gamma}_m^{[k]})^T, \quad (6.18)$$

where  $\mathbf{\Gamma}_m^{[k]} \in \mathbb{R}^{N_S \times N_S}$  represents the  $k$ -rank approximation of  $\mathbf{\Gamma}_m$ , while  $\lambda > 0$  is the regularization parameter to control the influence of the regularization on the solution.

## 7. Data Driven FV Estimation

### 7.1. Network Inputs and Output

In order to determine the inputs and the outputs of a DNN, it is imperative to understand the properties of the datasets. For our UT scenario, there are three important properties. First, since our measurement system registers the measurements at a given measurement coordinate, samples are only available on the grid points. In other words, our data are lattice data. Second, we are interested in computing predictions batch-wise. This means that the available lags in a single batch are limited with  $N_h$  and fixed, allowing us to express all the lags as a vector. As a consequence, we can obtain the desired distribution of the available lags *a priori*, which can be compared to the actual distribution of the available data, providing an insight regarding the spatial statistics of the data.

Moreover, since now the lags can be represented as a vector of finite size, we only need to know the frequency variogram values for these particular lags. In other words, there is no need to estimate a set of continuous functions for frequency variograms, which requires to estimate the model order of the modified Bessel function. This opens up the possibility to estimate frequency variograms non-parametrically, which goes well with the use of DNN. There is yet a problem with such non-parametric estimates, i.e experimental frequency variograms. Although the data are discrete, the underlying process is continuous over space, thus the resulting frequency variogram should be continuous over the lags as well. As a solution, we propose to smooth those estimates with respect to lags. In this way, we can obtain a smooth frequency variogram, while keeping its computation tractable.

The last aspect to consider is the frequency response itself. Although there is no correlation in Fourier coefficients of two different frequencies, each of them exhibits some kind of spatial statistical structure. This indicates that there should be similarities in the structure of their frequency variograms, not in their values. Figure 7.1 illustrates this notion, showing that while their amplitude may differ significantly, the spatial structures of frequency variograms are indeed quite similar once they are normalized. The structures shown here can be interpreted as follows. For the frequencies within the signal range, their frequency variogram values change slowly over the lags, demonstrating strong spatial correlations. On the other hand, the frequency variogram values of the noise components change abruptly, indicating there is only little spatial correlation. Such characteristics in structure is unique and can be learned by a Deep Neural Networks (DNN). However, it is highly likely that the amplitude cannot be estimated without further knowledge about the underlying process.

Based on these considerations, the following quantities are selected for the network

inputs and the output. We set the goal of the network to estimate the structure of the frequency variogram for a single frequency,  $\omega_m$ . For this purpose, we chose two different inputs: (1) distribution of the available lags within the sampling positions and (2) the smoothed experimental frequency variograms of  $\omega_m$  and its two neighboring bins, i.e.  $\omega_{m-1}$  and  $\omega_{m+1}$ . Both inputs are represented as vectors of size  $N_h$ . As the output, the network should generate a vector of the same size  $N_h$ , representing the estimate of the frequency variogram structure for  $\omega_m$ . This can be seen as a vector-valued nonlinear regression problem, which has been proven to be a suitable task for DNN [29].

The advantage of this selection is its flexibility. Since the network should perform an estimate only for a single frequency bin, it remains unaffected even if measurement parameters, such as central or sampling frequency, are changed. This also enables the same network to be used for other frequencies, allowing parallel computing. Moreover, as the lags are not required for estimation, it can be used for scenarios with different spatial spacing, as long as the total number of lags remains the same as that of the training.

## 7.2. Network Architecture

In this section, the selected network architecture is described, which is illustrated in Figure 7.2. In general, data should be preprocessed before being fed into a network. In our case, there are two inputs which require different processing. The first one, denoted as  $\Phi_{\text{hist}}(\cdot) : \mathbb{R}^{N_h} \mapsto \mathbb{R}^{N_h}$ , is a transformation of sampling positions into the histogram of their available lags relative to the desired one.

The other one,  $\Phi_{\text{EFV}}(\cdot) : \mathbb{C}^M \mapsto \mathbb{R}^{N_h}$ , is to transform Fourier coefficients of a single frequency bin into a smoothed experimental frequency variogram. Here, the frequency responses of two neighboring bins are transformed as well to utilize the similarity in structures, suggesting in total of three frequency responses,  $\Pi_m = [\pi_{m-1} \pi_m \pi_{m+1}] \in \mathbb{C}^{K \times 3}$ , are used for frequency variogram estimation.

There are two operations within  $\Phi_{\text{EFV}}(\cdot)$ : (1) compute method-of-moments estimate for each lag value and (2) smooth this estimate with respect to lags via polynomial curve fit and normalize with its largest values. The first operation corresponds to (4.65) to ensure that the network output accounts for the spatial statistics of the available data. The second operation, on the other hand, is introduced rather empirically. Since there may be missing lags within the available sampling positions, leaving those points empty led to unsuccessful training. We suppose this is because such empty inputs fail to limit the degree of freedom in the training space, making the problem very difficult for the network to learn. Hence, we decided to provide values for those missing points, which are obtained through smoothing.

After these transformations, the two different inputs, i.e.  $\Phi_{\text{hist}}(S_N) \in \mathbb{R}^{N_h}$  and  $\Phi_{\text{EFV}}(\cdot) \in \mathbb{R}^{N_h}$ , are scaled and concatenated to be fed into the network. This yields three vectors  $\mathbf{x}_1^{(0)}$ ,  $\mathbf{x}_2^{(0)}$  and  $\mathbf{x}_3^{(0)}$  of size  $\mathbb{R}^{2N_h}$  as the network inputs, each of which contains the histogram and the experimental frequency variogram of each frequency bin. In the network, the inputs are first fed into a one dimensional CNN. This process can be

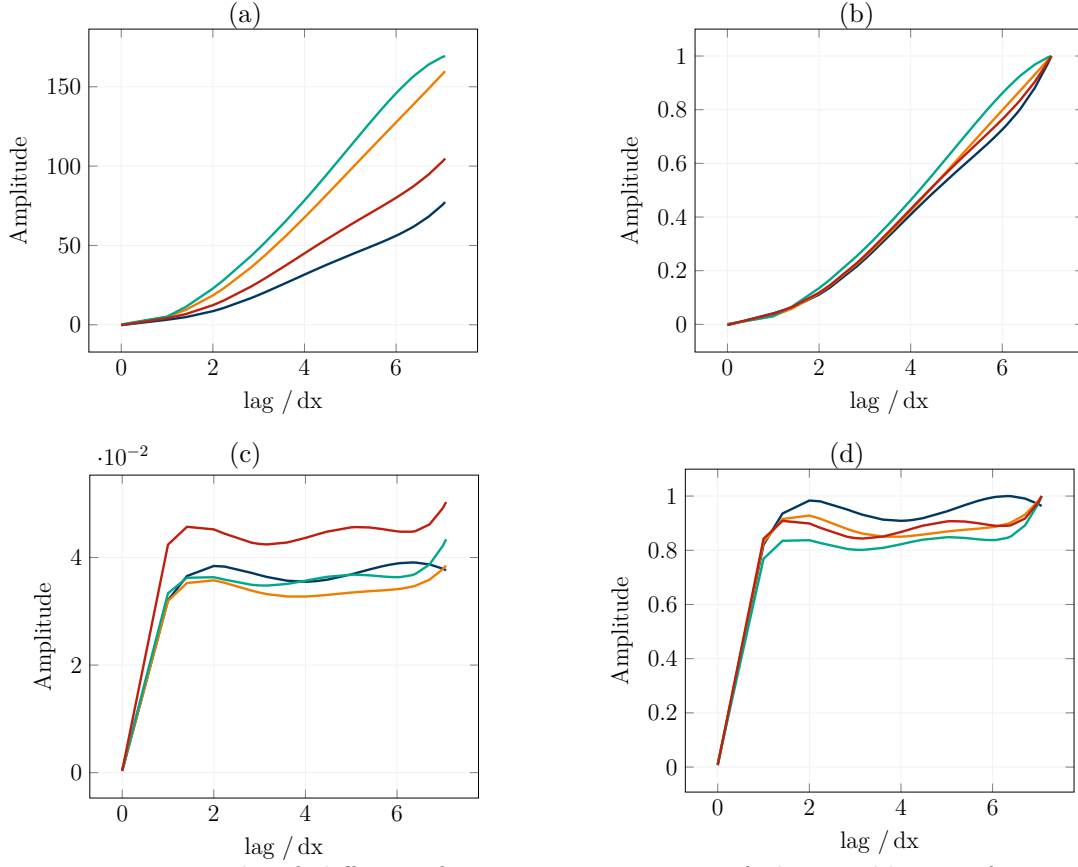


Figure 7.1.: Example of different frequency variograms of the neighboring frequency bins: frequency variograms on the left-hand side are with its true amplitude whereas those on the right-hand side are normalized with the maximal value of each frequency. Figure 7.1a and 7.1b are for the frequency components within the signal frequency range, which exhibits the strong spatial correlation. On the other hand, Figure 7.1c and 7.1d are for the frequencies outside the significant range, where the spatial correlation is small.

seen as a similar operation to weighted window averaging in (4.65) to incorporate the information of the neighboring bins. This follows  $L - 1$  layers of MLP, which in the end produces the output of size  $\mathbb{R}^{N_h}$ . Finally the obtained outputs are scaled back with the original largest values, which is the estimate of the frequency variogram for  $\omega_m$  denoted as  $\hat{\gamma}_P^{[\omega_m]} \in \mathbb{R}^{N_h}$ .

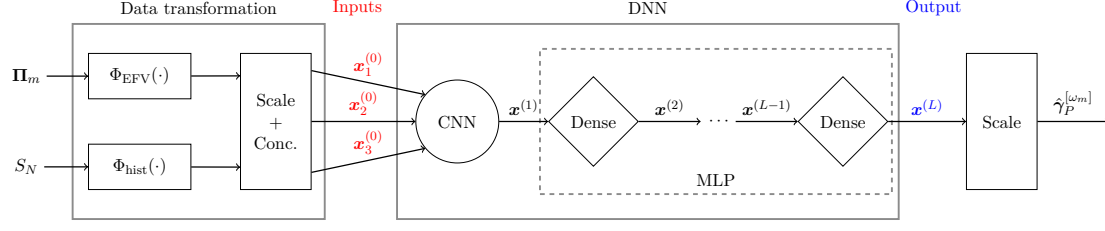


Figure 7.2.: Illustration of the network architecture for estimating frequency variogram of a single frequency (FVnet)

### 7.3. Training Scheme

Training of the network was conducted using sets of synthetic UT data, which are based on the forward modeling (2.8) with parameter values listed in Table 7.1 . Each synthetic data has a dimension of  $N_x \times N_y \times N_z = 30 \times 30 \times 512$ , within which at least one scatterer resides. To diversify the datasets, we varied the values of some forward model parameters, which are summarized in Table 7.2. From these synthetic datasets, a block of size  $N_x \times N_y \times N_z = 10 \times 10 \times 512$  is selected and subsampled. This subsampled data is then transformed into smoothed experimental frequency variograms, which serve as training inputs. As the training output, we selected the corresponding smoothed experimental frequency variogram which is obtained from fully sampled data.

	Parameter	Value/range
Specimen / Measurements	Material	Steel
	Speed of sound $c_0$	5900 m/s
	Dimension (L $\times$ D $\times$ H)	45 mm $\times$ 45 mm $\times$ 90 mm
	Spacing along $x$ -axis dx	0.5 mm
	Spacing along $y$ -axis dy	0.5 mm
	Sampling frequency $f_S$	80 MHz
	Temporal interval dt	12.5 ns
	Spacing along $z$ -axis dz	36.875 $\mu$ m
UT pulse	Transducer beam spread	25°
	Pulse model	Gabor
	SNR	20 dB

Table 7.1.: Parameter values for forward modeling, fixed for all datasets



	Parameter	Value/range
Synthetic data	ROI (I) ( $30 \times 30 \times 512$ )	$x = 10 \dots 35$ mm $y = 10 \dots 35$ mm $z = 24 \dots 43$ mm
	ROI (II) ( $30 \times 30 \times 512$ )	$x = 10 \dots 35$ mm $y = 10 \dots 35$ mm $z = 33 \dots 52$ mm
Scatterer configuration	Number of scatterers in ROI	[2, 5, 10]
	Reflectivity of each scatterer	0.1 ... 1
UT pulse	Center frequency $f_C$	$3.36 \pm 0.34$ MHz
	Bandwidth (−6 dB)	$1.0 \pm 0.1$ MHz
	Wavelength $\lambda$	$\approx 1.756$ mm
Spatial subsampling	Block size ( $N_x \times N_y \times N_z$ )	$10 \times 10 \times 512$
	Number of scans / block	10 ... 100
	Scan distribution	Uniform or random walk

Table 7.2.: Parameter values for forward modeling, varied with each dataset

	Parameter	Value/range
Data transformation	Block size ( $N_x \times N_y \times N_z$ )	$10 \times 10 \times 512$
	Maximal lag $h_{\max}$	$5\sqrt{2} \text{ dx} \approx 7.071 \text{ dx}$
	Number of lags $N_h$	25
	Polynomial order for smoothing	5
	Scaling for $\Phi_{\text{hist}}(\cdot)$	20
DNN training	Library	tensorflow
	Optimizer	Adam
	Loss function	Mean squared error
	Validation metrics	Mean squared error
	Batch size	128
	Number of training data	720000
	Number of validation data	180000
DNN model	Activation function	Rectified Linear Unit
	Kernel regularization with factor	L1 regularizer $10^{-8}$
	CNN layer	1
	neurons	32
	filter size	3
	MLP layer	4
	neurons	32
	Total parameters to be trained	9,881

Table 7.3.: Parameter values related to training of FVnet

## 8. Summary of Part II

In Part II, the proposed data driven hybrid preprocessing scheme is presented. For its development, we seek to establish an interpolation scheme which is tractable in computation and at the same time keeps the interpretability of the entire process. This is achieved through three factors. One is reducing the dimensionality of the data by computing the predictions in space-frequency domain. Another one is to perform prediction batch-wise instead of the entire ROI at once. The last one is to decompose the interpolation task into two subtasks, one of which is performed via DNN. In this way, the prediction is based on the geostatistical approach, SF-Kriging, while we estimate the spatial statistics of the frequency responses using DNN, benefitting from its fast execution. This allows us to understand the role DNN plays in our interpolation scheme, avoiding to rely solely on its black-box nature.

For this purpose, UT data are first modeled as a CSRF, which demonstrates that spatial intrinsic stationarity can hold true for UT data as well. This follows the formulation of block SF-Kriging, where multiple time series data are predicted at the same time. Here, a possible solution is proposed for keeping the neighboring range small to avoid the destructive impacts from the samples farther away.

To determine the network architecture and its training scheme, the property of our UT data are carefully examined. As our UT data is lattice data, the possible lags are limited within a batch, allowing us to estimate only a vector-valued frequency variograms. This enables us to formulate the estimation problem of frequency variograms as a vector-valued regression problem, which is well-known to be a suitable task for DNN. Based on these considerations, the network architecture as well as training scheme are described.

## Part III.

# Reconstruction of Irregularly Subsampled Ultrasonic Testing Data

In order to evaluate the proposed preprocessing scheme, two sets of simulations are conducted. With these simulations, we aim to examine (1) the estimation capability and accuracy of the trained FVnet and (2) effectiveness of incorporating a physically-aware preprocessing scheme into the measurement system, 3D SmartInspect. Along with the performance results of the proposed method, its limitations are discussed to determine where further improvements are possible.

## 9. Simulation Setting

### 9.1. Specimen and Its Raw Measurement Data

For conducting simulations, a set of automatic UT measurement data is employed. This data set is obtained from a steel object, denoted as the MUSE specimen in this work. Although the proposed method is intended to be incorporated into manual measurement systems, for this work we simulated manual measurements by subsampling the fully sampled MUSE data, which serve as a ground truth to compare with the obtained simulation data. This ground truth data is sampled with equal distance of 0.5 mm over the entire surface of the specimen. The geometry of the specimen is illustrated in Figure 9.1 and the specific values of the relevant parameters are summarized in Table 9.1.

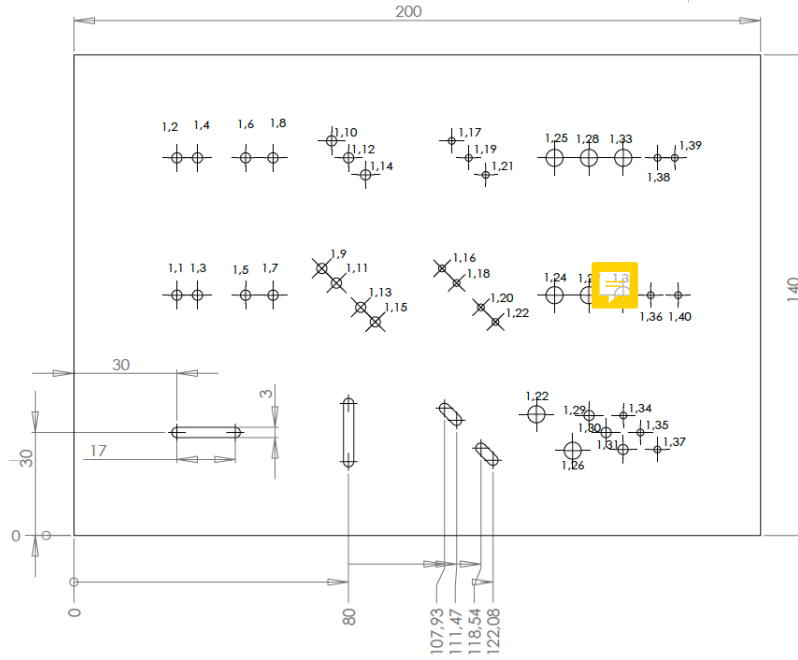


Figure 9.1.: Illustration of the geometry of MUSE specimen

### 9.2. Interpolation Parameters

For interpolation, we fixed the batch size and the maximal lag which limits the neighboring region for predictions at each prediction point. The batch size is selected to be

	Parameter	Value/range
Specimen	Material	Steel
	Speed of sound $c_0$	5900 m/s
	Dimension ( $L \times D \times H$ )	200 mm $\times$ 140 mm $\times$ 90 mm
UT Probe	Manufacturer	KARL DEUTSCH
	Model	P 1462.1
	Diameter (transducer element)	10 mm
	Center frequency $f_C$	$4.0 \pm 0.4$ MHz
	Bandwidth ( $-6$ dB)	$2.6 \pm 0.5$ MHz
Measurements	Wavelength $\lambda$	$\approx 1.475$ mm
	Spacing along $x$ -axis $dx$	0.5 mm
	Spacing along $y$ -axis $dy$	0.5 mm
	Sampling frequency $f_S$	80 MHz
	Temporal interval $dt$	12.5 ns
	Spacing along $z$ -axis $dz$	36.875 $\mu$ m

Table 9.1.: Parameter values for MUSE dataset

small to ensure the spatial intrinsic stationarity within the batch. In addition to those parameters, we set two other parameters for SF-Kriging. One is the minimal number of samples within the neighboring range. This reinforces a certain level of prediction accuracy by preventing to compute predictions for the positions where there are only few samples available. The other one is the regularization factor for SF-Kriging weights defined in (6.18).

	Parameter	Value/range
For all schemes	Batch size ( $N_x \times N_y \times N_z$ )	$10 \times 10 \times 512$
	Maximal lag $h_{\max}$	$5\sqrt{2} dx \approx 7.071 dx$
For SF-Kriging	Minimal number of points	5
	Weights regularization $\lambda$ in (6.18)	1.0

Table 9.2.: Parameter values for interpolation

## 10. Simulation I: Batch Interpolation

In this chapter, the results of the first set of simulations, Simulation I, are presented. In an attempt to demonstrate the performance of the trained FVnet, a randomly selected batch of the MUSE data is subsampled and interpolated by different schemes. The obtained results are compared in terms of their predicted A-Scans .

### 10.1. Simulation Parameters and Scenario

For the simulation of batch interpolation, following parameters and scenario is selected. The dataset is a part of the MUSE data, hence the specimen and its measurement parameters are the same as the ones summarized in Table 9.1. Moreover, we chose the same values for the parameters regarding frequency variograms and SF-Kriging listed in Table 9.2. The ROI of this simulation is limited to a single batch, and the sampling positions are selected randomly with the size of 15 points illustrated in 10.1. This is equivalent to 1.305 samples per squared wavelength.

As interpolation schemes, we compare three different approaches. One is Inverse Distance Weighting (IDW), where weights are calculated only in terms of the space. This means that a constant weight is applied to each A-Scan. Other two perform interpolation using SF-Kriging, yet their frequency variograms are differently estimated. One is computed based on all A-Scans in the batch, representing the ground truth estimate of frequency variograms. The other is estimated by the trained FVnet method only from the available data.

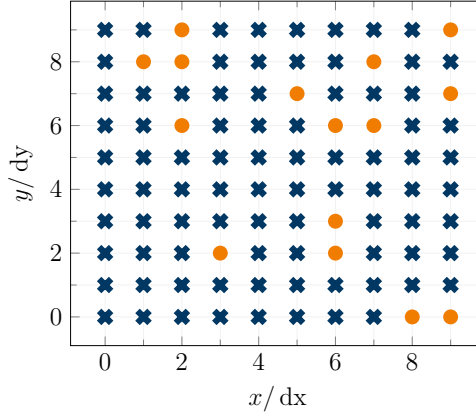


Figure 10.1.: Illustration of the sampling positions (in orange circles) and the prediction positions (in blue crosses).

## 10.2. Interpolation Results and their Evaluation

In this section, the obtained results of Simulation I are presented in Figure 10.2 to 10.5. Overall, IDW performs the worst compared to SF-Kriging. As shown in Figure 10.2, IDW tends to averaging out over the available samples, generating very similar results for two separated prediction positions. This confirms its sensitivity to irregular sampling, which is inevitable for our UT scenario and many other applications.

On the contrary, SF-Kriging exhibits its robustness against the clustering and irregular samples. Even for the position where the variance is high due to the lack of samples, their predictions are well representative compared to that of IDW (Figure 10.2b, 10.2d and 10.2f). This confirms that SF-Kriging is general and not confined to geoscientific data, and can indeed be employed very well with UT data. There are some points where no prediction can be made due to the lack of samples within their neighboring range, which is whitened in Figure 10.4. It is also noteworthy to see the accuracy of its error estimation. Aside from the unpredicted positions, the points with high variance are indeed higher in their actual prediction error. This is a very good indication to be used as a feedback for suggesting the regions to be sampled more intensively.

In terms of frequency variogram estimation, however, there is some obvious deviations compared to the ground truth shown in Figure 10.5. While there is almost no difference between the estimated and the ground truth frequency variogram for the strongest frequency component, FVnet cannot yield an adequate estimate for another frequency which is still within the signal range. This is due to the fact that FVnet only estimates the normalized frequency variograms, which are then scaled based on the sample mean of the second moments. If some subbands are not present in all over the ROI, making the sample mean biased. As a consequence, the results of FVnet based SF-Kriging may deviate more in amplitude than performing SF-Kriging with the perfect field knowledge as seen in Figure 10.3.



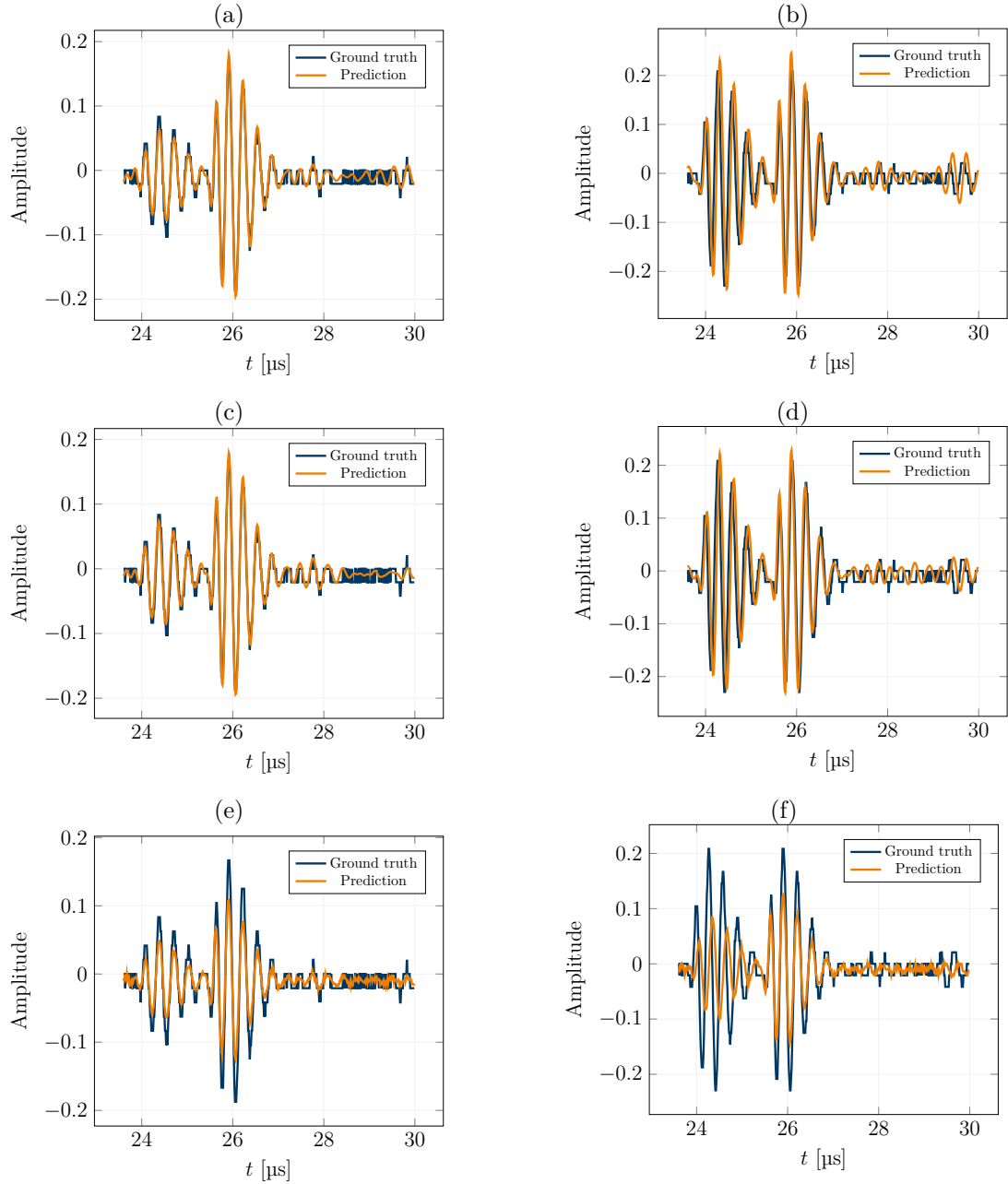


Figure 10.2.: Obtained A-Scan predictions (in orange) for two different points compared to their ground truth (in blue). On the left-hand side, the predictions for  $\mathbf{s}_1 = (5, 5)$  are displayed, whereas on the right-hand side, the predictions for  $\mathbf{s}_2 = (0, 2)$  are shown. The results of different approaches are presented in each row: SF-Kriging with the ground truth frequency variogram (top), SF-Kriging with the FVnet estimated frequency variogram (middle) and IDW (bottom).

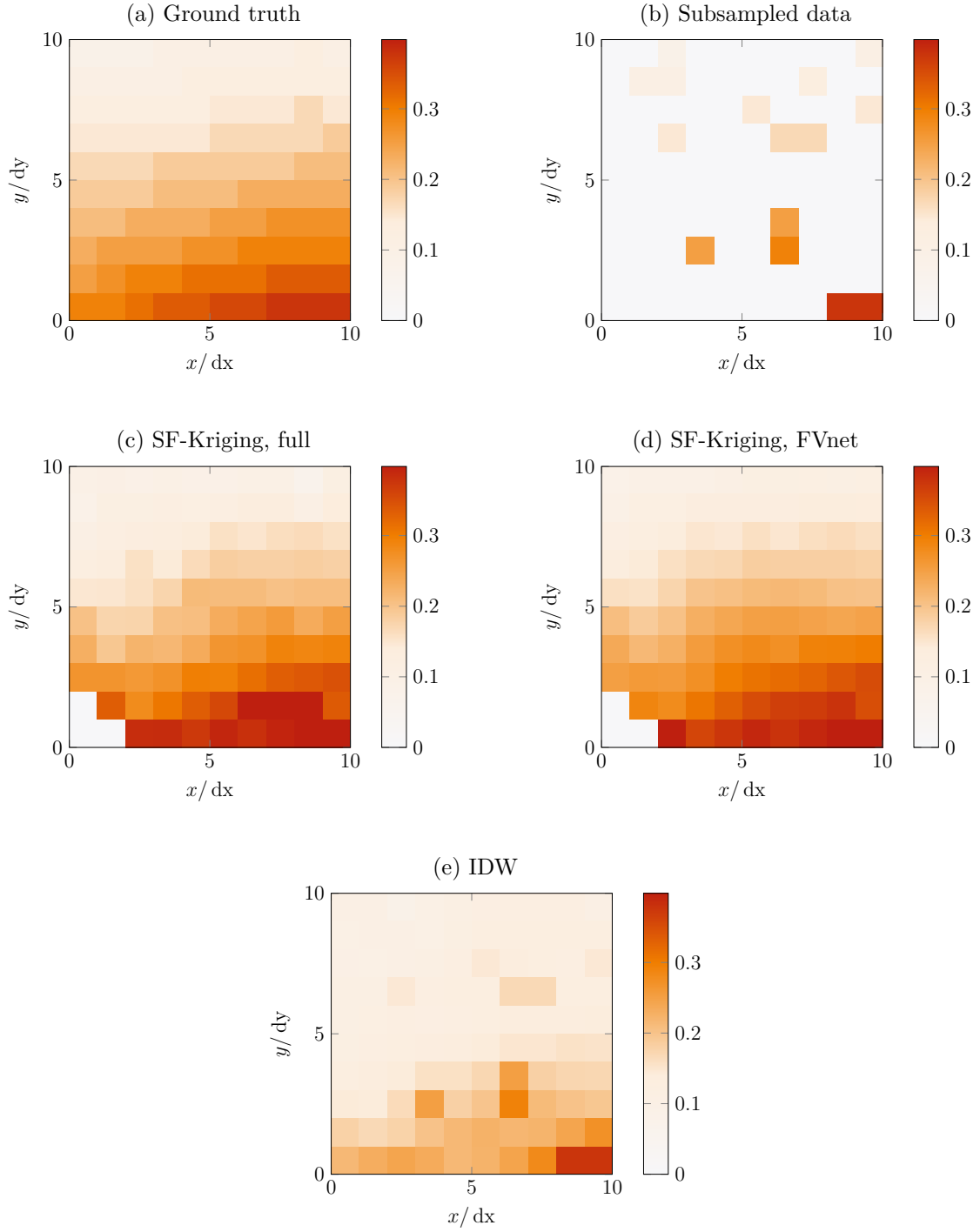


Figure 10.3.: C-Scan of the batch-wise interpolation results: 10.3a actual batch data, 10.3b subsampled data, 10.3c SF-Kriging results based on the ground truth frequency variograms, 10.3d SF-Kriging results based on the FVnet estimated frequency variograms and 10.3e IDW results. Whitened areas in (c) and (d) indicate that there are not enough samples available and thus left unchanged.

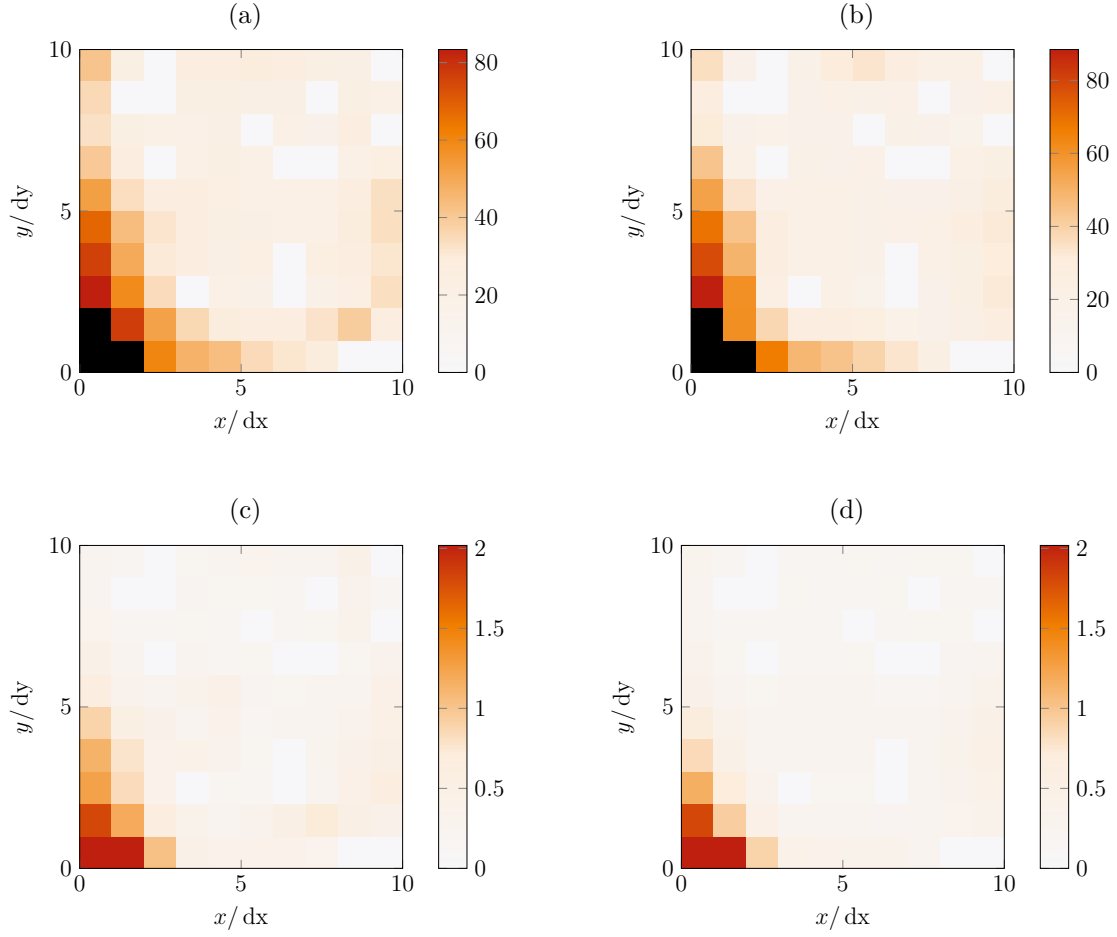


Figure 10.4.: C-Scan representation of SF-Kriging prediction error: Figures on the right-hand side are the ones based on the ground truth frequency variograms, whereas those on left-hand side are based on the FVnet estimates of frequency variograms. The upper two display the estimated prediction error, calculated as the sum of variance, while the bottom two show the actual prediction error for both methods. Black areas in the estimates indicate that there are not enough samples available, thus no prediction is made.

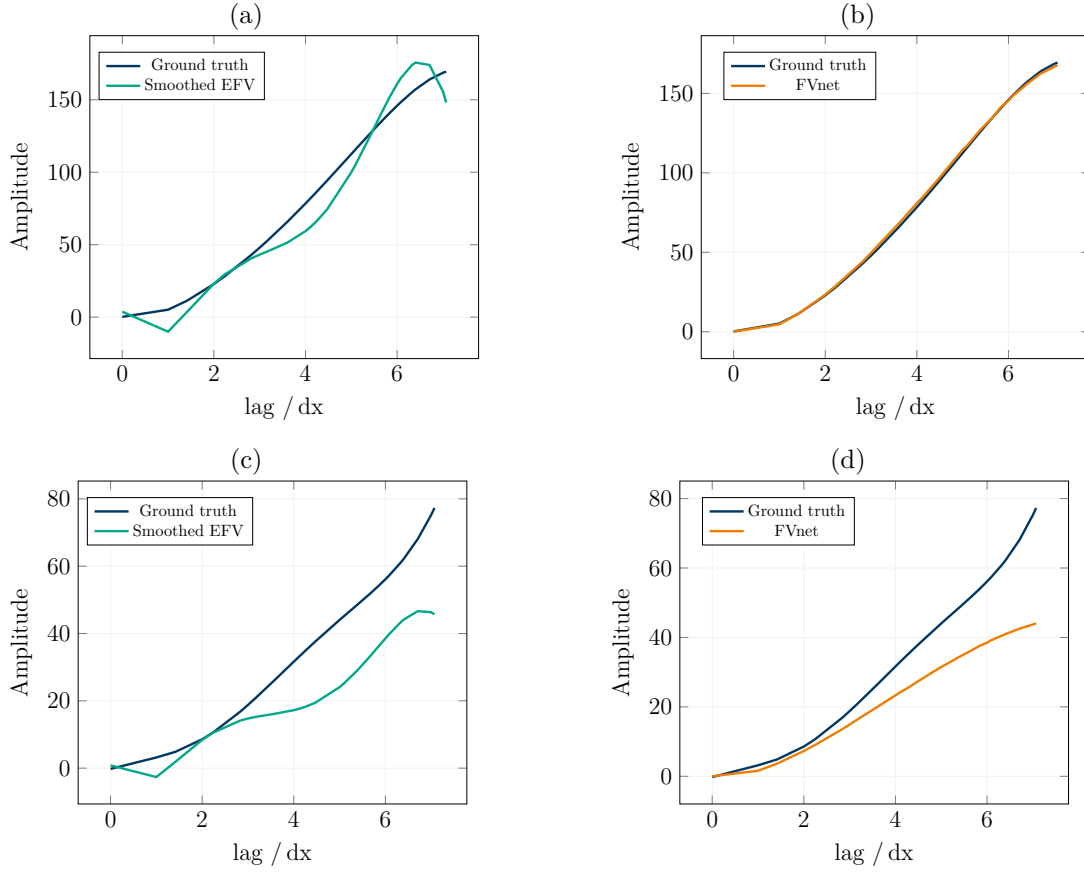


Figure 10.5.: Comparison of frequency variogram estimation with respect to the ground truth (in blue): smoothed method-of-moments estimator (left) and the estimates obtained with the trained FVnet (right), both shown in orange. In each row, frequency variogram of different frequencies are displayed. Here, frequency variogram of two different frequencies,  $f_{21}$  and  $f = f_{19}$ , are displayed.  $f_{21}$  is the strongest frequency component of the signals within the ROI .

# 11. Simulation II: Reconstruction of Subsampled UT Data

In this chapter, the results of the second set of simulations, Simulation II, are presented. The goal of these simulations is to investigate the effectiveness of the proposed preprocessing scheme depicted in Figure 5.1. For this purpose, a part of the MUSE specimen is subsampled and interpolated using different approaches. To perform predictions of missing points while keeping the computation tractable, interpolations are performed in a sliding window manner. This means that the ROI is divided into multiple batches, in each of which batch-wise prediction is performed. These batches are overlapped with the rate of 50% and the results are averaged accordingly. The obtained interpolation results are then reconstructed via SAFT.

In order to evaluate the performance of the proposed scheme, the imaging quality of the reconstruction results are examined. Here, number of samples are varied to investigate how many samples are required to obtain reasonable results. Since the proposed preprocessing scheme can provide estimate of prediction errors, which can be used as feedback to suggest the operator where he or she should take measurements next, this is also of our interest to see how such feedback can affect the resulting reconstructions. Hence, we introduced resampling of the ROI after the initial predictions are performed. These resampled data are then fed into interpolation processes, whose results are reconstructed as well. Those results of resampling are also presented and discussed in this chapter.

## 11.1. Simulation Parameters and Scenario

For Simulation II, the following parameters and scenario are applied. Again, regarding the specimen, frequency variograms and SF-Kriging, their parameter values are summarized in Table 9.1 and 9.2.

The ROI for Simulation II is a part of the MUSE specimen, whose location is depicted in Figure 11.1. Its size is  $N_x \times N_y \times N_z = 50 \times 50 \times 512$ , which is divided into batches with the overlap of 50 %. This results in 81 batches within the ROI in total. The interpolation results are pooled in the measurement system until the entire region is processed, before they are reconstructed. This means that the reconstructions are performed for the entire ROI instead of batch-wise. As interpolation schemes, we again selected IDW and SF-Kriging based on frequency variograms estimated using the trained FVnet approach.

Since the simulations involve random spatial sampling of the ROI, Monte-Carlo simulations are conducted, where for each realization a different set of scan positions is

selected. The number of scan positions  $N_S$  are varied within the range of 5 to 17% of all the grid points in the ROI, which is equivalent to 0.44 to 1.48 samples per squared wavelength. For each coverage, we compute total of 80 realizations, and scan positions are selected randomly, yet focused around the defects where the intensity of the raw measurements is high.

On the other hand, resampling was done based on the interpolation scheme. For interpolation using IDW, the resampling positions are randomly selected as well, since it cannot provide any feedback regarding its prediction accuracy. On the contrary, for interpolation via SF-Kriging, the resampling positions are selected based on the estimates of the prediction errors  $\sigma^2$ , which are the sum of the prediction variance defined in (4.64). Here we introduced a threshold to determine which points are of high variance and thus should be sampled. This threshold was set to  $2\mu_e$ , where  $\mu_e$  denotes the average of the error estimates  $\sigma^2$  within the ROI. The number of resampling positions is set to  $N_S$ , resulting in total of  $2N_S$  A-Scans available after resampling.

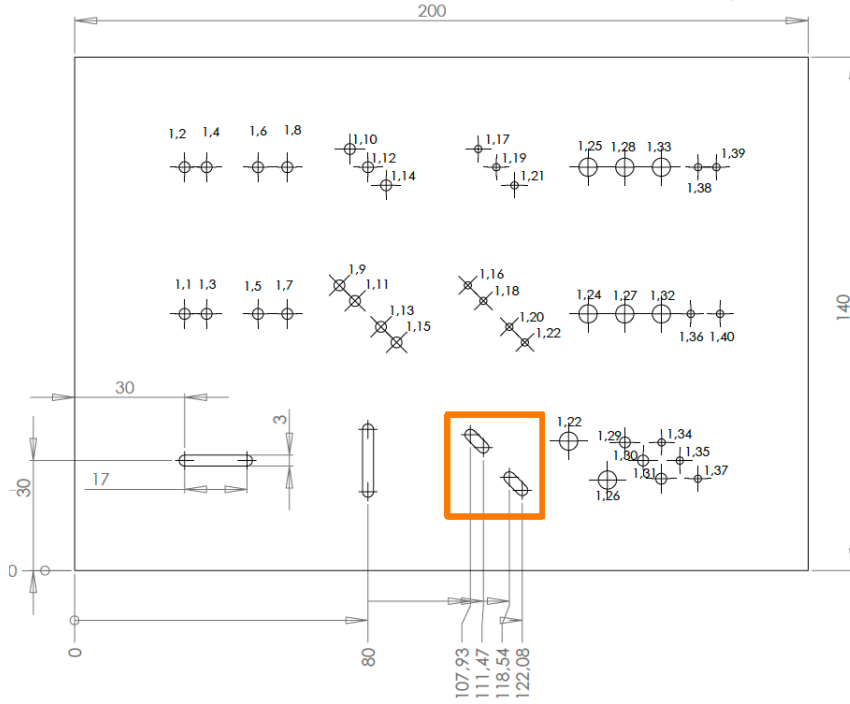


Figure 11.1.: Illustration of the ROI for Simulation II (highlighted by an orange square)

## 11.2. Evaluation Metrics

Since the imaging quality is strongly related to the human perception, providing satisfactory numerical evaluations is challenging. For this reason, in this work we use two different metrics: normalized root squared error  $SE^\dagger$  and *Generalized Contrast-to-Noise*

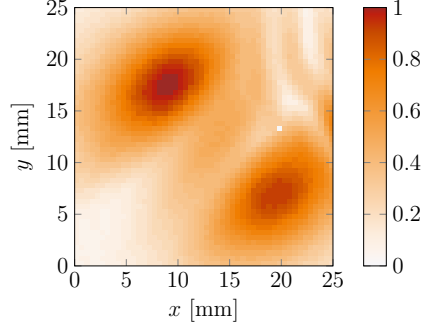


Figure 11.2.: C-Scan representation of the actual data in the ROI selected for Simulation II

*Ratio* (gCNR) [52]. As reference, the ground truth result, which are reconstruction results of the fully sampled measurement data, is selected.

The normalized root squared error  $SE^\dagger$  shows the difference between the obtained reconstruction results  $\hat{\mathbf{B}}$  and the reference  $\mathbf{B}$ . This is calculated with

$$SE^\dagger = \frac{\|\gamma \hat{\mathbf{B}} - \mathbf{B}\|_2}{\|\mathbf{B}\|_2}, \quad (11.1)$$

where  $\gamma$  is a normalization factor which can be obtained from

$$\gamma = \frac{\text{vec}\{\mathbf{B}\}^T \text{vec}\{\hat{\mathbf{B}}\}}{\text{vec}\{\hat{\mathbf{B}}\}^T \text{vec}\{\hat{\mathbf{B}}\}}. \quad (11.2)$$

On the other hand, the gCNR accounts more for human perception of the image contrast by considering the overlap of the probability density functions of inside and outside the target area [52]. For the given target area, the overlap of the probability density functions is calculated with

$$OVL = \int \min\{p_i(x), p_o(x)\} dx, \quad (11.3)$$

where OVL represents the overlap, while  $p_i$  and  $p_o$  are the probability density function of inside and outside the target area, respectively. Based on this, the gCNR is obtained from

$$\text{gCNR} = 1 - OVL, \quad (11.4)$$

revealing that the higher gCNR value indicates the better resolution of the image. Here we defined the target area from the reference where its envelope is above  $-3$  dB of its maximum value.

Since we conducted sets of Monte Carlo simulations, the average value of each data set, denoted as  $MSE^\dagger$  and  $MgCNR$ , is calculated and shown in the following sections.

### 11.3. Interpolation and Reconstruction Results

In this section, a set of exemplary results are presented along with their sampling positions. For each result, the corresponding gCNR values relative to that of the ground truth are listed as well to provide further insights into their imaging quality. The gCNR value of the ground truth is 0.8512.



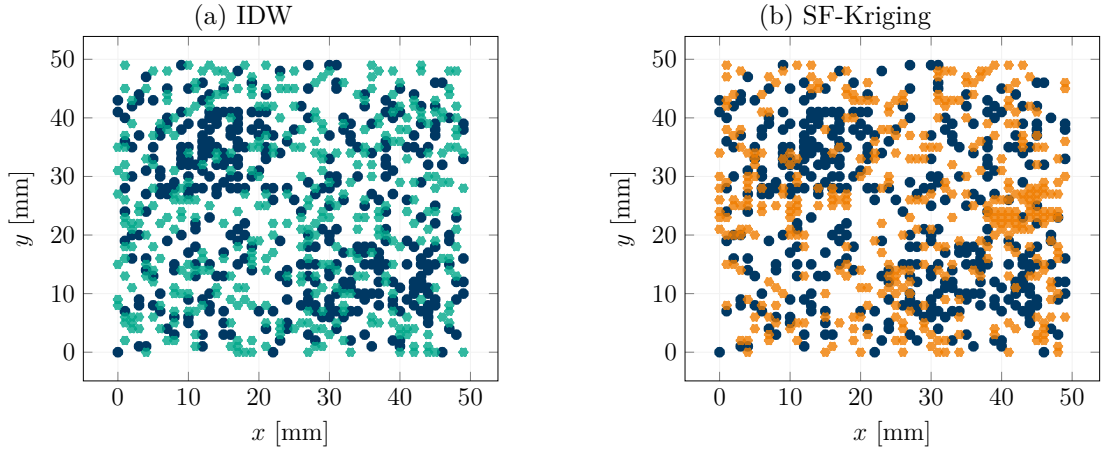


Figure 11.3.: Illustration of sampling (blue) and resampling positions based on interpolation schemes: IDW (left, green) and SF-Kriging with FVnet (right, orange). The initial coverage here is 15% of the all grid points.

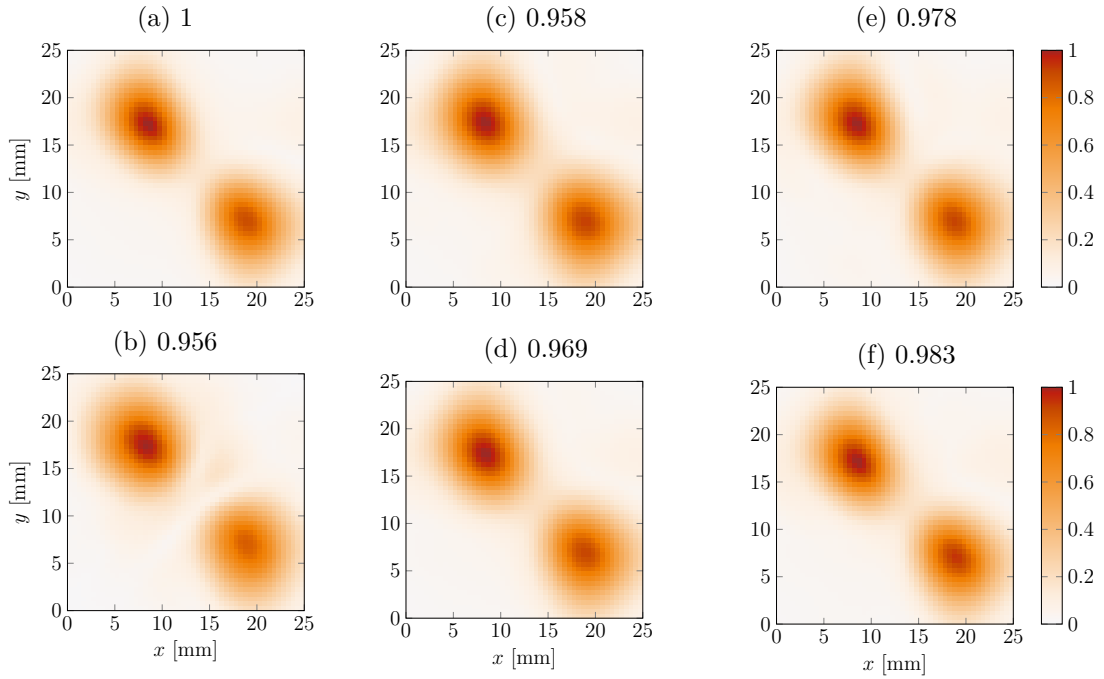


Figure 11.4.: C-Scan of the obtained simulation results (15 % coverage) with their relative gCNR values to that of the ground truth: 11.4a: ground truth, 11.4b: the result without any preprocessing, 11.4c: the IDW result only from the initial sampling positions, 11.4d: the IDW result after resampling, 11.4e: the SF-Kriging result only from the initial sampling positions and 11.4f: the SF-Kriging result after resampling,

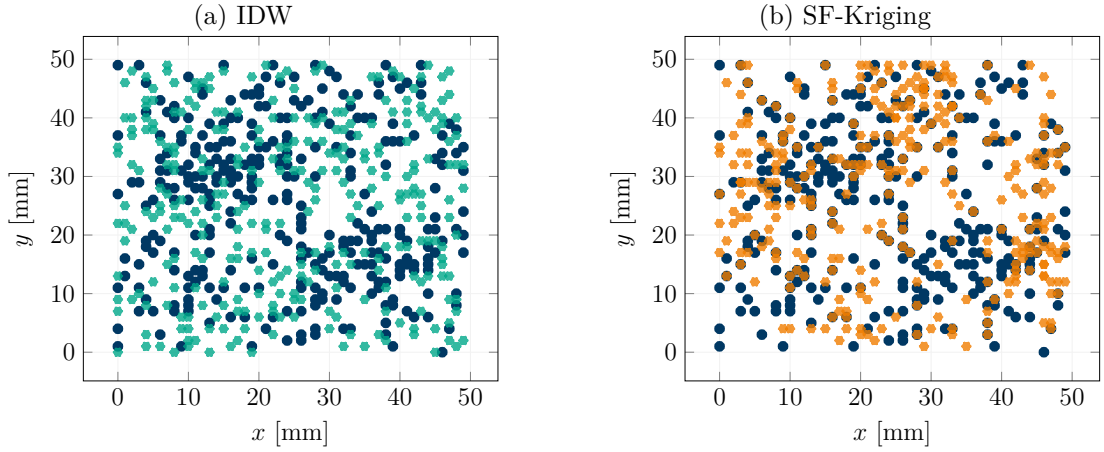


Figure 11.5.: Illustration of sampling (blue) and resampling positions based on interpolation schemes: IDW (left, green) and SF-Kriging with FVnet (right, orange). The initial coverage here is 12% of the all grid points.

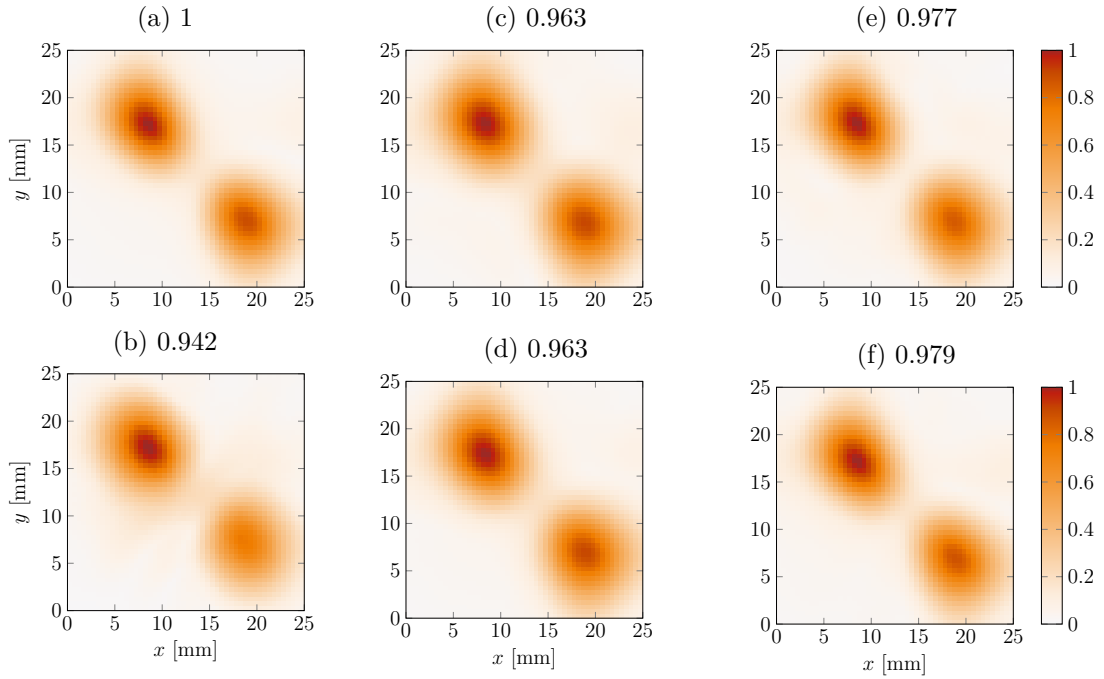


Figure 11.6.: C-Scan of the obtained simulation results (12 % coverage) with their relative gCNR values to that of the ground truth: 11.6a: ground truth, 11.6b: the result without any preprocessing, 11.6c: the IDW result only from the initial sampling positions, 11.6d: the IDW result after resampling, 11.6e: the SF-Kriging result only from the initial sampling positions and 11.6f: the SF-Kriging result after resampling,

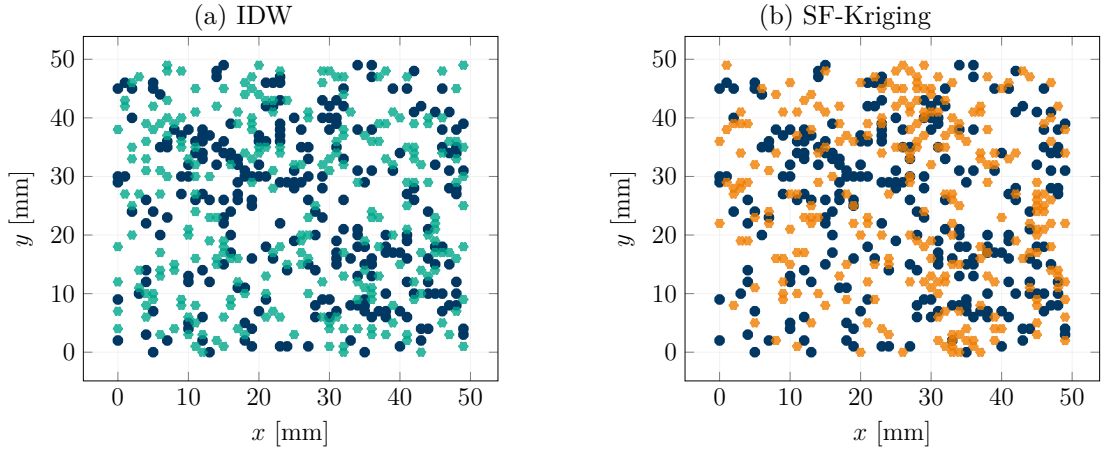


Figure 11.7.: Illustration of sampling (blue) and resampling positions based on interpolation schemes: IDW (left, green) and SF-Kriging with FVnet (right, orange). The initial coverage here is 10% of the all grid points.

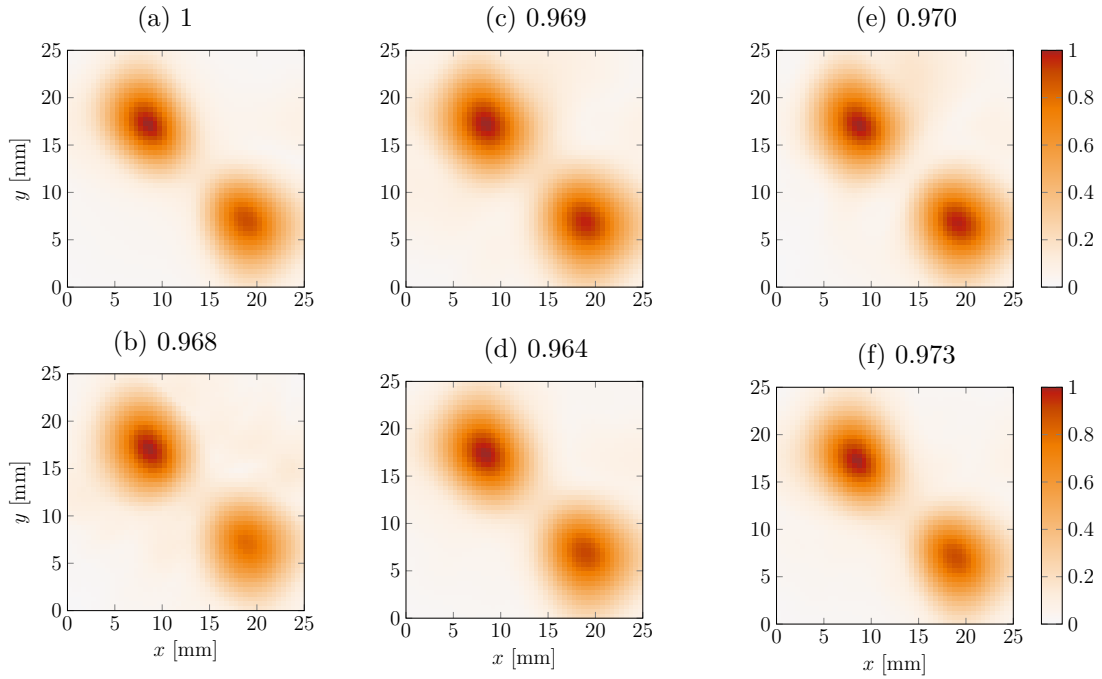


Figure 11.8.: C-Scan of the obtained simulation results (10 % coverage) with their relative gCNR values to that of the ground truth: 11.8a: ground truth, 11.8b: the result without any preprocessing, 11.8c: the IDW result only from the initial sampling positions, 11.8d: the IDW result after resampling, 11.8e: the SF-Kriging result only from the initial sampling positions and 11.8f: the SF-Kriging result after resampling,

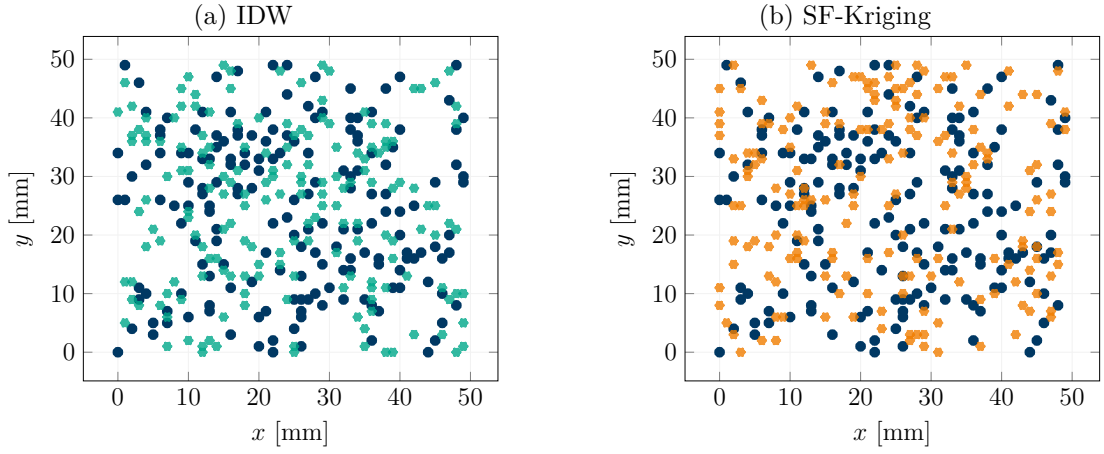


Figure 11.9.: Illustration of sampling (blue) and resampling positions based on interpolation schemes: IDW (left, green) and SF-Kriging with FVnet (right, orange). The initial coverage here is 7% of the all grid points.

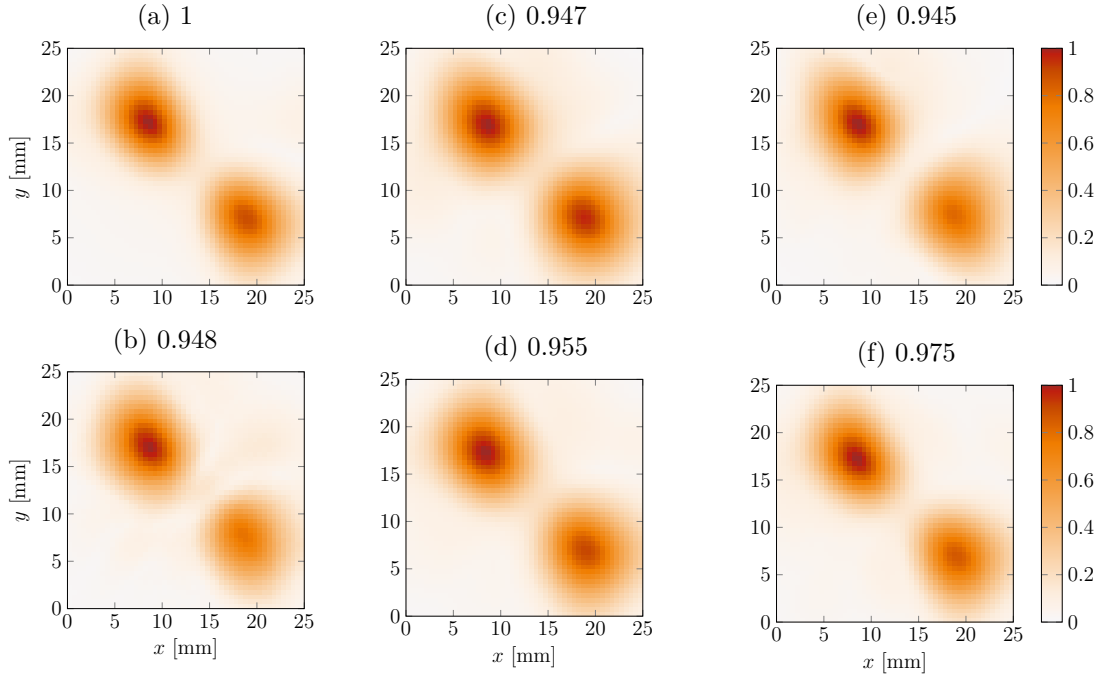


Figure 11.10.: C-Scan of the obtained simulation results (7% coverage) with their relative gCNR values to that of the ground truth: 11.8a: ground truth, 11.8b: the result without any preprocessing, 11.8c: the IDW result only from the initial sampling positions, 11.8d: the IDW result after resampling, 11.8e: the SF-Kriging result only from the initial sampling positions and 11.8f: the SF-Kriging result after resampling,

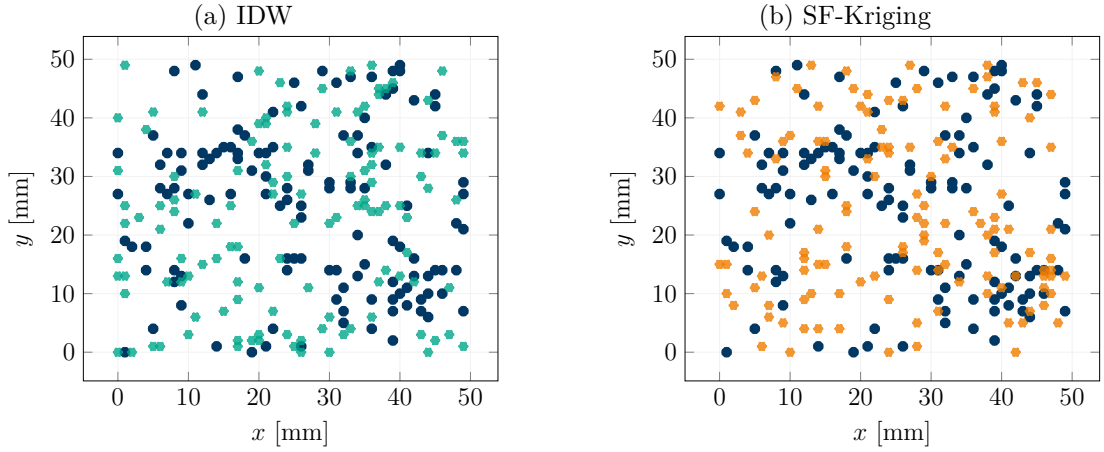


Figure 11.11.: Illustration of sampling (blue) and resampling positions based on interpolation schemes: IDW (left, green) and SF-Kriging with FVnet (right, orange). The initial coverage here is 5% of the all grid points.

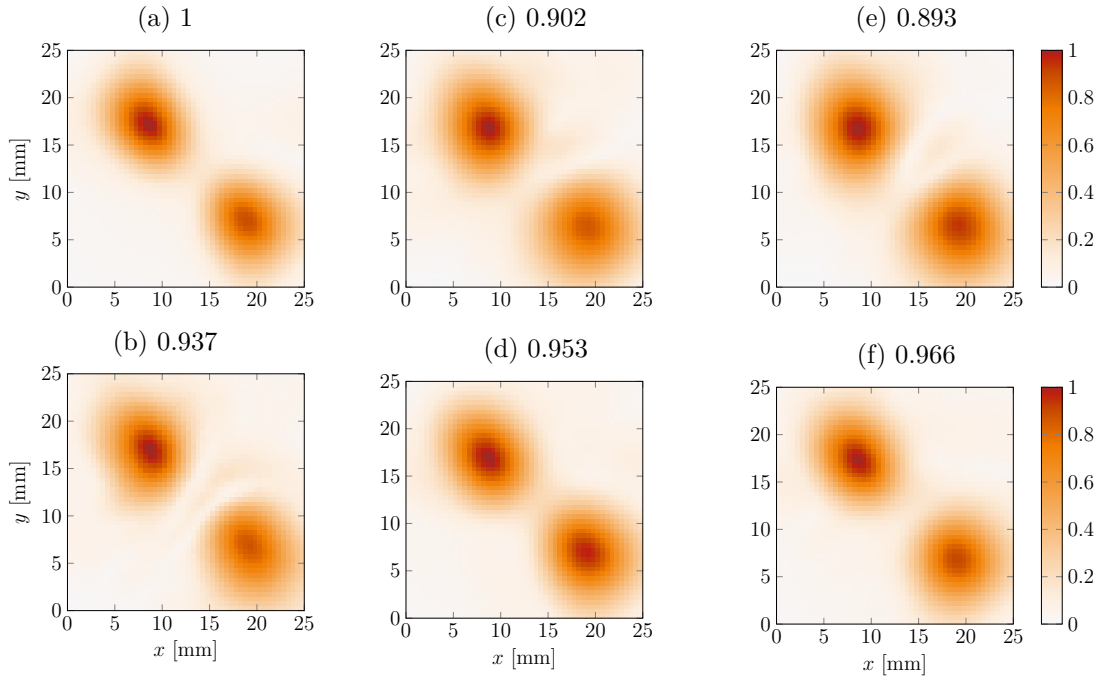


Figure 11.12.: C-Scan of the obtained simulation results (5% coverage) with their relative gCNR values to that of the ground truth: 11.12a: ground truth, 11.12b: the result without any preprocessing, 11.12c: the IDW result only from the initial sampling positions, 11.12d: the IDW result after resampling, 11.12e: the SF-Kriging result only from the initial sampling positions and 11.12f: the SF-Kriging result after resampling,

## 11.4. Evaluation of the Results

In this section, the resolution of the reconstructions are examined. Since the obtained results are in three dimensional, their C-Scan images are used for evaluation. In total, we compute 80 realizations for each coverage, and their evaluation results are averaged, which are shown in Figure 11.13. Considering that the number of A-Scans are doubled after resampling, we first focus on the results based only on the initial sampling positions, before discussing the impact of effective resampling schemes.

For a very low coverage less than 7%, which is equivalent to 0.61 samples per wavelength, both  $\text{MSE}^\dagger$  and  $\text{MgCNR}$  values suggest that interpolation does not contribute to enhance, or may even degrade, the SAFT resolution. Indeed, there is almost no difference in resolution shown in Figure 11.12. This implies that inaccurate interpolations are rather distractive, and the measurements are better left unprocessed. However, with increasing coverage, positive effects of interpolation can be observed. With a coverage higher than 10%, artefacts in the reconstructions of interpolated data are much less intense than that of unprocessed data (Figure 11.8 and 11.4).

Within the two considered interpolation schemes, there are also noticeable differences. Since IDW essentially averages out the available data if there are not sufficient data available or clustered, the resulting reconstructions tend to represent scatterers as circles (Figure 11.8c and 11.4c). This property is well presented in their exemplary results, indicating that it requires more number of scans to be able to image non-circular shapes.

On the other hand, as Figure 11.8 and 11.4 demonstrate, the reconstructions after SF-Kriging can capture the shape of non-circular scatterers better than IDW with the same coverage. It is worth mentioning that the possible deviation in amplitude, which we discussed in Section 10.2, appears to have insignificant effect on the reconstruction resolution. Furthermore, if the coverage is higher than 12%, which corresponds to 1.04 samples per squared wavelength,  $\text{MSE}^\dagger$  and  $\text{MgCNR}$  values of reconstructions after SF-Kriging are equivalent or better than that of IDW after resampling. Considering that number of scans are doubled after resampling, it is a significant sign that SF-Kriging can perform very well in terms of the reconstruction resolution.

Let us now consider how we can benefit from a better sampling scheme. Since resampling positions are also randomly selected for IDW, their results shown in Figure 11.13 are essentially the results of IDW with higher coverage, ranging from 10% to 30%, i.e. 0.87 to 2.61 samples per squared wavelength. On the contrary, resampling after SF-Kriging is based on the prediction variance, which can be seen as rather clustered sampling in Figure 11.3, 11.7, 11.9 and 11.11. For a lower coverage, there is almost no difference in reconstruction results between random and variance-based resampling. This is due to the lack of initial samples in many batches, which makes variances unattainable for the majority of the points. As a result, only few points are selected tactically (Figure 11.11), resulting in very similar results as that of random sampling. Yet, with increasing coverage the variance based resampling scheme outperforms its random counterpart, demonstrating that strategic sampling leads to enhanced resolution in reconstructions.

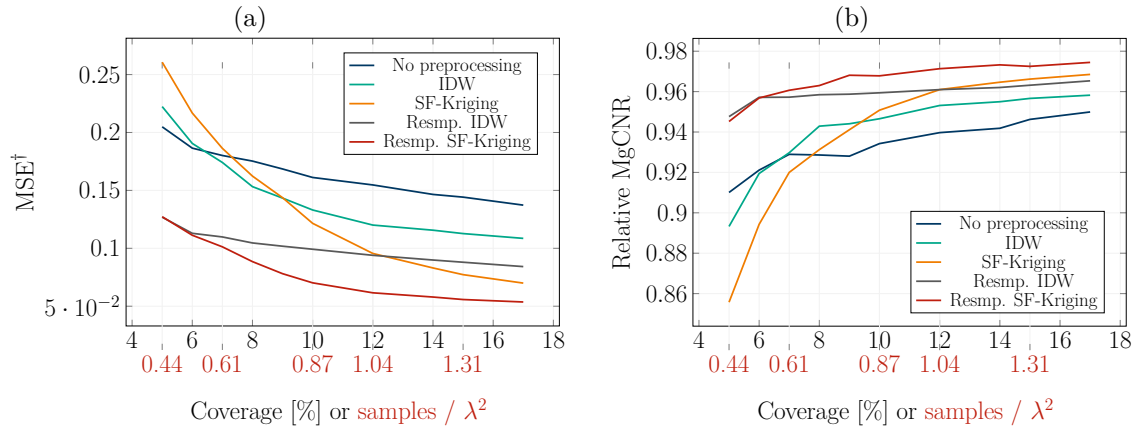


Figure 11.13.: Comparison among reconstruction results in terms of  $MSE^\dagger$  (left) and MgCNR (right) over coverage of the ROI (in black) or samples per squared wavelength (in red). Reconstruction results are computed with two different sampling schemes. One is based on the initial random subsampling, which yields reconstructions of raw subsampled data (blue), preprocessed data with IDW (green) and preprocessed data with SF-Kriging. The other two results are computed using the resampling positions depending on the interpolation methods: reconstructions of randomly resampled data preprocessed with IDW (gray) and variance based resampling data preprocessed with SF-Kriging (red).

## 12. Summary of Part III

In Part III, two sets of simulations are conducted to examine the performance of the proposed interpolation method and the effectiveness of the entire preprocessing scheme. For the simulations, a set of automatic UT measurements, MUSE data, are irregularly subsampled. In this manner, we have the reference which serves as the ground truth to compare to the obtained simulation results.

### **Silumation I: Batch Interpolation**

The first set of simulations, Simulation I, is intended to investigate two things: (1) the estimation accuracy of FVnet and (2) the validity of SF-Kriging for UT data. For this purpose, a randomly selected batch of MUSE data is irregularly subsampled and then interpolated. In order to ensure the spatial intrinsic stationarity within the batch, the batch size is selected as  $N_x \times N_y = 10 \times 10$ . As interpolation schemes, three approaches are employed: IDW and SF-Kriging with two different estimates of the frequency variogram. Frequency variograms are estimated either based on the perfect field knowledge, i.e. with the fully sampled batch data, or via the proposed data driven method, FVnet, from the subsampled data. The frequency variogram estimates using the perfect field knowledge is regarded as the ground truth to evaluate the estimation accuracy of FVnet.

The obtained results show that SF-Kriging overall outperforms IDW, presenting the effectiveness of SF-Kriging for UT data. For 1.305 samples per squared wavelength, FVnet yields comparable prediction results for the missing A-Scans in the batch. When it comes to the frequency variogram estimation, however, the trade-off to keep the flexibility of FVnet may become obvious. Since FVnet only estimates the frequency variogram structures, it may fail to properly scale the estimated frequency variograms. As a consequence, there may be non-negligible deviation in amplitude between the ground truth based prediction and that of FVnet.

### **Simulation II: Reconstruction of Subsampled UT Data**

The second set of simulations, Simulation II, concerns the entire preprocessing scheme for assisted manual UT measurements depicted in Figure 5.1. Here, we seek to investigate the two aspects of the proposed preprocessing scheme in terms of the resulting reconstruction resolution: (1) whether interpolating the missing A-Scans leads to an enhanced resolution and (2) if providing the feedback based on the prediction variance can increase the resolution compared to random sampling.



The simulations are conducted firstly by subsampling a selected ROI of the MUSE data, which are then interpolated before being fed into post-processing along with the predicted A-Scans. To keep the computation tractable and ensure the spatial intrinsic stationarity within the interpolation region, the ROI is divided into multiple batches, and the interpolation is performed in a sliding window manner. The size of a single batch is  $N_x \times N_y = 10 \times 10$ , and the overlap rate is selected as 50%. We obtained in essence three different types of results. One is the reconstruction results of the unprocessed subsampled data. The other two are reconstructions of the preprocessed data via either IDW or FVnet supported SF-Kriging. To illustrate the impact of the data size, number of samples within the ROI is varied from 0.44 to 1.48 samples per squared wavelength.

Overall, we found that interpolation can contribute to an enhanced resolution, if there is adequate amount of samples available, in our case more than 0.61 samples per squared wavelength. While IDW tends to average out the missing A-Scans, the proposed method leads to a better imaging of the scatterer structure if the coverage is higher than 0.87 samples per squared wavelength. Indeed, if the coverage is higher than 1.04 samples per squared wavelength, our method outperforms IDW based on the double amount of the samples, suggesting its effectiveness for predicting UT data. Moreover, the simulation results also demonstrate that strategical sampling based on the prediction variance yields an increased reconstruction resolution. This implies that providing the feedback to the inspection operator has a positive impact on the overall measurement quality.

# Conclusion

In an attempt to ultimately improve their reconstruction results, in this thesis a data-driven hybrid preprocessing scheme is presented for irregularly subsampled UT measurements. As a part of preprocessing, we aimed to interpolate the missing measurements, which is in essence nonlinear spatio-temporal interpolation. The main goal of this thesis is to establish an interpretable DNN scheme to benefit from both theory-based approaches and DNN, which makes use of its nonlinear mapping capability and fast on-site computing. The primary work of this thesis is formulation of our research problem such that we understand the role DNN plays in our preprocessing scheme.

In the first part of this thesis, measurement systems and data modeling for manual UT measurements are first presented to describe the considered scenario for this thesis. This follows an overview of conventional DNN and its state of the art approaches for spatio-temporal interpolation. Here, the main concern of traditional DNN is highlighted, which is its black box nature. This prevents us to understand what the network does and learns, leaving it ungeneralizable and requiring a huge amount of training data to compensate. In the last chapter of the first part, fundamental idea of spatio-temporal interpolation is extensively studied. As a well-established scientific field in spatial data analysis, geostatistics offer profound insight with respect to handling of spatio-temporal data. The emphasis here is placed on the fact that spatio-temporal data cannot be treated as just an extension of spatial data and requires careful examination of their statistical aspects. Through literature survey it becomes more prominent that there are growing interests in spatio-temporal data analysis, yet it still remains a very challenging task.

In the second part of this thesis, the proposed preprocessing scheme is presented. Based on the examination of our needs and the nature of UT data, the preprocessing of UT data is divided into two phases: (1) estimation of field statistics and (2) prediction of measurements. In this way, we can combine geostatistical interpolation technique with a help of DNN to obtain better estimates of field statistics. For this purpose, UT data are formulated as a geostatistics based model, demonstrating that our problem can be indeed considered as a geostatistical problem, which is one of the contribution of this work. Considering the properties of our measurement data, the estimation process via DNN is formulated as a vector-valued nonlinear regression problem, which is well known to be a suitable task for DNN. This follows the description of network architecture, training scheme and dataset.

The last part of this thesis is dedicated to examine the validity of the proposed preprocessing scheme. For this purpose, two sets of simulations are conducted. The first set of simulations, Simulation I, focuses on the prediction accuracy of SF-Kriging and the estimation performance of FVnet compared to the one we could obtain with fully

sampled measurements. The obtained results show that SF-Kriging overall outperforms IDW, demonstrating its general effectiveness for UT data. In view of A-Scan prediction, the performance of FVnet is comparable to that of ground truth based prediction. Yet, FVnet estimates only the structure of frequency variograms, resulting in possible fluctuation of frequency variogram values. As a result, there may be non-negligible deviation in amplitude between the ground truth based prediction and that of FVnet. This shows as the price we paid by prioritizing the flexibility over including more domain specific knowledge, such as spectral shape.

With the second set of simulations, Simulation II, the effectiveness of the entire preprocessing scheme was investigated. The main purpose here is to evaluate the proposed preprocessing scheme in terms of the resolution of resulting SAFT reconstructions. Overall, performing interpolation appears to contribute to enhanced resolution, unless the subsampling is extremely low. For a very low coverage, in our case less than 0.61 samples per squared wavelength, the results show that keeping the raw data unprocessed yields better resolution in reconstructions. With increasing number of samples, however, there is observable difference in reconstruction resolution between processed and unprocessed data. Here as well FVnet supported SF-Kriging outperforms IDW for the subsampling with the same or even higher coverage. This also confirms its applicability to UT data. Moreover, the obtained results indicate that the concern regarding deviation in amplitude does not have much effect on the reconstruction results. As a part of Simulation II, the effect of strategic sampling was also investigated. Compared to random sampling of the same coverage, the results demonstrate that strategic sampling leads to enhanced resolution in reconstructions. This implies that providing the feedback has indeed positive impact on the measurement quality.

Overall, the proposed preprocessing scheme exhibits its effectiveness for our UT scenario. Yet, there are still rooms for improvement, which can be further investigated. First, amplitude of frequency variograms can be estimated as well by modeling them parametrically. In this way, the trained network includes the SPDE based domain knowledge, leading to more physically aware outputs. Second, the reconstruction results could be incorporated into the loss function during the training phase, which makes the network also aware of our ultimate goal. Another possibility would be tuning the weight regularization parameter of SF-Kriging. Currently a fixed value is used for all frequency components, yet this does not have to be, and adjusting this regularization parameter individually may yield better prediction results. Toward actual online applications, the processing time can be accelerated by efficient parallel computing or progressive approach.

# Bibliography

- [1] P. Cawley. Non-destructive testing - current capabilities and future directions. *Proceedings of the Institution of Mechanical Engineers, Part L*, vol 215, 2001.
- [2] F. Krieg and J. Kirchhof and S. Koderä and S. Lugin and A. Ihlow and T. Schwen-der and G. Del Galdo and F. Römer and A. Osman. SAFT processing for manu-ally acquired ultrasonic measurement data with 3D smartInspect. *Insight - Non-Destructive Testing and Condition Monitoring*, 61(11):663–669(7), November 2019.
- [3] B. Valeske, A. Osman, F. Römer, and R. Tschuncky. Next Generation NDE Sensor Systems as IIoT Elements of Industry 4.0. *Research in Nondestructive Evaluation*, 31(5-6):340–369, 2020. <https://doi.org/10.1080/09349847.2020.1841862>.
- [4] L. Ericsson and M.G. Gustafsson. Perception and Entropy Inspired Ultrasonic Grain Noise Suppression using Noncoherent Detector Statistics. In *ECNDT*, 1998.
- [5] T.E. Hall, L.D. Reid, and S.R. Doctor. The SAFT-UT Real-Time Inspection System - Operational Principles and Implementation. Technical report, Pacific Northwest Laboratory, 1988.
- [6] J. Krautkrämer and H. Krautkrämer. *Ultrasonic Testing of Materials*. Springer-Verlag, 1990.
- [7] F. Lingvall. *Time-domain Reconstruction Methods for Ultrasonic Array Imaging*. PhD thesis, Uppsala University, 2004.
- [8] Rick Pandey. Robust Signal Processing of Manually Acquired Single-Channel Syn-thetic Aperture Ultrasound Measurement Data. Master’s thesis, Technische Uni-versität Ilmenau, 2020.
- [9] A. Karpatne, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar. Machine Learning for the Geosciences: Challenges and Opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8):1544–1554, 2019.
- [10] P. Revesz. *Spatiotemporal Interpolation Algorithms*, pages 2736–2739. Springer US, 2009. [https://doi.org/10.1007/978-0-387-39940-9\\_803](https://doi.org/10.1007/978-0-387-39940-9_803).
- [11] G. Atluri and A. Karpatne and V. Kumar. Spatio-Temporal Data Mining: A Survey of Problems and Methods. *ACM Computing Surveys*, 2018.

- [12] N. Moradi and M. Dousty and R. C. Sotero. Spatiotemporal Empirical Mode Decomposition of Resting-State fMRI Signals: Application to Global Signal Regression. 2019. <https://www.frontiersin.org/articles/10.3389/fnins.2019.00736/full>.
- [13] N. A. C. Cressie and C. K. Wikle. *Statistics for SPATIO-TEMPORAL DATA*. WILEY, 2011.
- [14] M. Sherman. *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*. WILEY, 2010.
- [15] F. Sigrist and H.R. Künsch and W. A. Stahel. Stochastic partial differential equation based modelling of large space-time data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, page 3–33, 2014. <http://dx.doi.org/10.1111/rssb.12061>.
- [16] T. S. Rao and G. Terdik. A New Covariance Function and Spatio-Temporal Prediction (Kriging) for A Stationary Spatio-Temporal Random Process. *Journal of Time Series Analysis*, 2017. <https://onlinelibrary.wiley.com/doi/full/10.1111/jtsa.12245>.
- [17] A. Karpatne and G. Atluri and J. H. Faghmous and M. Steinbach and A. Banerjee and A. Ganguly and S. Shekhar and N. Samatova and V. Kumar. Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. *IEEE Transactions on Knowledge and Data Engineering*, 29(10):2318–2331, 2017. [=http://dx.doi.org/10.1109/TKDE.2017.2720168](http://dx.doi.org/10.1109/TKDE.2017.2720168).
- [18] J. Willard and X. Jia and S. Xu and M. Steinbach and V. Kumar. Integrating Physics-Based Modeling with Machine Learning: A Survey, 2020. <https://arxiv.org/abs/2003.04919>.
- [19] V. Monga, Y. Li, and Y. C. Eldar. Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing. *IEEE Signal Processing Magazine*, 38(2):18–44, 2021.
- [20] T. S. Rao and S. Das and G. N. Boshnakov. A Frequency domain approach for the estimation of parameters of spatio-temporal stationary random processes. *Journal of Time Series Analysis*, 35(4):357–377, 2014.
- [21] T. S. Rao and G. Terdik. ON THE FREQUENCY VARIOGRAM AND ON FREQUENCY DOMAIN METHODS FOR THE ANALYSIS OF SPATIO-TEMPORAL DATA. *Journal of Time Series Analysis*, 2017. <https://onlinelibrary.wiley.com/doi/full/10.1111/jtsa.12231>.
- [22] Ahmad Omira. Real-time Reconstruction of Manually Measured Compressed Ultrasound Synthetic Aperture Measurement Data, 2021. Bachelor’s thesis.

- [23] R. Demirli and J. Saniie. Asymmetric Gaussian Chirplet Model for Ultrasonic Echo Analysis. *IEEE International Ultrasonics Symposium Proceedings*, 2010.
- [24] J. Kirchhof, F. Krieg, F. Römer, A. Ihlow, A. Osman, and G. Del Galdo. Speeding up 3D SAFT for ultrasonic NDT by sparse deconvolution. In *IEEE IUS*, 2016.
- [25] T. G. Kolda and B. W. Bader. *Tensor Decompositions and Applications*. 2009.
- [26] W. W. Hsieh. *Machine learning methods in the environmental sciences: Neural networks and kernels*. Cambridge university press, 2009.
- [27] K. He, X. Zhang, S. Ren, and J. Sun. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In *IEEE International Conference on Computer Vision (ICCV)*, 2015.
- [28] Ž. Ivezić and A. J. Connolly and J. T. VanderPlas and A. Gray. *Statistics, data mining, and machine learning in astronomy: a practical Python guide for the analysis of survey data*. Princeton University Press, 2019.
- [29] N. Buduma and N. Lacascio. *Fundamentals of Deep Learning*. O’Reilly Media, 2017.
- [30] Y. A. LeCun and L. Bottou and G. B. Orr and K. Miller. *Efficient BackProp*, pages 9–48. Springer, 2012. <http://yann.lecun.com/exdb/publis/pdf/lecun-98b.pdf>.
- [31] S. Hochreiter and J. Schmidhuber. LONG SHORT-TERM MEMORY. *Neural Computation*, 1997.
- [32] S. Avalos and J. M. Ortiz. Geological modeling using a recursive convolutional neural networks approach, 2019. <https://arxiv.org/abs/1904.12190>.
- [33] X. Chai and H. Gu and F. Li and H. Duan and X. Hu and K. Lin. Deep learning for irregularly and regularly missing data reconstruction, 2021. <https://www.nature.com/articles/s41598-020-59801-x>.
- [34] Y. Wang and M. Long and J. Wang and Z. Gao and P. S. Yu. PredRNN: Recurrent Neural Networks for Predictive Learning using Spatiotemporal LSTMs. In *31st Conference on Neural Information Processing Systems (NIPS 2017)*, 2017.
- [35] Y. Long and X. She and S. Mukhopadhyay. HybridNet: Integrating Model-based and Data-driven Learning to Predict Evolution of Dynamical Systems. In *2nd Conference on Robot Learning (CoRL 2018)*, 2018.
- [36] R. Sen and H. F. Yu and I. Dhillon. Think Globally, Act Locally: A Deep Neural Network Approach to High-Dimensional Time Series Forecasting. In *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, 2019.

- [37] Z. Bai and B. Cai and W. Shangguan and L. Chai. Deep Learning Based Motion Planning For Autonomous Vehicle Using Spatiotemporal LSTM Network, 2019. <https://arxiv.org/abs/1903.01712>].
- [38] J. Ma and Y. Ding and J. C. P. Cheng and F. Jiang and Z. Wan. A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM2.5. *Journal of Cleaner Production*, 2019.
- [39] M. Reichstein and G. Camps-Valls and B. Stevens, M. Jung and J. Denzler and N. Carvalhais and Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature*, 566:195–204, 2019. <https://www.nature.com/articles/s41586-019-0912-1>.
- [40] Y. Wu and D. Zhuang and A. Labbe and L. Sun. Inductive Graph Neural Networks for Spatiotemporal Kriging, 2020. <https://arxiv.org/abs/2006.07527>.
- [41] T. Bai and P. Tahmasebi. Accelerating geostatistical modeling using geostatistics-informed machine Learning. *Computers and Geosciences*, 2019.
- [42] G. Christakos. *Random Field Models in Earth Sciences*. Academic Press, 1992. <https://www.sciencedirect.com/book/9780121742300/random-field-models-in-earth-sciences>.
- [43] S. Shekhar and Z. Jiang and R. Y. Ali and E. Eftelioglu and X. Tang and V. M. V. Gunturi and X. Zhou. Spatiotemporal data mining: A computational perspective. *ISPRS International Journal of Geo-Information*, 2015.
- [44] J. Li and A. D. Heap. A review of spatial interpolation methods for environmental scientists. *Geoscience Australia*, page 137pp, 2008. [http://planet.botany.uwc.ac.za/NISL/BCB\\_BIM\\_honours/Course\\_Documents\\_2016/Spatial\\_Interpolation\\_statistics\\_methods.pdf](http://planet.botany.uwc.ac.za/NISL/BCB_BIM_honours/Course_Documents_2016/Spatial_Interpolation_statistics_methods.pdf).
- [45] L. Li, T. Lossner, C. Yorke, and R. Piltner. Fast Inverse Distance Weighting-Based Spatiotemporal Interpolation: A Web-Based Application of Interpolating Daily Fine Particulate Matter PM2.5 in the Contiguous U.S. Using Parallel Programming and k-d Tree. *International Journal of Environmental Research and Public Health*, September 2014. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4199009/>.
- [46] N. A. C. Cressie. *Statistics for Spatial Data*. WILEY, 1993.
- [47] A. Bårdossy. Introduction to geostatistics. Institute of Hydraulic Engineering, University of Stuttgart.
- [48] M. Mälicke and E. Möller and H. D. Schneider and S. Müller. mmaelicke/scikit-gstat: A scipy flavoured geostatistical variogram analysis toolbox, 2021. <http://doi.org/10.5281/zenodo.4835779>.

- [49] D. R. Brillinger. *Time Series: Data Analysis and Theory*. SIAM: Society for Industrial and Applied Mathematics, 1981.
- [50] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. Society for Industrial and Applied Mathematics, 1998.
- [51] P.C. Hansen and T. Sekii and H. Shibahashi. THE MODIFIED TRUNCATED SVD METHOD FOR REGULARIZATION IN GENERAL FORM. *SIAM J. Sci. and Stat. Comput*, 1991. <https://epubs.siam.org/doi/abs/10.1137/0913066?journalCode=sijcd4>.
- [52] A. Rodriguez-Molares, O. M. H. Rindal, J. D’hooge, S. Måsøy, A. Austeng, M. A. L. Bell, and H. Torp. The generalized contrast-to-noise ratio: a formal definition for lesion detectability. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2019.
- [53] F. Krieg and S. Lugin and J. Kirchhof and A. Ihlow and T. Schwender and G. del Galdo and F. Römer and A. Osman. SAFT processing for manually acquired ultrasonic measurement data with 3D smartInspect. In *SHM-NDT*, 2018.
- [54] Data presentation. <https://www.nde-ed.org/EducationResources/CommunityCollege/Ultrasonics/EquipmentTrans/DataPres.htm>, 2014. Online : last accessed 09.01.2019.



# Notation and Symbols

## Variable Types, array access and functions

$\boldsymbol{v}$	Vector
$\boldsymbol{v}_i$	Access to the $i$ -th scalar element of $\boldsymbol{v}$
$\boldsymbol{U}$	Matrix
$\boldsymbol{U}_{ij}$	Access to the scalar element in $\boldsymbol{U}$ , where $i$ and $j$ denote its column and row, respectively
$\text{vec}\{\boldsymbol{U}\}$	Vectorization of a matrix $\boldsymbol{U}$

## Frequently used terms and symbols

$c_0$	Speed of sound in a medium
$f_S$	Sampling frequency
$\lambda$	Wave length
$\text{dx}$	Spatial discretization step along the $x$ -axis
$\text{dy}$	Spatial discretization step along the $y$ -axis
$\text{dz}$	Spatial discretization step along the $z$ -axis
$\text{dt}$	Temporal resolution of the measurement
$N_x$	Number of samples along the $x$ -axis in reconstruction and ground truth
$N_y$	Number of samples along the $y$ -axis in reconstruction and ground truth
$N_z$	Number of samples along the $z$ -axis in reconstruction and ground truth
$N_t$	Number of temporal samples of the measurements

## Notations for SRF and STRF

### Spatial Random Field (SRF)

SRF depends on the position  $\boldsymbol{s}_i$ :

$$Y(\boldsymbol{s}_i).$$

For second-order stationary SRF: its spatial correlation function depends solely on the lag vector  $\boldsymbol{h}$ :

$$\varphi_Y(\boldsymbol{h}) = \text{Corr} \{Y(\boldsymbol{s}), Y(\boldsymbol{s} + \boldsymbol{h})\}.$$

For intrinsic stationary SRF:

Spatial incremental process for a fixed  $\mathbf{h}_{ij}$ , which depends on the position  $\mathbf{s}_i$ :

$$\begin{aligned} X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i) &= Y(\mathbf{s}_i) - Y(\mathbf{s}_j) \\ &= Y(\mathbf{s}_i) - Y(\mathbf{s}_i + \mathbf{h}_{ij}). \end{aligned}$$

Since  $Y(\mathbf{s}_i)$  is intrinsic stationary,  $X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i)$  is zero-mean spatial second-order stationary. This leads its spatial correlation function with a fixed  $\mathbf{h}_{ij}$  to be expressed as a function of the lag vector  $\boldsymbol{\vartheta}$  independent of  $\mathbf{s}_i$ :

$$\varphi_X^{[\mathbf{h}_{ij}]}(\boldsymbol{\vartheta}) = \text{Corr} \left\{ X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i), X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i + \boldsymbol{\vartheta}) \right\}.$$

### Spatio-Temporal Random Field (STRF)

STRF depends on the position  $\mathbf{s}_i$  and time:

$$Y(\mathbf{s}_i, t).$$

This spatio-temporal process can also be described in terms of the (temporal) frequency as

$$Z_Y(\mathbf{s}_i, \omega) = \mathcal{F}_t \{ Y(\mathbf{s}_i, t) \}$$

For spatio-temporal second-order stationary STRF: its spatial correlation depends on the spatial lagvector  $\mathbf{h}$

$$\varphi_Y^{[t]}(\mathbf{h}) = \text{Corr} \{ Y(\mathbf{s}, t), Y(\mathbf{s} + \mathbf{h}, t) \},$$

and its spatio-temporal correlation depends also on the temporal lag  $\tau$  as

$$\varphi_Y(\mathbf{h}, \tau) = \text{Corr} \{ Y(\mathbf{s}, t), Y(\mathbf{s} + \mathbf{h}, t + \tau) \}.$$

For a spatio-temporal intrinsic stationary STRF:

For a fixed  $\mathbf{h}_{ij}$ , its spatial incremental process varies with the position  $\mathbf{s}_i$

$$X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i, t) = Y(\mathbf{s}_i, t) - Y(\mathbf{s}_j, t).$$

Analog to the purely spatial case, the incremental process  $X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i, t)$  is zero-mean spatio-temporal second-order stationary. This means that for a fixed lag vector  $\mathbf{h}_{ij}$  its spatial correlation depends only on the lag vector  $\boldsymbol{\vartheta}$  and independent of the position  $\mathbf{s}_i$

$$\varphi_X^{[\mathbf{h}_{ij}, t]}(\boldsymbol{\vartheta}) = \text{Corr} \left\{ X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i, t), X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i + \boldsymbol{\vartheta}, t) \right\}.$$

This incremental process can also be expressed in the frequency domain as

$$Z_X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i, \omega) = \mathcal{F}_t \left\{ X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i, t) \right\},$$

which depends on the position  $\mathbf{s}_i$ .

For a fixed lag vector  $\mathbf{h}_{ij}$  of the increment and a single frequency component  $\omega_m$ , the spatial correlation of the incremental frequency response is defined as

$$\varphi_{Z_X}^{[\mathbf{h}_{ij}, \omega_m]}(\boldsymbol{\vartheta}) = \text{Corr} \left\{ Z_X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i, \omega_m), Z_X^{[\mathbf{h}_{ij}]}(\mathbf{s}_i + \boldsymbol{\vartheta}, \omega_m) \right\} \quad \forall \quad \omega_m.$$

# Glossary

**3D SmartInspect** Measurement assistance system developed by Fraunhofer Institute for Nondestructive Testing IZFP [3, 53]

**A-Scan** The simplest form of presentation for UT data, where the received signal is displayed over the time axis [54]

**autoencoder** An unsupervised learning algorithm to recover the significance in the inputs, such as extracting the clean underlying image structure from its noisy input.

**C-Scan** Top view of 3D data, where the absolute largest values of each position are illustrated as the intensity of color [54]. In this work, this way of presentation is employed for visualization, based on which the results are evaluated.

**frequency variogram** One possibility to express the second-order spatial statistics of a STRF. It is defined in space-frequency domain and can be used for obtaining MMSE estimate of frequency responses of the desired time series data. [16, 20, 21].

**FVnet** Proposed data-driven approach to estimate a frequency variogram for a single frequency (Chapter 7)

**intrinsic stationarity** a weaker, thus more general, stationarity assumption than second-order stationarity. It states spatial WSS for the increments of an SRF [14, 46]. This is called SRF with homogeneous increments in [42]. See definition in (4.10).

**Kriging** one of the mostly used inference methods in geostatistics. Under proper setting, it derives a best linear unbiased estimator [14, 42, 46].

**second-order stationarity** a stationarity assumption often made in geostatistics which is the equivalent notion of spatial WSS [14, 46]. This is called homogeneous SRF in [42]. See definition in (4.8).

**variogram** the crucial parameter of geostatistics, which is used to predict unknown values of an SRF via Kriging [14, 42, 46]. See definition in 4.10 (ii).

# Acronyms

<b>AWGN</b>	Additive White Gaussian Noise
<b>CNN</b>	Convolutional Neural Networks
<b>CSPDE</b>	Complex Stochastic Partial Differential Equation
<b>CSRF</b>	Complex Spatial Random Field
<b>DFT</b>	Discrete Fourier Transform
<b>DL</b>	deep learning
<b>DNN</b>	Deep Neural Networks
<b>IDW</b>	Inverse Distance Weighting
<b>LSTM</b>	Long Short Term Memory
<b>ML</b>	Maximum Likelihood
<b>MLP</b>	Multilayer Perceptrons
<b>MMSE</b>	Minimum Mean Squared Error
<b>MSE</b>	Mean Squared Error
<b>MTSVD</b>	Modified Truncated Singular Value Decomposition
<b>NDT</b>	Nondestructive Testing
<b>NN</b>	Neural Network
<b>PDE</b>	Partial Differential Equation
<b>RNN</b>	Recurrent Neural Networks
<b>ROI</b>	Region of Interest
<b>SAFT</b>	Synthetic Aperture Focusing Technique
<b>SF-Kriging</b>	Space-Frequency Kriging
<b>SPDE</b>	Stochastic Partial Differential Equation

**SRF** Spatial Random Field

**STRF** Spatio-Temporal Random Field

**ToF** Time-of-Flight

**UT** Ultrasonic Testing

**WSS** Wide Sense Stationarity

# Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

Ilmenau, .....2021

---

Sayako Koderä