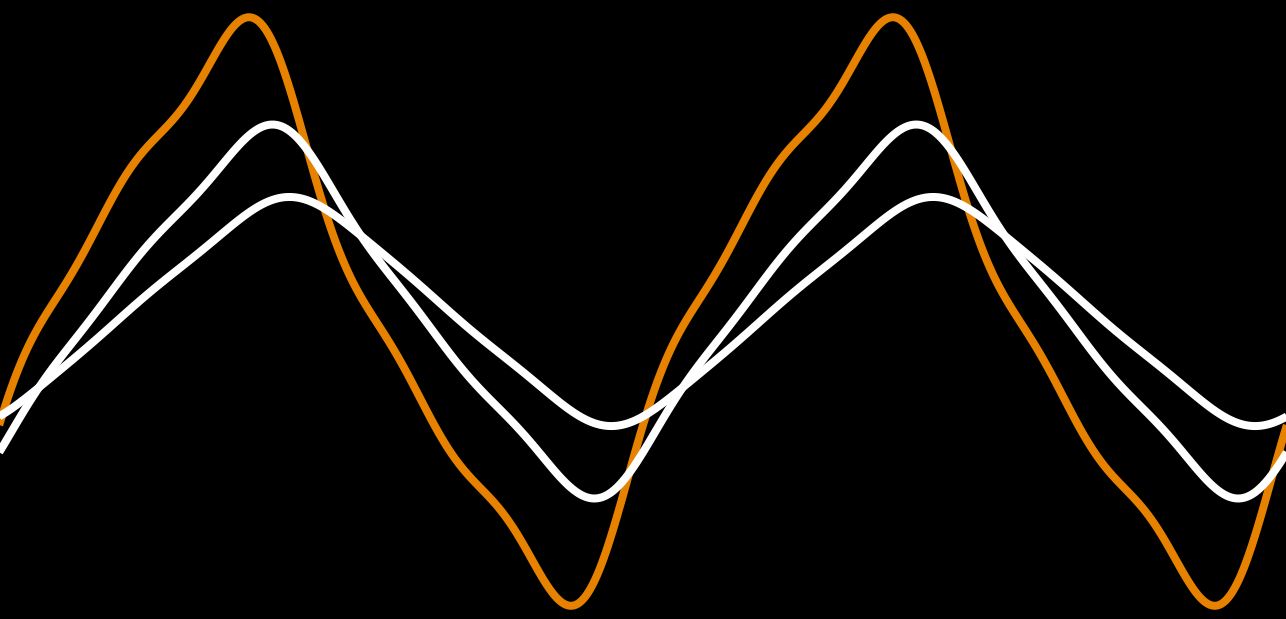
Frequency domain estimation of spatially varying transport coefficients

Frequency domain estimation of spatially varying transport coefficients



Ricky van Kampen

Frequency domain estimation of spatially varying transport coefficients

Ricky van Kampen

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 - EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union not the European Commission can be held responsible for them.

The author has successfully completed the educational program of the Graduate School of the Dutch Institute of Systems and Control (DISC).

A catalogue record is available from the Eindhoven University of Technology Library.
ISBN: 978-90-386-5703-5

Typeset by the author using the pdf LaTeX documentation system.
Cover design: Ricky van Kampen & Shari Bruijstens
Reproduction: Gildeprint - Enschede

Frequency domain estimation of spatially varying transport coefficients

Proefschrift

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus prof.dr.ir. F. P. T. Baaijens, voor een
commissie aangewezen door het College voor
Promoties, in het openbaar te verdedigen
op maandag 03 april 2023 om 16:00 uur

door

Ricky Jan Robert van Kampen

geboren te Spijkenisse

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr.ir. P. D. Anderson
1e promotor:	prof.dr. H. J. Zwart
2e promotor:	prof.dr. S. Weiland
co-promotor:	dr.ir. M. van Berkel (DIFFER)
leden:	prof.dr. M. Huysmans (Vrije Universiteit Brussel)
	prof.dr.ir. P. M. J. van den Hof
	prof.dr.ir. K. J. Keesman (Wageningen University & Research)
adviseur:	dr. T. Tala (VTT Technical Research Centre of Finland)

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

"*Je n'ai fait celle-ci plus longue que parce que je n'ai pas eu le loisir de la faire plus courte.*"
- Blaise Pascal

"*I made this one longer only because I did not have time to make it shorter.*"
- Blaise Pascal

Summary

Frequency domain estimation of spatially varying transport coefficients

The quantification of transport phenomena that describe the exchange of energy, charge, mass, or momentum, plays a key role in natural sciences and engineering. The generalized mathematical model that describes these transport phenomena is called the generic scalar transport model and belongs to the class of parabolic partial differential equations (PDEs). Although transport models are often derived from first principles, the corresponding quantities representing the key mechanisms of transport are uncertain or unknown. As this is a common and important estimation problem, various methods have been developed in the literature. Depending on the field of application, the methodologies range from estimating constant coefficients using the analytical solution of the PDE in the frequency domain, to estimating spatially varying coefficient functions using numerical PDE approximation methods. This thesis expands on both these approaches by extending and comparing different frequency domain estimation approaches. The choice for frequency domain methods is deliberate as it has numerous advantages over time-domain approaches. One of the most important being the reduction of the PDE to a complex-valued ordinary differential equation. Moreover, the information carrying spectral content of perturbation experiments for PDEs is typically sparse which is exploited to its full extend.

The first part of this thesis focuses on the extension of a maximum likelihood estimator (MLE) for constant parameters in one spatial dimension (1D), which allows to compare the effect of choosing different boundary conditions for the analytical solution, namely the standard semi-infinite domain versus a locally bounded domain. To be specific, this thesis extends both the locally bounded and semi-infinite approach to allow for multiple arbitrary noisy measurements. Consequently, the same (number of) measurements can be used in both approaches allowing for fair comparison between the two. These methods applied to both simulations and real experimental data from hydrology showing the merit of multi-measurement estimation with local domains over the semi-infinite domain.

The second and largest part of this thesis focuses on developing methodologies for estimating spatially varying coefficients for parabolic PDEs. The standard approach is to minimize the output error, i.e. the difference between the measurements and model value at the measurement locations. If the coefficients are unconstrained, the estimated coefficients often contain large (non-physical) oscillatory artifacts due to noise amplifications. Classically this is resolved by regularization, i.e. constraining the unknown coefficients. This thesis avoids this (often) arbitrary regularization on the unknown coefficients by minimizing the equation error, i.e. the difference between the measured or estimated state and the state of the model. The approach is as follows: 1) a discretization is applied to the spatial coordinate; 2) a linear parameterization is applied to the unknown coefficients. The difference between the degrees of freedom of this parameterization and the considered number of spatial points directly reflects the amount of regularization applied to the estimation problem. Moreover, the linear parameterization in combination with the equation error results in a linear least squares problem that has a known closed-form solution. This allows 3) direct calculation of the global optimum by solving a linear matrix equality resulting in a low computational cost. This is an important improvement as the estimation of spatially varying coefficients is normally associated with high computational cost due to necessary iterative optimization methods. Hence, this new methodology allows to quickly estimate and compare different coefficient parameterizations in a machine learning like manner. Moreover, it allows for estimating the spatially varying coefficients in multiple spatial dimensions, which is typically computational heavy.

As further improvement of the 1D estimation method, the case where all measurements contain noise is considered. Consequently, the closed-form solution will become biased. Therefore, the methodology is extended by first determining the probability density function of the state with Gaussian process regression and then a MLE to find the transport coefficients.

The proposed methodologies are applied within two different fields. In the field of hydrology, it used to estimate the mixing of groundwater and surface water causing the exchange of contaminants affecting the water quality. While in the field of nuclear fusion, it is used to estimate the heat and particle transport that dictate the reactor efficiency. It is expected that these methods find their way to more applications and help users to identify their transport models, and give rise to a new perspective in transport coefficients estimation.

Contents

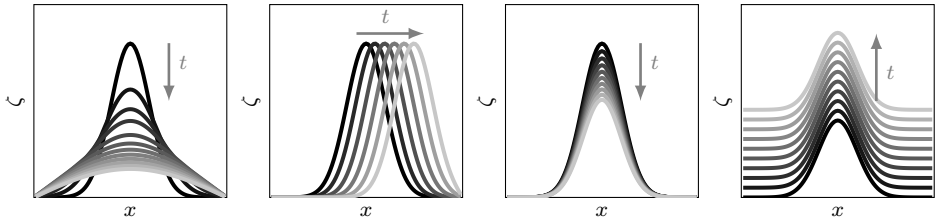
1 Introduction

Abstract - The quantification of transport phenomena plays a key role in many disciplines within natural sciences and engineering as transport phenomena, in its many manifestations in physics, often affects the performance of (engineered) systems. This thesis deals with the quantification of transport phenomena by inferring the key exchange rates from measurements of the system dynamics. More specifically, the focus is on spatially varying coefficient functions in the generic transport equation for a single scalar variable, e.g., temperature. As the standard scalar transport equation is a parabolic partial differential equation (PDE), the developed methods are generic. This will be demonstrated by applying the methods to two examples from completely different domains that are described by very similar equations, namely the heat transport in river stream-beds and heat and particle transport in nuclear fusion reactors.

1.1 Transport phenomena

The study of transport phenomena concerns the exchange of energy, charge, mass, species, or momentum from one region (spatially defined location) to another within systems that are constituted by a very large number of particles, such as molecules and atoms (Mauri 2015). Transport phenomena play a key role in many disciplines within natural sciences and engineering as transport phenomena, in its many manifestations in physics, often affects the performance of (engineered) systems. Examples are the transport in nuclear fusion reactors of particles, heat, and momentum which determine the reactor's efficiency (Ryter et al. 2010); the mixing of groundwater and surface water that cause the exchange of contaminants affecting the water quality (Boano et al. 2014; Schneidewind et al. 2016); medical treatments such as local hyperthermia therapy that use heat to damage and destroy cancer cells while causing minimal damage to normal tissues (Deuflhard et al. 2012); industrial mixing of fluids such as polymers and food by (laminar) flow reorientation to improve mixing/heating rates (Paul et al. 2003); movement of heat and moisture through buildings that can result in mold growth and decay of the building (Havinga and Schellen 2019). Where engineers are mainly focused on manipulating the systems using the transport models, scientist are more interested in understanding the underlying processes causing transport phenomena, e.g., complex flow patterns caused by turbulence.

The most accurate way to describe these transport phenomena is by mathematical models that capture the underlying conservation principles. As a result, models across different disciplines share a very similar mathematical framework such that tools in the analysis of one field can directly applied to others. However, the common challenge among all disciplines considering transport is identifying the (physical) coefficients in these mathematical models representing the different exchange rates. Even though the models are often derived from first principles, the physical coefficients, in most applications, are uncertain, unknown, or only partially known. As the coefficients often depend on many other system variables and its environment, the coefficient cannot be isolated and measured separately. Therefore, the exact coefficients for these mathematical models need to be simultaneously identified from the system evolution over time and space or calculated with high fidelity modeling taking fundamental physical mechanisms into account. This is possible as each of the coefficients affects the evolution of the system differently. The effect of the four most important exchange rates (diffusion, convection, reaction/damping, and source/sink) on the evolution of the considered transport variable ζ, e.g. temperature or density, is shown in Figure 1.1. Hence, by measuring the evolution of the transport variable, either during a designed experiment or natural operation, it is possible to identify the different coefficients.



(a) diffusion D **(b)** convection V **(c)** damping K **(d)** source P

Figure 1.1. The effect of the different (constant) transport coefficients on the transport variable ζ as a function of time t and space x. For a fixed initial condition ζ_0, the response of ζ as a function of space is shown at different time stamps. The arrow and change in gray scale indicate the evolution of time.

1.1.1 Perturbation studies to analyze transport

Transport can be either analyzed in steady state, i.e., without temporal changes, or by perturbing the transport variable. When studying transport in steady state, the different transport coefficients, such as diffusion and convection, contributing to the steady state cannot be distinguished. Only for some special case, e.g., diffusion only or with prior knowledge of other transport coefficients, the transport coefficient can be determined (Lopes Cardozo 1995; Anibas et al. 2009). As this is generally not the case, temporal dynamics are required to provide the additional information to quantify the different transport phenomena. To measure the temporal dynamics, the transport variable is perturbed such that it deviates from its steady state. This is often done by modulating the (boundary) input of the system. If this perturbation is sufficiently small, the dynamics of a nonlinear model can be accurately described by a linear model that is obtained after linearization around the point of interest. The general equation for the linearization of a multivariable function $f(\boldsymbol{\zeta})$ at the operating point $\boldsymbol{\zeta}_0$ is

$$f(\boldsymbol{\zeta}) \approx f(\boldsymbol{\zeta}_0) + \nabla_{\boldsymbol{\zeta}} f \,|_{\boldsymbol{\zeta}_0} \cdot (\boldsymbol{\zeta} - \boldsymbol{\zeta}_0), \tag{1.1}$$

where $\boldsymbol{\zeta} \in \mathbb{R}^p$ is the vector of p transport variables, e.g., density and temperature. Here, only the perturbation $\tilde{\boldsymbol{\zeta}} = \boldsymbol{\zeta} - \boldsymbol{\zeta}_0$ is of our interest as this represents the dynamics close to $\boldsymbol{\zeta}_0$ and can be described by a linear model. For transport, this often results in a parabolic partial differential equation of the following form (Lopes Cardozo 1995):

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{\zeta}} = A\nabla^2_{\boldsymbol{x}}\tilde{\boldsymbol{\zeta}} + B\nabla_{\boldsymbol{x}}\tilde{\boldsymbol{\zeta}} + C\tilde{\boldsymbol{\zeta}} + \boldsymbol{S}_{\tilde{\boldsymbol{\zeta}}}, \tag{1.2}$$

where $\tilde{\boldsymbol{\zeta}}$ is a function of the spatial variable $\boldsymbol{x} \in \mathbb{X} \subset \mathbb{R}^n$ and the temporal variable $t \in \mathbb{T} := [t_0, \infty)$. The gradient with respect to $\boldsymbol{x}$ is given by $\nabla_{\boldsymbol{x}}$. The sink/source vector $\boldsymbol{S}_{\tilde{\boldsymbol{\zeta}}}$ can be time and space dependent and the matrices A, B and C contain

the transport coefficients. The matrix elements can also be functions depending on many different terms including the values at the operating point of the transport variables and their derivatives. Generally, only one transport variable is perturbed and studied, i.e., one element from the vector of transport variables $\boldsymbol{\zeta}$. Therefore, the focus is on the quantification of transport coefficients for a single scalar transport variable.

1.1.2 Generic scalar transport

For scalar transport, i.e., one element from the vector of transport variables $\tilde{\boldsymbol{\zeta}}$ represented by the scalar variable $\tilde{\zeta} : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$, the transport equation is often expressed as

$$\frac{\partial}{\partial t}\tilde{\zeta} = -\nabla_{\boldsymbol{x}} \cdot \left(\tilde{\boldsymbol{V}}\tilde{\zeta} - \tilde{D}\nabla_{\boldsymbol{x}}\tilde{\zeta}\right) + \tilde{K}\tilde{\zeta} + \tilde{P}\tilde{\phi}, \tag{1.3}$$

which is also known as the *generic scalar transport model* or *equation* (Banks and Kunisch 1989; Baukal et al. 2000; Kuzmin et al. 2012). The expression above is equivalent to (1.2), but the coefficients have a similar but different interpretation. Depending on the field, different variations of the transport equation (1.2) are preferred, where (1.3) is the most common. As the transport equation can be the result of a linearization, the physical transport coefficients $\tilde{D}$, $\tilde{V}$, $\tilde{K}$, and $\tilde{P}$ may depend on the operating point $\boldsymbol{\zeta}_0$ of the transport variables and their spatial derivatives, i.e., $\tilde{D}$, $\tilde{V}$, $\tilde{K}$, and $\tilde{P}$ become functions of $\boldsymbol{\zeta}_0$, $\nabla_{\boldsymbol{x}}\boldsymbol{\zeta}_0$, etc. However, in practice the underlying dependencies may be unknown and are the subject of study. Treating the coefficients as spatially dependent, i.e., $\tilde{D} : \mathbb{X} \to \mathbb{R}_{>0}$, $\tilde{V} : \mathbb{X} \to \mathbb{R}^n$, $\tilde{K} : \mathbb{X} \to \mathbb{R}$, and $\tilde{P} : \mathbb{X} \to \mathbb{R}$, allows us to avoid dealing with all these possible dependencies. Furthermore, the spatial distribution of the sink/source $\tilde{P}$ is considered to be unknown, but the actuation signal $\tilde{\phi} : \mathbb{T} \to \mathbb{R}$ is to be either known or measured. Finally, for well-posedness, this partial differential equation is constrained by some linear boundary conditions that act on the edge of the spatial domain $\partial\mathbb{X}$ and an initial condition at time t_0 that is compatible with the model and its boundary conditions. The boundary conditions reflect how the transport variable $\tilde{\zeta}$ is connected to the environment. For example, a homogeneous Neumann boundary condition can represent a plane of symmetry or a zero heat flux.

In the literature and the remainder of this thesis the tilde notation is dropped and the operating points are not explicitly considered or mentioned. However, all the considered (linear) models only hold in a sufficiently small region around the considered operating point, which is commonly accepted and can be verified (Berkel et al. 2017a; Berkel et al. 2017b; Slief et al. 2022).

1.2 Applications of transport phenomena covered in this thesis

In this thesis, two different applications will be used to demonstrate the developed methods: heat transport in river stream-beds and heat and particle transport in nuclear fusion reactors. This section will describe the background of these applications in detail.

1.2.1 Transport phenomena in the hyporheic zone

The hyporheic zone is the dynamic interface between an aquifer and a surface water body, such as a stream or river. It includes the streambed, banks and part of the riparian floodplain, and in perennial rivers (i.e., rivers carrying water throughout the whole year) it is often water-saturated. A visualization is given in Figure 1.2. The hyporheic zone is a habitat for a variety of interstitial organisms, a spawning ground for fish and a rooting zone for certain aquatic plants (Schneidewind 2016). Active mixing of groundwater and surface water leads to constant exchange of water, heat, oxygen, carbon, nutrients and pollutants between the aquifer and the stream. These exchange processes have consequences for the water quantity and quality as well as for the stream ecology and health (Buss et al. 2009). Many studies have been devoted to modeling the fluid dynamical and biogeochemical mechanisms occurring in the hyporheic zone (Boano et al. 2014). Transport in the hyporheic zone is inherently three dimensional (3D) and variable in time. One of the most important quantities describing the exchange between groundwater and streams is the water flow per unit area and time, often called the Darcy flux or specific discharge. Direction and magnitude of this flux depend ultimately on differences in pressure between the groundwater and the connected stream section. These pressure heads can vary on a regional scale over many square kilometers (i.e., hydrostatic pressure head differences) as well as on smaller scale, e.g., due to objects on the streambed such as large cobbles or woody debris or due to bedforms such as pool-riffle-pool sequences. These regional and local flow components often superimpose each other. Studies have shown that regional flow entering a stream often has a much stronger vertical component than local flow and discharges predominantly near the stream center (Cuthbert and Mackay 2013; Shanafield et al. 2010). Flow closer to the streambanks has in turn often much stronger horizontal components (Lautz 2010; Reeves and Hatch 2016; Roshan et al. 2014).

A variety of field techniques has been developed to directly or indirectly determine the Darcy flux, including the use of heat as a tracer (Kalbus et al. 2006). Available temperature sensors are generally inexpensive, robust, and reliable compared to other tracers such as saline solutes, fluorescent dyes, radioactive isotopes, etc. (Schneidewind 2016; Leibundgut et al. 2009). Moreover, the natural temperature fluctuations, e.g., the day-night cycle and seasonal cycle, act as a natural

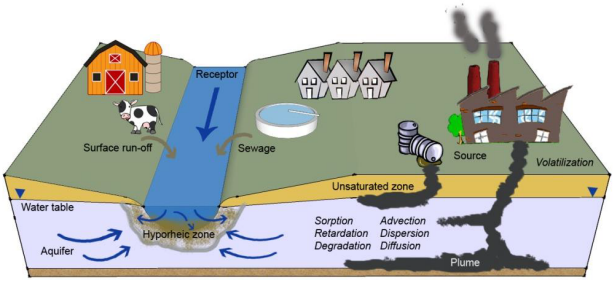


Figure 1.2. A visualization of the hyporheic zone of a typical perennial lowland river where active mixing of surface water and groundwater occurs, eventually resulting in the exchange of oxygen, carbon, nutrients, and contaminants between the stream and the aquifer and leading to a hydrologically and biogeochemically coupled system. Source: Schneidewind 2016.

perturbation signal in the surface water relative to the groundwater temperature. The resulting heat transfer between the surface water and groundwater can then be used to determine the Darcy flux. A simple illustrative example is given in Figure 1.3. The temperature perturbation present in the stream will be better visible deeper in the sediment if the stream is losing water than when the groundwater (without perturbation) is flowing into the stream. To quantify the Darcy flux, a model describing the heat transfer is required. While fully developed 3D modeling codes such as HydroGeoSphere (Brunner and Simmons 2011) that couple surface and subsurface flow and transport processes have recently gained more attention, in perennial streams with considerable vertical flow components, a 1D approach is often applied as it is much less complex and data-intensive.

A standard modeling assumption is that the heat transport through the streambed is governed by the processes of convection and diffusion (Stallman 1965; Rau et al. 2014; Anderson 2005). Hence, by using the convection-diffusion model and temperature measurements at different depths as a function of time, the convection coefficient that is linked to the Darcy flux can be estimated. The considered convection-diffusion model for the temperature $T(x,t)$ is given by the following parabolic partial differential equation (PDE) (Domenico and Schwartz 1997; Boano et al. 2014; Anderson 2005)

$$\frac{\partial}{\partial t}T = D_T\nabla^2 T - \nu_T\nabla T, \tag{1.4}$$

where ν_T and D_T represent the convective and diffusive coefficient, respectively. This equation follows from factoring out the generic transport equation (1.3) under the assumption of a constant flow $\nabla\nu_T = 0$. The Darcy flux q_T is linked to the convection coefficient ν_T via

$$\nu_T = q_T\frac{\rho_w c_w}{\rho c}, \tag{1.5}$$

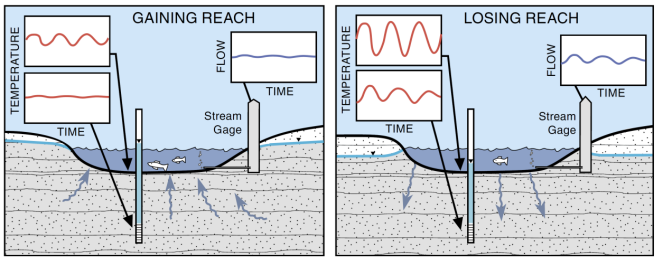


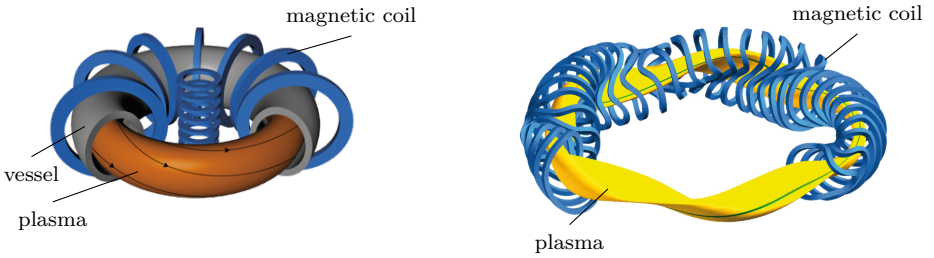
Figure 1.3. The temperature perturbation present in the stream is better visible deeper in the sediment if the stream is losing water (right) then when the ground water (without perturbation) is flowing into the stream (left). Note that in both cases the temperature perturbation is present at both measurement locations and the amplitude decreases deeper in the sediment. Source: Schneidewind 2016; Stonestrom and Constantz 2003; Rau et al. 2014.

where $\rho_w c_w$ and ρc are material properties, i.e., $\rho_w c_w$ is the volumetric heat capacity of water and ρc is the volumetric heat capacity of the water-sediment mixture. The volumetric heat capacity is either determined by taking the product of the specific heat capacities c_w, c and the densities ρ_w, ρ, or from literature or lab testing (Stonestrom and Constantz 2003; Luce et al. 2017). After obtaining the volumetric heat capacity, the Darcy flux q_T can be determined with a simple transformation of the convection coefficient ν_T.

However, in most environmental settings, the thermal properties are heterogeneous due to the different sediment layers, or due to strongly convergent or divergent flows. For such cases, a natural model extension is to include multi-dimensional transport with heterogeneous coefficients, as is hinted in (Rau et al. 2012; Cuthbert and Mackay 2013; Boano et al. 2014; Roshan et al. 2014). Therefore, this thesis will also consider multi-dimensional transport with heterogeneous transport coefficients.

1.2.2 Transport phenomena in nuclear fusion

To use nuclear fusion for power generation, the goal is to fuse two hydrogen isotopes, deuterium and tritium, into helium and a neutron. The difference in mass before and after the reaction results in energy, following Einstein's famous equation $E = mc^2$. The most common fusion reactor designs are based on magnetic confinement in which the fully ionized hydrogen gas (plasma) is confined in a toroidal shape, see Figure 1.4. The reactivity of the plasma inside the reactor is directly related to the number of deuterium and tritium collisions. This is related to the temperature, which allows for the two positively charged ions to overcome the coulomb repulsion, and the density which determines the collision frequency. The third important parameter is the rate at which the plasma loses energy by transporting energy and particles to the plasma periphery. The typical timescale at which the plasma holds its energy is called the energy confinement time τ_E. In the context of this work,



(a) Tokamak. Source: EUROfusion, Reinald Fenke, www.euro-fusion.org

(b) Stellerator. Source: Max Planck institute for Plasmaphysics, www.ipp.mpg.de

Figure 1.4. The two most common magnetic confinement fusion reactor designs, the tokamak (1.4a) and the stellerator (1.4b).

we note that the energy confinement time is associated with transport notably $\tau_E \propto D^{-1}$. The triple product of plasma density n, temperature T, and confinement time τ_e balances the reactivity with the energy losses. If the triple product is above some threshold L, the fusion reaction provides net power (Wesson 2004),

$$nT\tau_e \geq L. \tag{1.6}$$

Due to the density limit in magnetic confinement reactors machines (Greenwald 2002) and the optimal temperature range between 150 to 200 million degrees Kelvin, the only free parameter to increase the triple product is the energy confinement time τ_E that depends on transport. Furthermore, the threshold L also depends on the density and temperature distributions. Hence, transport phenomena determine the performance and efficiency of nuclear fusion reactors. Therefore, transport is intensively studied within the field of nuclear fusion (Lopes Cardozo 1995; Ryter et al. 2010).

The plasma organizes itself in toroidal surfaces of constant magnetic flux and temperature. Therefore, transport of heat and particles is perpendicular to these surfaces and can be described by partial differential equations with a single spatial (radial) coordinate $\rho \in [0, 1]$, see Figure 1.5. However, the heat and particle transport are coupled as the particle flux Γ, electron thermal flux q_e, ion thermal flux q_i, and current density j are driven by the gradients of the electron density n_e, electron temperature T_e, ion temperature T_i, and the electric field potential E (Bourdelle 2005; Citrin et al. 2017; Garbet et al. 2004; Ida 1998; Köchl 2009; Ryter et al. 2010; Mantica et al. 2008). As these relations can be nonlinear and depend on the variables themselves, linearized models are used.

For example, consider the electron particle flux Γ described by the diffusion-convection model

$$\Gamma = -D\nabla n_e - V n_e, \tag{1.7}$$

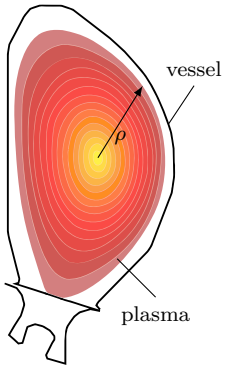


Figure 1.5. Cross section of a magnetically confined plasma and the normalized radius $\rho \in [0, 1]$, that is zero in the center and one at the last closed flux-surface, which is approximately the edge of the plasma. Adopted from Berkel 2015.

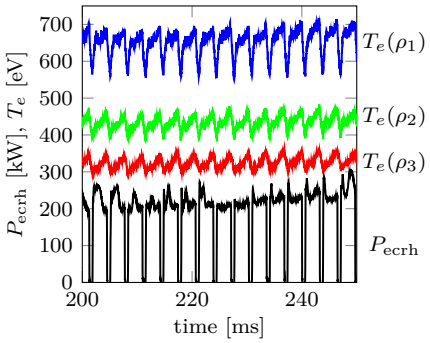


Figure 1.6. ECE measurements of the electron temperature $T_e(\rho)$ from an ECRH perturbation experiment in the RTP-tokamak.

where the transport coefficients D, V are assumed to depend on n_e, ∇n_e, T_e, and ∇T_e, i.e., D and V are functions. The linearization of the particle flux $\tilde{\Gamma}(\rho, t) = \Gamma(\rho, t) - \Gamma_0(\rho)$ around the operating functions $\mathcal{O} = \{\Gamma_0, n_{e,0}, T_{e,0}, \ldots\}$ is given by

$$\begin{aligned} -\tilde{\Gamma} = {} & D|_{\mathcal{O}} \nabla \tilde{n}_e + V|_{\mathcal{O}} \tilde{n}_e \\ & + \left(\nabla \tilde{n}_e \left.\frac{\partial D}{\partial \nabla n_e}\right|_{\mathcal{O}} + \tilde{n}_e \left.\frac{\partial D}{\partial n_e}\right|_{\mathcal{O}} + \nabla \tilde{T}_e \left.\frac{\partial D}{\partial \nabla T_e}\right|_{\mathcal{O}} + \tilde{T}_e \left.\frac{\partial D}{\partial T_e}\right|_{\mathcal{O}} \right) \nabla n_{e,0} \\ & + \left(\nabla \tilde{n}_e \left.\frac{\partial V}{\partial \nabla n_e}\right|_{\mathcal{O}} + \tilde{n}_e \left.\frac{\partial V}{\partial n_e}\right|_{\mathcal{O}} + \nabla \tilde{T}_e \left.\frac{\partial V}{\partial \nabla T_e}\right|_{\mathcal{O}} + \tilde{T}_e \left.\frac{\partial V}{\partial T_e}\right|_{\mathcal{O}} \right) n_{e,0}. \end{aligned} \tag{1.8}$$

The electron particle flux and the electron temperature are coupled as both depend on each other. Assuming that the coupling is weak and using sufficiently small perturbations in the density, these contributions can be neglected, thus $\tilde{T}_e \approx 0$ and $\nabla \tilde{T}_e \approx 0$. Furthermore, the particle flux cannot be measured directly, but the electron density can, e.g., with Thomson scattering, (far infrared) interferometry, or reflectometry (Hutchinson 2002). Consequently, the considered model describes the electron density rather than the flux. The evaluation of the electron density is described by the equation for the conservation of particles (Gentle 1988; Bishop and Connor 1990; Haas et al. 1991) that yields

$$\frac{\partial}{\partial t} \tilde{n}_e = -\nabla \tilde{\Gamma} + \tilde{S}_p, \tag{1.9}$$

where $\tilde{S}_P$ is a particle source/sink. Substituting the particle flux (1.8) into (1.9)

with $\tilde{T}_e = 0$ and $\nabla\tilde{T}_e = 0$ results in

$$\begin{aligned}\frac{\partial}{\partial t}\tilde{n}_e = -\nabla\Bigg(\Bigg(D|_{\mathcal{O}} + \frac{\partial D}{\partial \nabla n_e}\bigg|_{\mathcal{O}} \nabla n_{e,o} + \frac{\partial V}{\partial \nabla n_e}\bigg|_{\mathcal{O}} n_{e,o}\Bigg)\nabla\tilde{n}_e \\ + \Bigg(V|_{\mathcal{O}} + \frac{\partial D}{\partial n_e}\bigg|_{\mathcal{O}} \nabla n_{e,o} + \frac{\partial V}{\partial n_e}\bigg|_{\mathcal{O}} n_{e,o}\Bigg)\tilde{n}_e\Bigg) + \tilde{S}_p,\end{aligned} \tag{1.10}$$

which can be simplified to the generic scalar transport equation

$$\frac{\partial}{\partial t}\tilde{n}_e = -\nabla\left(\tilde{D}\nabla\tilde{n}_e + \tilde{V}\tilde{n}_e\right) + \tilde{S}_p \tag{1.11}$$

with the new diffusion and convection coefficients $\tilde{D}$ and $\tilde{V}$. Note that even when D and V are constant, the dependency on the other variables makes $\tilde{D}$ and $\tilde{V}$ spatially varying functions. So again, $\tilde{D} : \mathbb{X} \to \mathbb{R}_{>0}$ and $\tilde{V} : \mathbb{X} \to \mathbb{R}$. Hence, estimating $\tilde{D}$ and $\tilde{V}$ as spatially varying functions simplifies the estimation problem by eliminating the dependency of the other variables (n_e, ∇n_e, T_e, and ∇T_e). Moreover, by estimating $\tilde{D}$ and $\tilde{V}$ around different operating points, a new estimation problem can be formulated to determine the dependency on the other variables. In this way, the complicated task of estimating all the transport coefficients and its dependencies is split into separate smaller problems. This thesis only deals with the estimation of the coefficients as space dependent parameters and does not consider the dependencies.

The derivation of the particle transport in (1.11) can equivalently be performed for the (electron) heat transport (Gentle 1988; Bishop and Connor 1990; Haas et al. 1991; Lopes Cardozo 1995) which results in

$$\frac{\partial}{\partial t}\tilde{T}_e = -\nabla\left(\tilde{D}\nabla\tilde{T}_e + \tilde{V}\tilde{T}_e\right) + \tilde{S}_q. \tag{1.12}$$

For electron heat transport, the perturbation required to estimate the coefficients is introduced by modulating a heat source $\tilde{S}_q$. There are different methods for heating, of which electron cyclotron resonance heating (ECRH) is used most commonly for perturbative experiments. The reason is that ECRH provides localized heating, of which the deposition is often assumed to be well known from theory. The resulting temperature evolution is then measured at different spatial position via electron cyclotron emission (ECE) (Wesson 2004; Hutchinson 2002). As this heating is applied locally, there are two common approaches to determine the transport coefficients. Either, it is assumed that the source is placed completely outside the domain of interest and the resulting temperature perturbation is considered a boundary input, i.e., $\tilde{S}_q = 0$ (Lopes Cardozo 1995; Ryter et al. 2010). Alternatively, the source is placed within the domain of interest and needs to be estimated simultaneously with the other transport coefficients. The simultaneous estimation of the source and transport coefficient is particularly important as various errors and plasma conditions can lead to source distributions that are broader than predicted (Kirov et al. 2002; Chellaï et al. 2018; Chellaï et al. 2021; Brookman

et al. 2021; Slief et al. 2022; Slief et al. 2023). Examples are plasma instabilities such as edge turbulence and misalignment errors originating from the design and uncertainty of the launching system. For the estimation of the source distribution, it is assumed that the source is spatially fixed (with the fixed spatial source profile $\tilde{P} : \mathbb{X} \to \mathbb{R}$), and that the intensity of the source can be varied over time by the modulation signal $\tilde{\phi} : \mathbb{T} \to \mathbb{R}$, i.e., $\tilde{S}_q = \tilde{P}\tilde{\phi}$. This is a consequence of the linearity assumption by applying a perturbation around an operating point.

Both heat and particle transport adhere to the generic scalar transport description introduced in (1.3). Therefore, this thesis deals with the (simultaneous) estimation of the transport coefficients $\tilde{D}$, $\tilde{V}$, $\tilde{K}$ and the source $\tilde{P}$. The reaction/damping term $\tilde{K}$ is not considered for these applications, as $\tilde{K}$ has a negligible effect on the transport variable (Boorn 2021) and is therefore intrinsically difficult to estimate. However, for completeness this coefficient is considered as it may be a more dominant term in other fields.

1.3 Methods for estimating transport coefficients

In contrast to the transport variables, e.g., temperature, transport coefficient can generally not be measured directly. To accurately describe transport phenomena with a mathematical model, the corresponding transport coefficients need to be identified. Typically, the identification method use spatio-temporal measurements of the transport variable. This is a generic problem. As researchers across many different disciplines face the same challenge, many different methods for this identification have been developed.

Below, we present an overview the different identification methodologies in the time domain methods and frequency domain.

1.3.1 Time domain methods

The classic overview of time domain methods for the estimation of transport coefficients is given in the monograph by (Banks and Kunisch 1989) and references therein. However, these methods often consider constant coefficients, and known initial and boundary conditions. Therefore, different methodologies have been developed to deal with spatially varying parameters (Kravaris and Seinfeld 1985; Banks and Lamm 1985; Kunisch and Peichl 1991; Mochi et al. 1999), unknown initial conditions (Eason 1976; Roques et al. 2014; Zhao et al. 2016), and unknown boundary conditions (Eason 1976; Aihara and Bagchi 1988; Rossi et al. 2004; Ebrahimian et al. 2007). In general, all the methodologies take a similar approach. The transport coefficient estimation problem is posed as an optimization problem with the goal to minimize the difference between the measurement data and the model values at the dedicated set of sample points. In some special cases, such as constant parameters, the analytical solutions are known and can be used. When the analytical solution is unknown, the model is solved using numerical methods. To do so, both the spatial as well as the temporal coordinate are discretized.

For the spatial coordinate, the most common approaches are finite difference, finite volumes, finite elements, or spectral decomposition (Banks and Kunisch 1989; Eason 1976; Belov 2002; Vogel 2002; Tarantola 2005; Kirsch 2011). This reduces the problem to a system of coupled ordinary differential equations which is solved using standard numerical methods. However, as both the solution to the model and optimization problem are solved iteratively, the computational cost can increase rapidly. Another problem that arises in case of non-constant coefficients are highly oscillatory artifacts due to noise amplifications (Vogt et al. 2010). Hence, regularization plays an important role for these methods (Banks and Kunisch 1989; Vogel 2002; Ito and Kunisch 2008).

New advancements in the field of machine learning brings alternative promising methods to estimate transport coefficients (Benosman and Farahmand 2017; Raissi et al. 2017; Rai and Tripathi 2019; Raissi et al. 2020). Where traditional time-domain methods first need to compute the solution to the model and then determine the error by comparing it to the measurement data, machine learning approaches interpolate the input and output data where the model acts as a constraint between the two interpolations. In this way, the solution to the model and the best parameter fit are solved simultaneously. Although, the above four references mainly cover constant parameters, it should also be applicable for spatially varying coefficients.

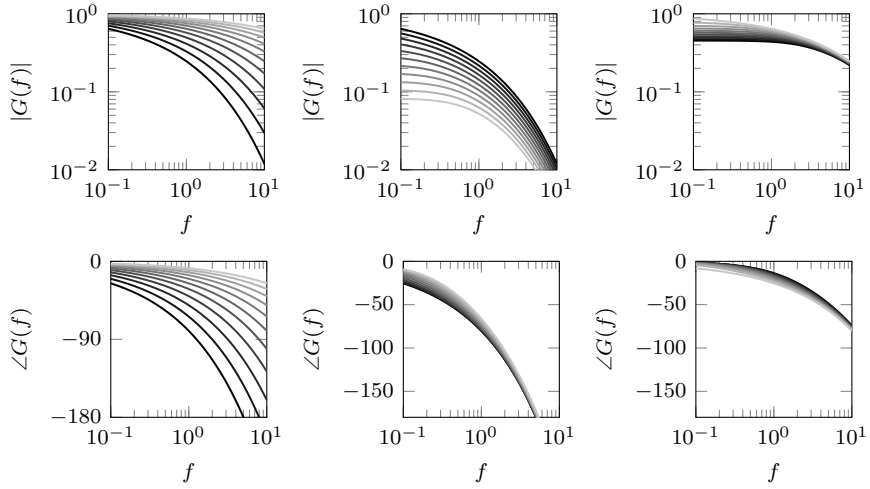
The main disadvantages of time-domain methods is that they either need to estimate the time derivative that is very sensitive to noise or they have to compute the solution to the model for each step in the iterative optimization routine when estimating the transport coefficients. Transforming the problem to the frequency domain has many advantages, including not having to deal with the time evolution or time derivative.

1.3.2 Frequency domain methods

By applying the Fourier transform to the linear or linearized model as well as to the measurement data, the model is cast into the frequency domain. In this way, the partial differential equation describing the transport phenomena (1.3) becomes a complex-valued ordinary differential equation (ODE)

$$i\omega\tilde{Z} = -\nabla_{\boldsymbol{x}} \cdot \left(\tilde{\boldsymbol{V}}\tilde{Z} - \tilde{D}\nabla_{\boldsymbol{x}}\tilde{Z}\right) + \tilde{K}\tilde{Z} + \tilde{P}\tilde{\Phi}, \tag{1.13}$$

with angular frequency ω and the Fourier transformed variables $\tilde{Z} = \mathcal{F}\{\zeta\}$ and $\tilde{\Phi} = \mathcal{F}\{\tilde{\phi}\}$, where $\mathcal{F}$ is defined as the (discrete) Fourier transform. The most well known frequency domain approaches exploit this by using the analytical solution to the model by considering a constant diffusion and convection term, no damping term, no source term, and one boundary as an input where the perturbation is diminished at infinity (Lopes Cardozo 1995; Luce et al. 2013). The two (constant) coefficients are then estimated using the amplitude and phase difference between two spatial points at a single frequency. To put it differently, the frequency response function $G(\omega) = \frac{Z(x_2,\omega)}{Z(x_1,\omega)}$ is estimated using a single frequency. In this way the



(a) diffusion D **(b)** convection V **(c)** damping K

Figure 1.7. The effect of the different (constant) transport coefficients on the frequency response function $G(f) = \frac{Z(x_2,f)}{Z(x_1,f)}$ on a semi-infinite domain. The gray scale indicates the value of the coefficient, where a lighter color means a larger value. The used ranges are $D \in [0.1, 10]$, $V \in [0, 1]$ and $K \in [-10, 0]$.

number of equations and unknowns are equal causing these methods to be very susceptible to noise. On the other hand, these frequency response functions based on the semi-infinite domain clearly illustrates how the different transport coefficients manifest in the frequency domain (see Figure 1.7).

Therefore, these methods have been put forward by including the damping term, estimating the uncertainty, using a maximum likelihood estimator, using multiple frequencies (Berkel et al. 2013; Vandersteen et al. 2015; Sohn and Harris 2021), and multiple spatial measurement locations (Vandersteen et al. 2015). In this way, the parameter estimation problem is overdetermined and less prone to noise. However, these methods all have the underlying semi-infinite domain assumption, i.e., the applied perturbation signal that drives the system away from it's steady state can only travel in one direction and the parameters are assumed to be constant from the measurement interval to infinity. To isolate the coefficient estimation problem to the measured (local) domain, (Berkel et al. 2014c; Schneidewind et al. 2016; Berkel et al. 2019) uses three measurements, two as boundary conditions and one as an output. By considering constant parameters in this isolated local domain, analytical transfer functions in combination with a maximum likelihood estimator are used to determine the coefficients.

Analyzing spatially varying coefficients, there is no generic analytic transfer available. Therefore, (Das et al. 2019) uses a numerical approximation of the transfer function by parameterizing the unknown coefficients including the source,

and the measurements as boundary conditions. The spatially varying coefficients are estimated by minimizing the (weighted) least squares error between the measurements and the outputs generated by the transfer functions. Similar to the time domain methods, this methodology suffers from highly oscillatory artifacts caused by noise amplifications.

All the previously mentioned frequency domain methods use transfer functions to estimate the transport coefficients. Another approach is to consider the transport equation after integration over the spatial coordinate (Escande and Sattin 2012; Sattin et al. 2012; Brookman et al. 2021; Slief et al. 2022). This removes one of the spatial derivatives. Consequently, the model does no longer include spatial derivatives of the coefficients. Under the assumption that the source is either known or outside the domain, the spatially varying coefficients can be determined separately at each spatial point by solving a linear regression problem (Escande and Sattin 2012; Sattin et al. 2012). For the simultaneous estimation of the coefficients and the source, the unknown coefficients are first parameterized and then determined by minimizing a least squares criterion that considers multiple frequencies and spatial measurement locations (Brookman et al. 2021; Slief et al. 2022).

Generally speaking, frequency domain methods for constant parameters and noisy measurements are amply available in literature (Lopes Cardozo 1995; Luce et al. 2013; Berkel et al. 2013; Vandersteen et al. 2015; Berkel et al. 2014c; Schneidewind et al. 2016; Berkel et al. 2019; Sohn and Harris 2021), while for more complex models, e.g., spatially varying coefficients, literature is scarce, especially when considering measurement uncertainty.

1.3.3 Time versus frequency domain

In this thesis the considered models and its dynamics are linear, therefore the time and frequency domain are equivalent. However, the access to information differs from one domain to the other. This is discussed in detail in (Pintelon and Schoukens 2012, pp. 522-527). Here, we briefly highlight a few of the main (dis)advantages for estimating transport coefficients. Naturally, for real-time applications, the time domain methods have numerous advantages over the frequency domain. However, the focus of this thesis is on parameter identification of recorded experiments instead of real-time applications. The use of recorded experiments allows noncausal data processing methods before the coefficient identification step, and gives working in the frequency domain the following advantages over the time domain:

1) The information carrying spectral content of perturbation experiments for PDEs is typically sparse. By only considering the information carrying frequencies, i.e., with a significantly high signal-to-noise ratio, the resulting datasets are smaller and reduce the required computational effort during the identification step.

2) As the information carrying spectrum is typically sparse and the added

stochastic uncertainty is spread over the complete spectrum, the effect of noise on the information carrying spectrum is significantly reduced.

3) The standard assumption of circular complex normal distributed noise on the Fourier coefficients is a weak assumptions on time domain noise distribution as the central limit theorem shows that several different additive noise distributions result in circular complex normal distribution of the Fourier coefficients (Billingsley 2012).

4) The PDE reduces to a limited set of complex-valued ordinary differential equations, which significantly reduces the analysis computation. This property is commonly used in the standard frequency domain methods for constant parameter (Luce et al. 2013; Lopes Cardozo 1995).

Although there are many advantages of working in the frequency domain, literature on frequency methods to estimate spatially varying coefficients is scarce. This in contrast with time domain methods that have received a lot of attention throughout the years. Also in the last years, recent advancements regarding machine-learning are brought to the field of transport coefficient identification in time domain, but not in the frequency domain. Hence, the aim of this thesis is to close the gap between time and frequency domain methods for the identification of spatially varying coefficients and explore new avenues for machine-learning approaches in the frequency domain.

1.4 Objective and contributions

As explained in the previous sections, estimation of spatially varying coefficients in a generic scalar transport model is crucial for many applications and can be the first step for identifying coefficients that depend on other variables. As these coefficients cannot be measured directly, they need to be inferred from limited measurement data that can be polluted with (or otherwise affected by) noise. This noise may be seen as stochastic uncertainty. Moreover, it is important to note that our goal is to quantify coefficients that have a physical meaning rather than minimizing the modeling error. Therefore, the research question for this thesis is formulated as follows:

> *How can we accurately estimate the physical spatially varying coefficients in a generic scalar transport model in the presence of stochastic uncertainty?*

The following sub-research questions are formulated to provide the path towards answering the main research question:

- How does data processing and modeling choices affect the estimated coefficient?

- When does the estimation of spatially varying coefficients require regularization and how does this affect the estimates?
- How to deal with stochastic uncertainty when estimating spatially varying coefficients?

The first sub-question: **How does data processing and modeling choices affect the estimated coefficient?** is answered via the following two contributions and are covered in Chapters 2 and 3.

Contribution I *In this thesis it is demonstrated how advanced data processing can significantly improve the coefficient estimates.*

For the coefficient estimation process it is important that the model and experimental data are compatible. For example, the frequency domain models only consider the forced response of the system, while the measurements can also contain transients and noise. To make the model and measurements compatible, the unwanted contributions, in our case transients and noise, are removed by the local polynomial method (Pintelon et al. 2010a; Pintelon et al. 2010b; Pintelon and Schoukens 2012). As the importance and effect of correct data processing is not widely available within the different communities, we published work in the nuclear fusion and hydrology community showing the quality difference between processed and non-processed experimental datasets (Berkel et al. 2020; Berkel et al. 2023). Moreover, we show how this difference affects the coefficient estimation process (Berkel et al. 2023).

Contribution II *This thesis provides an extension of a maximum likelihood estimator for constant transport coefficients to include multiple arbitrary sensor locations. This extension enables the comparison on how different choices for the boundary conditions affect the coefficient estimates.*

In the field of hydrology, the semi-infinite domain is still the leading model choice that is used to estimate transport coefficients. With the semi-infinite domain, one measurement is used as boundary condition while all other measurements are considered as outputs, implying that all parameters are constant for all spatial values up to infinity. The method introduced in (Berkel et al. 2014c; Schneidewind et al. 2016) uses two measurements as boundary conditions while considering only one measurement as output. In this way, the estimation problem is isolated to the measured domain where the coefficients are only assumed to be constant over this locally bounded domain instead of over an unbounded spatial domain. For a fair comparison with the semi-infinite model, the method in (Berkel et al. 2014c; Schneidewind et al. 2016) has been extended such that it can consider any number of outputs. In this way, a fair comparison between the semi-infinite and

bounded domain model assumption can be made and it demonstrates that the bounded-domain model is the safer model choice. With a simulation example this thesis shows that the semi-infinite model can estimate the groundwater flow with a wrong direction. Application on an experimental dataset from the field of hydrology, also given in this thesis, shows that locally bounded domain results in better and more trustworthy estimates. The extension of the bounded domain model and the effect of the semi-infinite domain model are published in (Kampen et al. 2022d).

The next sub-question: **When does estimating spatially varying coefficients require regularization and how does this affect the estimates?** is answered by the following contributions and covered in Chapter 4.

Contribution III *This thesis contributes in formulating the estimation of spatially varying transport coefficients as a linear least squares problem where the global optimum is given by a closed-form solution. It is shown that this closed-form solution makes the methodology fast and reliable.*

The standard approach towards spatially varying coefficient estimation is to minimize the output error, i.e., the difference between the measurements and model value at the measurement locations. Typically, the models are high-order approximations of the underlying infinite dimensional model. If the coefficients or model values in-between measurement locations are unconstrained, the estimated coefficients often contain large (non-physical) oscillatory artifacts. Classically this is resolved by regularization, i.e., constraining the unknown coefficients. By minimizing the equation error, i.e., the difference between the measured or estimated state and the model state, the estimation problem can be formulated as a linear least squares problem that allows direct calculation of the global optimum by solving a linear matrix equality against a low computational cost.

Contribution III.A *In addition to contribution III, it is shown that the least squares optimization technique for coefficient estimation gives more insight in how regularization affects the coefficient estimation.*

The fast computation of the optimum allows to test different parameterizations of the coefficient functions and investigate the effect on the estimated coefficients in a machine learning like manner. In this way, it is shown how different parameterizations of the coefficients act as regularization and how it affects the estimates. Moreover, using equation error formulation in the frequency domain clearly shows where en when regularization is applied and the nature of the oscillatory artifacts.

These results are published in (Kampen et al. 2021c).

Contribution III.B *In addition to contribution III, it is shown that the least squares optimization technique can be extended such that it can deal with coefficients for transport in multiple spatial dimensions.*

For nD transport coefficients, the number of variables increases rapidly. Having access to a method that can estimate the transport coefficients against a low computational cost while guaranteeing the (global) optimal solution is important. Extending the methodology that considers the estimation of 1D spatially varying coefficients as a linear least squares problem such that it is applicable for the estimation of multidimensional transport coefficients, results in a closed-form solution that has a low computational cost and guarantees the optimal solution. This extension is published in (Kampen et al. 2023b).

The last sub-question: **How to deal with stochastic uncertainty when estimating spatially varying coefficients?** is answered by the following contributions and covered in Chapter 5.

Contribution IV *This thesis contributes in adding state estimation via Gaussian process regression and maximum likelihood estimation of the coefficients to deal with stochastic uncertainty present in the measurements.*

In this thesis the estimation of spatially varying coefficient is posed as a linear least-squares problem. In case the measurements are polluted with noise, i.e., with stochastic uncertainties, the closed-form solution of this method is biased. To deal with these uncertainties in an optimal way, Gaussian process regression is used to determine the probability density function of the state. Then, a maximum likelihood estimator is used to determine the spatially varying coefficients that are most likely to result in the estimated probability density function of the state. This work is published in (Kampen et al. 2023a).

Together these contributions help answering the research question **How can we accurately estimate the physical spatially varying coefficients in a generic scalar transport model in the presence of stochastic uncertainty?** This is reflected by the last contribution:

Contribution V *This project delivered MATLAB® implementations of all the proposed methodologies that are applied and verified on both, simulated and experimental data.*

The working of the developed methods has been verified using simulations by

implementing them in MATLAB®. Thereafter, these implementations have found their way to practice as they have been used to estimate transport coefficients from experimental data within two different fields, namely hydrology and nuclear fusion (Kampen et al. 2022d; Schneidewind et al. 2021; Slief et al. 2022; Slief et al. 2023).

1.5 Outline

The outline of the thesis follows the contributions as mentioned above. First, the data processing is covered in Chapter 2. Thereafter, in Chapter 3, the constant parameter case with stochastic measurement uncertainty is covered, including the effect of different model assumptions regarding the boundary conditions. Then, in Chapter 4, the focus is on a new closed form solution to estimate spatially varying coefficients. Starting with the 1D problem without uncertainty, to gain more insight in the effect of highly oscillatory artifacts and how to overcome them. In Chapter 5, the closed form solution is extended to make it applicable for transport in multiple spatial dimensions. In Chapter 6, the closed form solution is replaced with a maximum likelihood estimator such that it can deal with stochastic uncertainties in an optimal way. This maximum likelihood estimator requires probability density function of the state that is determined via Gaussian process regression. Finally, in Chapter 7, the research questions are answered followed by a discussion on future work. As the chapters are based on individual publications, all chapters can be read individually.

2

A framework for reliable coefficient estimation

Abstract - Monitoring hyporheic flow and exchange fluxes between groundwater and surface water is vital for integrated water resource management, analysis of solute transport and water chemistry, as well as predictions of ecological implications. The last decade has seen an increased interest in field techniques that use analysis of temperature time series at two or more depths in streambed sediment to estimate hyporheic and exchange fluxes because of their low cost, robustness, and ease of installation. The steps necessary to systematically address this estimation problem are often intermingled in the literature. This conflation of steps has perhaps limited advances in better characterizing the fluxes and their uncertainty such as the faux requirement for single-sinusoid forcing. Therefore, the intent of this work is to clearly outline the steps and choices inherent to the analysis of temperature time-series from the viewpoint of a system theory perspective. Thereby, we clarify potential error sources at each step, and present signal processing and estimation methods in hope of clarifying misconceptions and providing a sound theoretical basis from which the field can progress on the flux estimation challenge. To support this clarification, signal processing and parameter estimation choices and their quantitative consequences are illustrated using a laboratory dataset.

The content of this chapter is based on: Berkel, M. van, C. H. Luce, R. J. R. van Kampen, et al. (2023). "A framework to improve the reliability of temperature-based estimates of flow and diffusion in streambeds with frequency domain examples". In: *in preparation*.

2.1 Introduction

Inverse modeling approaches using temperature measurements in streambeds are some of the key techniques used in quantifying exchange fluxes between rivers and groundwater (Anibas et al. 2016; Bryant et al. 2020; Irvine et al. 2017; Jensen and Engesgaard 2011; Zhou et al. 2017). Understanding the exchange fluxes between groundwater and surface water is essential for evaluating stream metabolism and biogeochemical cycling in river corridors (Boano et al. 2014; Boulton 2007; Boulton et al. 1998; Boulton et al. 2010; Brunke and Gonser 1997; Brunke and Gonser 1997; Buffington and Tonina 2009; Harvey and Gooseff 2015; Krause et al. 2011; Krause et al. 2017) and to better understand the ecological functioning of hyporheic zones and their role and interactions with river floodplains and the wider catchment (Baxter and Hauer 2000; Boulton et al. 2010).

A number of different methodological approaches have been developed in recent years, allowing increased convenience in estimation of fluid fluxes and thermal properties in streambeds from point observations of temperature time series (Ford et al. 2021; Irvine et al. 2015; Koch et al. 2016; Munz and Schmidt 2017). Many of these approaches are based on the initial assumptions framed by (Stallman 1965) and later extended by (Goto et al. 2005; Hatch et al. 2006; Keery et al. 2007). At the same time, perceived limitations of Stallmans approach have been discussed as well (e.g., (Briggs et al. 2014; Chen et al. 2018; Lautz 2010; Lautz 2012; Rau et al. 2010)). Here, we introduce a generalization of the model of Stallman in terms of system theory representation, which allows the connection between Stallman's original work and multi-frequency approaches in terms of system theory. This requires a clear interpretation of the different components in the estimation procedure, such as the roles of model framing relative to choices in data processing, and how, when they are interchangeable and when not. Although not demonstrated here, this extension is also important for heterogenous and multi-dimensional parameter estimation challenges. Therefore, this work discusses an integrated framework for the processes of estimating parameters using heat as a tracer in the following three steps:

1) Model selection (Section 2.2);

2) Signal processing (Section 2.3);

3) Estimation methods (Section 2.4).

Evaluating these components and showing their links will allow scientists to improve their decisions about method selection and understand the consequences of the individual decisions. We begin with a discussion of the model choices and how they frame and constrain solution methods and sensitivities to different errors. We follow this by quantitatively demonstrating the consequences of multiple choices available for both signal processing and parameter estimation based on the signal processing results. This allows quantitative evaluation of alternative signal processing and parameter estimation approaches within the range of suitable problem framings.

The signal processing and parameter estimation techniques that we evaluate in this work are only available for some model forms. When these signal processing and parameter estimation techniques are performed in a way that improves precision and provides information about uncertainty, there is, in turn, capacity to evaluate the appropriateness of the model form, closing the loop.

An important contribution in this paper is the rigorous and independent demonstration of the merit of specific choices in the signal processing and parameter estimation steps in the context of real measurements, which we have not seen in the literature previously and which has been missing from our previous publications. As such, this work is closely linked to our multi-frequency work (e.g., (Kampen et al. 2022d; Vandersteen et al. 2015)) and choices made therein, which in turn is based on system theory considerations (Antsaklis and Michel 1997; Curtain and Zwart 1995; Hespanha 2018; Khalil 2002; Oppenheim et al. 1997; Pintelon and Schoukens 2012; Söderström and Stoica 1989). This effort is supported by applying the method steps to experimental data from a laboratory set-up taking it away from a pure mathematical analysis while keeping the outcomes comprehensible and demonstrating the merit of specific choices for a known model. This shows that spectral components contribute important system information, i.e., its parameters, and the merit of methodologies to optimally extract them from the data and process them to achieve model parameter estimates. This also allows to compare different combinations of spectral information showing that even in total absence of strong sinusoidal component such as the diurnal cycle or a self-induced forcing, still the intrinsic parameters can be estimated from the data which resembles more noise than signal.

2.2 Model structure choices and their consequences

A model (taking a broad sense of the word) defines a relationship between the data and the parameters that we want to estimate. Model selection is consequently the starting point of every (model) parameter estimation exercise. In the context of estimating surface water-groundwater interaction using heat as a tracer, the model we consider is the advection-diffusion equation (ADE), which is based on the energy balance, and the parameters of interest are the thermal properties and fluid velocity. As such, inflow and outflow of energy are defined via the boundaries and the initial state of energy, i.e., boundary conditions and initial conditions. The 1D model that is commonly used to determine specific discharge in and across streambeds (e.g., (Anibas et al. 2018; Gariglio et al. 2013; Jensen and Engesgaard 2011; Lautz 2012)) is given by,

$$\frac{\partial T(z,t)}{\partial t} = \frac{k}{c_m \rho_m} \frac{\partial^2 T(z,t)}{\partial z^2} - q_f \frac{c_f \rho_f}{c_m \rho_m} \frac{\partial T(z,t)}{\partial z}, \tag{2.1}$$

where T is the temperature, t is time, z is the vertical coordinate into the bed, k is the thermal conductivity of water and sediment, q_f is the flux of water per

unit area, c is the specific heat of either the mixture (m subscript) or water (f subscript) and ρ is the density. Furthermore, (2.1) is incomplete without defining the initial condition at $T(z, t = 0)$ and the boundary conditions on the domain $[z_1, z_2]$, e.g., $T(z_1, t) = f_1(t)$ and $T(z_2, t) = f_2(t)$.

The ADE as written in (2.1) represents the specific case where everything is uniform in space and invariant in time with the exception of $T(z, t)$. Further, in this common usage, there is no dependency of the model parameters on temperature or any other parameter. It is important to analyze some other representations of the ADE, both in terms of parabolic partial differential equations (PDEs) and boundary conditions, to recognize the choices we make when selecting a particular solution method.

2.2.1 Variations on the (linear) advection-diffusion equation

To illustrate how the representation of (2.1) influences parameter estimates, we analyze the ADE in more general terms

$$\frac{\partial \left(c_m \rho_m T(z,t)\right)}{\partial t} = \frac{\partial}{\partial z}\left(k \frac{\partial T(z,t)}{\partial z}\right) - \frac{\partial \left(c_f \rho_f q_f T(z,t)\right)}{\partial z}, \tag{2.2}$$

where dependencies of the model parameters on temperature and other varying model parameters (can) exist. This results in a number of different models depending on how the parameters (c_m, $_m$, k, q_f) are expected to vary in time and space, which are summarized in Table 2.1. We show how these models belong to different model classes and that the methods we can use to solve them are different. For simplicity, we only consider here one spatial dimension (1D) and assume unknown linear boundary conditions and initial conditions (see Section 2.2.3). A 1D approach places constraints on the mass balance, allowing simplifications that would not exist under multi-dimensional flow conditions.

There are a few points worth noting from Table 2.1 with respect to selecting solution methods. Both (2.1) and (2.3) are linear, time invariant models and have well-defined Fourier representations (similar to (Stallman 1965)). Whereas (2.4) and (2.5) belong to the class of time-varying models and as such do not have a well-defined Fourier representation. They should be analyzed in the time domain or using joint Fourier-time representations (Lataire and Pintelon 2011), the simplest being a Fourier transform of a sliding window (Luce et al. 2013). They can also be handled by fitting parameters for finite difference implementations of (2.4) when supporting information is available (e.g., (Koch et al. 2016; McAliley et al. 2022)). Moreover, analyzing parameter dependencies in this framework shows that the estimated model parameters will unlikely be the averaged model parameters over time, which is empirically supported by (DeWeese et al. 2017). Systematic treatments of (2.4) and (2.5) are a field of ongoing research.

A more recent class of systems that is being considered in heat transport (Zhang et al. 2021) is based on fractional models (Benson et al. 2000; Schumer et al. 2001; Schumer et al. 2003a; Schumer et al. 2003b) which are, for instance, used when there

Table 2.1. Partial differential equations describing 1D heat transport given particular assumptions about parameter variations in time and space. Subscripts m, s, f denote mixture, solid, and fluid (water) fractions.

coefficient properties	**model variations**	
uniform time-invariant	$\frac{\partial T}{\partial t} = \frac{k}{c_m \rho_m} \frac{\partial^2 T}{\partial z^2} - q_f \frac{c_f \rho_f}{c_m \rho_m} \frac{\partial T}{\partial z}$	(2.1)
non-uniform time-invariant	$\frac{\partial T}{\partial t} = \frac{k(z)}{c_m(z) \rho_m(z)} \frac{\partial^2 T}{\partial z} + \left(\frac{\partial k(z)}{\partial z} - c_f \rho_f q_f(z) \right) \frac{1}{c_m(z) \rho_m(z)} \frac{\partial T}{\partial z}$	(2.3)
uniform time-varying	$\frac{\partial T}{\partial t} = \frac{k(t)}{c_m(t) \rho_m(t)} \frac{\partial^2 T}{\partial z^2} - q_f(t) \frac{c_f \rho_f}{c_m(t) \rho_m(t)} \frac{\partial T}{\partial z}$	(2.4)
non-uniform time-varying	$\frac{\partial T}{\partial t} = \frac{1}{c_m(z,t) \rho_m(z,t)} \left[k(z,t) \frac{\partial^2 T}{\partial z^2} + \left(\frac{\partial k(z,t)}{\partial z} - c_f \rho_f q_f(z,t) \right) \frac{\partial T}{\partial z} \right]$	(2.5)

is heterogeneity in the fluid velocity field. The classification of fractional models depends on the definition of fractional derivatives in these models. In case they have a suitable definition in the Fourier domain and/or Laplace domain, such as Riesz and Caputo fractional derivatives (Zhang et al. 2021), they generally belong to the class of linear time invariant models (Magin 2004) and all the signal processing tools described in Section 2.3 apply. The advantage of fractional descriptions is that they no longer need rational forms of the Laplace variable s or $\sqrt{s}$ (appearing in the Stallman solution), but now also fractions can be used, i.e., s^α with α being some non-integer (Zhang et al. 2021). Hence, fractional descriptions provide much more freedom. The downside is that these models are more difficult to interpret as often the physical relationship to the model parameters is reduced.

2.2.2 Boundary conditions for the advection-diffusion equation

The conceptual model includes not only the ADE in (2.1), but also the corresponding boundary and initial conditions. The initial conditions are generally ignored in the frequency representation of (2.1) but will be considered in more detail in Section 2.3.

For the problem in (2.1) to be well-posed, a unique solution must exist, and corresponding boundary conditions need to be defined. Consequently, inverse solutions for the estimation of parameters in the ADE usually apply a linear boundary condition of some form. For our ADE proposed by (Stallman 1965) and explored by others since (Goto et al. 2005; Hatch et al. 2006; Keery et al. 2007; Luce et al. 2013; Vandersteen et al. 2015), a Dirichlet type top boundary condition (temperature directly specified over time) is used (e.g., $T(z = z_1, t) = f(t)$) while the lower boundary condition is the semi-infinite boundary condition with temperature specified as a constant as z goes to infinity. Hence, the semi-infinite

boundary condition implicitly assumes uniform conditions to substantial (infinite) depth, which is rarely the case. Alternatively, a stated bottom boundary condition at a specified depth can be used (e.g., $T(z = z_2, t) = f_2(t)$) (Schneidewind et al. 2016), and its importance with respect to the correct estimation of model parameter has been demonstrated by (Kampen et al. 2022d). Other boundary conditions may perhaps be more appropriate for specific ADEs, e.g., specified spatial derivatives of temperature (Neumann boundary conditions), or specified fluxes at the boundaries (Robin boundary conditions) under upwelling conditions in rivers (Caissie and Luce 2017). In summary, it is important to realize the impact of boundary condition choices on the solution and the resulting model. Simplifying the boundary conditions to create a linear time-invariant form amenable to solutions using a Fourier transform may give acceptable results if the influence of the boundary condition (in part due to the small forcing) is small on the domain of analysis or if the assumption is realistic. This is further elaborated when considering non-linear boundary conditions as discussed in the next section.

2.2.3 Non-linearities in the advection-diffusion equation

In the previous two sections, we discussed the simple 1D representations with linear boundary conditions. In reality, we can encounter multi-dimensional flow (Cranswick et al. 2014; Ghysels et al. 2021; Reeves and Hatch 2016; Roshan et al. 2012; Shanafield et al. 2010) and non-linear dependencies of both the model parameters and boundary conditions. These result in differences with the linear 1D representations briefly discussed here.

Equations presented in Table 2.1 only consider 1D flow. When flow bends in space, energy flows and diffuses in multiple dimensions, and the mass balance is no longer strictly satisfied by changes in porosity, which is a necessity if flow is only 1D. Numerical solutions exist if adequate data are available. However, if one only knows temperatures in a 1D array and tries to interpret them without consideration of the bending flows, errors arise (Cuthbert and Mackay 2013). Note that this is not an issue where flowlines are parallel and not directly vertical (e.g., in the z direction), rather is the two-dimensional consideration important when flow curves, and the degree of error is a function of the degree of curvature (Cuthbert and Mackay 2013).

Even in one dimension, nonlinearities exist that so far have been poorly explored in terms of how they affect parameter identification. For example, the fluid viscosity generally depends on temperature, yielding an ADE where q_f is a function of the state variable T,

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} - q_f(T) \frac{c_f \rho_f}{c_m \rho_m} \frac{\partial T}{\partial z}. \tag{2.6}$$

In (2.6), κ is the thermal diffusivity which in some representations includes a (non-linear) function of diffusivity (dispersivity) and fluid velocity (Roshan et al. 2012), e.g., the square of the fluid velocity (Rau et al. 2012). The latter

would strictly speaking still be a linear time invariant ADE. However, when it is compounded with other non-linear dependencies such as the temperature dependency in (2.6), complexity grows quickly. The common solution is the management of the perturbation size in forward modeling and experimental design, e.g., amplitude at the top, which can keep the non-linear effects in check while manifesting some measurable (linearizable) process. For the inverse problem, advanced methods to deal with such non-linearities also exist but require a more specific experimental design and more advanced estimation methods (Nuij et al. 2008; Berkel et al. 2017b). Non-linear boundary conditions can also give rise to considerable non-linearities in the ADE. For example, consider (2.1) again with $q_f = 0$ and a non-linear boundary condition of the form

$$\kappa_{\text{outside}} \left. \frac{\partial T}{\partial_z} \right| z_2 = f_{\text{nl}} \left(T(z_2, t) \right), \tag{2.7}$$

which describes that the heat transfer at the boundary depends non-linearly on the temperature. An example is a pool-boiling system consisting of a metal heater which is interfaced with a pool of water at the boundary. The heater is described by a linear ADE with constant parameters as shown in (2.1), one linear boundary condition, and a non-linear boundary condition, (2.7) which is connected to a water pool, outside the domain. The water outside the domain has a thermal diffusivity depending non-linearly on the temperature. At room temperature the non-linearity is weak, but it is severe near the evaporation temperature (liquid-gas transfer). Perturbations around the evaporation temperature lead to switching between high heat transfer at the boundary below the evaporation temperature (liquid) and low heat transfer above the evaporation temperature (gas). This results directly in rich non-linear dynamics including bifurcations, even though we have the simple (linear) ADE describing our heater domain (see for details van (Gils et al. 2014). Hence, (weakly non-)linear boundary conditions are a strong requirement needed to apply the methods discussed in this work (and many others).

2.3 Signal processing for parameter estimation

Framing signal processing more broadly as a step in a process highlights the range of available choices, some less limiting than others. Frequency domain analysis of time-series has found some application in hydrology in general (e.g., (Acworth et al. 2014; Acworth et al. 2016; Graf et al. 2014; Schweizer et al. 2021; Vereecken et al. 2016; Wörman et al. 2006; Wu et al. 2020)) and in analyzing temperature data to delineate ground water surface water exchange in particular (e.g. (Luce et al. 2017; Onderka et al. 2013; Schneidewind et al. 2016; Sohn and Harris 2021; Vandersteen et al. 2015; Wörman et al. 2012)). However, while the application of signal processing methods is well developed in many engineering fields (e.g., (Curtain and Zwart 1995; Oppenheim et al. 1997; Pintelon and Schoukens 2012)) its application in hydrology, and especially in identifying groundwater-surface water exchange is

sparse. An important step missing in usual groundwater analyses is understanding that temperature signals observed at different sediment depths are connected by the model (system), including its parameters, e.g., flux, diffusion, depth. Hence, exploiting the idea that signals are not independent entities but connected on all levels, i.e., originating from a system offers significant improvements in processing (Pintelon et al. 2010b). This section summarizes ideas and methods from signal and system analysis with a specific focus on frequency domain techniques (Oppenheim et al. 1997; Pintelon and Schoukens 2012). The temperature tracing models described in (2.1) and (2.3) are some of the most commonly used models for determining the specific discharge via an inverse solution and belong to the class of linear time-invariant models (Hespanha 2018). As such, appropriate procedures and methods have been outlined in the scientific literature for signal processing (Oppenheim et al. 1997) and estimation (Keesman 2011; Ljung 1999; Söderström and Stoica 1989) including Fourier transforms. In this section, we describe signal processing techniques associated with frequency domain and linear time invariant models. Other model classes generally require other solution techniques (see e.g., (Rijlaarsdam et al. 2017)). As signal processing is always performed in the discrete time domain, we will only consider the discrete finite Fourier domain in this section.

2.3.1 Why signal processing for parameter estimation?

Here, we illustrate the application of signal processing using temperature data from two depths collected in a laboratory sand tank experiment as described by (Luce et al. 2017). A key point we wish to illustrate is that there are different ways of representing the same data Figure 2.1.

Both the time domain and frequency domain representations are well-known and are shown in Figures 2.1a to 2.1c for measurements at two different depths, respectively. To calculate these, we have used a rectangular window to select the data and applied the standard discrete Fourier transformation (FFT-algorithm) (Bracewell 1986) without additional signal processing. The only selection we have made is choosing the rectangular window such that it corresponds approximately to an integer number of periods of the dominant oscillation. We note that in the Fourier spectra, we see multiple peaks, with one being related to the natural diurnal cycle and the others to the forcing applied (2h-cycle) in the laboratory. Moreover, we see many apparent noisy signal components. We add to these familiar representations, Figures 2.1d and 2.1e processed frequency domain signals. At first sight, this does not make a significant change to the dataset as the spectra look very similar. However, when analyzing the amplitude ratio and phase difference Figures 2.1f and 2.1g, i.e., a systems representation of the frequency response function (FRF), the unprocessed FRF (raw) representation and the processed FRF differ significantly. The unprocessed FRF appears very noisy, while the processed FRF is a smooth function over the frequency as expected for semi-infinite domain solutions. For this reason, the processed FRF will be used for parameter estimation, explained in detail in Section 4. Note that, representations shown in Figure 2.1

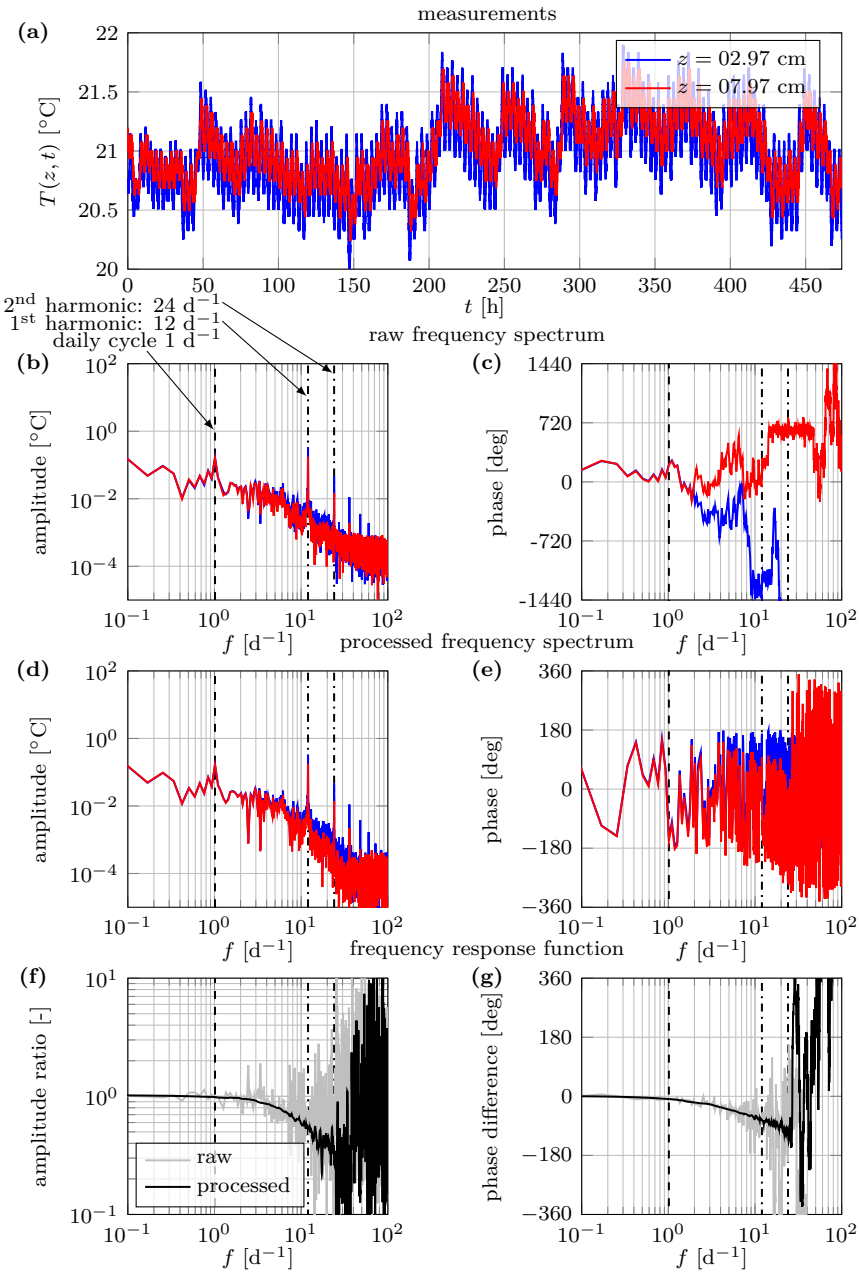


Figure 2.1. a) Temperature time-series collected at two different depths in a lab experiment with a 2-hourly (12 d-1) sawtooth shaped input temperature excitation as described in (Luce et al. 2017). Additional thermal variation is caused by daily variations in the building temperature. The (raw) amplitude b) and phase spectra c) obtained via fast Fourier transform (FFT) are further processed by the local polynomial method (LPM) resulting in new amplitude (d) and phase spectra e), which is introduced in Section 2.3.4. The amplitude f) and phase g) of the corresponding frequency response functions between the temperature time series (input $z = 2.97$ cm, output $z = 7.97$ cm, in gray) are clearly improved after processing the data with the LPM (in black).

can in essence be calculated for any signal.

To understand why the representation in Figure 2.1 is important, we relate back to Section 2.2. We know that the analytical inverse solution to (2.1) with semi-infinite lower boundary conditions and specified upper temperatures is given by (Luce et al. 2013)

$$\kappa = \frac{\omega \Delta z^2}{\Delta \phi^2 \left(\frac{1}{\eta} + \eta \right)}, \tag{2.8}$$

$$q_f = \frac{c_m \rho_m}{c_f \rho_f} \frac{\omega \Delta z}{\Delta \phi} \left(\frac{1 - \eta^2}{1 + \eta^2} \right), \tag{2.9}$$

$$\eta = \frac{\Delta \left(\ln(A) \right)}{\Delta \phi}, \tag{2.10}$$

with $\kappa = \frac{k}{c_m \rho_m}$. This relationship holds for any forcing frequency. Without going into detail, we observe that the angular frequency $\omega = 2\pi f$ and distance Δz between measurement points can be considered given. Hence, there are only two quantities that we need to determine to find κ and q_f: the phase difference $\Delta\phi = \phi_2 - \phi_1$ and the logarithmic amplitude difference $\Delta \left(\ln(A) \right) = \ln \left(\frac{A_2}{A_1} \right)$ between measurement points. Determining these two quantities correctly, and with high accuracy, is key to obtaining good estimates of κ and q_f. That these two components are the most important quantities to represent and analyze linear time invariant models, such as the ADE, is well known in the literature (Franklin et al. 2015; Oppenheim et al. 1997; Pintelon and Schoukens 2012), and this important ratio (see (2.11)) has been formalized in the concept of the frequency response function (FRF). Note that an FRF can be also obtained by evaluating the Laplace variable s in a transfer function on the imaginary axis, i.e., $s = i\omega$.

The FRF, $G(\omega)$, is defined as the ratio as the ratio of the (complex-values) Fourier coefficients or spectra between two signals with the Fourier transformed (FT) temperature defined as $T \overset{FT}{\rightarrow} \hat{T}$:

$$G(\omega) = \frac{\hat{T}_2(\omega, z_2)}{\hat{T}_1(\omega, z_1)} = \frac{A_2(\omega) e^{i\phi_2(\omega)}}{A_1(\omega) e^{i\phi(\omega)}} \tag{2.11}$$

or equivalently

$$\ln(G) = \ln \left(\frac{A_2}{A_1} \right) + i \left(\phi_2 - \phi_1 \right) \tag{2.12}$$

with $\frac{A_2}{A_1}$ the amplitude ratio and $\Delta\phi$ the phase change of the frequency response function (or transfer function).

The concept of transfer function or FRF follows from the notion of linearity: if we multiply the forcing signal with a factor, or one or multiple sinusoidal components, linearity states that the temperature at depth is scaled proportionally, and in taking the ratio, this proportionality factor is divided out. Note also that

linearity together with time invariance are coupled to the Fourier representation, e.g., for linear time invariant systems all the frequency components can be treated individually and independently (Oppenheim et al. 1997).

The goal of signal processing is now to determine the best $G(\omega)$ from measurements. In Figures 2.1f and 2.1g, we show a comparison between the standard discrete Fourier transform (minimally processed (raw) frequency response function) and the frequency response function after further processing shown in Figures 2.1d and 2.1e using the local polynomial method (LPM) which removes errors in estimating the Fourier coefficients (explained in Section 2.3.4). As we will show in Section 2.4.1, the empirical $G(\omega)$ is a good representation of the theoretical $G(\omega)$ corresponding to (2.1) in the entire range of $0.1 - 10$ d^{-1} for this specific example. Hence, not only do the obvious peaks, 1, 12 and 24 d^{-1}, convey information on the ADE, but also all the other frequencies can contribute to the estimation of κ and q_f, emphasizing there is no need to use or isolate a single sinusoidal forcing. The reason is that some forcing is present at most frequencies, so the resulting temperature measurements at depth will have a contribution from these multiple-frequency forcing terms, which is useful information.

2.3.2 Signal components

In Section 2.3.1, we have established that to estimate our model parameters, $\Delta\phi$ and $\ln\left(\frac{A_2}{A_1}\right)$ need to be determined with high accuracy, which is equivalent to determining $G(\omega)$ with high accuracy. Signal processing can be performed on two levels, i.e., on the individual signal level (Section 2.3.3) or on a system level (Section 2.3.4). For any individual signal, we distinguish generally among three different components being (i) periodic components, (ii) trend, and (iii) noise. These components are shown in Figure 2.2 for both the time domain and frequency domain.

Particularly, transients play an important role in signal processing as they are not part of the (periodic) model representation $G(\omega)$ of our ADE. Transients are a collection term for signal components that are not periodic and result from several causes. The strongest transient observed in the signals at depth originates from the system state (temperature) that is not in equilibrium at the start of the window of observation, which is associated with the initial condition. Others could be sensor (calibration) drifts or more complicated drifts associated with non-linearities as described by (2.5) in Section 2.2.3. Regardless of the cause, such signal components need to be removed as the frequency representation of the ADE assumes non-periodic signal components to be zero. Note that the steps explained here could also be used in time domain approaches with the additional operation of the discrete inverse Fourier transform (IFFT) to bring the signal back to the time domain (Berkel et al. 2018b).

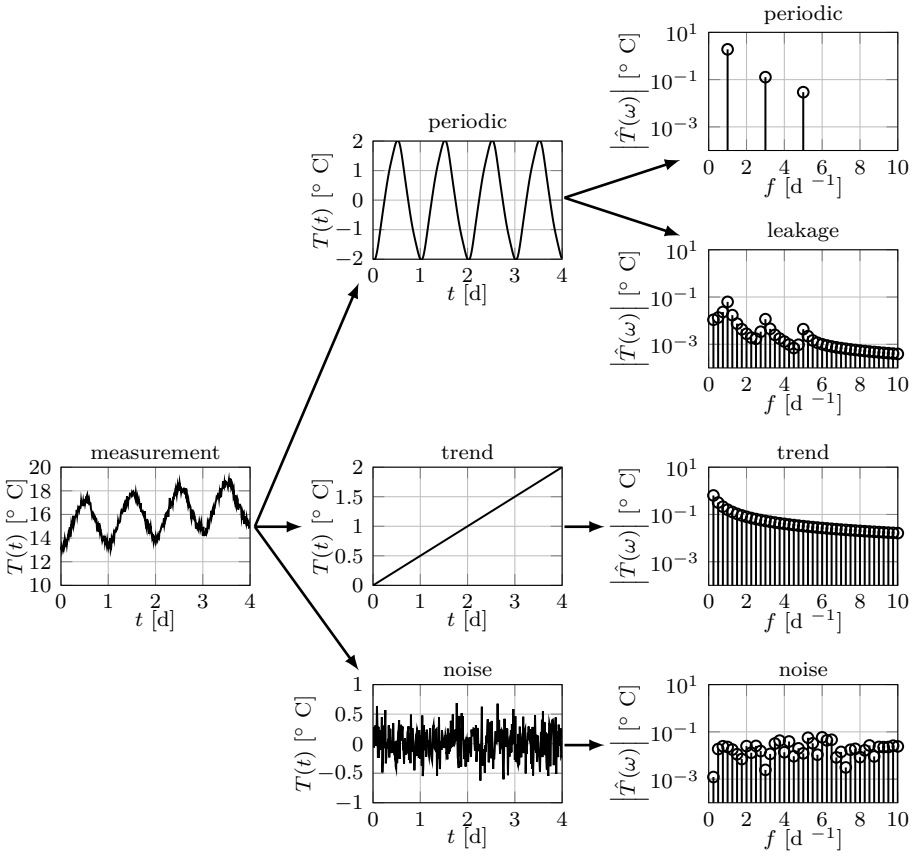


Figure 2.2. Overview of a typical signal and its three main components: trends due to sensor drifts and the initial conditions; the wanted periodic signal part used for estimating advection and convection; and unwanted sensor noise. In the frequency domain, we have an additional spectral leakage term due to noise and possibly non-integer period selection.

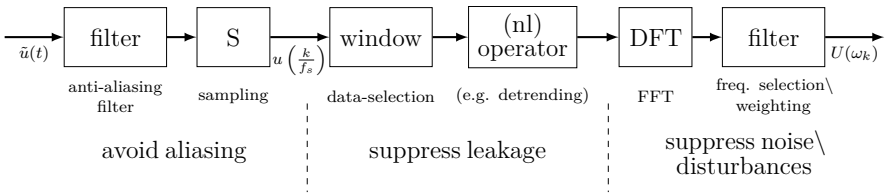


Figure 2.3. Overview of standard signal processing steps from a continuous time signal to a discrete frequency representation of the signal. Note that $\tilde{u}(t)$ is the original continuous signal whereas the others are discretized with index k using sampling frequency f_s (DFT stands for discrete Fourier transform). The term (nl) specifies a non-linear operation in the detrending whereas the other steps are generally considered linear operations.

2.3.3 Signal processing on individual signals

Three general steps Figure 2.3 are applied to extract periodic components $G(\omega)$ from signals by reducing transients and noise and avoiding aliasing, which is a non-physical signal artifact introduced due to sampling; A) avoid aliasing by applying the correct sampling process; B) suppress transients which is equivalent to spectral leakage in the frequency domain; and C) suppress noise and/or disturbance.

A) Aliasing is easily avoided by adequate sampling and if necessary, properly resampling the data at a later stage. Most (commercial) loggers these days have anti-aliasing filters installed or use more advanced methods to avoid aliasing, e.g., based on sigma-delta converters (Baker 2007; Baker 2008). Consequently, the set-up takes care of the aliasing and as long as the user does not resample the data, this part rarely poses a problem.

B) Due to transients and noise, the starting and end points of a time series do not have the same value. In that case, taking the Fourier transform results in (spectral) leakage. Additionally, another spectral leakage term could exist if the number of periods is non-integer, even though the original signal might be noiseless and perfectly periodic. Both leakage contributions would manifest in contributions spread over frequencies most strongly at low frequencies or as side-lobes in the frequency spectrum, respectively. The three standard remedies to suppress transients (spectral leakage) are detrending, windowing, and selecting a time window that corresponds to an integer number of signal periods.

   Detrending addresses the root cause of spectral leakage due to transients as it removes non-periodic trends directly from the data. Alternatively, windowing (Hamming, von Hann, etc.) differently from the square window can be applied to down-weight the end points of a signal, thus reducing spectral leakage. However, this is a much more delicate operation, as modifying signals in this way can also negatively impact their quality. This is especially true when signals are nearly periodic or if we have long time traces because applying a

Table 2.2. System representation of signals in linear time invariant systems at three different depths. We use ω_k to denote the k-th Fourier coefficient as we are dealing with discrete signals here.

measure-ment		system connection		drifts		leakage		noise
$\hat{T}(z_0,\omega_k)$	$=$	$R(z_0,\omega_k)$	$+$	$T_r(z_0,\omega_k)$	$+$	$T_n(z_0,\omega_k)$	$+$	$N(z_0,\omega_k)$
$\hat{T}(z_1,\omega_k)$	$=$	$\underbrace{H(z,\omega_k)R(z_0,\omega_k)}_{U(z_1,\omega_k)}$	$+$	$T_r(z_1,\omega_k)$	$+$	$T_n(z_1,\omega_k)$	$+$	$N(z_1,\omega_k)$
$\hat{T}(z_2,\omega_k)$	$=$	$\underbrace{G(z,\omega_k)R(z_0,\omega_k)}_{Y(z_2,\omega_k)}$	$+$	$T_r(z_2,\omega_k)$	$+$	$T_n(z_2,\omega_k)$	$+$	$N(z_2,\omega_k)$

window on a periodic dataset will make it aperiodic again, and for long time traces (compared to the frequencies studied) leakage might only have a small impact on the frequencies of interest.

Although detrending and windowing are useful in improving results from signal processing, other methods based on the relationship between signals (Schneidewind et al. 2016; Kampen et al. 2022d; Vandersteen et al. 2015) can yield more reliable improvements as will be discussed in Section 2.3.4.

C) Noise suppression is done naturally using the discrete Fourier transform where periodic components will remain at the same frequency (bin or point) and the noise is spread over frequency (bins). Then, spectral signal components only containing noise (or dominant disturbances) can be removed or simply ignored. Both strategies are common and implicitly applied when selecting the diurnal sinusoid. As noise, even after extensive signal processing, will be present on all frequencies including on the diurnal cycle, it needs to be properly quantified by, e.g., analyzing the (co-)variances.

2.3.4 Signal processing using system properties

As we are interested in the parameters of the ADE, i.e., the system, rather than the individual signals, we suggest using different signal processing techniques, or better, system processing techniques, which allow for more accurate estimates of $G(\omega)$. In a system, signals are not considered as separate entities, rather they are connected. In practice, the temperature time-series at depth are related to the forcing signal and the signals at other depths and are not separate entities. In such cases $G(\omega)$ represents one part of the relationship between the signals, however, the components that need to be separated are still the same as in Figure 2.2. As an example, a system of temperature measurements in the frequency domain ($\hat{T}(z_0,\omega_k)$, $\hat{T}(z_1,\omega_k)$, $\hat{T}(z_2,\omega_k)$) at three depths (0, 1, 2) is shown in Table 2.2.

We start with some arbitrary forcing $R(z_0,\omega_k)$ at some top channel where measurements $\hat{T}(z_0,\omega_k)$ can contain two types of leakage (discussed previously), i.e., due to, (sensor) drifts $T_r(z_0,\omega_k)$ and due to spectral leakage when applying the

discrete Fourier transform to a signal $T_n(z_0, \omega_k)$. Moreover, we may have spatially- and time (colored) correlated noise components on the different channels $N(z_0, \omega_k)$, $N(z_1, \omega_k)$, $N(z_2, \omega_k)$. In case of non-periodic forcing (Figure 2.1), we cannot remove the transient from $G(\omega)$ (perfectly) in contrast to non-sinusoidal periodic forcing, where this is possible. However, by assuming the top channel z_0 to be noiseless, we can determine $G(\omega)$ perfectly, at least in theory (see (Pintelon and Schoukens 2012) for a formal proof). To assume $N(z_0, \omega_k) = 0$ is a reasonable (weak) assumption, as the noise on the top channel is small compared to the signal. Hence, by neglecting the noise $N(z_0, \omega_k)$) at the top channel we only introduce a small error compared to a true noiseless top channel. Assuming that we can fully eliminate the transients in all steps, we obtain the actual signals $R(z_0, \omega_k)$, $U(z_1, \omega_k)$, $Y(z_2, \omega_k)$ with some uncertainty based on the noise $N(z_i, \omega_k)$. This provides the information required to calculate $H(z, \omega_k)$ and $G(z, \omega_k)$ and, consequently we obtain the parameters in the ADE at every frequency as shown in (2.8) to (2.10).

The removal of transients represents a challenging task. This can be achieved by separating smooth (non-periodic) signal components from non-smooth (periodic) components as a function of frequency. It is well established that (arbitrary) forcing is never smooth as function of frequency, where a purely sinusoidal forcing is an extreme example as it presents itself as just a spike in the frequency domain. However, $\hat{T}(z, \omega_k)$ will consist of both smooth and non-smooth components. Both $H(z, \omega_k)$ and $G(z, \omega_k)$ are smooth functions of frequency and are multiplied with non-smooth functions of the frequency $R(z_0, \omega_k)$ and $U(z_1, \omega_k)$, resulting in again non-smooth functions. Additionally, the aliasing contributions $T_n(z, \omega_k)$ will always be smooth (see (Pintelon and Schoukens 2012) for a formal proof). This means that if we create (local) smooth fits over frequency, we can separate the transient contributions from the systems response by subtracting them from the signals $\hat{T}(z_0, \omega_k)$, $\hat{T}(z_1, \omega_k)$, $\hat{T}(z_2, \omega_k)$ to arrive at signals $R(z_0, \omega_k)$, $U(z_1, \omega_k) + N(z_1, \omega_k)$, and $Y(z_2, \omega_k) + N(z_2, \omega_k)$. Other possibilities to make these fits include rational polynomial functions (McKelvey and Guérin 2012) or interconnected polynomials (Gevers 2005). Using these methods, we can also quantify the noise by analyzing the small residual fluctuations on the locally smooth fits over frequency which correspond to the noise (co-)variances. From these signals, we can significantly improve the calculation of $G(z, \omega_k)$ including its uncertainties as shown in Figure 2.1.

As discussed previously, detrending is the removal of trends from the time signal (measured data) using linear or quadratic functions. If we were able to determine the (specific) trend that does not belong to the periodic measurements, we could remove the actual (non-linear) trend, which is exactly what the estimation of the smooth curve over frequency tries to achieve.

Also, for completeness, we want to mention that in specific cases, applying the LPM method to remove transients can be non-ideal. Transients that are unrelated to spectral leakage can also carry information of the system and its parameters but are extremely difficult to be distinguished from spectral leakage. Hence, the removal of the transient is generally necessary, but can also mean that useful data is removed. In comparison single-sinusoidal approaches remove

all transient information, except at the diurnal cycle and also remove all other frequency components. Although for significant natural forcing this discussion is largely irrelevant as there is no way to distinguish useful transients from non-useful transients, however, for perfect periodic forcing and dominant transient forcing this claim needs to be nuanced.

2.3.5 A discussion of related single sinusoidal approaches

In the previous section, we presented practical signal analysis methods applicable for an entire system and showed that regardless of how one reduces spectral leakage or noise, the correct mathematical framing is that the relationship $G(\omega)$ (transfer function) needs to be extracted from the data, which holds for all frequencies. The estimate of $G(\omega)$ is denoted by $\hat{G}(\omega)$. Nevertheless, a number of practices have evolved and grown around a perfect sinusoidal theme (of the diurnal), i.e., that $y(t)$ (e.g., temperature at depth) and $u(t)$ (e.g., surface temperature) must each be approximated by a sinusoid(al) representation (often separately), applying for instance, regression techniques based on Kalman filtering (Young et al. 1999) used in (Keery et al. 2007) and (Irvine et al. 2015). These methods give the impression of signal improvement but actually extract a signal (sinusoid) of reduced quality compared to the original signal when used for parameter estimation. The problem then is twofold:

- The resulting amplitude and phase (of the pure sinusoids) are best sinusoidal fits of the signals and not of the ADE-system, i.e., do not represent the correct/best estimate of $G(\omega)$ and consequently the parameter estimates.
- All the other useful information at all the other frequencies is lost, which can be significant and often, more useful information is lost than is actually used (see Section 2.4 for more explanation).

In other words, regardless of which method is applied, the ratio $\frac{Y(\omega)}{U(\omega)} = G(\omega)$ needs to be retained (and improved), but specifically not the ratios of the best sinusoidal fits of the individual signals as sometimes is proposed.

2.4 Estimation methods

Many of the commonly applied methods for estimation use only a small portion (mostly the diurnal signal) of the information available in temperature signals to estimate water flux, and, sometimes, thermal diffusivity (e.g., (Hatch et al. 2010; Irvine et al. 2015; Irvine et al. 2017)). When uncertainty is reported, it is usually framed on reported sensor uncertainty (Glose et al. 2019; Luce et al. 2013). A more broadly supported and implemented approach to uncertainty estimation and validation is through the use of overdetermination, i.e., using more frequencies or more locations than strictly necessary to obtain a single parameter value. This section follows the latter approach and exploits the overdetermination (in the

Table 2.3. Parameter estimates at different frequencies of the raw and signal processed data presented in Figure 2.1 with σ_κ and σ_{q_f} as the standard deviation of the parameter estimates.

	raw	**processed**		**raw**	**processed**	
f [d^{-1}]	κ [$cm^2\ s^{-1}$]	κ [$cm^2\ s^{-1}$]	σ_κ [$cm^2\ s^{-1}$]	q_f [$cm\ s^{-1}$]	q_f [$cm\ s^{-1}$]	σ_{q_f} [$cm\ s^{-1}$]
1	0.02321	0.00870	0.00286	0.00260	0.00256	0.00009
12	0.00557	0.00556	0.00004	0.00215	0.00218	0.00004
24	0.00602	0.00586	0.00014	0.00209	0.00202	0.00006

frequency domain) to its maximum. This starts by understanding that the ADE has two characteristic equations in the frequency domain, namely amplitude and phase (or real and imaginary parts), and two unknown parameters, namely flow and diffusivity. At each frequency, one can calculate a unique solution for κ and q_f. However, simply averaging over these estimates per frequency (Berkel et al. 2014a) would not converge to a correct κ and q_f as is demonstrated below. Hence, a different unbiased (efficient) approach is necessary, known as maximum likelihood estimation. In this context, we also discuss approaches that better aid in identifying parameter values and uncertainty using multiple frequencies, and multi-sensor estimation.

2.4.1 Single frequency estimates and their uncertainty

Coefficients at each frequency are calculated individually using the data of $\hat{G}(\omega)$ as shown in Figure 2.1, which is presented in Figure 2.4. In Figures 2.4a and 2.4b, the raw discrete Fourier transform is used, while results applying the LPM-processed signals are shown in Figures 2.4c and 2.4d, where information about noise at each frequency is used to estimate confidence bounds based on the linear transformation of the variance of the Fourier coefficients, which is explained in detail in Sections 2.4.2 and 2.4.3. We see that the result is erratic over the frequency range but also that in some frequency ranges the estimates are less erratic. This is also supported by Table 2.3 where we see that both for the raw and processed data the estimates are comparable at some specific frequencies.

These high signal-to-noise ratios are closely related to the lab set-up used for this experiment, with a primary forcing at 12 [d^{-1}]. If we compare the raw estimates per frequency (Figures 2.4a and 2.4b) to those when using the LP-processed data (Figures 2.4c and 2.4d) as discussed in Section 2.3, we see that the results are significantly improved between the frequency range 1 [d^{-1}] and 10 [d^{-1}]. After LPM-processing, we have precise estimates for nearly every frequency. The added processing provides good estimates not only at frequencies where forcing is strong, but at many other frequencies, giving a better estimate of the shape of $G(\omega)$. Adding many (hundreds) of frequency points can help improve parameter estimates (see Section 2.4.4) significantly. The estimates are reliable not only for the perturbed

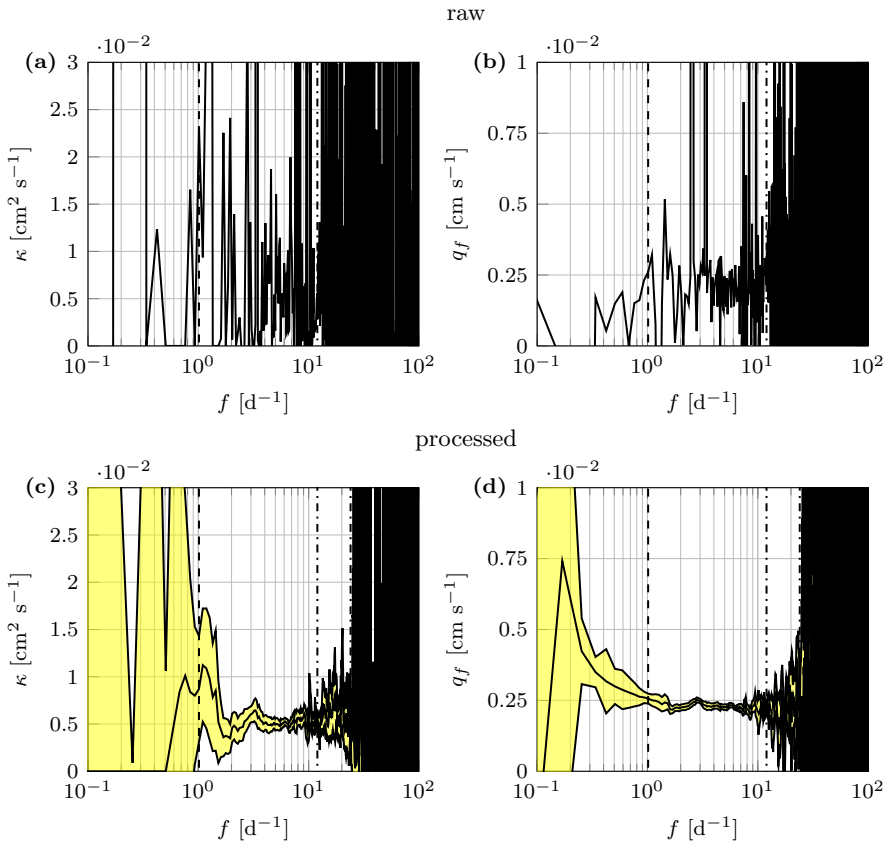


Figure 2.4. Parameter estimates of diffusivity (a, c) and flow (b, d) for the raw (a, b) and signal processed (c, d) data presented in Figure 2.1. They are estimated per single frequency. The confidence bounds (yellow shading) are calculated based on the variances of the Fourier coefficients determined by the LPM (at each frequency) using linearly approximated 95% confidence bounds based on the MLE method in domains. The ratio for the calculation from the advection to q_f is given by $\frac{c_m \rho_m}{c_f \rho_f} = 1.0449$. Some corresponding values can be found in Table 3 showing the parameter values with their standard deviation for the main forcing components.

frequencies (1 [d^{-1}], 12 [d^{-1}], etc. in this particular experiment), but over a large frequency range. While the daily frequency is certainly a high-power frequency for stream temperature in many natural settings, most low-frequency forcing and most high-frequency harmonics, especially those of the daily frequency, hold considerable information.

2.4.2 Quantifying noise at single frequency

In most parameter identification experiments, the general approach to estimate the uncertainty is quantifying the (stochastic) noise through overdetermination. The most realistic measure for noise is based on the data itself and generally not sensor uncertainty, as the latter is small for modern measuring devices compared to the variation of the data over time. We propose to use instead the variance of Fourier coefficients over periods calculated using the LPM (Pintelon and Schoukens 2012) (explained in Section 2.A). When we assume Gaussian (time-domain) noise, the resulting distribution function is a complex(-valued) circular normal distribution (CCND) (Goodman 1963). This is a bi-variate Gaussian distribution (Goodman 1963) where the variance of imaginary and real parts is the same (circular) but the variance will change over frequency (color of noise). On the other hand, the resulting spectrum is independent over frequency (linearity). This is shown in Figure 2.5.

The histogram based on a numerical calculation presented in Figure 2.5, shows that there is indeed a Gaussian distribution with equal variances in real and imaginary values. That the confidence circles do not have the same radius, i.e., variance, is shown in Figure 2.5a. These distribution functions are used to construct multi-frequency estimators.

2.4.3 Confidence bounds at a single frequency estimate

Here, we show how to translate the frequency domain distribution (CCND) to a confidence bound calculation for a single frequency estimate of dispersity κ and water flux q_f. As the single frequency is defined by a two-dimensional distribution function (Figure 2.5), mapping it onto two parameters, κ and q_f, results also in a two-dimensional distribution which defines the confidence bounds. This is shown in Figure 2.6.

Figure 2.6 shows the steps that are taken to calculate the joint distribution function of κ and q_f using two (spatial) channels at a single frequency (18 d^{-1}). As an intermediate step, we have also included the distribution of the transfer function $G(\omega)$. Using (2.8) to (2.10) results in a unique transformation from the single frequency at two locations to κ and q_f. However, (2.8) to (2.10) is a (strongly) non-linear transformation, hence, the distribution function of κ and q_f is the inverse non-Gaussian χ^2-distribution (Berkel et al. 2014a). This can also clearly be seen by the joint distribution and its cross-sections (Figures 2.6d, 2.6e and 2.6f). The confidence bounds are theoretically non-symmetric and only for a

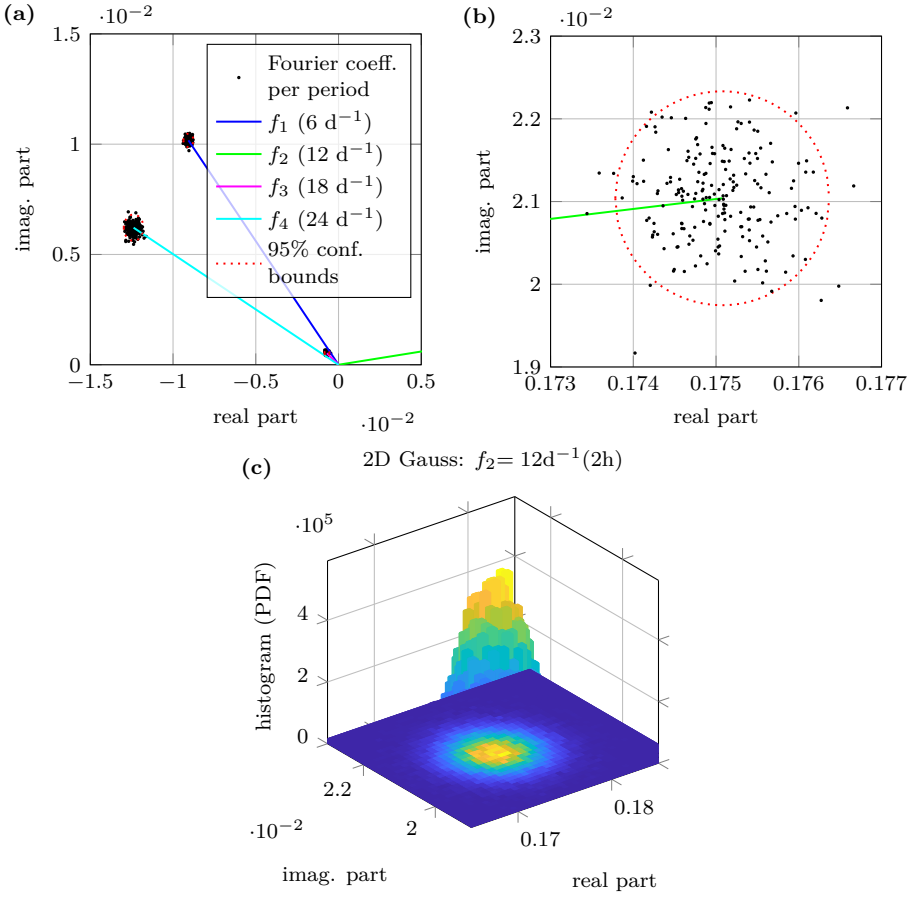


Figure 2.5. Statistical representations of frequencies 6, 12, 18, 24 [d^{-1}] of the temperature time traces presented in Figure 2.1a. Figure (b) shows a zoom of the Fourier coefficients per period (200 samples total) at 12 d^{-1} including the 95% confidence interval, all represented by circles. Figure (c) shows the bi-variate Gaussian distribution function based on a numerical histogram calculation (10000 samples total) (see (Berkel et al. 2014a) for details). This distribution is known as complex circular normal distribution (CCND) as the variances for real and imaginary parts are the same.

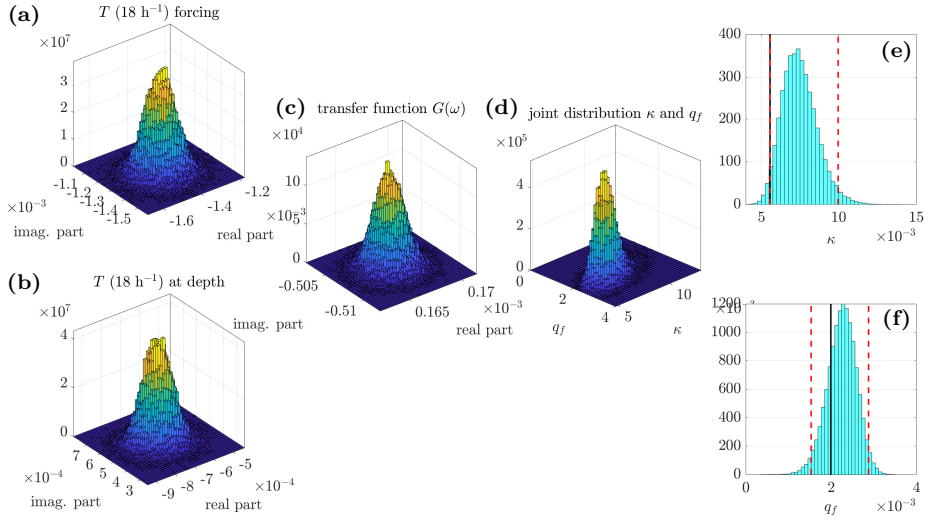


Figure 2.6. Distribution functions (histogram or probability density function (PDF)) all at frequency 18 [d^{-1}] of the different steps in calculating the confidence bounds on κ and q_f where: (a) the CCND for the forcing $U(\omega_k) = \hat{T}(z_1, \omega_k)$; (b) the CCND at depth $Y(\omega_k) = \hat{T}(z_2, \omega_k)$; (c) the CCND of the transfer function $\hat{G}(\omega)$ at 18 [d^{-1}]; (d) the non-Gaussian distribution of κ and q_f; and (e, f) the cross-sections including in (red-dashed) lines the 95% confidence bounds and the approximate value of on κ and q_f at 18 [d^{-1}]. Note, that all distributions are calculated based on Monte-Carlo calculations and hence no prior information on their shape, e.g., Gaussian, is used except for the Gaussian time domain samples used to generate these distributions.

high signal-to-noise ratio does the distribution of κ and q_f converge to a Gaussian distribution (and diverge for a decreasing signal-to-noise ratio). Most apparent is the long tail of the distribution of κ, leading to a large upper bound. Moreover, an important aspect is that κ and q_f are not independent but correlated quantities. This is rarely reported as cross-correlation (Pearson coefficient) but describes to what extent κ and q_f can be independently estimated.

To obtain the confidence bounds on κ and q_f, we take an intermediate step by calculating the variance of the transfer function $\hat{G}(\omega)$ introduced in (2.11) and its distribution shown in Figure 2.6c. There is an analytic expression nicely reflecting all the important components with $Y(\omega_k) = \hat{T}(z_2, \omega_k)$ and $U(\omega_k) = \hat{T}(z_1, \omega_k)$ as defined in (Pintelon and Schoukens 2012)

$$\hat{\sigma}_G^2\left(G(\omega_k)\right) \approx \left|\hat{G}(\omega_k)\right|^2 \left(\frac{\sigma_Y^2(\omega_k)}{\left|Y(\omega_k)\right|^2} + \frac{\sigma_U^2(\omega_k)}{\left|U(\omega_k)\right|^2} - 2\mathrm{Re}\left(\frac{\sigma_{YU}^2(\omega_k)}{\left|Y(\omega_k)\overline{U(\omega_k)}\right|^2} \right) \right). \tag{2.13}$$

The variance on the transfer function $\hat{\sigma}_G^2$ depends on the cross-correlation between two channels $\sigma_{YU}^2(\omega_k)$, e.g., caused by electrical equipment noise. The variance $\hat{\sigma}_G^2$ also depends inversely on the signal-to-noise ratio and on the transfer function $\hat{G}(\omega_k) = \frac{Y(\omega_k)}{U(\omega_k)}$ itself. We have to realize that the exact input and output and true $G(\omega_k)$ are unknown, such that Y, U, and $\hat{G}$ need to be based on the calculated Fourier coefficients resulting in an approximation of the variances, which also holds for the variance estimates and confidence bounds on κ and q_f. The calculation of the confidence bounds for single frequencies of κ and q_f is problematic as the distributions are non-symmetric and consequently no analytic expression exists. Generally, for single frequencies, field measurements result in significant non-Gaussian distributions when compared to our laboratory measurements due to our high signal-to-noise ratios even for the 18 d^{-1}. Consequently, the asymmetric bounds need to be approximated using Monte-Carlo calculations (as shown in Figures 2.6e and 2.6f) or the signal-to-noise ratio needs to be increased allowing for a Gaussian approximation, which in practice can only done by averaging over frequencies, which is the topic of the multi-frequency analysis. A Monte-Carlo analysis is used in this work to process generated samples (Fourier coefficients) from a distribution (CCND) and analyze the outcomes, e.g., κ and q_f for which the variance and the (non-symmetric) confidence bounds are calculated.

2.4.4 Multi-frequency model parameters estimation (maximum likelihood)

The challenge in multi-frequency analysis is to search for solutions for which with increasing data, here the number of frequencies, the model parameter estimates constantly improve. In practice, this is tightly linked to using (multivariate) Gaussian distributions both in estimating the model parameters and in confidence

calculation. This section attempts to make multi-frequency model parameter estimation more understandable. However, the deep statistical interpretations and calculations will not be derived here (see for this, e.g., (Keesman 2011; Ljung 1999; Söderström and Stoica 1989)).

To better understand multifrequency analysis, let us first consider some standard approaches. These will also be compared in Table 2.4 based on the four frequencies presented in Figure 2.5 (and Figure 2.8 in Section 2.B). The most naïve choice is to average over the estimated model parameters, which can be denoted as

$$\hat{\theta} = \frac{1}{F}\sum_{k=1}^{F}\theta(\omega_k), \tag{2.14}$$

with $\theta = \begin{bmatrix}\kappa & q_F\end{bmatrix}$ and F as the total number of frequencies used. Alternatively, one could also weight the variance or standard deviation

$$\hat{\theta} = \frac{\sum_{k=1}^{F}\frac{\theta(\omega_k)}{\sigma_\theta(\omega_k)}}{\sum_{k=1}^{F}\frac{1}{\sigma_\theta(\omega_k)}}. \tag{2.15}$$

However, one needs to first acquire the variances to use (2.15). As the previous section shows, the distribution functions for most $\theta(\omega_k)$ are non-Gaussian. Hence, acquiring the variances $\sigma_\theta^2(\omega_k)$ is a rather challenging endeavor as we need to acquire them via Monte-Carlo simulations with sufficient sampling, see Section 2.4.3. One could select by hand a minimum signal-to-noise ratio of those frequencies for which the parameters can be well approximated by a Gaussian distribution (time-consuming).

For both methods, there is no guarantee that adding more data results in a better overall estimate due to the non-Gaussian nature of the distribution functions. Moreover, the cross-correlation that exist between κ and q_f is not being considered here resulting in an increase in error. In conclusion, the direct approach in (2.15) causes many computational difficulties, does not guarantee convergence to the true value and can only be used for extremely high signal-to-noise ratios.

Alternatively, standard least-squares algorithms can be used which also do not assure that the estimate becomes better with an increasing number of frequencies, although this result is better conditioned than (2.14) in terms of retaining Gaussian distributions. Consequently, it offers a path to understanding how to approach the multi-frequency estimation challenge. The standard least-squares estimator in terms of transfer functions is the cost function

$$\hat{\theta} = \arg\min_{\theta}\frac{1}{F}\sum_{k=1}^{F}\left|\hat{G}(\omega_k) - G(\theta,\omega_k)\right|^2, \tag{2.16}$$

where the measured transfer function $\hat{G}(\omega_k)$ was calculated in Figure 2.1 and its analytical counterpart, the transfer function $G(\theta,\omega_k)$ following from the ADE derived in (2.8) to (2.12).

Notation-wise, arg min of θ states that we minimize this least-squares criterium by changing $\theta = \begin{bmatrix}\kappa & q_F\end{bmatrix}$ such that the sum of the squared error over frequencies is minimal. However, note that there is a clear difference between (2.14) and (2.16) statistically. As Figure 2.6c demonstrates, the distribution functions of $\theta(\omega_k)$ are non-Gaussian and as such averaging them results in a non-Gaussian distribution for the resulting multi-frequency distribution of $\theta = \begin{bmatrix}\kappa & q_F\end{bmatrix}$ with all its associated problems, whereas the real and imaginary parts of $\hat{G}(\omega_k)$ shown in Figure 2.6 are (near) Gaussian and as such (2.16) yields a (near) Gaussian distribution. The problem with this representation is that it does not account for the quality, i.e., the uncertainty, of the estimates of $\hat{G}(\omega_k)$. Combining the ideas from (2.15) and (2.16), gives rise to the weighted least-squares estimator. However, this does (again) not always guarantee that the estimate will improve with increasing data. Therefore, we have to show that the estimator is efficient (Pintelon and Schoukens 2012), i.e., that the parameters would converge to the exact value if we went towards an infinite amount of data with minimum variance. Technically, the lower bound of the minimum variance (unbiased estimate) is known as the Cramer-Rao lower bound (Pintelon and Schoukens 2012), which is the best we can achieve given the data. Estimators that come extremely close to this property are maximum likelihood estimators (MLEs). In transfer function representation the MLE is given in (e.g., (Vandersteen et al. 2015))

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{F} \sum_{k=1}^{F} \frac{\left|\hat{G}(\omega_k) - G(\theta, \omega_k)\right|^2}{\hat{\sigma}_G^2(\theta, \omega_k)}, \tag{2.17}$$

which can be seen as weighting the least-squares estimator with $\hat{\sigma}_G^2(\theta, \omega_k)$. The estimator variance $\hat{\sigma}_G^2(\theta, \omega_k)$ for the 2-point estimator is given in (2.13) where the estimate $\hat{G}(\omega_k)$ is replaced by the true $G(\theta, \omega_k)$ (for which we need to determine θ). Hence, the weight also depends on the parameters to be estimated.

The maximum likelihood estimator does not need user interaction with the exception of non-model related disturbances, e.g., the electrical grid disturbance. Consider the simplified case as in (e.g. (Vandersteen et al. 2015)) where $\hat{\sigma}_G^2(\omega_k) = \frac{\sigma_Y^2(\omega_k)}{|U(\omega_k)|^2}$, then the bad estimates of $\hat{G}(\omega_k)$ are suppressed (high noise variance $\hat{\sigma}_Y^2(\omega_k)$ or small signal $|U(\omega_k)|^2$ and good measurements are increased in weight, e.g., the amplitude of the forcing, $|U(\omega_k)|^2$, being large around the diurnal cycle. Hence, the user no longer needs to decide what is good data and what is bad data, the maximum likelihood estimator takes over this decision process. Also, the (near) Gaussian distribution is retained as $\hat{\sigma}_G^2(\omega_k)$ is a frequency-dependent scaling factor and the overall estimate also has lower variance converging to an ever more Gaussian estimate. The complexity of the MLE increases when taking more realistic conditions into account such as noise on the top channel (Schneidewind et al. 2016) and multiple channels (Kampen et al. 2022d). The only challenge is the optimization of the cost function. However, as only two parameters need to be estimated, and given the computational power of modern computers this is no longer a real issue.

Table 2.4. Multi-frequency parameter estimates of the frequencies 6, 12, 18, 24 d^{-1} belonging to the different multiple frequency approaches calculated using Monte-Carlo calculations. In italic the variances calculated using (linear) propagation of uncertainty (explained in Section 2.B).

method	eq	$\kappa \times 10^{-3}$ [$cm^2\ s^{-1}$]	σ_κ [$cm^2\ s^{-1}$]	$q_f \times 10^{-3}$ [$cm\ s^{-1}$]	σ_{q_f} [$cm\ s^{-1}$]	Pearson factor
average	(2.14)	6.0313	$3.02 \cdot 10^{-6}$	2.2413	$9.89 \cdot 10^{-7}$	−0.50
weighted average	(2.15)	5.6361	$6.26 \cdot 10^{-7}$	2.1885	$1.84 \cdot 10^{-7}$	−0.45
LSE	(2.16)	5.5678	$1.17 \cdot 10^{-9}$	2.1782	$9.89 \cdot 10^{-11}$	0.44
MLE	(2.17)	5.5696	$1.05 \cdot 10^{-9}$	2.1825	$8.81 \cdot 10^{-11}$	0.44
MLE no-MC	(2.17)	5.5690	$1.13 \cdot 10^{-9}$	2.1826	$9.37 \cdot 10^{-11}$	0.44

To go beyond explanation of the alternative estimates in (2.14) to (2.17), we also demonstrate the improvement towards the MLE estimates. We use the data presented in Figure 2.6 for four frequencies. For the variance calculations we use Monte-Carlo simulations. The results are presented in Table 2.4. When looking at Table 2.4 two important considerations have to be taken into account: 1) the exact value is unknown and we use experimental data, which can also randomly be on the exact value and 2) the data presented here is from high-quality laboratory experiments, hence, signal-to-noise ratios are high and differences between estimators are less distinct. The improvement, for instance, with the MLE increases significantly compared to the other estimators when the data quality decreases, i.e., with decreasing signal-to-noise ratio.

Table 2.4 clearly shows the increase in accuracy of the estimated parameters especially when we consider the variances of the different estimates. We also see that retaining Gaussian properties is very beneficial as all implicit estimators give good results. Moreover, for all estimates there is a strong correlation between the estimates meaning that changing the κ estimate will have a significant impact on the estimate of q_f, which is something that needs to be remembered (and reported).

2.4.5 Multi-frequency estimation of confidence bounds

As we now have a method which can deal with multiple frequencies, we can now also apply the MLE to our experimental processed data shown in Figure 2.1. Using the MLE with different sets of frequencies for estimation leads to the outcomes in Figure 2.7.

Three frequency regions are considered for estimation, using only frequencies lower than the primary excitation frequency, the diurnal (24h) signal, and the first three harmonics of the primary, 2-hr (12 d^{-1}), excitation. The results are summarized in Table 2.5.

While both multi-frequency estimates (MLE) produce similar results, the diurnal

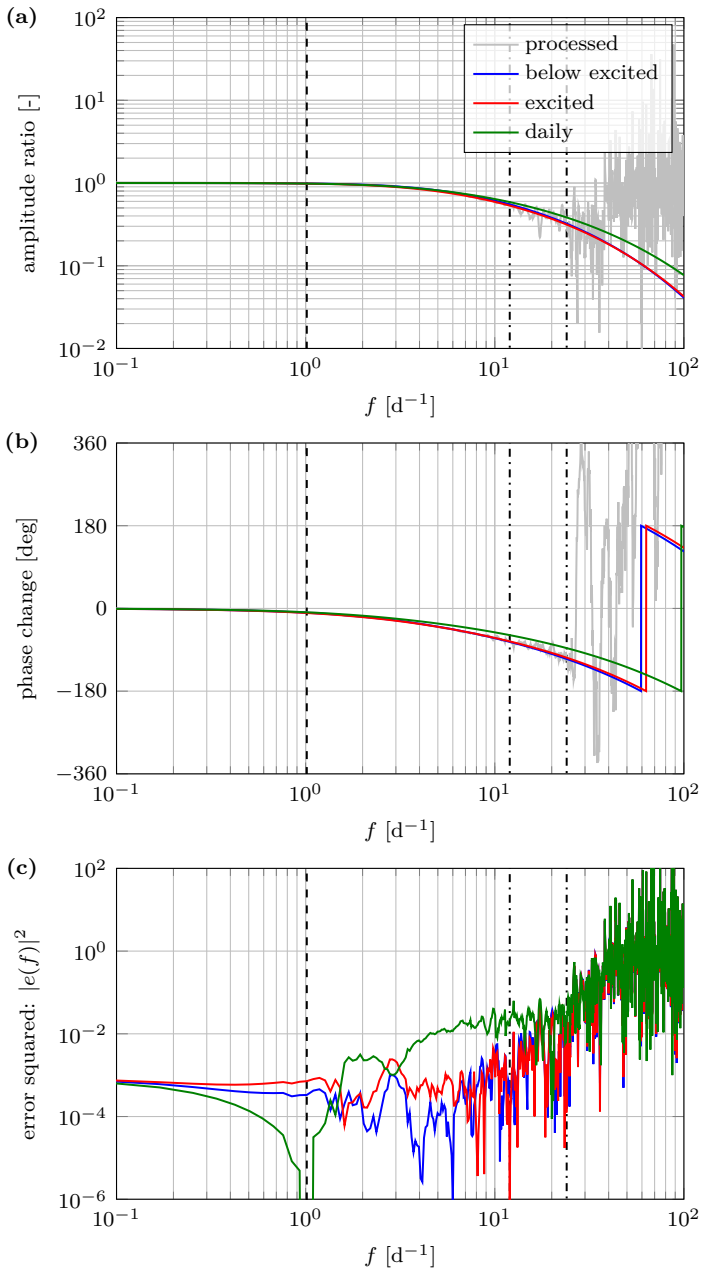


Figure 2.7. The FRF from the LPM for the interval (depths) and the three different fits: using all frequencies below the first excited frequency (12 d^{-1}) except the daily (blue); using the first three harmonics of the sawtooth perturbation (red); and only using the daily excitation (green). Here, a) is the amplitude ratio spectrum, b) the phase difference, and c) the error. The corresponding values can be found in Table 2.5.

Table 2.5. Parameter estimates of the diurnal frequency and multiple frequencies belonging to the MLE estimates in Figure 2.7.

f [d^{-1}]	κ [$cm^2\ s^{-1}$]	σ_κ [$cm^2\ s^{-1}$]	q_f [$cm\ s^{-1}$]	σ_{q_f} [$cm\ s^{-1}$]	Pearson factor
1	0.00870	0.00286	0.00256	0.00009	−0.103
$< 12 \backslash \{1 d^{-1}\}$	0.00518	0.00003	0.00230	0.00001	−0.286
12, 24, 36	0.00558	0.00003	0.00217	0.00001	−0.450

frequency alone does not produce a result as accurate, even though it is one of the strongest components in the signal. This is supported by the corresponding variance calculation based on propagation of uncertainty (explained in Section 2.B). Note that model errors (see Section 2.2) can also play a role, which results in the model parameters not being exactly within their confidence bounds as the variance does not account for deterministic errors related to the choice of underlying model used in the analysis (as is the case in Table 2.5 for the two bottom rows).

2.4.6 Model error identification

Several methodologies are available to identify model errors. Straightforwardly, model errors can be detected when the confidence bounds at different frequencies do not overlap. Alternatively, residual errors and their statistical properties can be assessed, which is especially powerful when using independent training and validation sets (machine learning). Also, whiteness residual tests are powerful in determining if dynamics is left in the data or the residual is white, i.e., only noise. These two methods are applied for the ADE, (2.1), in (Berkel et al. 2013). For high noise levels, many potential models could in theory fit within the error bounds. However, when the noise level is low, only a few of these models will fit, and we can start assessing where our (deterministic) mathematical model deviates from the actual real model. This approach is most powerful when applied across a large frequency range. A good example is the estimate purely of the diurnal cycle shown in Figure 2.7, which has an error significantly larger error than the other estimates for the entire frequency range (exception directly around the daily cycle).

If the above analysis is combined with the concept of linearization (sufficiently small perturbation) and corresponding linear time invariant model descriptions (both discussed in Section 2.2), we note that $G(\omega)$ is uniquely defined, which means in practice that we can confirm the degree to which the system is really linear time-invariant or has some other physical description. This is one of the most powerful methods to identify and confirm the physical description embodied in the model. This utility of multi-frequency analysis might be even more important than the multi-frequency estimation of model parameters.

2.4.7 Multi-depth analysis

A related approach to using multiple frequencies is using sensors at multiple depths (Vandersteen et al. 2015). It has been extended to also include noise on the top sensor used to determine the model parameters. This results in a matrix representation of the problem for both the noise (including (2.14)) and the transfer functions (matrix) which can be analytic for uniform domains (Kampen et al. 2022d). One problem is that measurements at depth will contribute less and less information (as was the case for higher frequencies) due to their lower signal to noise ratio. Hence, especially in the uniform case, the top sensors and low frequencies (at depth) will dominate the solution that is calculated.

2.5 A forward-looking perspective

Current analytical frameworks using heat as a tracer to quantify groundwater - surface water exchange fluxes are not immediately transparent about the origins of the limitations and uncertainties encountered when applying those approaches. Here, we provide a framework in which to consider and address some of those complexities. In part, we do that by separating among model conceptualization, data analysis, and parameter quantification to help identify, isolate, and resolve some of the sources of error. In part, we further the aims of improving estimates by noting some recent advances in signal processing and estimation methods that reduce the uncertainty from how we process the data so that we can report uncertainties in estimated parameters of interest, typically fluxes and sediment-water thermal properties. Our evaluations of methods against laboratory data show that choices of model structure, signal processing technique, and parameter estimation method have profound effects on parameter and parameter uncertainty estimates. Perhaps of similar importance, previously unwanted signal components can reveal information relevant to model structure and parameter uncertainty. We propose that separating the analysis of decisions in the different steps of application will help the community identify the limitations of individual choices made at each step, leading to more reliable interpretation of the results and their uncertainties. We know based on approaches in system theory that such a systematic framework is necessary to achieve reliable parameter estimates in more complex challenges such as multi-dimensionality, heterogeneity, and in the case of non-linear dependencies in the parameters. Consequently, this may also lead to the possibility to assess the model structure uncertainties systematically with respect to heterogeneity and nonlinear processes. There is a parallel here with modeling frameworks being generated in other areas of hydrology (e.g., (Clark et al. 2015), where an initial decomposition into components of a solution can aid in systematic improvements.

Appendix

2.A Variance calculation in the frequency domain

In an ideal case, noise can be quantified by removing the deterministic (periodic) components from the signal and calculating the variance(s) of the remaining data, which then consists purely of noise without transients. We cannot achieve this in practice as the deterministic signal is unknown, so we seek to determine the noise level based on overdetermination. In the frequency domain, this is conceptually done by calculating the variance over periods shown in Figure 2.8.

Figure 2.8 shows that starting from a deterministic periodic signal to which stationary Gaussian distributed noise is added, results in a time-independent spectrum. When this noisy signal is transformed into the frequency domain, it results in a variation of the Fourier coefficient per period around the deterministic Fourier coefficient similar to time samples varying around the deterministic signal. In practice, this means we transform a 1D distribution function to a 3D distribution function in terms of the real-, imaginary-, and frequency axis. As Figure 2.5 shows, the resulting distribution function when we assume Gaussian (time-domain) noise is complex (-valued) with circular normal distribution (CCND) (Goodman 1963). In practice, with arbitrary forcing, the variance calculation procedure presented in Figure 2.8 is slightly different, but conceptually the same because the same distributions in Figure 2.5 hold. For the calculation of the variances, transient removal is required, for which we use the LPM as explained in Section 2.3.4. Then, the variance level is estimated over a small frequency range instead using individual frequencies as in Figure 2.8. The latter method has been shown to be reliable and has been cross-validated against other methods (Monteyne et al. 2013).

2.B Confidence bound calculation

Calculating confidence bounds using Monte-Carlo simulations is computationally expensive and for a large number of frequencies, infeasible; also explaining why (only) four multi-frequency analyses are performed in Table 2.4. However, as we have attempted to keep the estimates as Gaussian as possible and achieving near minimum variance based on the data for the MLE, it becomes acceptable to use

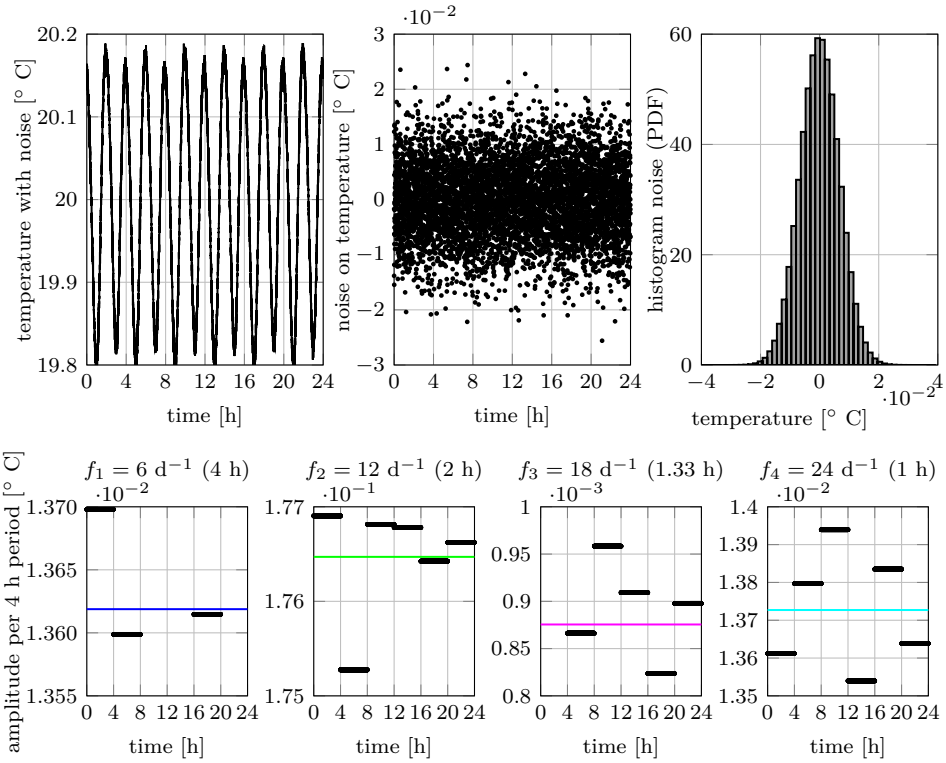


Figure 2.8. Graphical representation of the variance calculation of a single signal used to determine confidence bounds: (top) shows the signal, the (additive) noise, and its Gaussian distribution (histogram); (bottom) the Fourier coefficients per period at different (independent) frequencies with in color the average Fourier coefficient which is by definition equal to that of the entire time-window.

the standard technique of *propagation of uncertainty* to determine the confidence bounds. Propagation of uncertainty is based on the Taylor expansion around a (differentiable) non-linear function, assuming that we can approximate from a Gaussian distribution to a (multivariate) Gaussian distribution with different mean and variance.

For a function $f(x)$ depending on one parameter, x, this results in $\sigma_f^2 \approx \left|\frac{\partial f}{\partial x}\right|^2 \sigma_x^2$. Many of us also recognize the standard two-dimensional formula for, $f(x,y)$ gives $\sigma^f \approx \left|\frac{\partial f}{\partial x}\right|^2 \sigma_x^2 + \left|\frac{\partial f}{\partial y}\right|^2 \sigma_y^2 + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\sigma_{xy}^2$. This is also exactly used in (2.13) (for complex-values) to calculate the variance of the transfer function based on the variance of the Fourier coefficients at the two distinct channels. Using the same two-dimensional formula for our function writing $f(x,y) = G(\kappa, q_f)$ for two frequencies, hence, results in matrix form (with the $\frac{1}{2}$ compensating for the complex circular variance transformation to real-valued variance):

$$\frac{1}{2}\begin{bmatrix}\sigma_G^2(\omega_1) & 0\\ 0 & \sigma_G^2(\omega_2)\end{bmatrix} = J_G^{\mathsf{H}}\begin{bmatrix}\sigma_\kappa^2 & \sigma_{\kappa q_f}^2\\ \sigma_{q_f\kappa}^2 & \sigma_{q_f}^2\end{bmatrix} J_G, \tag{2.18}$$

with

$$J_G = \frac{\partial G(\hat{\theta})}{\partial \theta} = \begin{bmatrix}\frac{\partial G(\hat{\theta},\omega_1)}{\partial \kappa} & \frac{\partial G(\hat{\theta},\omega_2)}{\partial \kappa}\\ \frac{\partial G(\hat{\theta},\omega_1)}{\partial q_f} & \frac{\partial G(\hat{\theta},\omega_2)}{\partial q_f}\end{bmatrix} \tag{2.19}$$

To calculate the covariance matrix of κ and q_f denoted by $\mathrm{Cov}\left(\hat{\theta}\right)$, we need to invert this relationship. This can be done by exploiting the diagonality of the $\mathrm{Cov}\left(G(\hat{\theta},\omega_k)\right)$ such that it can be combined with J_G and as the derivative to $\frac{\partial \hat{G}(\omega_k)}{\partial \theta} = 0$, we can calculate the real-valued $\mathrm{Cov}(\hat{\theta})$ using the inversion

$$\mathrm{Cov}(\hat{\theta}) = \left[\mathrm{Re}\left(2J_\theta^{\mathsf{H}} J_\theta\right)\right]^{-1}, \tag{2.20}$$

with

$$J_\theta = \frac{\partial}{\partial \theta}\left(\frac{\hat{G}(\omega_k) - G(\hat{\theta},\omega_k)}{\hat{\sigma}(\hat{\theta},\omega_k)}\right), \tag{2.21}$$

which has special significance as J_θ is used in most gradient-descent based algorithms to estimate $\hat{\theta}$ and hence does not need to be calculated extensively. The variances calculated based on this method are also compared to the variances from the Monte-Carlo analysis in Table 2.4 showing the applicability of this calculation. The confidence can be calculated using the p value, e.g., $\pm 1.96\sigma_\kappa \approx 95\%$ confidence bounds.

3

Estimating locally constant transport coefficients taking uncertainty into account

Abstract - This chapter presents the LPMLEn, a new method to estimate vertical flux and thermal diffusivity from streambed temperature time series using the frequency domain. The main advantages of this new method are: (a) the use of multiple frequencies and multiple sensors for the parameter estimation; (b) noise/uncertainty handling in an optimal way; (c) the possibility to estimate the parameters with both semi-infinite and bounded domain models; and (d) the compensation for temperature drifts in the data known as transients. The capabilities of the LPMLEn are demonstrated using both synthetic and field data, highlighting the advantages of the bounded domain model over the semi-infinite domain model in the parameter estimation process.

This chapter is published as: Kampen, R. J. R. van, U. Schneidewind, C. Anibas, et al. (2022d). "LPMLEn A Frequency Domain Method to Estimate Vertical Streambed Fluxes and Sediment Thermal Properties in Semi-Infinite and Bounded Domains". In: *Water Resources Research* 58.3 (Feb. 2022), e2021WR030886. DOI: 10.1029/2021wr030886.

3.1 Introduction

The use of heat as a tracer has become an important and frequently applied tool to quantify water fluxes across streambeds (Abbott et al. 2016; Irvine et al. 2017) since available temperature sensors are inexpensive, robust, and reliable. Numerous studies have explored the theory behind and application of heat as a tracer to quantify water flow in and across riverbed sediments (Constantz 2008; Rau et al. 2014) whereas others have focused on the development of methods and devices to collect temperature time series at multiple streambed depths and locations (Banks et al. 2018; Briggs et al. 2012; Schmidt et al. 2014). To estimate vertical fluxes from these now more and more abundantly available, temperature-time series data, a variety of model codes and software tools (Ford et al. 2021; Irvine et al. 2015; Koch et al. 2016; Munz and Schmidt 2017) have been devised that make use of analytical (Goto et al. 2005; Hatch et al. 2006; Keery et al. 2007; Luce et al. 2013) or numerical solutions (Lapham 1989) of the 1D heat transport equation.

The vast majority of previous studies interested in the quantification of the vertical exchange flux across streambeds (often also called Darcy flux) from temperature-time series data has made use exclusively of the diurnal temperature signal (Fanelli and Lautz 2008; Hatch et al. 2010; Jensen and Engesgaard 2011). Although, the diurnal frequency represents the strongest temperature signal in many field settings,(Wörman et al. 2012) demonstrated how additional information contained in the temperature signal gained from assessing multiple frequencies can potentially improve the estimation of flux and thermal diffusivity. Subsequently, methods were put forward by (Vandersteen et al. 2015; Schneidewind et al. 2016; Sohn and Harris 2021), that solve the 1D heat transport equation after (Carslaw and Jaeger 1959) in the frequency domain and explicitly consider multiple frequencies during flux estimation (see Table 3.2 in Section 3.A for a concise overview).

As not all frequencies are equally informative, multi-frequency analysis requires that noise (i.e., uncertainties) at individual frequencies be accounted for (Berkel et al. 2014a). (Vandersteen et al. 2015; Schneidewind et al. 2016; Sohn and Harris 2021) all handle the noise in an optimal way when estimating flux and thermal diffusivity with a maximum likelihood estimator (MLE). Differences in their methods are found in how these uncertainties are determined, how the model domain is delineated and the used information content. For example, (Sohn and Harris 2021) use the multitaper method while (Vandersteen et al. 2015; Schneidewind et al. 2016) apply the local polynomial (LP) method to estimate uncertainties and reduce leakage. Where (Sohn and Harris 2021) as well as (Vandersteen et al. 2015) considered the subsurface as semi-infinite, (Schneidewind et al. 2016) developed an early bounded (finite) domain model with the idea to estimate vertical flux for distinct streambed sections. This early bounded model (LPMLE3) utilized information from three (noisy) sensors, two for the input (top and bottom boundaries) and the third for the output.

Here, we introduce LPMLEn, which builds upon works from (Vandersteen et al. 2015; Schneidewind et al. 2016) by extending the bounded LPMLE3 method to

include temperature data from n sensors in the parameter estimation process and by considering noisy input and output data for both semi-infinite and bounded domains. The possibility to use n sensors better constrains parameter uncertainty whereas the lower boundary condition better insulates the parameter estimation process from spatial and temporal heterogeneities outside the probed domain. This is especially helpful when flux and thermal diffusivity are estimated within small vertical subsections of the streambed to detect streambed heterogeneity. The advantage of using n sensors within bounded domains is beneficial in the analysis of flux within heterogenous sediment when data are collected with temperature probes containing many sensors such as multilevel temperature lances (Munz and Schmidt 2017) or FO-DTS (Selker et al. 2006).

3.2 The LPMLE*n* method

The LPMLE*n* combines the LP method with an MLE to estimate 1D vertical streambed fluxes and thermal diffusivities using data from n temperature sensors. It operates in the frequency domain and can use multiple frequencies and sensors simultaneously for the parameter estimation process. The LP method is applied to reduce leakage (transients) and estimate the noise levels at the selected frequency components of the measured temperature signals. The LPMLE*n* is provided here with two models: (a) the semi-infinite domain model where only an upper temperature boundary condition is specified to estimate the parameters and (b) a bounded (finite) domain model where an additional lower local temperature boundary condition is assigned to estimate the parameters for a distinct section of the streambed.

Like other 1D models, LPMLE*n* assumes local thermal equilibrium (or a slow-moving equilibrium without significant change of diffusivity and vertical flux), steady water flow and constant thermal parameters over the model domain. While it provides information on the vertical flux, we would like to remind potential users that in natural settings water flow has vertical and nonvertical flow components (horizontal or lateral). While flowlines at the center of a stream are formed in such a way that they often indicate predominantly vertical flow (Cuthbert and Mackay 2013; Shanafield et al. 2010), it has been shown that nonvertical flow components can often become significant closer to streambanks or in areas of high hyporheic flow (Lautz 2010; Reeves and Hatch 2016; Roshan et al. 2014). Nonvertical flow does not invalidate 1D heat transport estimates, however curvature in the flow path may, depending on the degree of curvature (Cuthbert and Mackay 2013). Areas with high curvature in flow paths warrant the use of more complex and data-intensive 3D heat transport models (Ghysels et al. 2021; Karan et al. 2014).

Coupled vertical (1D) water flow and heat transport defined after (Stallman 1965) is given by the advection-diffusion equation

$$\frac{\partial}{\partial t}T = D\frac{\partial^2}{\partial z^2}T + V\frac{\partial}{\partial z}T, \tag{3.1}$$

where the temperature $T(z,t)$ [Θ] around an equilibrium is a function of depth z [L] and time t [T]. Here, time-invariant parameter V [LT^{-1}] represents advection given by

$$V = -q_z \frac{\rho_w c_w}{\rho c}, \tag{3.2}$$

with ρc [$\mathrm{ML}^{-1}\mathrm{T}^{-2}\Theta^{-1}$] as the volumetric heat capacity of the water-sediment mix, $\rho_w c_w$ [$\mathrm{ML}^{-1}\mathrm{T}^{-2}\Theta^{-1}$] the volumetric heat capacity of water, and q_z [LT^{-1}] as the vertical flux (a positive value means downward flow). The parameter D [$\mathrm{L}^2\mathrm{T}^{-1}$] represents the effective thermal diffusivity and can be described by

$$D = \frac{\kappa}{\rho c} + \gamma(\psi, q_z), \tag{3.3}$$

where κ [$\mathrm{MLT}^{-3}\Theta^{-1}$] is the bulk thermal conductivity and γ is a function based on q_z and the thermal dispersivity ψ [L]. Here, we use the LPMLEn to estimate D without prior knowledge of either thermal conductivity or dispersivity. If all thermal parameters in (3.3) were known (e.g., through prior field or lab measurements), the LPMLEn could also be used to only estimate the Darcy flux, but this is discouraged because D and V are coupled. However, this prior knowledge could be used to validate the estimation method and model choice (Luce et al. 2013).

As the solution to (3.1) is determined by its boundary conditions, we first cover the analytical solution for two common choices, (a) the semi-infinite domain and (b) the bounded domain with Dirichlet (temperature specified) boundary conditions. As we provide the analytical solutions in the frequency domain, but the measured data are time series, a transformation to the frequency domain and additional processing are required, for which we use the Fast Fourier Transform (FFT) and the LP method, respectively. The unknown parameters are then estimated using the MLE considering multiple frequencies, n sensors and their uncertainty (the MLEn part).

3.2.1 The models and their analytical solutions

For well-posedness, i.e., to obtain a unique solution to (3.1), two boundary conditions are required. The two choices commonly applied are: (a) the semi-infinite domain where the upper boundary condition at depth z_U is given by the function $T_U(t)$ and the lower boundary condition is set to approach zero at infinity (Hatch et al. 2010; Luce et al. 2013; Sohn and Harris 2021; Stallman 1965; Vandersteen et al. 2015) and (b) the bounded domain where the upper boundary condition at depth z_U is given by the function $T_U(t)$ and the lower boundary at depth z_L by the function $T_L(t)$ (Schneidewind et al. 2016). The semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ is given by

$$\mathcal{M}_{\mathrm{SI}} : \begin{cases} \frac{\partial}{\partial t} T & = D \frac{\partial^2}{\partial z^2} T + V \frac{\partial}{\partial z} T \\ T(z_U, t) & = T_U(t) \\ \lim_{z \to \infty} T(z,t) & = 0, \end{cases} \tag{3.4}$$

and the bounded domain model $\mathcal{M}_{\mathrm{BD}}$ by

$$\mathcal{M}_{\mathrm{BD}}: \begin{cases} \frac{\partial}{\partial t}T & = D\frac{\partial^2}{\partial z^2}T + V\frac{\partial}{\partial z}T \\ T(z_U,t) & = T_U(t) \\ T(z_L,t) & = T_L(t). \end{cases} \tag{3.5}$$

Using a transfer function notation, the solution to both models is given according to van (Berkel et al. 2013; Berkel et al. 2014c) by

$$\Theta(z,s) = G_U(z,s,\theta)U_U(s) + G_L(z,s,\theta)U_L(s), \tag{3.6}$$

where $\Theta(z,s)$ is the Laplace transform of $T(z,t)$, while $U_U(s)$ and $U_L(s)$ are the Laplace transforms of the functions describing the upper and lower boundary conditions, respectively. The transfer functions $G_U(z,s,\theta)$ and $G_L(z,s,\theta)$ for the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ are given by

$$G_U(z,s,\theta) = e^{\lambda_1(s,\theta)(z-Z_U)} \tag{3.7}$$

$$G_L(z,s,\theta) = 0, \tag{3.8}$$

and for the bounded (finite) domain model $\mathcal{M}_{\mathrm{BD}}$ by

$$G_U(z,s,\theta) = \frac{\xi(z_L,s,\theta)\zeta(z,s,\theta) - \zeta(z_L,s,\theta)\xi(z,s,\theta)}{\zeta(z_U,s,\theta)\xi(z_L,s,\theta) - \zeta(z_L,s,\theta)\xi(z_U,s,\theta)} \tag{3.9}$$

$$G_L(z,s,\theta) = -\frac{\xi(z_U,s,\theta)\zeta(z,s,\theta) - \zeta(z_U,s,\theta)\xi(z,s,\theta)}{\zeta(z_U,s,\theta)\xi(z_L,s,\theta) - \zeta(z_L,s,\theta)\xi(z_U,s,\theta)} \tag{3.10}$$

with

$$\xi(z,s,\theta) = e_1^{\lambda}(s,\theta)z \tag{3.11}$$

$$\zeta(z,s,\theta) = e_2^{\lambda}(s,\theta)z \tag{3.12}$$

$$\lambda_{1,2}(s,\theta) = -a \mp \sqrt{a^2+bs}, \tag{3.13}$$

and $\theta = [a,b]^{\mathsf{T}} = \left[\frac{V}{2D}, \frac{1}{D}\right]$ as a simplification of the parameters for the eigenvalues, described in detail in (Berkel et al. 2014c). The aim is to estimate $\theta = [a,b]^{\mathsf{T}}$, which then can be transformed back to D and q_z. However, to estimate the unknown parameters, the model needs to be linked to the temperature time series. In practice, the Laplace variable can only be measured on the imaginary axis, thus $s = i\omega$, where $i = \sqrt{-1}$ and ω is the angular frequency. Any measured temperature time series can be transformed to the frequency domain using the FFT such that it is a function of $i\omega$. However, the measured streambed temperature data are commonly not perfectly periodic due to natural conditions such as temperature fluctuations that are slower than the measurement period, introducing spectral leakage (Pintelon and Schoukens 2012). Furthermore, the measured temperatures can be subject to transients due to the initial condition, sensor drift and noise. To reduce spectral leakage and remove the other unwanted contributions to the temperature signal, we apply the LP method to assess and improve the quality of the dataset.

3.2.2 Processing the dataset with the LP method

A standard method to reduce spectral leakage is windowing, which is known to introduce systematic (bias) and random (noise leakage) errors when transforming data to the frequency domain (Pintelon and Schoukens 2012). The LP method (Pintelon et al. 2010a; Pintelon et al. 2010b) outperforms windowing techniques by assuming linear relationships (transfer functions) between a given noiseless reference signal and the measured response, i.e., the LP considers that the signal originates from a system rather than it being independent. The LP method splits the measured signal into a forced response, (circular complex normally distributed) noise, and transients (i.e., the signal part that cannot be explained by the reference signal). Similar to (Vandersteen et al. 2015; Schneidewind et al. 2016), we use the temperature measurement $\Theta(z_1, \omega)$ at the top of the streambed as the "noiseless" reference signal, as this signal contains the largest temperature perturbation and as such is least subjected to measurement noise. As a result, the LP returns the (estimated) forced response $\hat{\Theta}(z_m, \omega)$ and the corresponding covariance matrix $C_{\hat{\Theta}}(\omega)$ for the remaining $m = 2, \ldots, M$ sensors, which will be used for the parameter estimation process. This is in contrast to (Vandersteen et al. 2015) where the "noiseless" reference signal is included in the estimation of V and D. Furthermore, the obtained covariance matrix allows us to assess the quality of the dataset and to take the measurement uncertainty consistently into account during parameter estimation performed subsequently using the MLE.

3.2.3 The *n*-point MLE

The aim of the MLE is to find those parameters $\theta = [a, b]^\mathsf{T}$ for which the output of the selected model (in our case $\mathcal{M}_{\mathrm{SI}}$ or $\mathcal{M}_{\mathrm{BD}}$) is the most likely to generate the measured temperatures. An MLE has been used previously by (Berkel et al. 2013; Berkel et al. 2014c) to estimate the thermal transport coefficients considering slab geometries on a semi-infinite domain with two sensors (MLE2) and a bounded domain with three sensors (MLE3). They were later incorporated into the LPML (Vandersteen et al. 2015) and the LPMLE3 (Schneidewind et al. 2016), respectively. A cylindrical representation of the MLE3 has also been put forward (Berkel et al. 2019).

In a similar fashion, for the LPMLEn, we construct the log-likelihood cost function $\mathcal{V}_{\mathrm{ML}}(\theta, X)$ that considers the uncertain boundary inputs U_U and U_L as in (Schneidewind et al. 2016) and determines the likelihood for the n-outputs Y_n, which are the remnant measurements of the system. For the sake of clarity, we define the new variable $X(\omega_k) = \left[U_U(\omega_k), Y_n(\omega_k)^\mathsf{T}, U_L(\omega_k)\right]^\mathsf{T}$ that contains both the inputs (boundary conditions) and the outputs (other temperature measurements) and has the corresponding covariance matrix C_X. Thus, for $\mathcal{M}_{\mathrm{SI}}$: $\left[\hat{\Theta}(z_2, \omega_k), \ldots, \hat{\Theta}(z_M, \omega_k), 0\right]^\mathsf{T}$ and for $\mathcal{M}_{\mathrm{SI}}$: $\left[\hat{\Theta}(z_2, \omega_k), \ldots, \hat{\Theta}(z_M, \omega_k)\right]^\mathsf{T}$. Using this general notation, the cost function for the log-maximum likelihood (ML)

is given as in (Pintelon and Schoukens 2012) by

$$\mathcal{V}_{\mathrm{ML}}(\theta, X) = \sum_{k \in \mathbb{K}} e(\omega_k, \theta, X)^{\mathsf{H}} \left[C_e(\omega_k, \theta)\right]^{-1} e(\omega_k, \theta, X), \tag{3.14}$$

with the Hermitian transpose $^{\mathsf{H}}$, the angular frequencies that contain relevant information $\omega_k, k \in \mathbb{K}$, which contains a total number of F frequencies. The error is then given by

$$e(\omega_k, \theta, X) = \left[-G_{U,n}(\omega_k), I_n, -G_{L,n}\right] X(\omega_k), \tag{3.15}$$

with its corresponding error covariance matrix

$$\begin{aligned} C_e(\omega_k, \theta) = &\left[-G_{U,n}(\omega_k, \theta), I_n, -G_{L,n}(\omega_k, \theta)\right] C_X(\omega_k) \\ &\left[-G_{U,n}(\omega_k, \theta), I_n, -G_{L,n}(\omega_k, \theta)\right]^{\mathsf{H}}, \end{aligned} \tag{3.16}$$

where I_n denotes the identity matrix of size $n \times n$. The transfer function vectors $G_{U,n}$ and $G_{L,n}$ are given by

$$G_{U,n} = \left[G_U(\check{z}_1, \omega_k, \theta), \ldots, G_U(\check{z}_n, \omega_k, \theta)\right]^{\mathsf{T}} \tag{3.17}$$

$$G_{U,n} = \left[G_L(\check{z}_1, \omega_k, \theta), \ldots, G_L(\check{z}_n, \omega_k, \theta)\right]^{\mathsf{T}}, \tag{3.18}$$

where $\check{z}_j$, $j = 1, \ldots, n$ are the n depths of the sensors considered as output of the model. Note that the cost function (3.14) is constructed from the sum of squares of the error that is normalized by its (co)variance. Each of the $n \times F$ equations in the cost function will cause a χ^2 distribution with $n \times F$ degrees of freedom. Under the assumption that we have no modeling error, the expected value of the cost function is $n \times F - \frac{2}{2}$. Here, the correction factor $\frac{2}{2}$ originates from the two real parameters that are estimated.

All things considered, we have now defined the complete set of noise and transfer functions belonging to the n-point estimator with constant transport coefficients. Note that the LPML in (Vandersteen et al. 2015) results in a much simpler form of (3.15) and (3.16), i.e., $C_e(\omega_k, \theta) = [-0, I_n, 0]\, C_X(\omega_k)\, [0, I_n, 0]^{\mathsf{H}}$, and that in the LPMLE3 in (Schneidewind et al. 2016) the matrix products were written out considering only one of n sensors resulting in scalar expressions for (3.15) and (3.16), which were independently derived at the time.

As with the LPML and LPMLE3, the parameters are estimated by minimizing the cost function $\mathcal{V}_{ML}$

$$\hat{\theta} = \arg\min_{\theta} \mathcal{V}_{\mathrm{ML}}(\theta, X). \tag{3.19}$$

The cost function $\mathcal{V}_{\mathrm{ML}}$ is nonconvex and is therefore often optimized using nonlinear least squares minimization techniques such as Gauss-Newton or Levenberg-Marquardt methods (Fletcher 1980; Levenberg 1944; Marquardt 1963). As the number of parameters is limited and the analytical transfer functions are known,

the analytical Jacobian is determined and used to improve computational efficiency. Furthermore, the analytical Jacobian is used to determine the variance of the estimated parameters following:

$$C_{\hat{\theta}} = \left[2\mathrm{Re}\left\{J_{\mathrm{ML}}(\hat{\theta}, X)^{\mathsf{H}} J_{\mathrm{ML}}(\hat{\theta}, X)\right\}\right]^{-1}, \tag{3.20}$$

with its Jacobian

$$J_{\mathrm{ML}}(\hat{\theta}, X) = \sum_{k\in\mathbb{K}} \frac{\partial}{\partial\hat{\theta}} C_e(\omega_k, \hat{\theta})^{-\frac{1}{2}} e(\omega_k, \hat{\theta}, X), \tag{3.21}$$

where $C^{\frac{1}{2}}$ is a square root of the positive definite matrix C, i.e., $C = C^{\frac{1}{2}} C^{\frac{1}{2}^{\mathsf{H}}}$. Until this point, the variable $\hat{\theta} = [\hat{a}, \hat{b}]^{\mathsf{T}}$ is estimated while we are interested in the variable $\hat{\theta}' = [\hat{D}, \hat{q}_z]^{\mathsf{T}}$. To obtain $\hat{\theta}'$, $\hat{\theta}$ is transformed via

$$\hat{\theta}' = \begin{bmatrix}\hat{D}\\ \hat{q}_z\end{bmatrix} = \begin{bmatrix}\frac{1}{\hat{b}}\\ -2\frac{\hat{a}}{\hat{b}}\frac{\rho c}{\rho_w c_w}\end{bmatrix}, \tag{3.22}$$

where the uncertainty of the variable $\hat{\theta}'$ is obtained using propagation of uncertainty, i.e.,

$$C_{\hat{\theta}'} = J_{\hat{\theta}\to\hat{\theta}'} C_{\hat{\theta}} J^{\mathsf{H}}_{\hat{\theta}\to\hat{\theta}'} \tag{3.23}$$

with the Jacobian of the transformation given by

$$J_{\hat{\theta}\to\hat{\theta}'} = \begin{bmatrix}0 & -\frac{1}{\hat{b}^2}\\ -2\frac{1}{\hat{b}}\frac{\rho c}{\rho_w c_w} & 2\frac{\hat{a}}{\hat{b}^2}\frac{\rho c}{\rho_w c_w}\end{bmatrix}. \tag{3.24}$$

Hence, we can estimate the most likely vertical flux and thermal diffusivity including the corresponding covariance matrix for both parameters.

3.3 Application of the LPMLE*n*

In this section, we estimate the thermal diffusivity and vertical flux by applying the LPMLE*n* on three synthetic and one experimental dataset using both, the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ and the bounded domain model $\mathcal{M}_{\mathrm{BD}}$. The utilization of synthetic data in 1D heat transport modeling to test methodological limitations and better understand the modeling process is a common approach also used by others (Glose et al. 2021; Glose et al. 2021; Irvine et al. 2020; Lautz 2010).

Here, the synthetic datasets are used to show that the semi-infinite domain model can estimate the flow in the wrong direction even though thermal diffusivity and Darcy flow are constant on the domain where the measurements are taken. Moreover, the semi-infinite domain can be seen as a special case of the bounded domain, i.e., on a semi-infinite domain, the bounded domain will still be exact while the inverse is not true. This gives the LMPLE*n* the unique capability to

Table 3.1. Simulated and estimated transport parameters using the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ and bounded domain model $\mathcal{M}_{\mathrm{BD}}$ for a scenario with constant parameters and sensors placed similarly to the experimental setup.

dataset	I (semi-infinite domain)		II (bounded domain)	
	$D \times 10^{-6}$ ($\mathrm{m^2\ s^{-1}}$)	q_z ($\mathrm{mm\ d^{-1}}$)	$D \times 10^{-6}$ ($\mathrm{m^2\ s^{-1}}$)	q_z ($\mathrm{mm\ d^{-1}}$)
simulated parameters	1.200	200	1.200	200
estimated with $\mathcal{M}_{\mathrm{SI}}$	1.200	200	1.830	−17.64
estimated with $\mathcal{M}_{\mathrm{BD}}$	1.200	200	1.200	200

diagnose whether the use of a semi-infinite domain approach is justified for a given dataset and whether the estimated parameters can be deemed accurate. This is demonstrated by applying the LPMLEn on an experimental dataset and comparing the estimation results and model fits.

3.3.1 Synthetic datasets

We apply the LPMLEn with both models on four synthetic datasets. The first two synthetic datasets consist of five temperature signals at depths $z = [0.15, 0.17, 0.20, 0.25, 0.35]$ m, generated by evaluating (3.6) with the transfer functions for the semi-infinite domain model (3.7) and (3.8) resulting in dataset I, and the bounded domain model (3.9) and (3.10) resulting in dataset II. The used input signals are taken from an experimental dataset. Here, the lower boundary condition for dataset II (bounded domain) does not originate from a semi-infinite domain model. Stated differently, the temperature signal at the lower boundary condition contains information that cannot be explained by the semi-infinite domain model, e.g., by a contribution of hyporheic flow. For the sake of clarity, we did not add noise to these datasets and used the identity matrix as the covariance matrix. More information about the generation of the synthetic datasets I and II is presented in Section 3.B. The simulated and estimated transport parameters for both datasets are shown in Table 3.1. For completeness, the model fits are presented in Figure 3.7 in Section 3.B. The estimation results show that if the underlying physics-based model is the semi-infinite domain model (dataset I), both models will estimate the same transport parameters. On the contrary, for dataset II where the lower boundary condition contributes to the solution, the flux estimate obtained with the semi-infinite domain model is in the wrong direction.

datasets I and II are based on constant parameters over the entire model domain. For the semi-infinite domain, this means that diffusivity and flux need to be constant up to infinity and deviations from this assumption can profoundly affect the estimates. We show this by generating synthetic dataset III where the diffusivity is changed from 1.5×10^{-6} $\mathrm{m^2\ s^{-1}}$] to 0.8×10^{-6} $\mathrm{m^2\ s^{-1}}$ at $z = 0.275$ m. The dataset is generated using (3.6) with a numerical approximation of the transfer functions using a central finite difference scheme in MATLAB with $10,001$

points. The upper boundary condition at $z = 0$ m is taken from an experimental dataset and the lower boundary condition at $z = 10$ m is set to zero to mimic a semi-infinite domain. Details about the generation of this dataset can be found in Section 3.C and our MATLAB model is similar to a numerical model presented in (Das et al. 2019). Furthermore, the temperature is known, i.e., measured, every 0.05 m from 0.1 to 0.5 m.

A single set of transport parameters is estimated at each depth using three neighboring sensors. The estimation results are plotted in Figure 3.1 at the middle sensor location, e.g., at 0.15 m, the estimated transport parameters are shown for the estimate using $z = [0.10, 0.15, 0.20]$ m. In Figure 3.1, we clearly see that in shallower depths, above 0.25 m, a large discrepancy exists between the simulated values and the estimated values when using the semi-infinite domain model. This is in contrast to the bounded domain model that estimates the transport parameters well. Moreover, at 0.15 m, the shallowest location tested, the semi-infinite domain model estimated the flow in the wrong direction. The change in diffusivity affects the temperature at shallower depths and the assumptions for the semi-infinite domain are not satisfied. Therefore, the semi-infinite domain model implicitly averages to greater depths during the parameter estimation, and consequently the diffusivity and velocity estimates attempt to compensate for the change in parameters. In contrast, the bounded domain model can isolate individually homogeneous units such that parameter estimates are correct, and only generate small errors in averaging across the step in diffusivity.

In addition to synthetic dataset III and Figure 3.1, Figure 3.8 in Section 3.D contains the estimation results of the diffusivity change in the opposite direction (dataset IV), i.e., from 0.8×10^{-6} to 1.5×10^{-6} m^2 s^{-1}. The results are similar to Figure 3.1, but now with the semi-infinite domain model overestimating the transport parameters instead of underestimating them.

In conclusion, we have shown two scenarios where the LPMLEn can help us detect when estimates of the semi-infinite domain approach are inaccurate compared to using a bounded domain model. Additionally, the LPMLEn with the bounded domain model might prove extremely helpful analyzing the vertical flux of distinct small (local) domains with variations in streambed sediment taking multiple sensors into account, especially for temperature data with very high vertical spatial resolution (e.g., in the cm range) as is typically the case for data collected with FO-DTS where the fiber-optic cable is coiled around a PVC core (Briggs et al. 2012; Folegot et al. 2018; Vogt et al. 2010).

3.3.2 Experimental dataset

Here, the LPMLEn is demonstrated on a 90-day temperature time series (location ML1) from the Slootbeek, a small tributary to the Aa River, Belgium. The field site and flow system of the River Aa and Slootbeek have been described in detail in previous studies (Anibas et al. 2018; Anibas et al. 2018; Ghysels et al. 2021). Temperature was measured using a multilevel temperature lance (UIT, Dresden,

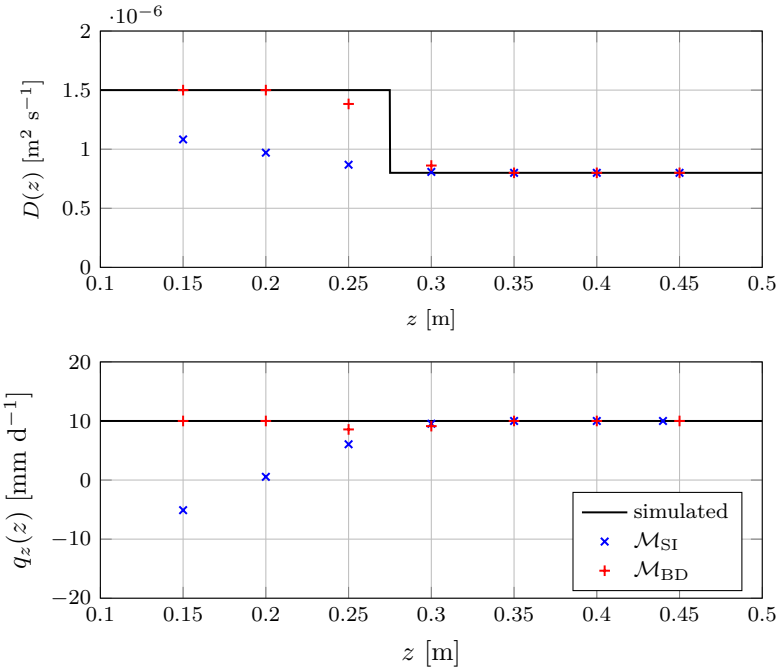


Figure 3.1. Simulated and estimated transport parameters using the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ and bounded domain model $\mathcal{M}_{\mathrm{BD}}$ for a scenario with a change in diffusivity, where the sensors are placed from 0.1 to 0.5 m, every 0.05 m and the parameters are estimated using a moving triplet of consecutive sensors for which the estimate is shown at the center of the triplet. Above 0.25 m, a significant discrepancy exists between the simulated values and the estimated values for the semi-infinite domain model but not for the bounded domain model. At 0.15 m, the semi-infinite domain model estimated the flow in the wrong direction.

Germany) from 17 February to 12 May 2012 every 10 min at the streambed top and six additional depths (see Figures 3.2a and 3.2b). To determine q_z following (3.22), we used $\rho_w c_w = 4.18 \times 10^6$ Jm^{-3} K^{-1} and $\rho c = 3.07 \times 10^6$ Jm^{-3} K^{-1}. The value for ρc is based on previous research from (Vandersteen et al. 2015) and closely represents a sandy loam (Ren et al. 2000; Stonestrom and Constantz 2003).

Similar to (Vandersteen et al. 2015), we analyze the experimental dataset using a rectangular 10-day window. To reduce the computational load during parameter estimation, only the relevant or informative frequencies should be included in the analysis. To select the relevant frequencies, the dataset is transformed to the frequency domain using the FFT and then processed using the LP method with sensor 1 as noiseless reference signal. As the LP method provides the covariance matrix of the processed Fourier coefficients for the other sensors, the ratio between the amplitude squared and the variance of the Fourier coefficients, i.e., signal-to-noise ratio (SNR), shown in Figure 3.2c for one 10-day window, can be used to select the informative frequencies for the parameter estimation process. Similar to (Vandersteen et al. 2015), we choose to include all frequencies up to 1.5 d^{-1}. While selecting the relevant frequencies and sensors for the parameter estimation is ultimately the choice of the modeler this selection will influence the parameter estimates. However, including noninformative frequencies in the parameter estimation process will have little impact on the estimates because the MLEn uses the error (co)variances as weighting, hence signals with a low SNR have a smaller contribution to the solution. They will influence (increase) the expected value of the cost function and delay the computation. On the other hand, discarding informative frequencies will influence the estimation depending on the (co)variance.

Comparatively, excluding a sensor has a more significant effect on the estimates than excluding noninformative frequencies as it also changes the domain on which the parameters are estimated. In (Vandersteen et al. 2015), the parameters are estimated using sensors 17, while we here estimate the parameters using sensors 26 as sensor 1 is used as noiseless reference signal for the LP method and sensor 7 is excluded due to the low SNR at some intervals. In addition, our bounded domain model $\mathcal{M}_{\text{BD}}$ uses the bottom sensor as boundary condition, i.e., $n = 3$, whereas the semi-infinite domain model $\mathcal{M}_{\text{SI}}$ considers this sensor as an output ($n = 4$). Consequently, the semi-infinite domain model is fitted over one additional sensor depth. For this reason, we also estimate the parameters with the semi-infinite domain model using sensors 25, such that the number of outputs between semi-infinite and bounded domain models is the same ($n = 3$). However, note that this estimate is then based on a reduced dataset compared to the bounded domain model. The diffusivity, vertical flux, and the (normalized) cost function value are obtained by moving the 10-day window along the dataset (Figure 3.3). As the number of equations nF is much larger than the correction factor $\frac{2}{2}$, the expected value $\frac{\mathcal{V}_{\text{ML}}}{nF} \approx 1$ if there are no modeling errors. Hence, the difference with this expected value reflects the modeling error.

From Figure 3.3c, we see that, in general, the bounded domain model has

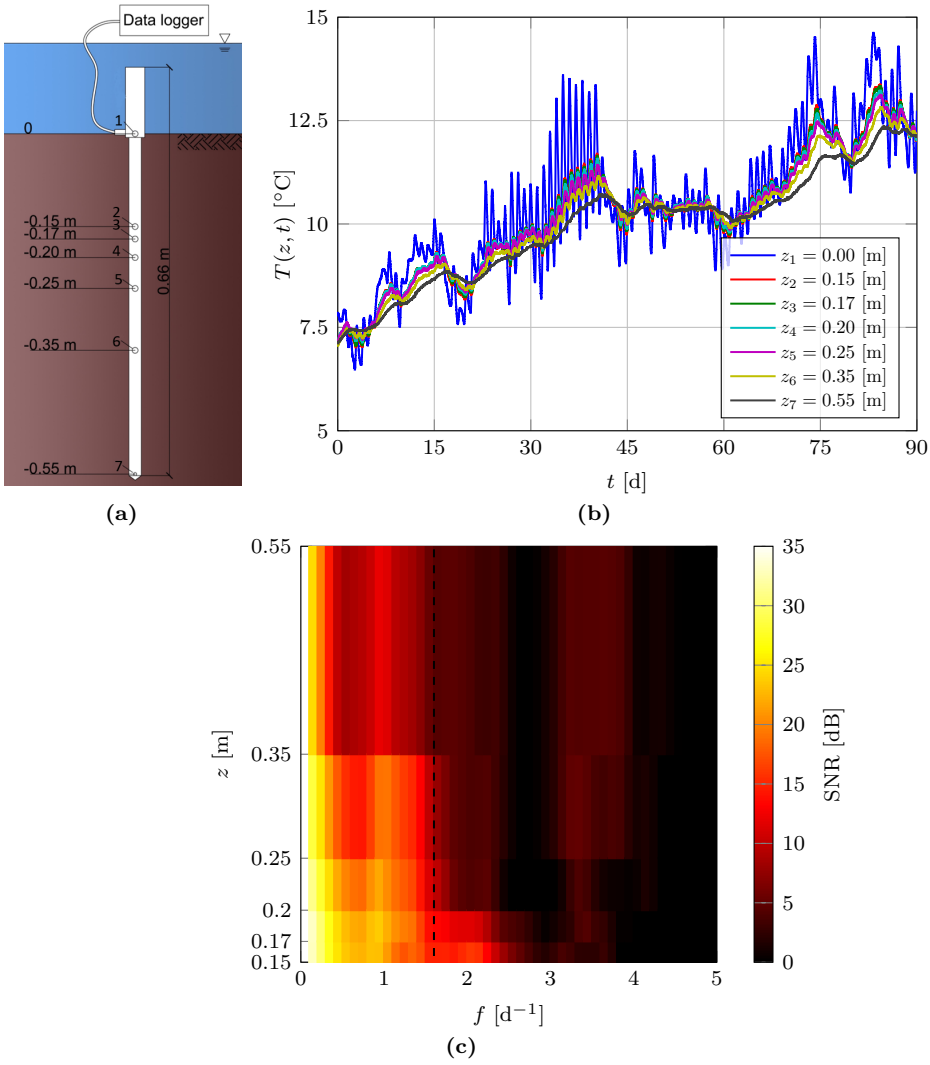


Figure 3.2. (a) Multilevel temperature lance used to collect data, (b) temperature data from the Slootbeek at ML1 (both modified from Vandersteen et al. 2015), (c) example of signal-to-noise ratio (SNR) plot for one 10-day window after using the local polynomial method. It can be used to select informative frequencies. We chose to include all frequencies up to 1.5 d^{-1}, indicated by the vertical dashed line.

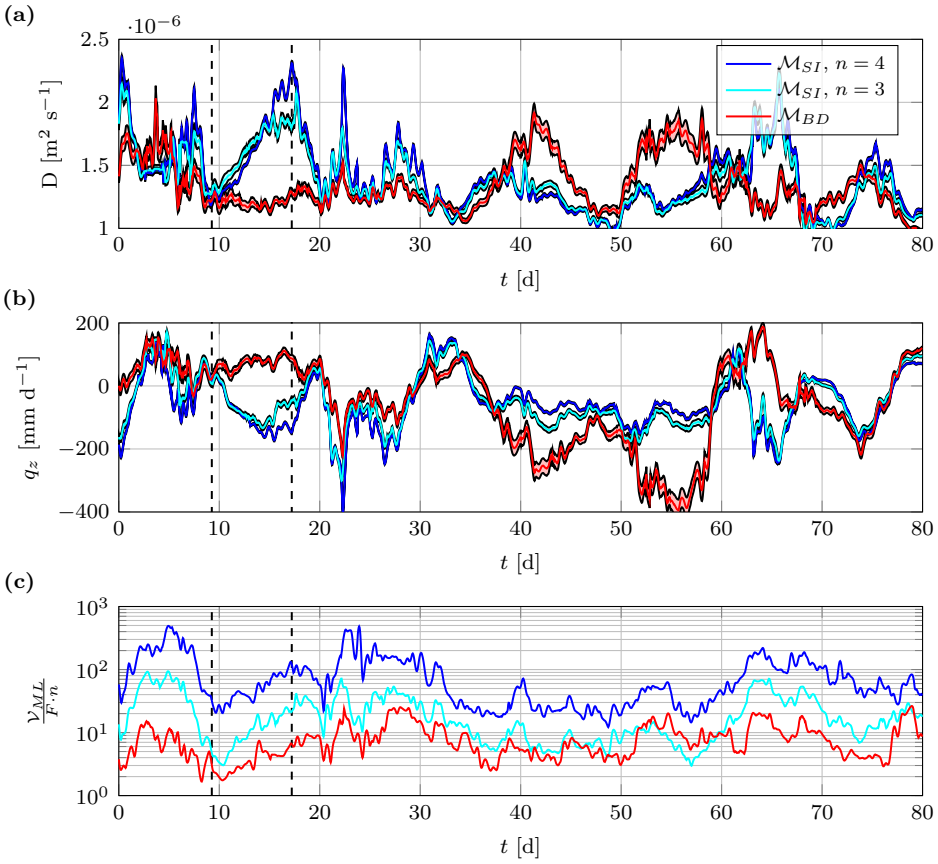


Figure 3.3. (a) Estimated diffusivity, (b) Darcy flux (positive means downwelling) with their 95% confidence bounds, and (c) (normalized) cost function value, shown at the start of the analyzed 10-day window for the semi-infinite domain $\mathcal{M}_{\mathrm{SI}}$ and bounded domain $\mathcal{M}_{\mathrm{BD}}$ models. The two vertical dashed black lines indicate the time for the model fits shown in Figure 3.4.

the smallest modeling error. Moreover, while on some intervals, estimates with the semi-infinite and bounded domain models agree, at other intervals the flow is estimated in opposite directions (Figure 3.3b, e.g., between day 10 and day 20). As shown with the synthetic dataset, if both models agree in their estimates, the estimates can be deemed trustworthy. However, in case the two models do not agree, the bounded domain model is more likely to estimate the correct flow. To verify this, we can look at how the models fit the data. For clarity, the model fits at $t_1 = 9.2569$ and $t_2 = 17.2153$ are shown in Figure 3.4 without the error bars on the measured data, while a figure with error bars is shown in Figure 3.9 in Section 3.E. In these figures, we see that at t_1 all models fit the data very similarly, while at t_2 the bounded domain model describes the phase behavior better than the semi-infinite domain models do.

Additional evidence is provided by looking at the joint variation of the estimated diffusivity and flux from Figure 3.3a and Figure 3.3b, presented in Figure 3.5. As mentioned previously, we estimated the thermal diffusivity instead of the bulk thermal conductivity as there is a physical basis to expect variation of the diffusivity with flux (3.3), namely, hydraulic dispersion. Two common notations of the thermal dispersivity are given by (Marsily 1986; Rau et al. 2012)

$$D = \frac{k}{\rho c} + \psi \left| \frac{\rho_w c_w}{\rho c} q_z \right|, \tag{3.25}$$

$$D = \frac{k}{\rho c} + \psi \left(\frac{\rho_w c_w}{\rho c} q_z \right)^2. \tag{3.26}$$

The estimates originating from the bounded domain model $\mathcal{M}_{\mathrm{BD}}$ follow that pattern much more strongly than the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$, suggesting that some of the variations estimated by the semi-infinite domain model are spurious. A linear least squares fit of (3.25) and (3.26) on the bounded domain estimates is shown along the estimates in Figure 3.5. As both fits do not perfectly fit the data, further research is necessary to establish more insight into the joint variation of the estimated parameters. However, one should note that the observed variations are not the result of small variations within the dataset, but originate from larger trends as small variations, e.g., estimates originating from the next and previous 10-day windows, would be close to each other (within each others 95% confidence ellipses, see Figure 3.10 in Section 3.F).

Combining the knowledge from the synthetic dataset, the model fits and the variation of the diffusivity with flux, indicates to us that the semi-infinite domain model estimates the flow in the wrong direction and that the estimates originating from the bounded domain model are more trustworthy. However, as with any model no absolute guarantee exists that all underlying physical processes in our natural system are described sufficiently well.

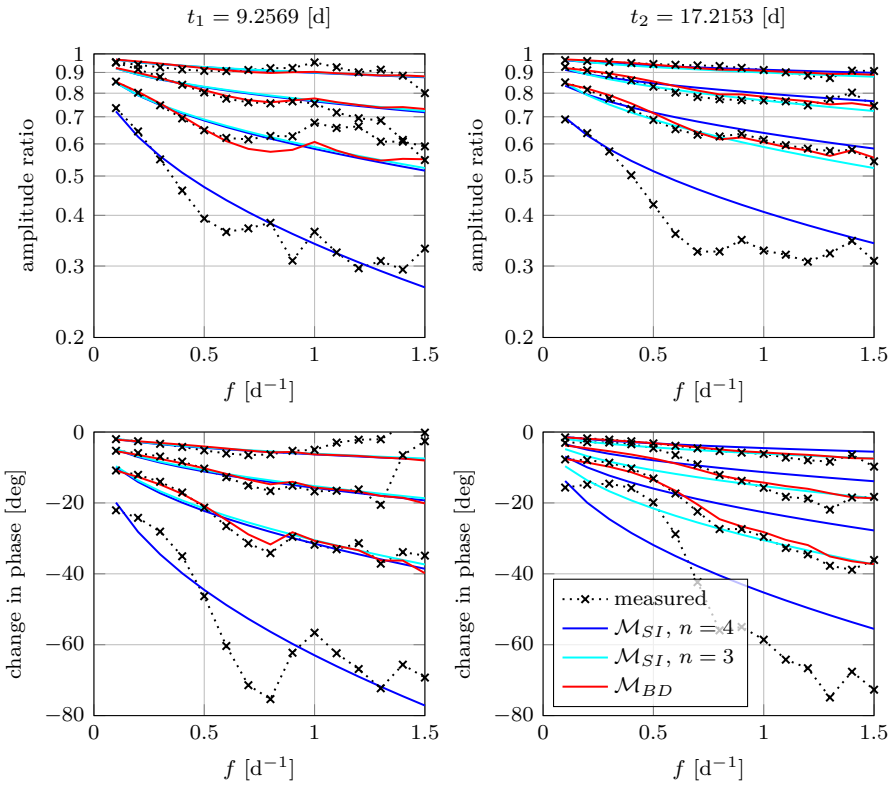


Figure 3.4. The amplitude ratio (top) and change in phase (bottom) of the measured frequency response functions and fitted semi-infinite domain model $\mathcal{M}_{\text{SI}}$ and bounded domain model $\mathcal{M}_{\text{BD}}$ at the two times shown with dashed lines in Figure 3.3. The shown amplitude ratios and changes in phase are from sensor 2 at 0.15 m to sensors 3, 4, 5, and 6 at 0.17, 0.20, 0.25, and 0.35 m, respectively. Note that sensor 6 is used as boundary input for the bounded domain model; thus, the fitted response would equal the measured response and is left out for clarity.

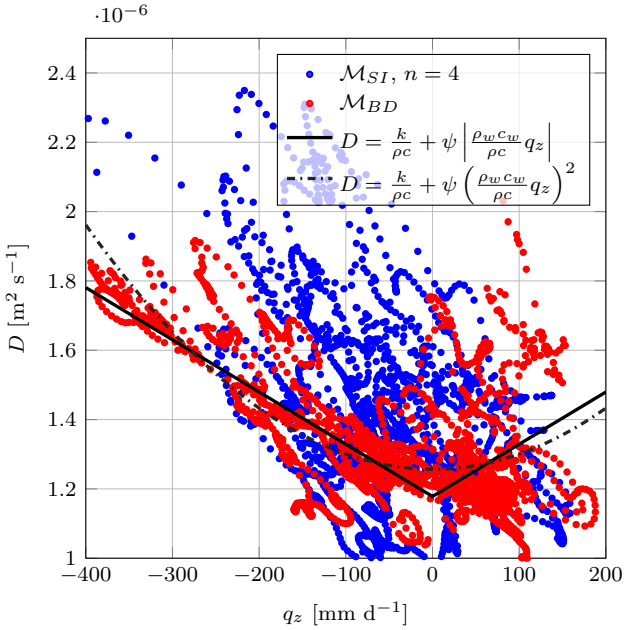


Figure 3.5. Diffusivity as a function of flux for the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ and bounded domain model $\mathcal{M}_{\mathrm{BD}}$ compared to two common notations that are fitted on the estimates from the bounded domain model.

3.4 Conclusion

The LPMLEn extends previous frequency domain approaches to estimate vertical streambed fluxes and thermal diffusivities by including temperature data from n sensors in the multifrequency parameter estimation process. It contains the commonly used semi-infinite domain model, and a more robust and physically based bounded (finite) domain model. Processing the data with the LP method and systematically taking the measurement uncertainty into account by minimizing the (log) ML cost function result in an estimate of the parameters and their corresponding uncertainties.

The application of the two models to the same dataset is crucial in verifying whether the use of the semi-infinite domain approach is justified and the estimated diffusivity and Darcy flux can be deemed accurate. The semi-infinite domain model is a special case of the more general, bounded (finite) domain model; hence, the semi-infinite domain model is more likely to estimate the flux erroneously whereas the bounded domain model has a higher likelihood to estimate the direction of the flux correctly as is demonstrated using synthetic datasets as well as field measurements. In case estimates obtained with both model types agree, model fits and cost function values are also very similar. Where the estimated parameters do not agree, the bounded domain model fits the data better and has a much lower cost function value. As such, the LPMLEn with the bounded model can be used to verify assumptions made for the semi-infinite domain model. Additionally, if the dataset allows for it, the bounded domain model can easily be used to estimate parameters for distinct vertical streambed sections, which might shed light on the contribution of shallow hyporheic flux and deeper groundwater upwelling to the overall flux estimate in future studies.

Appendix

3.A Summary of frequency domain methods in hydrology

Table 3.2. Summary of characteristics of frequency domain models used to estimate vertical streambed fluxes from temperature-time series.

name	source	model type	sensors	frequency	noisy output	noisy input
Luce et al.	Luce et al. 2013	semi-infinite	2	single	no	no
MLE2	Berkel et al. 2013	semi-infinite	2	multiple	yes	yes
MLE3	Berkel et al. 2014c	bounded	3	multiple	yes	yes
LPML	Vandersteen et al. 2015	semi-infinite	n	multiple	yes	no
LPMLE3	Schneidewind et al. 2016	bounded	3	multiple	yes	yes
Sohn and Harris	Sohn and Harris 2021	semi-infinite	2	multiple	yes	yes
LPMLE*n*	Kampen et al. 2022d	**semi-infinite or bounded**	**n**	**multiple**	**yes**	**yes**

3.B Synthetic datasets I and II (constant parameters)

The two synthetic datasets are generated using the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ and the bounded domain model $\mathcal{M}_{\mathrm{BD}}$, respectively. The used input parameters are shown in Table S2 and are inspired by the experimental dataset presented in the manuscript. Moreover, we use two (processed) signals from the experimental dataset presented in the manuscript as input for our synthetic dataset, namely the signal at $z = 0.15$ m for the upper boundary condition and at $z = 0.55$ m for the lower boundary condition. We consider the complete 90-day signal, taking all frequencies up to 2.5 d^{-1}. The synthetic temperature signals are directly determined in the frequency domain, using equations (3.6) and the corresponding transfers functions (3.7) and (3.8) for the semi-infinite domain model, and (3.9) and (3.10) for the bounded domain model in the manuscript. The synthetic datasets are than created by taking the input temperature signal from the upper boundary

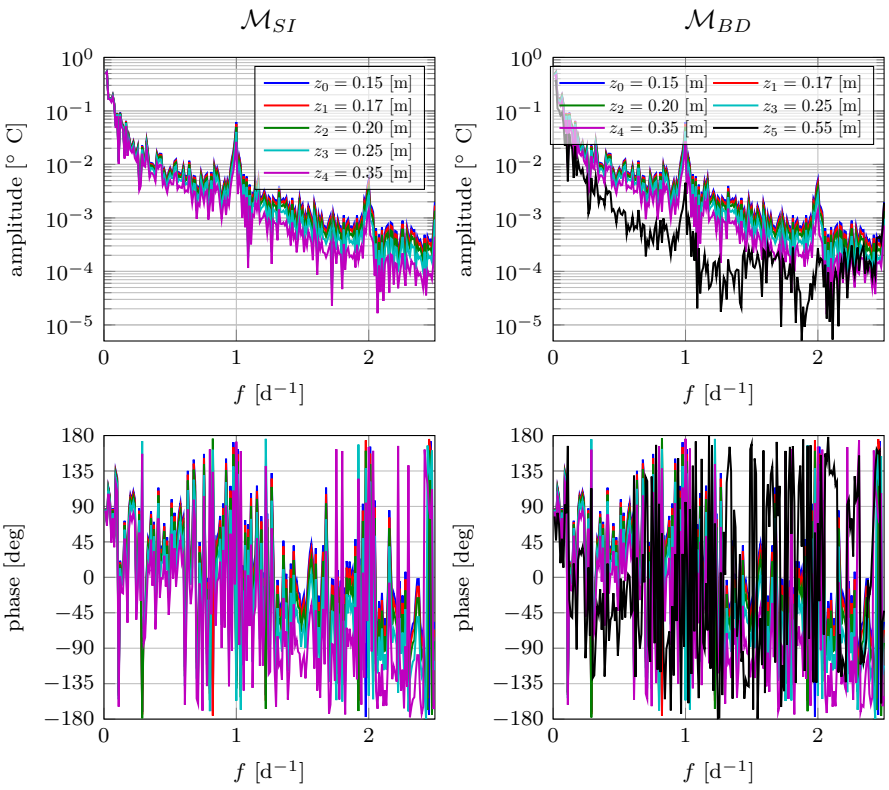


Figure 3.6. The amplitude (top) and phase (bottom) of synthetic datasets I & II generated by the semi-infinite domain model $\mathcal{M}_{\text{SI}}$ (left) and bounded domain model $\mathcal{M}_{\text{BD}}$ (right), respectively

condition at z_U, and the temperatures at z_1, z_2, z_3 and z_4. Note that we leave the temperature signal at z_L out of both synthetic datasets. The two synthetic datasets are visualized in Figure 3.6, where for completeness the lower boundary condition is shown in black. The estimated parameters are depicted in Figure 3.1 and the corresponding estimated transfer function are fitted through the synthetic datasets in Figure 3.7. For the semi-infinite domain model this fit is given by $\frac{\Theta(z_m,f)}{\Theta(z_U,f)} = G_U(z_m, f, \theta)$ while for the bounded domain model, the fit is shown by $\frac{\Theta(z_m,f)}{\Theta(z_U,f)} = G_U(z_m, f, \theta) + G_L(z_m, f, \theta)\frac{\Theta(z_L,f)}{\Theta(z_U,f)}$. Note that for the synthetic dataset generated with $\mathcal{M}_{\text{SI}}$ both models fit the data perfectly, while in case of the synthetic dataset generated with $\mathcal{M}_{\text{BD}}$, only the $\mathcal{M}_{\text{BD}}$ model fits the data perfectly. The parameter fit with the $\mathcal{M}_{\text{SI}}$ merely captures the average of the amplitude behavior, and completely fails to describe the correct trend in the phase behavior.

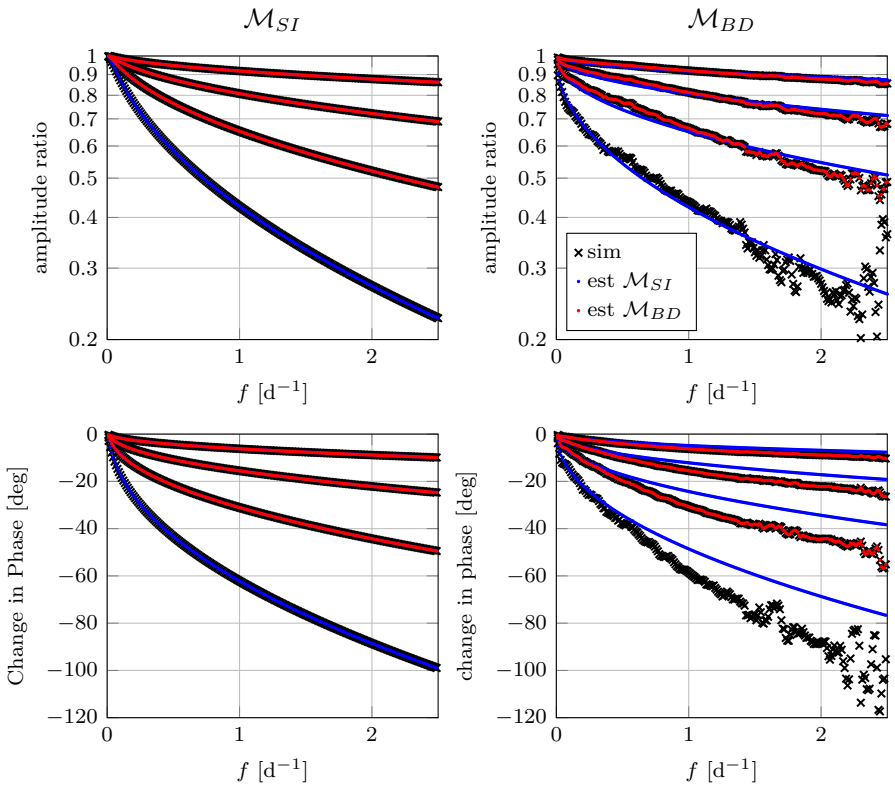


Figure 3.7. The amplitude ratio (top) and change in phase (bottom) of the measured frequency response functions and fitted semi-infinite domain model $\mathcal{M}_{\text{SI}}$ and bounded domain model $\mathcal{M}_{\text{BD}}$ for synthetic datasets I & II generated by the semi-infinite domain model $\mathcal{M}_{\text{SI}}$ (left) and bounded domain model $\mathcal{M}_{\text{BD}}$ (right), respectively. The shown amplitude ratio and change in phase are from sensor 1 at 0.15 m to sensors 2, 3, 4 and 5 at 0.17, 0.20, 0.25 and 0.35 m, respectively. Note that sensor 5 is used as boundary input for the bounded domain model; thus, the fitted response would equal the measured response and is left out for clarity.

3.C Synthetic dataset III (change in parameters)

As there is no analytical solution known for a model with a change in parameters, we approximate the partial differential equation by applying a semi-discretization using a central finite difference scheme to discretize the spatial coordinate

$$\begin{aligned}\frac{\partial}{\partial t}T(z_j,t) = D(z_j)\frac{(T(z_{j+1},t) - 2T(z_j,t) + T(z_{j-1},t)))}{(\Delta z)^2} \\ + V(z_j)\frac{(T(z_{j+1},t) - T(z_{j-1},t)))}{2\Delta z},\end{aligned} \tag{3.27}$$

where j denotes the discretization index and Δz the vertical spacing. This equation can directly be transformed to the frequency domain

$$\begin{aligned}i\omega\Theta(z_j,\omega) = D(z_j)\frac{(\Theta(z_{j+1},\omega) - 2\Theta(z_j,\omega) + \Theta(z_{j-1},\omega)))}{(\Delta z)^2} \\ + V(z_j)\frac{(\Theta(z_{j+1},\omega) - \Theta(z_{j-1},\omega)))}{2\Delta z}.\end{aligned} \tag{3.28}$$

In this way, we can determine the spatial temperature at the measurement points for each frequency individually using a two-boundary condition. Our MATLAB® implementation is based on the work of (Das et al. 2019).

3.D Synthetic dataset IV (change in parameters)

Complementary to synthetic dataset III, where the diffusivity is changed from $1.5 \cdot 10^{-6}$ m^2 s^{-1} to $0.8 \cdot 10^{-6}$ m^2 s^{-1} at $z = 0.275$ m, we created an additional dataset where the diffusivity is changed from $0.8 \cdot 10^{6}$ m^2 s^{-1} to $1.5 \cdot 10^{-6}$ m^2 s^{-1}. Equal to Figure 3.1, we estimate a single set of transport parameters at each depth using three neighboring sensors. The estimation results are plotted in Figure 3.8 at the middle sensor location, e.g., at 0.15 m, the estimated transport parameters are shown for the estimate using $z = [0.10, 0.15, 0.20]$ m.

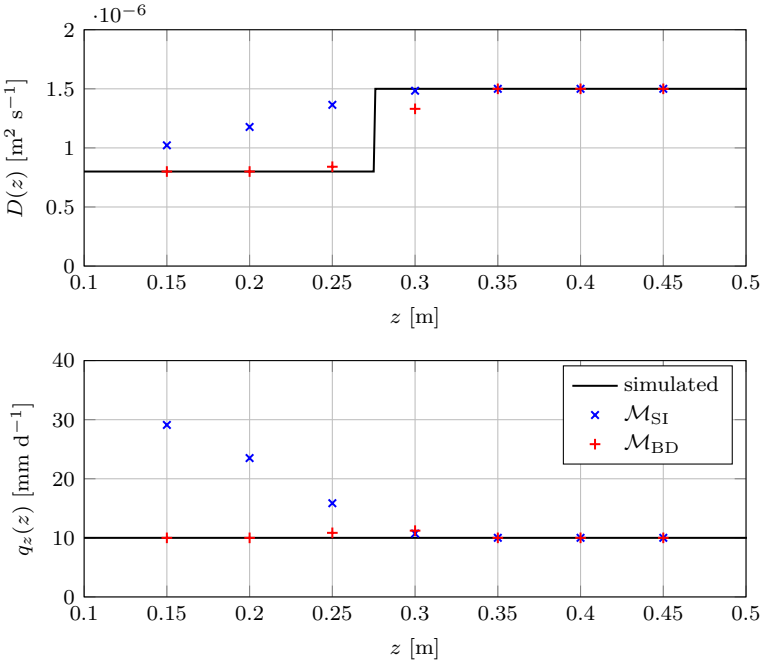


Figure 3.8. Simulated and estimated transport parameters using the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ and bounded domain model $\mathcal{M}_{\mathrm{BD}}$ for a scenario with a change in diffusivity, where the sensors are placed from 0.1 to 0.5 m, every 0.05 m and the parameters are estimated using a moving triplet of consecutive sensors for which the estimate is shown at the center of the triplet.

3.E Model fits on experimental dataset

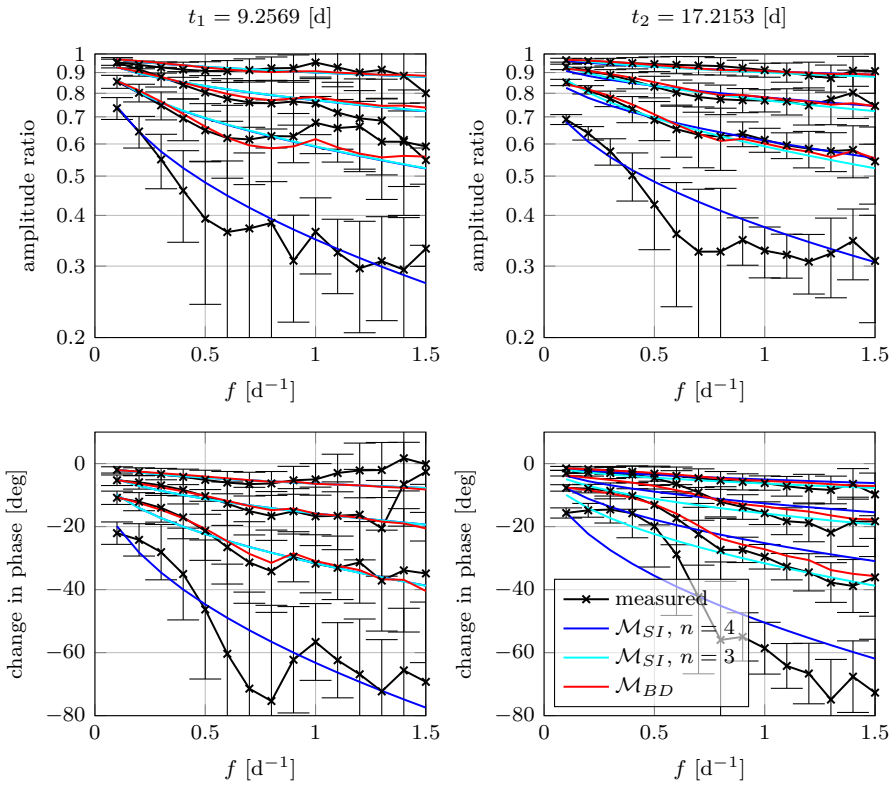


Figure 3.9. The amplitude ratio (top) and phase difference (bottom) of the measured frequency response functions with their 95% confidence interval and the fitted semi-infinite domain model $\mathcal{M}_{\text{SI}}$ and bounded domain model $\mathcal{M}_{\text{BD}}$ at the two times shown with dashed line in Figure 3.3. The shown amplitude ratio and change in phase are from sensor 2 at 0.15 m to sensors 3, 4, 5 and 6 at 0.17, 0.20, 0.25 and 0.35 m, respectively. Note that sensor 6 is used as boundary input for the bounded domain model; thus, the fitted response would equal the measured response and is left out for clarity.

3.F Thermal dispersivity fits on experimental dataset

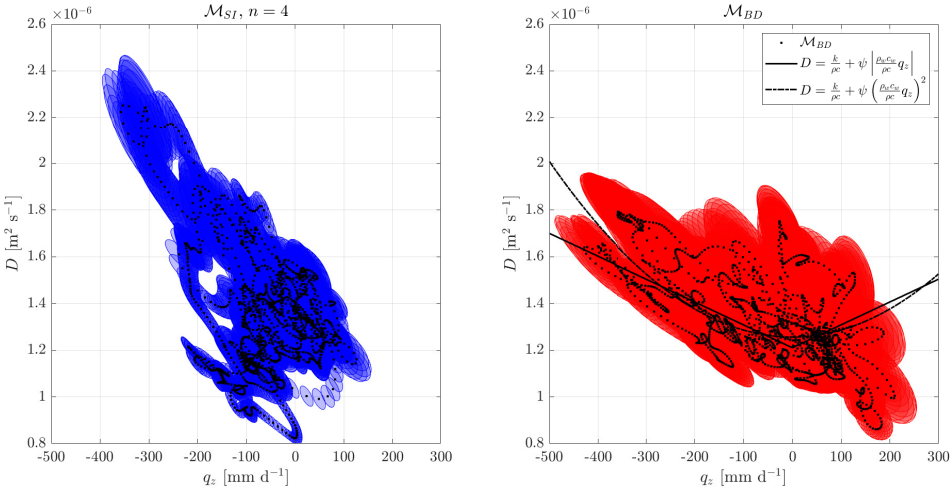


Figure 3.10. Diffusivity as a function of flux for the semi-infinite domain model $\mathcal{M}_{\mathrm{SI}}$ (left) and bounded domain model $\mathcal{M}_{\mathrm{BD}}$ (right) with their 95% confidence ellipses. For the bounded domain model, the two common notations for dispersion are fitted.

4

Estimating 1D spatially varying coefficients

Abstract - This chapter presents a closed-form solution to estimate space-dependent transport parameters of a linear one dimensional diffusion-transport-reaction equation. The infinite dimensional problem is approximated by a finite dimensional model by 1) taking a frequency domain approach, 2) linear parameterization of the unknown parameters, and 3) using a semi-discretization. Assuming full state knowledge, the commonly used output error criterion is rewritten as the equation error criterion such that the problem results in linear least squares. The optimum is then given by a closed-form solution, avoiding computational expensive optimization methods. Functioning of the proposed method is illustrated by means of simulation.

This chapter is published as: Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021c). "A Closed-Form Solution to Estimate Spatially Varying Parameters in Heat and Mass Transport". In: *IEEE Control Systems Letters* 5.5 (Nov. 2021), pp. 1681–1686. DOI: 10.1109/lcsys.2020.3042933.

4.1 Introduction

Heat and mass transport phenomena are widely studied in the domain of physics and chemistry. Examples include, but are not limited to, transport of thermal energy through nuclear fusion reactors (Berkel et al. 2019), the study of the groundwater-surface interaction systems (Schneidewind et al. 2016), and heat or moisture transport in buildings (Irsyad et al. 2017). These phenomena are typically modeled around an operating point by linear parabolic partial differential equations (PDEs), commonly known as diffusion-transport-reaction equations (Incropera et al. 2006). In most physical diffusion-transport-reaction systems the -exact- parameters are often unknown. Hence, data-driven estimation of the unknown physical parameters is necessary to determine a model which can be used for simulation, analysis, prediction, and control.

Historically, the utilization of measured data to determine an estimate of -physical- parameters is known as an inverse problem (Tarantola 2005; Kirsch 2011); in contrast to forward problems where the model is used to generate data. Specifically, determining unknown physical parameters in dynamical systems is often considered as an optimization problem where an error criterion based on the mismatch between model and measurements is minimized. To this end, there are typically two sub-fields in PDE estimation: 1) grey-box identification where unknown parameters are estimated by minimizing the difference between a pre-selected model (class), e.g., a PDE with unknown diffusion, and measured data (Banks and Kunisch 1989; Vogel 2002); 2) parameter learning which resorts to machine learning like techniques to avoid pre-selection of model-class (Xun et al. 2013; Tartakovsky et al. 2018). In both cases the underlying infinite dimensional models are generally approximated by finite dimensional models.

The standard method to estimate the unknown parameters is the output error criterion, i.e., taking the -weighted- sum of the squared error between the measurements and model output, primarily solved with iterative optimization methods (Banks and Kunisch 1989; Bock et al. 2013; Vogel 2002; Kirsch 2011; Mechhoud et al. 2015). These suffer from two problems: 1) they are solved iteratively, which is time consuming and in many cases the problem is non-convex, i.e., no guarantee for convergence to an optimal solution; 2) as only the output error is optimized and the unknown parameters are unconstrained, the parameters can start oscillating even when the output error is zero due to spatial aliasing of the state, which is estimated simultaneously. The latter is generally resolved by regularizing the unknown parameters (Ito and Kunisch 2008). However, this regularisation is often artificial, as there is usually no *a priori* information on how the unknown parameters change as function of space. Hence, here we propose a different approach by separating the two problems. First, we estimate the solution of the state over space based on a finite number of measurements. This allows us to transform the output error in an equation error criterion. By transforming this criterion into the frequency domain, we can derive a closed-form expression for the unknown parameters based on the estimated states, which is the novelty

of the presented method. Hence, we can uniquely and directly -without iteration- determine the diffusivity, convectivity, reactivity, and the source -as function of space-, simultaneously. Combining this with advances in 1) modern frequency domain signal processing to reduce noise and removing the initial condition (Pintelon and Schoukens 2012; Berkel et al. 2020); 2) recent innovations in dealing with experimentally unknown boundary conditions (Berkel et al. 2014c); 3) a wide variety of functions to spatially parameterize transport (Das et al. 2019). This results in a highly versatile and fast method to acquire reliable estimates of the spatially varying parameters. This is in strong contrast to direct solutions proposed in the literature which are based on (piecewise) constant parameters (Berkel et al. 2014b) and often only consider diffusion (Ewing and Lin 1988; Liao et al. 2009).

4.2 Problem formulation

The estimation of space-dependent physical parameters in heat and mass transport is performed based on the following specifications.

Model-class A class of linear parabolic PDEs is considered in a one dimensional bounded spatial geometry to model the spatio-temporal dynamics of heat and mass transport. For all $x \in \mathbb{X} := [x_b, x_e] \subset \mathbb{R}$ and $t \in \mathbb{T} := [t_0, \infty) \subseteq \mathbb{R}_{\geq 0}$, the class of PDEs is defined as

$$\frac{\partial z}{\partial t} = D(x)\frac{\partial^2 z}{\partial x^2} + V(x)\frac{\partial z}{\partial x} + K(x)z + P(x)s(t). \tag{4.1}$$

Here, the state $z : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ is a multi-variable function (that can describe temperature or mass concentration) and (4.1) is understood point-wise in $x \in \mathbb{X}$ and $t \in \mathbb{T}$ with z evaluated as $z(x,t)$. The physical transport parameters are diffusivity $D : \mathbb{X} \to \mathbb{R}_{>0}$, convectivity $V : \mathbb{X} \to \mathbb{R}$, and reactivity $K : \mathbb{X} \to \mathbb{R}$. The external input is denoted by $s : \mathbb{T} \to \mathbb{R}$. Furthermore, the spatial distribution of the input is given by the function $P : \mathbb{X} \to \mathbb{R}$.

For well-posedness, the PDE is constrained by two boundary conditions at locations x_b and x_e. Moreover, the initial condition $z(\cdot, t_0)$ is assumed compatible with the model and its boundaries.

Measured data Corresponding to the heat and mass transport phenomena, location specific values of the state function z are measured over time and available as data. Let there be $M > 0$ sensors that measure z at the locations given by the set of points $\mathbb{X}_M := \{\check{x}_1, \check{x}_2, \ldots, \check{x}_M\} \subset \mathbb{X}$. The measured output signals are $y^m(t) = z(\check{x}_m, t)$ for all $t \in \mathbb{T}$, with $m \in \{1, \ldots, M\}$. The input $s(t)$ is assumed to be known or measured for all $t \in \mathbb{T}$.

Problem formulation In practical applications, the explicit definitions of space-dependent transport parameters $\{D(x), V(x), K(x), P(x)\}$ are often not available.

Moreover, the boundary conditions that constrain heat and mass transport phenomena may depend on the spatially varying transport parameters, and, hence, remain unknown. In a similar fashion, in practice, the -exact- initial condition $z(\cdot, t_0)$ may also be unknown. Therefore, we take the extremum measurements as boundary inputs, however, other linear boundary conditions are also allowed.

To complete the model in (4.1), the spatially varying transport parameters $\{D(x), V(x), K(x), P(x)\}$ have to be estimated based on measured data. This results in the following estimation problem:

Problem 4.1. *Given the pre-processed data-set (see remark 4.2)*

$$\mathbb{D} := \{y^1(t), \dots, y^M(t), s(t) \mid t \in [t_0, \infty)\},$$

estimate the unknown function $\gamma : \mathbb{X}_E \to \mathbb{R}^4$,

$$\gamma(x) := \operatorname{col}\big(D(x), V(x), K(x), P(x)\big), \tag{4.2}$$

by minimizing a cost function $\mathcal{V}\big(\mathbb{D}, z(x,t;\gamma)\big)$ *-defined in Sec IV- over* γ *such that the solution* $z(x,t;\gamma)$ *satisfies the model* (4.1) *with parameters* γ *in the sense that*

$$\frac{\partial z}{\partial t} = \operatorname{row}\left(\frac{\partial^2 z}{\partial x^2}, \frac{\partial z}{\partial x}, z, s\right)\gamma, \tag{4.3}$$

subject to the boundary conditions and initial condition

$$z(\check{x}_1, t) = y^1(t), \quad z(\check{x}_M, t) = y^M(t), \quad z(x, t_0) = 0.$$

Remark 4.1. *Using the extremum measurements as Dirichlet boundary conditions reduces the estimation domain to* $\mathbb{X}_E := [\check{x}_1, \check{x}_M]$ *(Berkel et al. 2014c). Moreover, this allows to set-up the parameter estimation problem without the need of full knowledge about the actual boundary conditions.*

Remark 4.2. *With -advanced- signal processing techniques, the measured signal can be split into a transients/drift signal (non-steady-state behavior, e.g., from the initial condition), a forced response (from the excitation) and additive (filtered) noise (Pintelon and Schoukens 2012; Berkel et al. 2020). By removing the transient and noise terms from the original signal, only the forced response remains in the filtered dataset* $\mathbb{D}$*, which is equivalent to* $z(\cdot, t_0) = 0$.

4.3 Finite dimensional frequency domain problem

Problem 4.1 is infinite dimensional with no known analytic solution for $z(x,t)$. Therefore, this section approximates the infinite dimensional problem by creating a finite dimensional model following the methodology from (Das et al. 2019).

4.3.1 Frequency domain approach

Assuming that the discrete Fourier transformed input $S(k)$ has (excited) frequency bins $k \in \mathbb{K}$, problem 4.1 can be studied in the frequency domain without loss of information (Parseval's Theorem). Moreover, due to linearity of the model, $Z(x,k)$ is independent for each excited bin k. The frequency domain model of (4.3) is then given by

$$\begin{aligned} i\omega_k Z =& \text{row}\left(\frac{\partial^2 Z}{\partial x^2}, \frac{\partial Z}{\partial x}, Z, S\right)\gamma \\ Y(k) :=& \text{col}\left(Z(\check{x}_2,k),\ldots,Z(\check{x}_{M-1},k)\right) \end{aligned} \tag{4.4}$$

subject to the boundary conditions

$$Z(\check{x}_1,k) = Y^1(k), \quad Z(\check{x}_M,k) = Y^M(k),$$

with discrete Fourier transformed state $Z : \mathbb{X}\times\mathbb{K}\to\mathbb{C}$, input $S:\mathbb{K}\to\mathbb{C}$, output $Y(k)$, $i^2=-1$ and angular frequency ω_k corresponding to the k^{th}-bin.

Remark 4.3. *In excitation experiments only a finite number of bins are informative, i.e., those bins which are present in the input $S(k)$ and are above the noise level (Pintelon and Schoukens 2012). Hence, in practice only a -few- finite number of bins need to be considered (see (Berkel et al. 2018b) for details).*

4.3.2 Linear parameterization of the unknown functions

For estimation purposes, assume that γ belongs to a function space Γ that is parameterized by a surjective mapping $\Pi:\Theta\to\Gamma$ that is described by a finite sum of basis functions $B_r(x) := \text{diag}\left(B_r^D(x), B_r^V(x), B_r^K(x), B_r^P(x)\right)$,

$$\gamma(x;\theta) := [\Pi(\theta)](x) := \sum_{r=1}^{R} B_r(x)\theta_r, \tag{4.5}$$

with $\theta = \text{col}\left(\theta_1,\ldots,\theta_R\right)$ and $\theta_r = \text{col}\left(\theta_r^D,\theta_r^V,\theta_r^K,\theta_r^P\right) \in \Theta \subset \mathbb{R}^4$. With this parameterization, the estimation of γ amounts to estimating $\theta\in\mathbb{R}^{4R}$.

4.3.3 Semi-discretization

The infinite dimensional model (4.4) is approximated by a central finite difference scheme that converges to the exact solution for $N\to\infty$ (Quarteroni and Valli 1994). The finite dimensional model is

$$\begin{aligned} i\omega_k \mathbf{Z}(k) &= A(\theta)\mathbf{Z}(k) + B(\theta)U(k) \\ Y(k) &= C\mathbf{Z}(k) \end{aligned} \tag{4.6}$$

with state vector $\mathbf{Z}(k) := \text{col}\left(Z(x_2,k),\ldots,Z(x_{N-1},k)\right)$ at sample $x_j\in\mathbb{X}_d\subset\mathbb{X}$, $j\in\{1,\ldots,N\}$, extended input vector $U(k) = \text{col}\left(S(k),Y^1(k),Y^M(k)\right)$. Here,

$A(\theta)$ and $B(\theta)$ contain the boundary conditions, are linear affine in θ, and defined in (Das et al. 2019). The -observation- matrix C maps states to output. Altogether, the finite dimensional problem yields

Problem 4.2. *Given the pre-processed data-set*

$$\mathbb{D} = \{\mathrm{col}\left(Y^1(k), \ldots, Y^M(k)\right), S(k) \mid k \in \mathbb{K}\}.$$

Estimate θ by minimizing a cost function $\mathcal{V}\big(\mathbb{D}, \mathbf{Z}(k;\theta)\big)$ over θ such that $\mathbf{Z}(k;\theta)$ satisfies the model (4.6).

4.4 The inverse problem

This section describes our methodology to estimate the unknown weights. In the literature for distributed parameter systems, different error criteria are proposed to calculate the parameters. The commonly used criterion is the output error criterion (Banks and Kunisch 1989; Bock et al. 2013; Vogel 2002; Kirsch 2011; Das et al. 2019) and the rarely used equation error criterion (Banks and Kunisch 1989) which we will use in the frequency domain to derive a closed-form solution for the unknown parameters.

4.4.1 Output error criterion

The commonly used cost function is the output error criterion which is the sum of the squared error between solution $z(x,t;\gamma)$ for a given γ and measurements $y(t)$, i.e.,

$$\mathcal{V}_{\mathrm{oe}}(\theta) := \int_{t_0}^{\tau} \int_{x_b}^{x_e} \left|y(t) - \mathfrak{C}z(x,t;\gamma(x;\theta))\right|^2 \mathrm{d}x\mathrm{d}t \tag{4.7}$$

with observation map $\mathfrak{C}$ that maps state to output. This criterion is transformed into the frequency domain such that the output error criterion for problem 4.2 and (Das et al. 2019) is given by

$$\mathcal{V}_{\mathrm{oe}}(\theta) = \sum_{k\in\mathbb{K}} \left\| Y_k - C\left(i\omega_k I - A(\theta)\right)^{-1} B(\theta) U_k \right\|^2, \tag{4.8}$$

with shorthand notation $Y_k = Y(k)$ and $U_k = U(k)$. This criterion is nonlinear in θ due to the inverse and multiplication. As such, it needs to be iteratively optimized with often no guarantee for convergence to the global minimum. As this is generally the case for the output error criterion, we propose to use the equation error criterion for which a closed-form solution can be derived.

4.4.2 Equation error criterion

In the equation error criterion, the state of the model is replaced by measurements or estimates of the state $\hat{z}(x,t)$ (Banks and Kunisch 1989), such that for (4.3), the

equation error criterion is defined as

$$\mathcal{V}_{\text{ee}}(\theta) := \int_{t_0}^{\tau}\int_{\check{x}_1}^{\check{x}_M}\left(\frac{\partial\hat{z}}{\partial t} - \text{row}\left(\frac{\partial^2\hat{z}}{\partial x^2}, \frac{\partial\hat{z}}{\partial x}, \hat{z}, s\right)\gamma(x;\theta)\right)^2 \mathrm{d}x\mathrm{d}t. \quad (4.9)$$

This can be simplified in the frequency domain: the integral over time simplifies to a summation over the excited frequency bins $k \in \mathbb{K}$ without loss of information (Parseval's theorem). With the measured or estimated state vector in the frequency domain $\hat{\mathbf{Z}}_k$, the equation error for problem 4.2 simplifies to

$$\mathcal{V}_{ee}(\theta) = \sum_{k\in\mathbb{K}}\left\|i\omega_k\hat{\mathbf{Z}}_k - \left(A(\theta)\hat{\mathbf{Z}}_k + B(\theta)U_k\right)\right\|^2. \quad (4.10)$$

4.4.3 Derivation of the closed-form solution

Consider the equation error criterion in (4.10) and that $A(\theta)$ and $B(\theta)$ are linear affine in θ, then new matrices $\tilde{A}$, $\tilde{B}$ can be defined as a function of the data U_k and $\hat{\mathbf{Z}}_k$ such that $A(\theta)\hat{\mathbf{Z}}_k = \hat{A}(\hat{\mathbf{Z}}_k)\theta$ and $B(\theta)U_k = \hat{B}(U_k)\theta$ (see). As a result, (4.10) can be written as

$$\mathcal{V}_{\text{ee}} = \sum_{k\in\mathbb{K}}\left\|i\omega_k\hat{\mathbf{Z}}_k - \left(\hat{A}(\hat{\mathbf{Z}}_k) + \hat{B}(U_k)\right)\theta\right\|^2, \quad (4.11)$$

with the closed-form solution

$$\hat{\theta} := \left(\left(\bar{A}+\bar{B}\right)^{\mathsf{H}}\left(\bar{A}+\bar{B}\right)\right)^{-1}\left(\bar{A}+\bar{B}\right)^{\mathsf{H}}\bar{W}, \quad (4.12)$$

where $^{\mathsf{H}}$ denotes the Hermitian transpose, $\bar{A}$, $\bar{B}$, and $\bar{W}$ are the column concatenation of $\hat{A}(\hat{\mathbf{Z}}_k)$, $\hat{B}(U_k)$, and $i\omega_k\hat{\mathbf{Z}}_k$ for all excited bins $k \in \mathbb{K}$, respectively. Hence, the optimal weighting for θ, and thus $\gamma(x)$, are determined without using iterative optimization methods.

Remark 4.4. *The equation error in combination with different cost functions such as weighted and total linear least squares also have closed-form solutions (Markovsky and Huffel 2007).*

4.4.4 State estimation

The unique solution ((4.12)) requires knowledge of the state at each discretization point, i.e., a "space-"continuous measurement. If $N \to \infty$, the finite dimensional description converges to the true infinite dimensional solution and an exact solution can be found.

However, due to limited spatial measurements, the purpose is to estimate the full state based on these measurements. This raises the fundamental problem of unknown in-between sensor behavior. In case of regularized output error (Banks

and Kunisch 1989), the intermediate relationship in-between measurements is determined by the regularized functions for the transport parameters. This technique is reminiscent to the (spatial) NyquistShannon (NS) sampling theorem. Therefore, a consistent -spatial- signal reconstruction using the NS theorem is required and currently being worked out. Alternatively, the states can be inferred by interpolating the measurements using machine learning techniques -currently being further worked out- (Kampen et al. 2020; Ho et al. 2019); classic interpolation methods; or reducing the number of states in the model to match the measurements. Note that the selected method should be model free such that the problem remains affine in the unknown parameters. In the simulation section, we show the latter two approximation methods in combination with the closed-form solution.

4.5 Simulation results

In this section, two simulation scenarios are presented demonstrating the merit of the proposed methodology. As the here proposed method has a closed-form solution that only requires solving a linear matrix equality, it is expected to be significantly faster than iterative optimization methods, e.g., when compared to (Das et al. 2019). This can be exploited by considering two scenarios: *scenario 1* which estimates $\{D(x), P(x)\}$ testing different orders (weights) R of the basis functions and *scenario 2* which tests different interpolation density of the data when all four functions $\{D(x), V(x), K(x), P(x)\}$ are estimated using fixed set of basis functions $B_r(x)$ and their orders R.

4.5.1 Data generation

The simulation example is inspired by perturbative experiments in the field of nuclear fusion (Berkel et al. 2013). The heat transport, (4.1), is generally analyzed on the normalized domain $\mathbb{X} = [x_b, x_e] = [0, 1]$ of the -minor- plasma radius. Here x_b is at the center and x_e is at the edge of the plasma. The corresponding boundary conditions in the simulation are $\frac{\partial Z}{\partial x}(x_b, \cdot) = 0$ due to (axi)symmetry and Dirichlet boundary condition $Z(x_e, \cdot) = 0$ due to a significant temperature difference between core plasma ~170 million °C and edge plasma ~1 million °C. Typical functions used in nuclear fusion are

$$
\begin{aligned}
D^{\text{sim}}(x) &= 5x^3 - 0.005x + 5, \\
V^{\text{sim}}(x) &= 15x^2 - 0.005, \\
K^{\text{sim}}(x) &= -3x, \\
P^{\text{sim}}(x) &= 0.2 + \frac{7}{\sqrt{\pi}} e^{\frac{-(x-0.35)^2}{(0.1)^2}} + \frac{5.6}{\sqrt{\pi}} e^{\frac{-(x-0.6)^2}{(0.1)^2}}.
\end{aligned}
$$

For the perturbation of the plasma temperature a microwave source is used, where the excitation signal $S(\omega)$ is a block-wave of $\omega_0 = 50\pi$ with a 70% duty cycle. Here, only the first five harmonics $\omega_k = k\omega_0$, $k = 1, \ldots, 5$ have a significant contribution

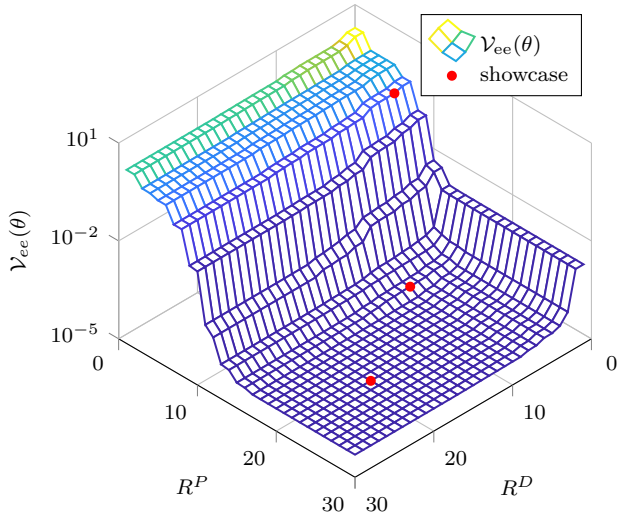


Figure 4.1. The fitting error map $\mathcal{V}_{\mathrm{ee}}(\theta)$ for scenario 1 by approximating the parameters $\{D(x), P(x)\}$ with different number of basis-functions, R^D, R^P, respectively. Here, the • annotates the locations $(R^D, R^P) = (3, 8)$, $(9, 16)$, and $(21, 23)$ shown in Figure 4.2.

and are used for the estimation. The temperature data is generated by a simulation with a central finite difference grid of $N = 801$ sample points.

4.5.2 Estimation of $\{D(x), P(x)\}$ with unknown order R

In *scenario 1*, the goal is to estimate $\{D(x), P(x)\}$ without knowledge on the correct orders (R^D, R^P) in (4.5), which practically means we do not know the shape of $\{D(x), P(x)\}$. For ease of explanation, we set $V(x) = K(x) = 0$ for this example. As the number of sensors plays an important role, we choose here $M = 22$ sensors that are located at $\check{x}_m = 0.05 + 0.04125(m - 1)$, with $m = 1, \ldots, M$. This corresponds to the electron cyclotron emission (ECE) diagnostic that measures in a medium sized fusion reactor.

In principle, any basis function can be used and several can be tested simultaneously. Here, we use -arbitrarily chosen- Chebyshev polynomials for both parameters (Quarteroni and Valli 1994).

We vary the orders R^D, $R^P \in \{1, \ldots, 30\}$ and use a discretization grid that equals the measurement grid $\mathbb{X}_d = \mathbb{X}_M$, thus $C = I$. Naturally, the cost (4.10) decreases for increasing R^D and R^P as shown in Figure 4.1. For three different combinations, i.e., $(R^D, R^P) = (3, 8)$, $(9, 16)$, and $(21, 23)$ the resulting estimates are shown in Figure 4.2.

Figure 4.2 shows that when the order (R^D, R^P) is too low, e.g., for $(3, 8)$, significant errors occur in the estimates. If the order is sufficiently high, e.g., for the pairs $(9, 16)$ and $(21, 23)$, the estimates have the correct value at the sensor

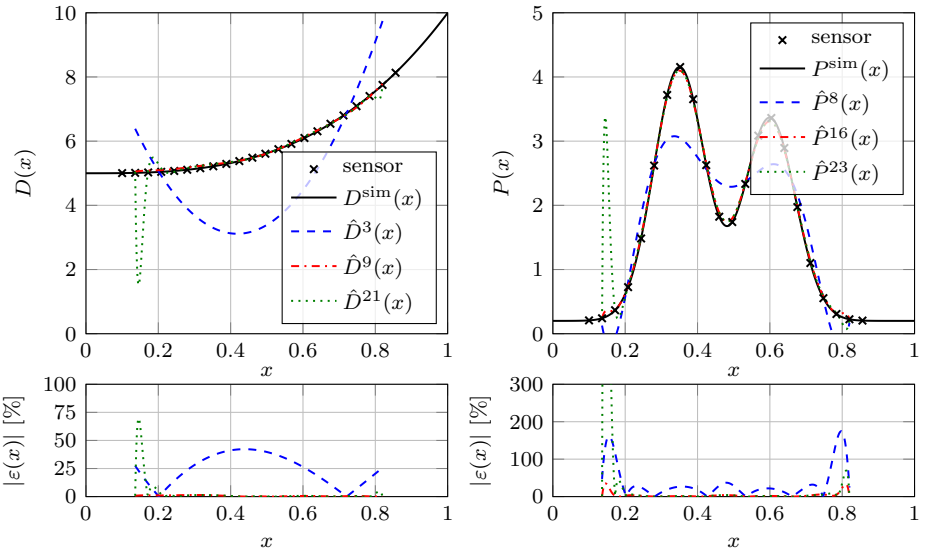


Figure 4.2. The estimated parameters $\{D(x), P(x)\}$ over $\mathbb{X}_E$ for scenario 1 with the number of basis-function $(R^D, R^P) = (3, 8)$, $(9, 16)$ and $(21, 23)$ and their relative error $\varepsilon(x)$.

locations as is imposed by the equation error. Consequently, only the values of $\gamma(x)$ at the sensor locations should be considered as a correct result given that the combination of basis functions and order give sufficient freedom.

Remark 4.5. *Although the intermediate points between sensor locations should not be considered in this methodology, $\gamma(x)$ is defined at these points and plotted for completeness. As the comparison of the orders $(9, 16)$ and $(21, 23)$ shows, the finite dimensional approximation does not pose a unique solution for $\gamma(x)$ at the intermediate points. This can be resolved by increasing the discretization grid, e.g., by interpolating the data, which is further investigated in the following section.*

4.5.3 Estimation using spatial interpolations of the temperature

In the simulation for *scenario 2*, we investigate the effect of interpolations on the estimation. In *scenario 2* all the functions are non-zero and according to Section 4.5.1. Here, we use a sensor grid which corresponds to that of a larger fusion reactor, i.e., the temperature is measured by $M = 60$ sensors located at $\check{x}_m = 0.1625m$, with $m = 1, \ldots, M$.

In this scenario, the order of the basis functions is fixed and only the weights are estimated. Polynomial basis functions are used for $\{D(x), V(x), K(x)\}$ and B-spline functions for $P(x)$ to approximate Gaussians. The choice of order for

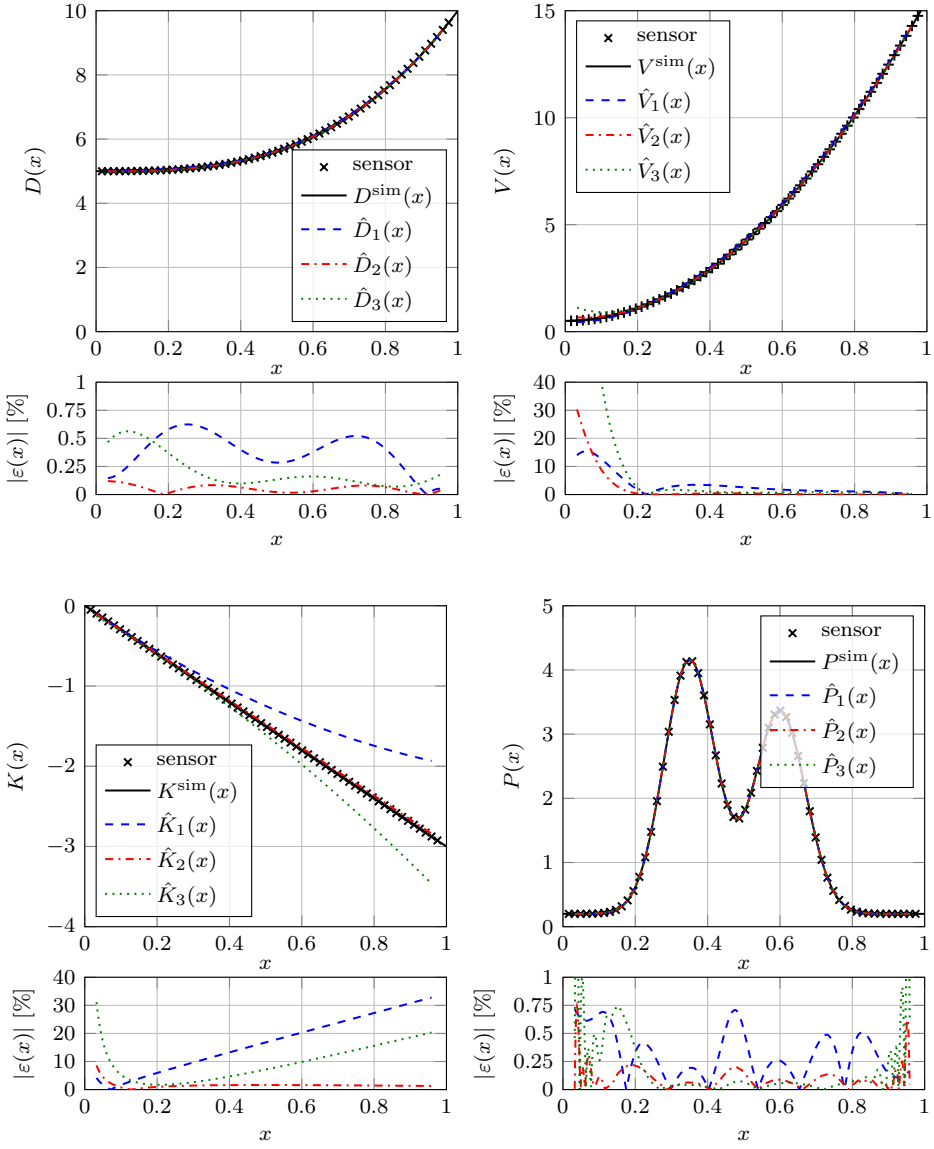


Figure 4.3. Estimated parameters $\{D(x), V(x), K(x), P(x)\}$ over $\mathbb{X}_E$ for scenario 2 and the relative error $\varepsilon(x)$ between the simulated parameter and the estimated parameter. The subscript 1, 2, 3 denote if the estimates are based on the measurements $N_1 = M$, or the interpolated measurements $N_2 = 2(M-1)+1$ or $N_3 = 10(M-1)+1$, respectively.

the monomial basis functions to estimate $\{D(x), U(x), K(x)\}$ is higher than the actual order and is chosen to be $R^D = 8$, $R^V = 6$ and $R^K = 3$ such that it is possible to find an exact description. The basis function for $P(x)$, $B_r^B(x)$ is designed using the De Boor's algorithm (Boor 1978) with 58 equally distributed control points, thus $R^P = 58$.

Three estimations are presented where, 1) the discretization grid equals the measurement grid $N_1 = M$, 2) the measurement data is interpolated over space to generate artificial spatial measurements, i.e., $N_2 = 2(M-1)+1$, and 3) more interpolation points, i.e., $N_3 = 10(M-1)+1$. The data is interpolated using cubic splines.

The estimation results are shown in Figure 4.3. Overall, the estimated parameters $\{D(x), P(x)\}$ closely match with the simulated parameters, while $\{V(x), K(x)\}$ are estimated with a significantly lower accuracy. Estimating $\{V(x), K(x)\}$ requires a higher accuracy of the discretization grid. Therefore, interpolating the data increases the accuracy (see results for N_2), however, over-interpolation can affect the estimation accuracy negatively (see results for N_3).

4.6 Conclusions and discussion

This chapter presents a novel method to efficiently estimate the unknown space-dependent transport parameters, based on a closed-form solution of the equation error criterion. The closed-form solution is formulated as a linear matrix equality such that high density grids can be solved computationally efficient. As a result, if the states $z(\cdot, t)$ are known, the parameters are estimated uniquely with the desired accuracy by taking a sufficiently dense discretization grid.

In practice, measurements are often only provided at a limited set of spatial locations. This is generally resolved by (i) applying regularization or restrictions on the unknown parameters or (ii) as we have done here by "interpolating" the measurements to increase the grid density. There are three reasons why we prefer approach (ii): 1) As we are estimating the unknown parameters, prior information on the -smoothness of- parameters is generally unavailable, whereas smoothness on the states is required due to the underlying parabolic PDE; 2) the states as function of space can easily be -visually- inspected for correctness and validated by taking additional spatial measurements; and 3) a closed-form solution can be used which significantly speeds up the process and avoids convergence to a local minimum.

Appendix

4.A Finite Difference Matrices

The matrices used in (4.11) are given by

$$\hat{A}\big(\hat{\mathbf{Z}}(k)\big) := \big[L_1^D, L_1^V, L_1^K, \mathbf{0}, \ldots, L_R^D, L_R^V, L_R^K, \mathbf{0}\big]\big(I_{4R} \otimes \hat{\mathbf{Z}}(k)\big), \tag{4.13}$$

$$\begin{aligned}\hat{B}\big(U(k)\big) := \big[&\mathbf{0}, g_1^D, h_1^D, \mathbf{0}, g_1^V, h_1^V, \mathbf{0}, \mathbf{0}, \mathbf{0}, f_1^P, \mathbf{0}, \mathbf{0}, \ldots, \\ &\mathbf{0}, g_R^D, h_R^D, \mathbf{0}, g_R^V, h_R^V, \mathbf{0}, \mathbf{0}, \mathbf{0}, f_R^P, \mathbf{0}, \mathbf{0}\big] \\ &\big(I_{4R} \otimes U(k)\big),\end{aligned} \tag{4.14}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{0}$ the zero vector/matrix of the appropriate size. The central finite difference matrices L with grid sample $\Delta_x > 0$ are

$$L_r^D := \frac{1}{(\Delta_x)^2}\tilde{B}_r^D \begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 \end{bmatrix}, \tag{4.15}$$

$$L_r^V := \frac{1}{2\Delta_x}\tilde{B}_r^V \begin{bmatrix} 0 & 1 & & \\ -1 & 0 & 1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 0 \end{bmatrix} \tag{4.16}$$

$$L_r^K := \tilde{B}_r^K. \tag{4.17}$$

Here, $\tilde{B}_r^D$, $\tilde{B}_r^V$ and $\tilde{B}_r^K$ are diagonal matrices of dimension $(N-2)\times(N-2)$ with the diagonal entries $B_r^D(x_j)$, $B_r^V(x_j)$, $B_r^K(x_j)$ evaluated at each grid point

x_j, $j \in \{2, \ldots, N-1\}$. The vectors for the input and boundary conditions are

$$g_r^D := \text{col}\left(\frac{B_r^D(x_2)}{(\Delta_x)^2}, 0, \ldots, 0\right) \tag{4.18}$$

$$g_r^U := \text{col}\left(-\frac{B_r^V(x_2)}{2\Delta_x}, 0, \ldots, 0\right) \tag{4.19}$$

$$h_r^D := \text{col}\left(0, \ldots, 0, \frac{B_r^D(x_{N-1})}{(\Delta_x)^2}\right) \tag{4.20}$$

$$h_r^U := \text{col}\left(0, \ldots, 0, \frac{B_r^V(x_{N-1})}{2\Delta_x}\right) \tag{4.21}$$

$$f_r^P := \text{col}\left(B_r^P(x_2), \ldots, B_r^P(x_{N-1})\right). \tag{4.22}$$

5

Estimating nD spatially varying transport coefficients and source function

Abstract - In the calculation of transport coefficients from experimental data precise knowledge of the source is usually assumed, while the identification of the coefficients focuses on specific geometries and one spatial variable. This chapter presents a method for the simultaneous estimation of both the distributions of transport coefficients as well as the source profile. A convex solution of the inverse problem is retained which makes the calculations highly computational efficient. Moreover, this allows for the estimation of multi-dimensional transport coefficients, source terms, and in the future the analysis of the effect of regularization on experimental data and transport coefficient distributions.

This chapter is published as: Kampen, R. J. R. van, J. de Vries, S. Weiland, et al. (2023b). "Fast simultaneous estimation of nD transport coefficients and source function in perturbation experiments". In: *Scientific Reports* 13.1 (Feb. 2023), p. 3241. DOI: 10.1038/s41598-023-30337-0.

5.1 Introduction

In many physical systems, generic scalar transport determines the behavior. Examples are transport of particles, heat and momentum in nuclear fusion reactors (Ryter et al. 2010), the hyporheic flow in groundwater-surface interaction systems (Boano et al. 2014), thermal transport for medical treatments such as local hyperthermia therapy (Deuflhard et al. 2012), etc. The most important physical quantities determining the transport are the transport coefficients representing the diffusivity, convectivity, reactivity/damping, and the sinks/sources in the system. Typically, these quantities cannot be measured directly but have to be inferred from measurements of the transport variable, such as the temperature for thermal transport or the concentration for particle transport. The evolution of the transport variable is described by the transport equation and is referred to as the state. Estimating the corresponding coefficients of the transport equation based on measurement is known as an inverse problem or referred to as grey box modelling.

In many cases, the geometry and the physics allow for the reduction of the transport model to a single spatial dimension. For example, in the context of magnetically confined plasmas there is a strong anistropy between transport processes along and perpendicular to the magnetic field lines. The toroidal geometery then allows for a 1D-approach in the main plasma. In this context, a 1D approach has been proposed multiple times (Krieger et al. 1990; Moret et al. 1993; Takenaga et al. 1998; Escande and Sattin 2012) and is sometimes referred to as the matricial approach (MA). As pointed out by Escande and Sattin (Escande and Sattin 2012), the main advantage of the MA is that after integration the exact value of two of the transport coefficients (using one frequency), namely diffusion and convection, can be determined without having to choose basis functions.

The MA approach is theoretically valid. However, experimental practicalities, such as the smoothing, extrapolation, and interpolation of the measurements, e.g., by splines, necessary to obtain acceptable spatial functions of the source, the state (and its derivatives) makes determination of the exact coefficients ambiguous. Smoothing operations on the (state) profile are equivalent to choosing lower order basis functions for the transport parameters in the estimation procedure. Hence, both are different forms of regularization and both can significantly impact the estimates of diffusion and convection. Moreover, the regularization of profile or transport parameters is a choice for optimization rather than an inherent modeling feature.

For estimation procedures, a strict distinction needs to be made between systems in which the source is at the system boundary and systems in which the source and sinks are within the domain. An example of the latter is a nuclear fusion reactor where there are various state dependent heating methods with state dependent source distributions. Here heating can stem from the nuclear fusion reaction as well as from neutral beam heating, ion cyclotron heating, electron cyclotron heating, lower hybrid heating, and Ohmic heating. To infer the transport in steady state, the power deposition for all these heating methods needs to be known in

detail. To complicate matters, the heating method can be subject to, or even drive, plasma instabilities that lead to source broadening (Baar et al. 1997; Ryter et al. 2001; Kirov et al. 2002; Chellaï et al. 2018; Chellaï et al. 2021; Brookman et al. 2021; Slief et al. 2022). Therefore, the deposition width of sources within the domain may be uncertain or subject to dynamic changes. Hence, simultaneous estimation of the transport parameters and source profile is crucial to achieve reliable estimates and to ensure a broad application of these estimation methods. As the aforementioned methods do not deal with the simultaneous estimation of sinks/sources and transport coefficients, we developed a new approach allowing for this simultaneous estimation while we also provide transparency in the applied regularization. By using high order basis functions, no additional regularization is applied to the coefficients. Moreover, regularization on the level of the state by smoothing or interpolation is still an option.

Similarly to (Escande and Sattin 2012), we formulate the problem as a linear regression problem which contains the source and damping. We use multiple harmonics because one of the strong advantages of estimation is over-determination, i.e., increasing the number of harmonics (equations) compared to the number of unknown coefficients reduces the uncertainty. To deal with this over-determination, we formulate the problem as a linear least squares problem for which we have a (unique) closed-form solution, thereby avoiding iterative optimization methods.

5.2 Methodology

We consider the convection-diffusion-reaction equation often following from linearization

$$\partial_t z = -\nabla \cdot (\mathbf{V} z - D \nabla z) + K z + P \phi, \tag{5.1}$$

describing the spatio-temporal evolution of the transport quantity $z(\mathbf{x}, t)$ where $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n$ denotes the spatial geometry. The transport parameters representing the diffusivity and reactivity are given by the functions $D(\mathbf{x})$ and $K(\mathbf{x})$, respectively. The diffusivity is a strictly positive function. For transport in multiple dimensions, the convectivity is a flow field given by the vector function $\mathbf{V}(\mathbf{x})$, while for a single dimension it is given by the function $V(\mathbf{x})$. Furthermore, we consider an unknown source profile $P(\mathbf{x})$ that is fixed in space but modulated over time by the known signal $\phi(t)$. In our problem we measure $z(\mathbf{x}, t)$, control $\phi(t)$, while the exact deposition $P(\mathbf{x})$ is unknown, and needs to be estimated in conjunction with the diffusivity D, flow velocity $\mathbf{V}$, and reactivity K. Linear, but unknown boundary conditions constrain the edge of the spatial domain $\mathbb{X}$ and determine the solutions of the PDE (well-posedness). It is assumed that the initial condition of the problem is compatible with the model and its boundaries, but that the response due to the initial condition has already diminished or is compensated for through advanced signal processing, see (Berkel et al. 2020).

To estimate the transport coefficients $\{D, \mathbf{V}, K, P\}$, we reformulate the methodology presented in (Kampen et al. 2021c) to be applied to a more general nD

setting. As the model, gradient and Laplacian operators are linear, we can rewrite (5.1) in the frequency domain as

$$i\omega Z = \left(\nabla^2 Z\right) D + \nabla Z \cdot \nabla D - \nabla Z \cdot \mathbf{V} - Z\left(\nabla \cdot \mathbf{V}\right) + ZK + \Phi P, \tag{5.2}$$

where $Z(\mathbf{x}, \omega)$ and $\Phi(\omega)$ are the Fourier transform of $z(\mathbf{x}, t)$ and $\phi(t)$, respectively. For now, consider that Φ, Z and all of its derivatives are known such that the problem is linear in the unknown transport parameters, thus the resulting least squares problem is quadratic in the parameters (Banks and Kunisch 1989). However, as the transport parameters are functions of the spatial variable, the problem is infinite dimensional and hard to solve. For this reason, we apply a linear parametrization to each of the unknown parameters, say $\gamma \in \{D, \mathbf{V}, K, P\}$. For this, introduce basis functions $B^\gamma(\mathbf{x})$ and weighting vector θ^γ of dimension R^γ such that $\gamma = \sum_{r=1}^{R^\gamma} B_r^\gamma \theta_r^\gamma$. This means that we only need to estimate $R = R^D + R^{\mathbf{V}} + R^K + R^P$ weights that parameterize the functions. Furthermore, this parametrization allows us to rewrite the problem by grouping the known terms in vectors

$$G^D = \begin{pmatrix} \left(\nabla^2 Z\right) B_1^D + \nabla Z \cdot \nabla B_1^D \\ \vdots \\ \left(\nabla^2 Z\right) B_{R^D}^D + \nabla Z \cdot \nabla B_{R^D}^D \end{pmatrix}^T \tag{5.3}$$

$$G^{\mathbf{V}} = \begin{pmatrix} -\nabla Z \cdot \mathbf{B}_1^V - Z\left(\nabla \cdot \mathbf{B}_1^V\right) \\ \vdots \\ -\nabla Z \cdot \mathbf{B}_{R^V}^V - Z\left(\nabla \cdot \mathbf{B}_{R^V}^V\right) \end{pmatrix}^T \tag{5.4}$$

$$G^K = \begin{pmatrix} ZB_1^K & \dots & ZB_{R^K}^K \end{pmatrix} \tag{5.5}$$

$$G^P = \begin{pmatrix} \Phi B_1^P & \dots & \Phi B_{R^P}^P \end{pmatrix}, \tag{5.6}$$

such that we can express (5.2) as an inner product

$$i\omega Z - \begin{pmatrix} G^D & G^{\mathbf{V}} & G^K & G^P \end{pmatrix} \theta = 0, \tag{5.7}$$

where θ stacks the vectors $\theta^D, \dots, \theta^P$ and should hold for all $(\mathbf{x}, \omega) \in \mathbb{X} \times \Omega$. As the user chooses the basis functions, the derivatives are known and (5.7) can be seen as a linear regression problem. Next, we approximate the equation error criterion (Banks and Kunisch 1989) by only considering the state Z at the measurement locations $\mathbf{x} \in \mathbb{X}_\mathbb{M} \subset \mathbb{X}$ for the relevant frequencies $\omega \in \Omega_\mathbb{M} \subset \Omega$ that are well above the noise level (i.e., excited/perturbed). The error criterion is then given by

$$\begin{aligned} \mathcal{V}_{ee} &= \int_\Omega \int_\mathbb{X} \left(i\omega Z - \begin{pmatrix} G^D & G^{\mathbf{V}} & G^K & G^P \end{pmatrix} \theta \right)^2 \mathrm{d}\mathbf{x}\mathrm{d}\omega \\ &\approx \sum_{\Omega_\mathbb{M}} \sum_{\mathbb{X}_\mathbb{M}} \left(i\omega Z - \begin{pmatrix} G^D & G^{\mathbf{V}} & G^K & G^P \end{pmatrix} \theta \right)^2. \end{aligned} \tag{5.8}$$

As the boundary conditions of the original problem are assumed to be unknown, the measurements on the edge of the measurement grid $\mathbb{X}_\mathbb{M}$ constitute Dirichlet

boundary conditions, but other choices are feasible and can be added to the equation error (Banks and Kunisch 1989).

Similar to (Escande and Sattin 2012; Kampen et al. 2021c), (5.8) can be seen as a linear regression model

$$\bar{Y} = \bar{G}\theta, \tag{5.9}$$

where $\bar{Y}$ and $\bar{G}$ are a concatenation of the considered data points for $i\omega Z$ and $\begin{pmatrix} G^D & G^{\mathbf{V}} & G^K & G^P \end{pmatrix}$, respectively. Then, the optimal solution in the least square sense is given by

$$\theta_{\text{opt}} = (\bar{G}^H \bar{G})^{-1} (\bar{G}^H \bar{Y}), \tag{5.10}$$

where $\bar{G}^H$ is the Hermitian transpose of $\bar{G}$. Hence, instead of iterative optimization procedures which may end up in local minima, we can simply evaluate the expression (5.10) to obtain the global minimum or use a linear equation solver. Moreover, due to linearity of the derived regression problem, different criteria than (5.8) such as weighted least squares also have a closed-form solution (Markovsky and Huffel 2007) and are equally easy to implement and use. In addition, due to the linearity it is straightforward to extend this to more advanced estimators such as a maximum likelihood estimator (Kampen et al. 2023a) which considers uncertainty in an optimal sense.

Until this point, we assumed that Z, ∇Z and $\nabla^2 Z$ are known at the measured locations, but note that in practice only Z is measured at a limited number of spatial locations. Similar to (Escande and Sattin 2012), the derivatives could be estimated by smoothing, interpolating and extrapolating the dataset. By doing so, one forces a specific solution of the transport parameters as they are directly related to the interpolation method and its derivatives. For this reason, we approximate $\nabla^2 Z$ and ∇Z using central finite difference (Sundqvist and Veronis 1970; Quarteroni and Valli 1994). For this semi-discretization it is known that increasing the number of (spatial) samples results in a convergence to the exact solution to the model (Quarteroni and Valli 1994). As we have a closed-form solution for the global optimum, our solution will thus also converge to the exact solution for an increasing number of (spatial) samples, under the assumption that the selected basis functions of γ have sufficient freedom. If the number of measurement locations is too limited such that it significantly affects the discretization accuracy, one can still resort to interpolation methods to increase the accuracy without requiring a complete continuous profile and therefore force a very specific solution.

Furthermore, the transport parameters are only considered at the data points that are included in the equation error. As a result, the estimated transport coefficients are exact at the considered data points if the provided derivatives are exact and the combination of basis functions gives sufficient freedom to describe the underlying function, while the values in-between data points can merely be used as interpolation of the estimated parameters. The maximum spatial variability of the transport coefficients is directly linked to the number of basis functions,

i.e., free parameters, and the number of spatial data points that constraint these free parameters. Hence, if the number of basis functions for a coefficient is lower than the number of spatial data points minus two (for the boundary conditions), regularization is applied naturally. This is in contrast to the classical transport codes where the choice of basis functions directly limits the solution space. Moreover, due to the low computational cost of the closed form solution, many different parametrizations can be evaluated, such that a machine learning approach (as shown in (Kampen et al. 2021c)) can be applied to find a suitable parametrization.

5.3 Simulation Examples

5.3.1 1D Simulation Example

In order to demonstrate the presented methodology, we start with a 1D simulation for a fusion relevant scenario, where the electron temperature Θ_e in cylindrical geometry is described by

$$i\omega\frac{3}{2}n_e\Theta_e = \frac{1}{\rho}\partial_\rho\left(\rho D n_e \partial_\rho \Theta_e + \rho V n_e \Theta_e\right) + P\Phi, \tag{5.11}$$

with spatially varying electron density $n_e(\rho)$. The temperature data is generated using a central finite-difference scheme with a grid $\mathbb{X}_\mathbb{M}$ of 1001 points and considering a block-wave modulation for ϕ with a 70% duty cycle at 5 Hz from which we only use the first three harmonics. The electron temperature is measured on an equidistant measurement grid consisting of 32 spatial points on which we assume the electron density and its gradient to be known. The simulated D and V are a third and second order polynomial, respectively, while the source profile P is a Gaussian. For the estimation procedure, B^D and B^V are sixth order polynomials and B^P is a third order B-spline. In this way, B^D and B^V have more parameters than required to describe D and V. By matching the number of control points for B^P with the spatial measurement grid minus two boundary conditions i.e., 30, no regularization is applied. The basis functions are multiplied with the required factors to account for the cylindrical geometry and spatially varying density, after which the unknown parameters are determined by evaluating (5.10). Figure 5.1 contains an overview of the used data and the estimation results. The estimated parameters are close to the simulated parameters despite the scarce finite difference grid to estimate the gradient. Only the estimate of V deviates somewhat at the end of the domain due to low amplitudes and change in density gradient.

As stated in (Escande and Sattin 2012), the source is crucial to the parameter estimation problem. Therefore, we want to show the importance of simultaneously estimating the transport parameters and the source profile, as a small mismatch between the true and used source profile can result in a large difference in estimated transport parameters. For the sake of clarity, we consider (5.11) with a constant density (to avoid cross-error analysis), diffusivity described by a third order polynomial, no convection, and a Gaussian source profile. The diffusivity is

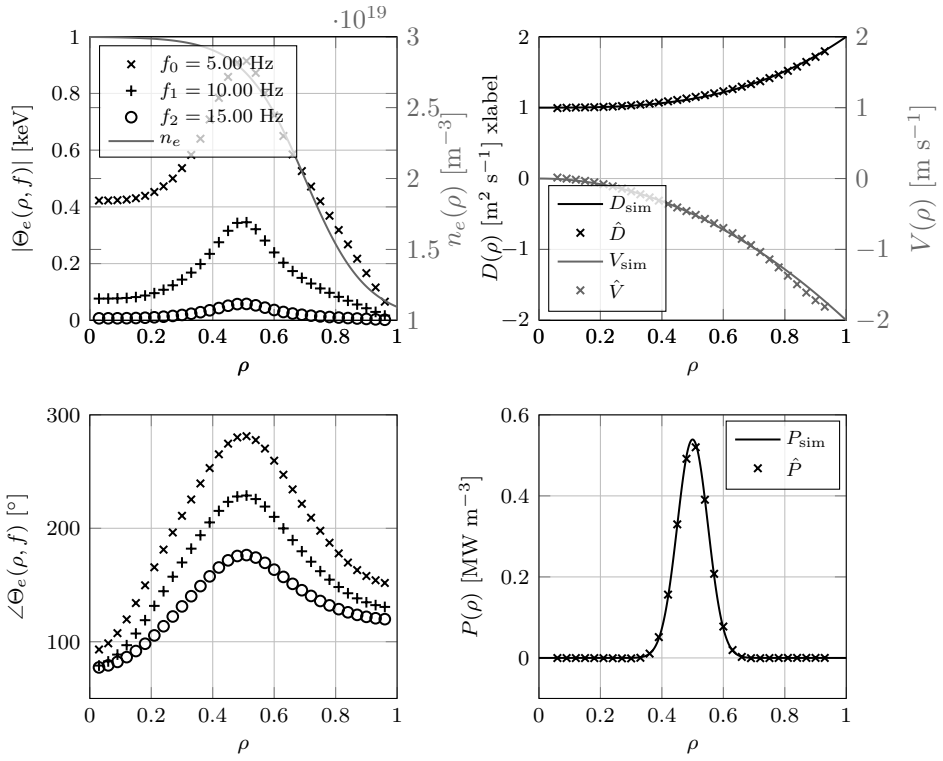


Figure 5.1. The density and simulated temperature profile (left) with the corresponding simulated and estimated transport parameters (right).

now estimated for two different scenario's: I) simultaneously with the source profile; and II) with a fixed source profile that is narrower than the simulated source profile. For the estimation, both diffusivity B^D and source B^P are third order B-spline with 30 control points such that there is no regularization. An overview of the used data and estimation results is given in Figure 5.2.

In Figure 5.2, scenario I), where the diffusivity and source are simultaneously estimated, there is only a small error in D at the boundary condition close to $\rho = 0$ due to low accuracy of the discretization grid. In Figure 5.2, scenario II), with the fixed used source that is smaller than the true source, there is a large error on the diffusion in the region with the source and a bias in the region without source.

5.3.2 2D Simulation Example

Although our methodology is applicable in nD, for the sake of clarity and presentation, we demonstrate it on a 2D scenario. The 2D scenario mimics a rotating fluid that is heated by a modulated uniform and Gaussian profile and uniform diffusion

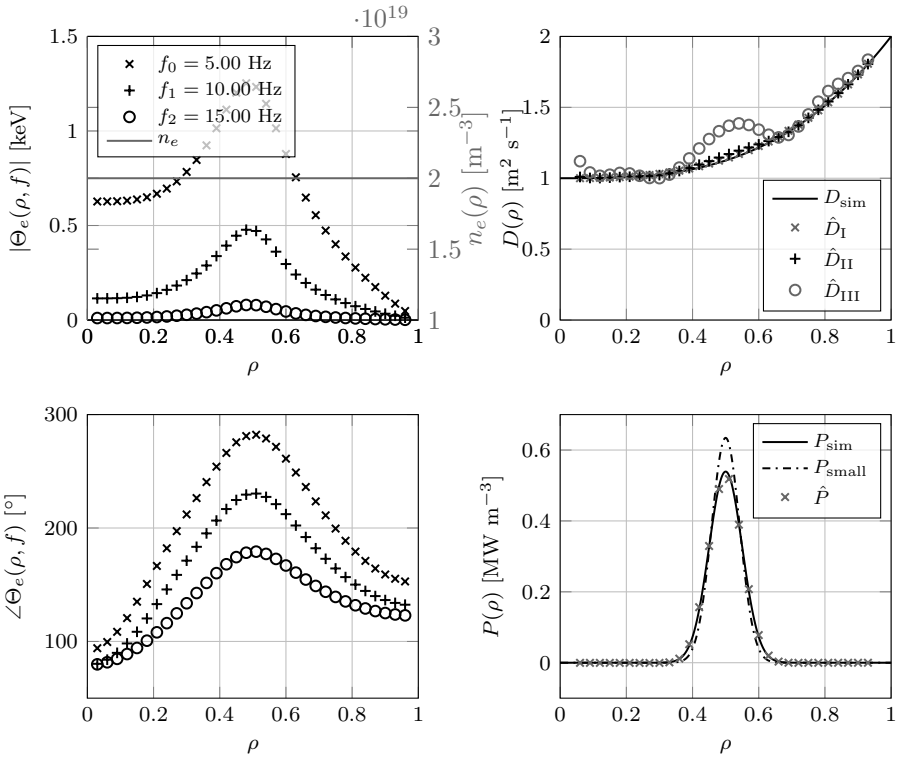


Figure 5.2. The density and simulated temperature profile (left) with the corresponding simulated and estimated transport parameters (right) for the two different scenario's: I) simultaneously estimate of the diffusivity $\hat{D}_{\text{I}}$ and source $\hat{P}_{\text{I}}$, and II) estimate of the diffusivity $\hat{D}_{\text{II}}$ with the sources that is too narrow P_{small}.

over the entire domain. The transport coefficients are given by the following functions:

$$D(x,y) = 1 \tag{5.12}$$

$$\mathbf{V}(x,y) = \begin{pmatrix} u \\ v \end{pmatrix} = 500 \begin{pmatrix} 0.5 - y \\ x - 0.5 \end{pmatrix} \tag{5.13}$$

$$P(x,y) = 1 + \frac{40}{\pi} e^{-\frac{(x-0.25)^2}{0.1^2}} e^{-\frac{(y-0.5)^2}{0.1^2}}. \tag{5.14}$$

The source is modulated by a 25 Hz block wave with a 50% duty cycle, where we only consider the first three harmonics (25 Hz, 75 Hz, 125 Hz). The temperature profile is generated by solving (5.2) using the finite volume-complete flux scheme (Thije Boonkkamp and Anthonissen 2011) with a grid of 1001×1001 points, and the measurements are taken on a 251×251 grid for the estimation procedure. The corresponding temperature profile of the first harmonic, the used convection and source are visualized in the four top subfigures in Figure 5.3.

The unknown 2D transport coefficients are parameterized by taking the product of two 1D basis functions for each $\gamma \in \{D, u, v, K, P\}$,

$$B_r^\gamma(x,y) = B_{r_x}(x) B_{r_y}(y) \tag{5.15}$$

with $r^\gamma = r_x^\gamma + (r_y^\gamma - 1) R_y^\gamma$, where R_y^γ is the total number of basis functions in the y direction. Here, B^D, B^u, B^v are 2×2 Chebyshev polynomials of the first kind and B^P a 40×40 Chebyshev polynomial. Hence, in total, we need to estimate 1612 unknown parameters, where evaluating (5.10) takes approximately 100 ms on a regular desktop. The estimation error on each function is also shown in Figure 5.3, which clearly shows that the parameters are well estimated and only small errors are made with respect to the function values. The error on D, u and v are low order polynomials due to the low order of the approximation, while the error on P is the result of Chebyshev polynomials that cannot perfectly describe the Gaussian deposition profile and compensation for the errors made by the finite-difference method for the gradient approximation. Note that we display the (log) relative error. As u and v cross zero at 0.5, the relative error tends to infinity and is therefore much large close to 0.5 while being small in the other regions.

5.4 Conclusion and Discussion

In conclusion, our proposed methodology provides new opportunities to study nD transport by estimating the sinks/sources and transport coefficients while providing transparency about regularization and interpolation with extremely low computational effort due to the closed-form solution that guarantees the global optimum for the selected optimization criteria. The low computational cost of the closed-form solution opens up a number of new significant opportunities: (i) machine learning like approaches to find the best parameterization of the

transport coefficients as demonstrated in Kampen et al. 2021c; (ii) fast estimation of multidimensional transport coefficients, for which we present the first results in this letter. Moreover, as the source no longer needs to be localized, since the transport parameters and sources/sinks can be estimated simultaneously, adding a source to regions with a low signal-to-noise can significantly improve the quality of the coefficient estimation.

Furthermore, it is possible to extend the current methodology to include more coefficients, e.g., derivatives to other variables following from linearization or cross-terms from coupled transport. However, there is no guarantee that these additional coefficients can be estimated by simply considering more harmonics as the solution to the inverse problem with additional coefficient functions is not necessary unique. This is in contrast to the considered diffusion-convection-reaction-source equation which is known to have a unique inverse solution. Hence, uniqueness of solution should be studied before considering more coefficients.

Finally, the proposed methodology does not consider noise. In case of noisy measurements, the estimation problem becomes an error-in-variables problem as the measurements are used in the regressor matrix. Therefore, estimates with the proposed least squares estimator will be biased where the bias depends on the signal-to-noise ratio. Moreover, the gradients in our proposed methodology are computed via a finite difference scheme. Therefore, the uncertainty of the estimated gradients is a combination of the uncertainty of the surrounding sensors and scales with the distance between sensors. Hence, denser measurement grids will increase the uncertainty on the estimated gradients.

The best way to deal with noise is to use an estimator that takes the noise into account, e.g., Bayesian or maximum likelihood. For transport 1D this is still computable (e.g., see Kampen et al. 2023a), whereas for nD transport this quickly becomes computational expensive due to the curse of dimensionality that applies to both the dataset and the number of free parameters. Therefore, a computational better, but less optimal way to deal with noise is to use a weighted (total) least squares criterion with a closed-form solution where the weights are based on the uncertainty. Furthermore, the effect of noise in the finite-difference gradient approximation can be mitigated by smoothing the measurements and using the smoothing function to estimate the gradients, e.g via spline interpolation which is common practice in some fields Escande and Sattin 2012; Kirov et al. 2002. In the estimation of nD transport coefficients there is clear tradeoff between accuracy and computational cost. The effect of noise on the estimates with different criterion and new methods to efficiently deal with noise in nD transport are part of our future research.

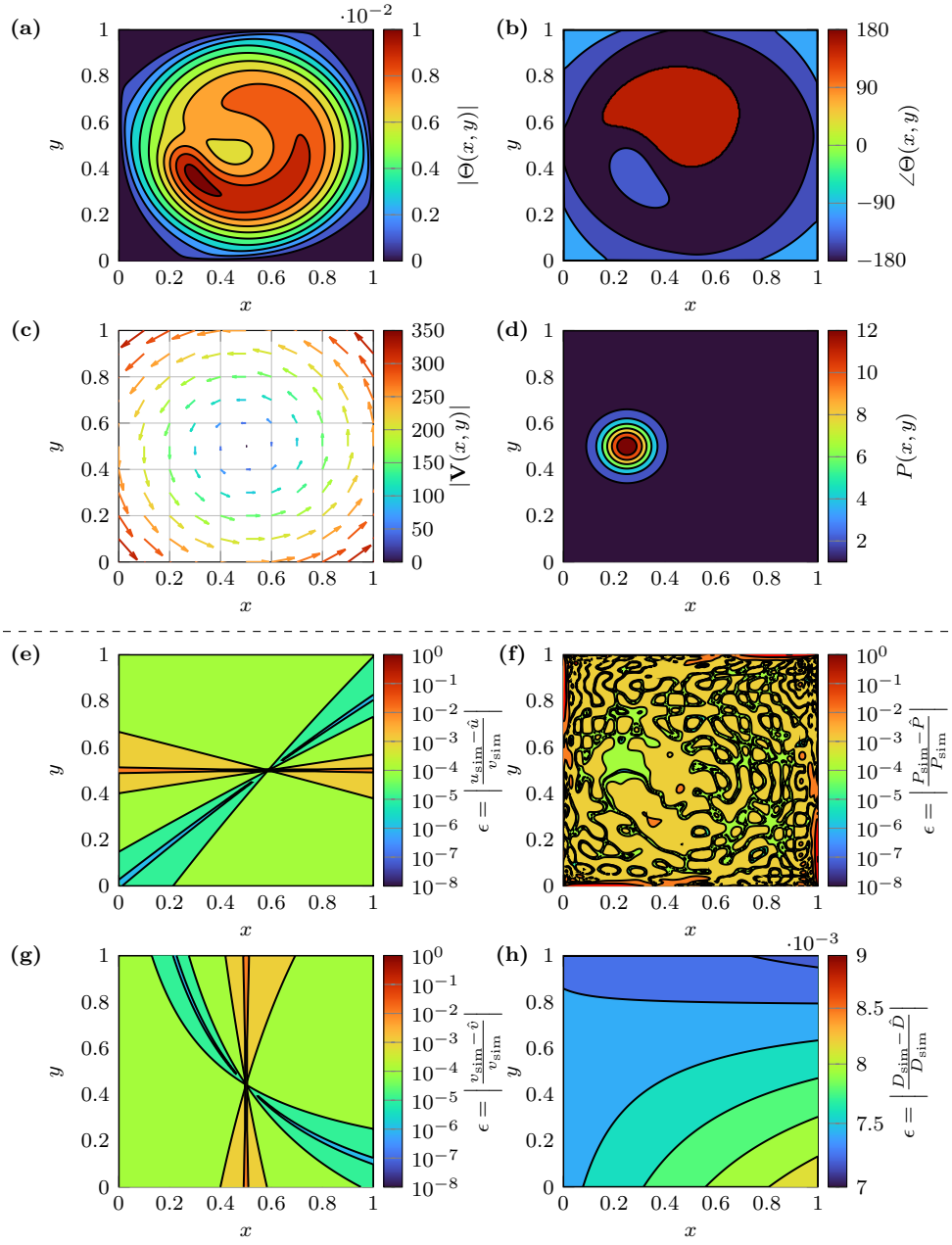


Figure 5.3. The 2D scenario with the absolute temperature profile of the modulation frequency at 25 Hz (a) and the phase (b), followed by the flow field (c) and source deposition profile (d). The relative estimation errors are shown for the x-component (e) and y-component (g) of the flow field, the source (f) and diffusivity (h).

6

Estimating 1D spatially varying coefficients taking uncertainty into account

Abstract - This chapter presents a closed-form solution to estimate space-dependent transport parameters of a linear one dimensional diffusion-transport-reaction equation. The infinite dimensional problem is approximated by a finite dimensional model by 1) taking a frequency domain approach, 2) linear parametrization of the unknown parameters, and 3) using a semi-discretization. Assuming full state knowledge, the commonly used output error criterion is rewritten as the equation error criterion such that the problem results in linear least squares. The optimum is then given by a closed-form solution, avoiding computational expensive optimization methods. Functioning of the proposed method is illustrated by means of simulation.

This chapter is published as: Kampen, R. J. R. van, M. van Berkel, and H. J. Zwart (2023a). "Estimating Space-Dependent Coefficients for 1D transport using Gaussian Processes as State Estimator in the Frequency Domain". In: *IEEE Control Systems Letters* 7 (2023), pp. 247–252. DOI: 10.1109/LCSYS.2022.3186626.

6.1 Introduction

Scalar transport (e.g., heat or mass) plays an important role in many different fields. For example, the efficiency of a nuclear fusion reactor is mainly determined by how much heat and particles the core plasma loses to the reactor wall (Hogeweij et al. 1998). Another example is given in the field of hydrology, where the goal is to identify hot spots for contaminants and nutrients in stream beds as a result of vertical ground water fluxes (Boano et al. 2014). Therefore, researchers require models which can be used for simulation, analysis, prediction, diagnosis and control of generic scalar transport. For most physical systems, these transport models are obtained using first principles, however the (exact) parameters to these models are unknown. Hence, data-driven estimation of the unknown physical parameters is necessary to complete the model, which is also known as an inverse problem or grey-box modeling.

The standard method to estimate the unknown parameters (diffusion, convection, reaction, and source/sink) in these typically infinite-dimensional models that describe the physical quantity, e.g., the temperature or density, is by minimizing the output error criterion, i.e., taking the (weighted) sum of the squared error between the measured data and the model output of a finite-dimensional approximation (Banks and Kunisch 1989; Kravaris and Seinfeld 1985). As this approach does not impose constraints on the state's variation in-between sensors, the state is allowed to oscillate (spatial aliasing) while the observed (output) error is zero. Even though these oscillations are small in the state, the errors in the estimated parameters are often significant, e.g., if they are based on the second derivative of the state (diffusion). This is generally resolved by regularizing the unknown parameters (Banks and Kunisch 1989; Kravaris and Seinfeld 1985; Ito and Kunisch 2008). As pointed out in (Kampen et al. 2021c; Escande and Sattin 2012), this regularization is often artificial as there is usually no *a priori* information on how the unknown parameters change as function of the spatial variable. Therefore, another option to add regularization is to first estimate the state as a function of the spatial variable using the measurements, and then perform the parameter estimation process (Banks and Kunisch 1989; Kampen et al. 2021c; Escande and Sattin 2012; Banks and Lamm 1985). This has the following advantages: (i) for an increasing number of samples the finite approximations used in these methods will converge towards the true solution, and so will their estimates; (ii) the availability of prior information on the state can be embedded in the state estimation process (such as smoothness due to the underlying model); (iii) the state estimate can be visually inspected for correctness and validated by taking additional (spatial) measurements; (iv) separating the state and parameter estimation process allows one to write the parameter estimation problem as a, often linear, regression problem (Banks and Kunisch 1989; Kampen et al. 2021c; Escande and Sattin 2012). Taking a frequency domain approach, (Kampen et al. 2021c) performs the state estimation via spline interpolation and exploits the linearity of the problem by deriving a closed-form solution for the global optimum using the ordinary least

squares criterion. However, for uncertain measurements this ordinary least squares solution is biased as both, the output and the regressor can contain error terms. Especially for the heterogeneous case, i.e., space-dependent parameters, very little noise already results in poor estimates. Under those circumstance this problem requires an errors-in-variables approach. Therefore, the two main contributions of this chapter are extending the methodology proposed in (Kampen et al. 2021c) by (i) determining the probability density function of the state as a function of the spatial variable using Gaussian process regression (GPR), and (ii) deriving a maximum likelihood solution such that it can deal with uncertainty in an optimal way when estimating the unknown parameters.

In contrast to other methods that use a Gaussian process (GP) to estimate the state, e.g., (Raissi et al. 2017; Rai and Tripathi 2019; Jidling et al. 2017), we perform the GPR in the frequency domain instead of the time domain. In this way, our method does not need to determine the time derivative, which is often hard to estimate under noisy conditions; and we can work with smaller datasets as in perturbation experiments only a limited number of frequencies are informative, i.e., excited and above the noise floor.

6.2 Problem formulation

The problem setup is specified as follows.

Model-class The estimation of the space-dependent physical coefficients is based on a one-dimensional linear parabolic partial differential equation (PDE) that describes the state, e.g., temperature or mass concentration, around an equilibrium point (Baukal et al. 2000):

$$\partial_t z = -\nabla \cdot (Vz - D\nabla z) + Kz + P\phi, \tag{6.1}$$

where $z : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ is a multi-variable function of a bounded space $x \in \mathbb{X} := [x_b, x_e] \subset \mathbb{R}$ and time $t \in \mathbb{T} := [t_0, t_e] \subseteq \mathbb{R}_{\geq 0}$. The state z is understood point-wise in x and t evaluated as $z(x,t)$. The physical transport coefficients are the diffusion $D : \mathbb{X} \to \mathbb{R}_{>0}$, convection $V : \mathbb{X} \to \mathbb{R}$, reaction $K : \mathbb{X} \to \mathbb{R}$, and the fixed spatially distributed source or sink $P : \mathbb{X} \to \mathbb{R}$. The source/sink is manipulated in time only by the external input $\phi : \mathbb{T} \to \mathbb{R}$. For well-posedness, the PDE is constrained at x_b and x_e by two (arbitrary) linear boundary conditions. Moreover, the initial condition $z(\cdot, t_0)$ is assumed to be compatible with the model and its boundaries. Due to linearity of the model, (6.1) can be considered in the frequency domain without loss of information (Pintelon and Schoukens 2012). Hence, the generic scalar transport equation in the frequency domain is

$$i\omega Z = -\nabla \cdot (VZ - D\nabla Z) + KZ + P\Phi, \tag{6.2}$$

with the Fourier transformed state $Z = \mathcal{F}(z) : \mathbb{X} \times \Omega \to \mathbb{C}$, input $\Phi = \mathcal{F}(\phi) : \Omega \to \mathbb{C}$, $i^2 = -1$, and angular frequency $\omega \in \Omega \subset \mathbb{R}$.

Measurement data The measurement signals $\mathbf{y}(t) := \operatorname{col}(y_1(t), \ldots, y_M(t))$ are assumed to be band-limited measurements of the scalar transport state z at $M > 2$ fixed locations given by the set $\mathbb{X}_M := \{\check{x}_1, \ldots, \check{x}_M\} \subseteq \mathbb{X}$ and are disturbed by noise $\varepsilon(t)$, i.e., $y_m = z(\check{x}_m, t) + \varepsilon_m(t)$. As we consider the problem in the frequency domain, we assume the discrete Fourier transform (DFT) spectra $\mathbf{Y}(k) = \mathcal{F}(\mathbf{y})$ and $\Phi(k)$ to be processed such that it only contains the forced response. This means that transients (non-steady-state behavior, e.g., due to the initial condition) have been removed by either waiting until the transient terms are diminished, or compensating for it using (advanced) signal processing techniques such as the local polynomial method (Pintelon and Schoukens 2012; Berkel et al. 2020) and hence, does not need to be considered further.

Additionally, we assume that the noise contributions are circular complex normally distributed in the frequency domain, which imposes weak assumptions on the time domain noise distribution (Pintelon and Schoukens 2012). As a result, the uncertainty of the (processed) Fourier spectrum is also circular complex normally distributed, i.e., $\mathbb{E}\{(\mathbf{Y}-\mathbb{E}\{\mathbf{Y}\})(\mathbf{Y}-\mathbb{E}\{\mathbf{Y}\})^{\mathsf{T}}\} = 0$ and $\mathbb{E}\{(\mathbf{Y}-\mathbb{E}\{Y\})(\mathbf{Y}-\mathbb{E}\{\mathbf{Y}\})^{\mathsf{H}}\} = C_{\mathbf{Y}}$, with the Hermitian transpose denoted as $^{\mathsf{H}}$. Here, $C_{\mathbf{Y}}$ is either known or can be estimated (Pintelon and Schoukens 2012). Moreover, due to linearity of the model, each (excited) angular frequency ω_k, with frequency bin $k \in \mathbb{K}$, is independent and only a finite number of the discrete angular frequencies are informative, i.e., those which are present in the (boundary) input and are above the noise level (Pintelon and Schoukens 2012). Therefore, we only consider those bins to be in $\mathbb{K}$, and excluded the DC and Nyquist frequency from the spectrum. All these signal and noise assumptions are standard (weak) assumptions for frequency domain analysis (Pintelon and Schoukens 2012).

Problem definition The goal is to estimate the set of space-dependent coefficients, i.e., functions, $\Gamma := \{D, V, K, P\}$ in (6.2), related to the physical quantities diffusion, convection, etc., using the (processed) spectra $\mathbf{Y}(k)$ and $\Phi(k)$. The (exact) boundary conditions that constrain the generic scalar transport are often unknown as they may depend on the equilibrium or unknown space-dependent parameters. Therefore, the extremum measurements are used as boundary inputs (Schneidewind et al. 2016; Das et al. 2019), which reduces the domain on which the space-dependent coefficients are estimated to $\mathbb{X}_E := [\check{x}_1, \check{x}_M]$. However, other linear boundary conditions are allowed. Hence, the formal problem definition is

Problem 6.1. *Given the processed data-set*

$$\mathbb{D} := \{\mathbf{Y}(k), C_{\mathbf{Y}}(k), \Phi(k) \mid k \in \mathbb{K}, \check{x}_m \in \mathbb{X}_M\},$$

estimate the unknown functions $\Gamma = \{D, V, K, P\}$ *by minimizing a cost function* $\mathcal{V}(\mathbb{D}, Z(x, k; \Gamma))$ *over* Γ *such that the solution* $Z(x, k; \Gamma)$ *satisfies the model* (6.2) *subject to the boundary conditions*

$$Z(\check{x}_1, k) = Y_1(k), \quad Z(\check{x}_M, k) = Y_M(k), \quad k \in \mathbb{K}.$$

6.3 State estimation by Gaussian process regression

As (Kampen et al. 2021c; Banks and Lamm 1985) shows, to have a unique solution for the coefficients, a continuous state Z is required. However, in general, the state is only measured at a limited number of spatial locations with uncertainty. To find the state Z as a continuous function of the spatial variable, we resort to GPR, because it takes the uncertainty into account and provides a probability density function of the state estimate which can then be used to determine the maximum likelihood solution of the unknown coefficients.

6.3.1 Gaussian process regression

A GP is defined as a collection of random variables, any finite number of which have a joint Gaussian distribution (Rasmussen and Williams 2005). Therefore, a GP $f(x)$ is completely defined by its mean function $\mu(x)$ and covariance function $\kappa(x, \hat{x})$, and is denoted as

$$f(x) \sim \mathcal{GP}\left(\mu(x), \kappa(x, \hat{x})\right). \tag{6.3}$$

If the process is observed under some Gaussian distributed noise ε with a known covariance C_ε, i.e., $\mathbf{y} = \operatorname{col}\left(f(\check{x}_1), \ldots, f(\check{x}_M)\right) + \varepsilon$, the joint prior distribution of the measured values $\mathbf{y}$ at $\check{\mathbf{x}}$ and the prior predictive distribution $\hat{\mathbf{f}}$ at $\hat{\mathbf{x}}$ is given by

$$\begin{bmatrix} \mathbf{y} \\ \hat{\mathbf{f}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}(\check{\mathbf{x}}) \\ \boldsymbol{\mu}(\hat{\mathbf{x}}) \end{bmatrix}, \begin{bmatrix} K(\check{\mathbf{x}}, \check{\mathbf{x}}) + C_\varepsilon & K(\check{\mathbf{x}}, \hat{\mathbf{x}}) \\ K(\hat{\mathbf{x}}, \check{\mathbf{x}}) & K(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \end{bmatrix} \right), \tag{6.4}$$

with the mean vector $\boldsymbol{\mu}$ and covariance matrix $K(\mathbf{x}, \mathbf{x})$ for which the v, w-th element has value $\kappa(x_v, x_w)$. The posterior predictive distribution after conditioning with the prior joint distribution is given by,

$$\hat{\mathbf{f}} \mid \check{\mathbf{x}}, \mathbf{y}, \hat{\mathbf{x}} \sim \mathcal{N}\left(\boldsymbol{\mu}_{\hat{\mathbf{f}}}, C_{\hat{\mathbf{f}}}\right), \tag{6.5}$$

$$\begin{aligned} \boldsymbol{\mu}_{\hat{\mathbf{f}}} &= \boldsymbol{\mu}(\hat{\mathbf{x}}) \\ &\quad + K(\hat{\mathbf{x}}, \check{\mathbf{x}})[K(\check{\mathbf{x}}, \check{\mathbf{x}}) + C_\varepsilon]^{-1}\left(\mathbf{y} - \boldsymbol{\mu}(\check{\mathbf{x}})\right), \end{aligned} \tag{6.6}$$

$$\begin{aligned} C_{\hat{\mathbf{f}}} &= K(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \\ &\quad - K(\hat{\mathbf{x}}, \check{\mathbf{x}})[K(\check{\mathbf{x}}, \check{\mathbf{x}}) + C_\varepsilon]^{-1} K(\check{\mathbf{x}}, \hat{\mathbf{x}}), \end{aligned} \tag{6.7}$$

which are the key equations for GPR (Rasmussen and Williams 2005). Typically, the covariance function, also called kernel, will have some free parameters α, called hyperparameters, and are tuned by maximizing the marginal likelihood, i.e., the likelihood that the given prior has generated the observed data. This is equal to minimizing the following cost function

$$\begin{aligned} \mathcal{V}_{\mathcal{GP}}(\alpha) = &-\left(\mathbf{y} - \boldsymbol{\mu}(\check{\mathbf{x}})\right)^{\mathsf{T}}\left[K(\check{\mathbf{x}}, \check{\mathbf{x}}, \alpha) + C_\varepsilon\right]^{-1}\left(\mathbf{y} - \boldsymbol{\mu}(\check{\mathbf{x}})\right) \\ &- \log\left(\det\left(K(\check{\mathbf{x}}, \check{\mathbf{x}}, \alpha) + C_\varepsilon\right)\right). \end{aligned} \tag{6.8}$$

6.3.2 Complex valued Gaussian process regression

As the measurements are complex valued, the GPR should be adopted accordingly. For the sake of simplicity, we consider the state Z at each frequency separately and assume it to be a zero mean circular complex GP which is a standard assumption (Lataire and Chen 2016; Hallemans et al. 2021). However, we pose the additional constraint that the covariance between the real and imaginary part is zero. This allows us to use standard kernels for real valued GPs while only limiting the information that is shared between the real and imaginary part. The complex valued GP is then given by the two joint real valued GPs

$$\begin{bmatrix} \mathrm{Re}\,(Z(x,k)) \\ \mathrm{Im}\,(Z(x,k)) \end{bmatrix} \sim \mathcal{GP}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \kappa(x,\hat{x},\alpha) & 0 \\ 0 & \kappa(x,\hat{x},\alpha) \end{bmatrix} \right), \tag{6.9}$$

denoted as $Z_{\mathrm{Re}} \sim \mathcal{GP}\,(0, \kappa_{\mathrm{Re}}(x,\hat{x}))$. The assumption of a zero mean is not a drastic restriction as it does not confine the mean of the posterior process to be zero (Rasmussen and Williams 2005). However, in the future, a mean could be included to increase interpretability of the model or add prior information. The following transformation is used to go from complex valued data to a real valued vector $\mathbf{Y}_{\mathrm{Re}}$ and covariance matrix $C_{\mathbf{Y}_{\mathrm{Re}}}$

$$\mathbf{Y}_{\mathrm{Re}} = \frac{1}{2} \begin{bmatrix} I & I \\ -iI & iI \end{bmatrix} \begin{bmatrix} \mathbf{Y} \\ \overline{\mathbf{Y}} \end{bmatrix}, \tag{6.10}$$

$$C_{\mathbf{Y}_{\mathrm{Re}}} = \frac{1}{4} \begin{bmatrix} I & I \\ -iI & iI \end{bmatrix} \begin{bmatrix} C_{\mathbf{Y}} & 0 \\ 0 & \overline{C_{\mathbf{Y}}} \end{bmatrix} \begin{bmatrix} I & I \\ -iI & iI \end{bmatrix}^{\mathsf{H}} \tag{6.11}$$

where $\overline{\mathbf{Y}}$ and $\overline{C_{\mathbf{Y}}}$ denote the complex conjugates. Now, the joint distribution of the real valued measurements and the real valued state prediction $\hat{\mathbf{Z}}_{\mathrm{Re}}$ follow from (6.4) while (6.5) to (6.7) are used to determine the (posterior) prediction $\hat{\mathbf{Z}}_{\mathrm{Re}}$ with mean $\boldsymbol{\mu}_{\hat{\mathbf{Z}}_{\mathrm{Re}}}$ and covariance matrix $C_{\hat{\mathbf{Z}}_{\mathrm{Re}}}$. To complete the state estimation process and go back to complex valued data, the following inverse transformation is used

$$\boldsymbol{\mu}_{\hat{\mathbf{Z}}} = \begin{bmatrix} I & iI \end{bmatrix} \boldsymbol{\mu}_{\hat{\mathbf{Z}}_{\mathrm{Re}}} \tag{6.12}$$

$$C_{\hat{\mathbf{Z}}} = \begin{bmatrix} I & iI \end{bmatrix} C_{\hat{\mathbf{Z}}_{\mathrm{Re}}} \begin{bmatrix} I & iI \end{bmatrix}^{\mathsf{H}}. \tag{6.13}$$

The kernel defines nearness or similarity between data points and can be used to embed prior information (Rasmussen and Williams 2005). As we can use standard kernels for real valued GPs, we choose the Matérn covariance function with $\nu = \frac{5}{2}$ as it enforces the GP, i.e., the state Z, to be twice (mean square) differentiable with respect to x. This corresponds to the solution of the underlying model (6.2) that is at least twice differentiable. In this way, we incorporate valuable prior knowledge of our system into the state estimation procedure via GPR. For completeness, the kernel is given by

$$\kappa(x,\hat{x},\alpha) = \sigma^2 \left(1 + \frac{\sqrt{5}|x-\hat{x}|}{\ell} + \frac{5(x-\hat{x})^2}{3\ell^2} \right) e^{-\frac{\sqrt{5}|x-\hat{x}|}{\ell}}, \tag{6.14}$$

with hyperparameters $\alpha = \mathrm{col}\,(\sigma, \ell)$ (Rasmussen and Williams 2005).

6.4 Finite-dimensional problem formulation

To create a finite-dimensional problem, we take the same approach as in (Kampen et al. 2021c; Das et al. 2019), starting with the linear parametrization of the unknown coefficients, followed by a finite difference scheme to approximate the spatial derivatives, resulting in a linear matrix equality. Therefore, the model (6.2) is reformulated such that it is linear in the state

$$i\omega_k Z = DZ'' + (D' - V)\, Z' + (K - V')\, Z + P\Phi, \tag{6.15}$$

where the prime ($'$) denotes the spatial derivative(s).

6.4.1 Parameterization of the unknown functions

To estimate the function $\gamma \in \Gamma = \{D, V, K, P\}$, we assume that each function can be described by a finite sum of basis functions B_r^γ weighted by $\theta^\gamma = \operatorname{col}(\theta_1^\gamma, \ldots, \theta_{R^\gamma}^\gamma)$,

$$\gamma(x, \theta^\gamma) = \sum_{r=1}^{R^\gamma} B_r^\gamma(x)\theta_r^\gamma. \tag{6.16}$$

Now, estimating the unknown coefficients is reduced to estimating $\theta = \operatorname{col}\left(\theta^D, \theta^V, \theta^K, \theta^P\right) \in \mathbb{R}^R$, with $R = R^D + R^V + R^K + R^P$. The basis functions B_r^γ should satisfy the model, which means that the derivative of B_r^D and B_r^V must exist. As the user chooses these basis functions, we consider these derivatives to be known.

6.4.2 Discretization procedure

The regression model requires spatial derivatives of the state. These can be obtained from the GPs (Ho et al. 2019). Alternatively, by approximating them using a numerical scheme, we can bypass the GPR by directly using the measurement data ($\hat{\mathbf{Z}} = \mathbf{Y}$). This allows to study the effect of the state estimation procedure on the estimated parameters and can help verifying the outcome. The spatial derivatives are approximated by a central finite difference scheme for non-equidistant grids (Sundqvist and Veronis 1970), as it is known that the finite-dimensional model will converge to the true infinite-dimensional model for an increasing number of points. For this we consider the discrete state vector $\mathbf{Z}(k) := \operatorname{col}(Z(x_1, k), \ldots, Z(x_N, k))$ that contains N spatial sample points $x_j \in \mathbb{X}_d \subseteq \mathbb{X}_E$, $j \in \{1, \ldots, N\}$, $x_1 = \check{x}_1$ and $x_N = \check{x}_M$. The derivatives for Z'', Z' and Z can be written using the matrices L_2, L_1, L_0, e.g, $\mathbf{Z}' = L_1\mathbf{Z}$. By excluding the derivatives on the boundaries, i.e., at x_1 and x_N, L_2, L_1 and L_0 are of size $N - 2 \times N$. Moreover, in this way, the boundaries are directly included as Dirichlet boundary conditions. Hence, (6.15) can be expressed as

$$i\omega_k L_0\mathbf{Z} = \hat{A}(\theta)\mathbf{Z} + \hat{B}(\theta)\Phi, \tag{6.17}$$

with

$$\hat{A}(\theta) = \sum_{r=1}^{R^D} (\tilde{B}_r^D L_2 + \tilde{B}_r^{D\prime} L_1)\theta_r^D - \sum_{r=1}^{R^V} (\tilde{B}_r^V L_1 + \tilde{B}_r^{V\prime} L_0)\theta_r^V + \sum_{r=1}^{R^K} \tilde{B}_r^K L_0 \theta_r^K \tag{6.18}$$

$$\hat{B}(\theta) = \sum_{r=1}^{R^P} (\tilde{B}_r^P \mathbf{1})\theta_r^P, \tag{6.19}$$

with $\tilde{B}_r^\gamma = \text{diag}\left(B_r^\gamma(x_2), \ldots, B_r^\gamma(x_{N-1})\right)$ as the diagonal matrix for each basis function $\gamma = D, V, K$ or P, and the column vector of ones denoted by $\mathbf{1}$. Hence, the problem is now written as a matrix equality that is bilinear in the unknown parameters θ and the discrete state vector $\mathbf{Z}$. Furthermore, note that other linear boundary conditions can be included by adapting $\hat{A}$ and $\hat{B}$ accordingly.

6.5 Maximum likelihood solution

For the maximum likelihood solution we assume that the state estimate or measurement $\hat{\mathbf{Z}}$ is circular complex normally distributed with a mean $\mu_{\hat{\mathbf{Z}}}$ and a covariance matrix $C_{\hat{\mathbf{Z}}}$. The cost function for the maximum likelihood is then given by (Pintelon and Schoukens 2012) as

$$\mathcal{V}(\theta, \hat{\mathbf{Z}}) = \sum_{k \in \mathbb{K}} \mathbf{e}(\theta, \mu_{\hat{\mathbf{Z}}}, k)^{\mathsf{H}} [C_{\mathbf{e}}(\theta, C_{\hat{\mathbf{Z}}}, k)]^{-1} \mathbf{e}(\theta, \mu_{\hat{\mathbf{Z}}}, k), \tag{6.20}$$

where the constant $F \ln(\pi) + \ln\left(\det\left(C_{\hat{\mathbf{Z}}}\right)\right)$ has been omitted as it does not affect the optima. The error vector from the model (6.17) yields,

$$\mathbf{e}(\theta, \mu_{\hat{\mathbf{Z}}}, k) = (i\omega_k L_0 - \hat{A}(\theta))\mu_{\hat{\mathbf{Z}}}(k) - \hat{B}(\theta)\Phi(k) \tag{6.21}$$

with the corresponding covariance matrix

$$C_{\mathbf{e}}(\theta, C_{\hat{\mathbf{Z}}}, k) = (i\omega_k L_0 - \hat{A}(\theta)) C_{\hat{\mathbf{Z}}}(k) (i\omega_k L_0 - \hat{A}(\theta))^{\mathsf{H}}. \tag{6.22}$$

The maximum likelihood solution is found by minimizing the cost function

$$\hat{\theta} = \arg\min_{\theta} \mathcal{V}(\theta, \hat{\mathbf{Z}}). \tag{6.23}$$

6.5.1 Minimizing the cost function

As the cost function is non-convex, we choose to solve it using iterative optimization methods such as Gauss-Newton or Levenberg-Marquardts (Nocedal and Wright 1999). For this we resort to the pseudo-Jacobian matrix which generally gives

faster convergence (Guillaume and Pintelon 1996). The parameter update $\Delta\theta$ for the iterative algorithms is found by solving the overdetermined set of equations

$$J_+(\theta, \hat{\mathbf{Z}}, k)\Delta\theta = \left[C_{\mathbf{e}}(\theta, C_{\hat{\mathbf{Z}}}, k)\right]^{-\frac{1}{2}} \mathbf{e}(\theta, \mu_{\hat{\mathbf{Z}}}, k), \tag{6.24}$$

for all $k \in \mathbb{K}$. The pseudo-Jacobian J_+ is given by

$$J_+(\theta, \hat{\mathbf{Z}}, k) = \begin{bmatrix} J_+^1(\theta, \hat{\mathbf{Z}}, k) & \dots & J_+^R(\theta, \hat{\mathbf{Z}}, k) \end{bmatrix}, \tag{6.25}$$

$$\begin{aligned} J_+^r(\theta, \hat{\mathbf{Z}}, k) = &\left[C_{\mathbf{e}}(\theta, C_{\hat{\mathbf{Z}}}, k)\right]^{-\frac{1}{2}} \left(\partial_{\theta_r}\mathbf{e}(\theta, \mu_{\hat{\mathbf{Z}}}, k)\right. \\ &\left. - \frac{1}{2}\partial_{\theta_r}C_{\mathbf{e}}(\theta, C_{\hat{\mathbf{Z}}}, k)\left[C_{\mathbf{e}}(\theta, C_{\hat{\mathbf{Z}}}, k)\right]^{-1}\mathbf{e}(\theta, \mu_{\hat{\mathbf{Z}}}, k)\right), \end{aligned} \tag{6.26}$$

where ∂_{θ_r} denotes the partial derivative to the weight θ_r and $C^{\frac{1}{2}}$ denotes the square root of the matrix (Pintelon and Schoukens 2012; Guillaume and Pintelon 1996). Note that these derivatives are easily obtained as the problem is formulated such that it is linear in the weights θ.

6.5.2 Calculation of the confidence intervals

The covariance matrix of the estimated weights $\hat{\theta}$ is given by (Pintelon and Schoukens 2012) as

$$C_{\hat{\theta}} = \left[2\mathrm{Re}\left(\sum_{k\in\mathbb{K}} J_+(\hat{\theta}, \hat{\mathbf{Z}}, k)^{\mathsf{H}} J_+(\hat{\theta}, \hat{\mathbf{Z}}, k)\right)\right]^{-1}. \tag{6.27}$$

As the transformation from $\hat{\theta}_\gamma$ to γ is linear, the $\mathfrak{p}$ confidence interval for the estimated functions $\gamma \in \Gamma$ is given by

$$C_\gamma(x, \hat{\theta}^\gamma, \mathfrak{p}) = \gamma(x, \hat{\theta}^\gamma) \pm \sqrt{2\sigma_\gamma^2(x)}\mathrm{erf}^{-1}(\mathfrak{p}), \tag{6.28}$$

where the variance of the function γ is determined using propagation of uncertainty

$$\sigma_\gamma^2 = \begin{bmatrix} B_1^\gamma(x) & \dots & B_{R^\gamma}^\gamma(x) \end{bmatrix} C_{\hat{\theta}^\gamma} \begin{bmatrix} B_1^\gamma(x) & \dots & B_{R^\gamma}^\gamma(x) \end{bmatrix}^{\mathsf{T}}, \tag{6.29}$$

where $C_{\hat{\theta}^\gamma}$ is a submatrix of the total covariance matrix $C_{\hat{\theta}}$.

6.6 Simulation results

The merit of the proposed methodology is demonstrated by generating a noisy dataset and estimating the transport coefficients using our early work (Kampen et al. 2021c) and the newly derived maximum likelihood estimator with and without the novel state estimation via GPR, including a noiseless dataset to focus on the differences GPR brings.

6.6.1 Data generation and state estimation

The simulation example is inspired by perturbative experiments in the field of nuclear fusion (Berkel et al. 2014c; Slief et al. 2022). The heat transport, (6.1), is generally analyzed on the normalized domain $\mathbb{X} = [x_b, x_e] = [0, 1]$ of the minor plasma radius. Here x_b is at the center and x_e is at the edge of the plasma. The corresponding boundary conditions in the simulation are $Z'(x_b, \cdot) = 0$ due to (axi)symmetry and Dirichlet boundary condition $Z(x_e, \cdot) = 0$ due to a significant temperature difference between core plasma ~170 million °C and edge plasma ~1 million °C. Typical parameter functions used in nuclear fusion are $D^{\text{sim}}(x) = 5x^3 - 0.005x + 5$, $V^{\text{sim}}(x) = -15x^2 + 0.005$, $K^{\text{sim}}(x) = 0$ and

$$P^{\text{sim}}(x) = \left(2 + \frac{7}{\sqrt{\pi}} e^{\frac{-(x-0.35)^2}{(0.1)^2}} + \frac{5.6}{\sqrt{\pi}} e^{\frac{-(x-0.6)^2}{(0.1)^2}}\right) \cdot 10^4,$$

which are equal to the functions in (Kampen et al. 2021c), except we set $K^{\text{sim}}(x) = 0$ as the reaction coefficient is intrinsically hard to estimate under noisy conditions compared to the other coefficients (Berkel et al. 2018a). Furthermore, we increased the power deposition with four orders of magnitude to be more realistic (Slief et al. 2022). For the perturbation of the plasma temperature, a microwave source is used, where the excitation signal $\Phi(\omega)$ is a block-wave of $\omega_0 = 50\pi$ with a 70% duty cycle. Here, only the first five harmonics $\omega_k = k\omega_0$, $k = 1, \ldots, 5$ are informative and used for the estimation. The temperature data is generated by simulation with a central finite difference grid of 1001 sample points where the state is measured at $M = 16$ spatial sensors locations, positioned at $\check{x}_m = \Delta x \cdot m$, with $\Delta x = 0.058$ and $m = 1, \ldots, M$. Noise from a circular normal distribution with covariance matrix $C_{\mathbf{Y}} = \sigma_\varepsilon(k)I$ is added to the measurements, where ε is chosen such that the maximum signal-to-noise ratio, $\text{SNR} = \frac{|Y_m(k)|}{\sigma_\varepsilon(k)}$, is 40 dB and lowers with 2.5 dB for each harmonic (due to transport). Hence, overall the SNRs range from 5 to 40 dB at the different data points. For the first two harmonics, the simulated and measured state with its $\mathfrak{p} = 0.95$ confidence interval are shown in Figure 6.1. Based on the measurements and their uncertainty, the hyperparameters of the GP are tuned by optimizing (6.8) for each frequency separately. Then, the state estimate and covariance matrix is determined at $N = 76$ points given by $x_n = 0.058 + 0.0116(n - 1)$, with $n = 1, \ldots, N$. The estimated state and its $\mathfrak{p} = 0.95$ confidence interval is also shown in Figure 6.1.

6.6.2 Parameter estimation

For the estimation procedure, we consider that there is some prior knowledge on the shape of the coefficients, i.e., that D and V are polynomials and P Gaussian. Therefore, the basis functions to estimate D and V are monomials $B_r^\gamma(x) = x^{(r-1)}$ with $R^D = 7$ and $R^V = 6$, respectively. The orders are significantly higher than the actual order such that it is possible to find an exact (noiseless) description of D and V. The basis functions for the source are B-splines, as it is linear in the

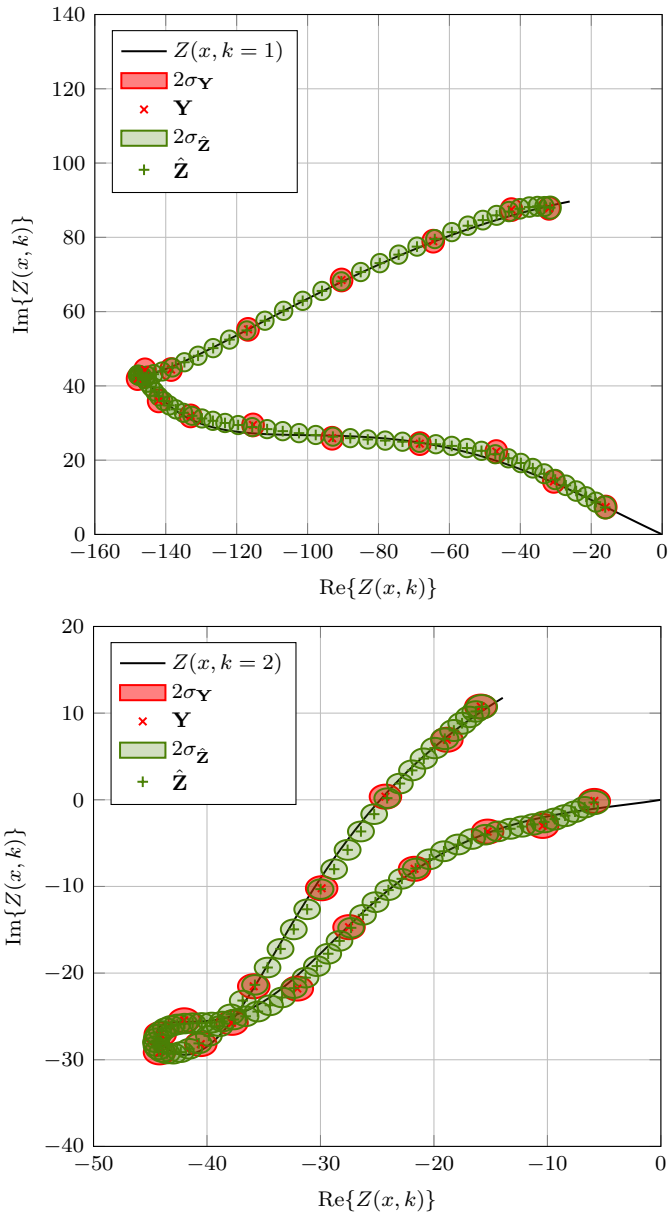


Figure 6.1. The simulated Z, measured $\mathbf{Y}$ and estimated state $\hat{\mathbf{Z}}$ shown in the complex plane for the first harmonic ($k = 1$) at $\omega = 50$ (top) and the second harmonic ($k = 2$) at $\omega = 100$ (bottom), including their $\mathfrak{p} = 0.95$ confidence interval.

unknown parameters and can describe smooth complex shapes. The B-splines are designed using the De Boor's algorithm (Boor 1978) with $R^P = 14$ control points and degree 3. Thus, in total, there are $R = 27$ free parameters. For a discussion on the effect of different basis functions, see (Kampen et al. 2021c).

The iterative optimization algorithm needs an initial starting point, which is set to $\theta = \mathbf{1}$. The result of the newly developed method with and without GPR is shown in Figure 6.2 along with the simulated coefficients. To show the improvement, the estimates using the methodology in (Kampen et al. 2021c) is also shown in Figure 6.2. Although the methodology in (Kampen et al. 2021c) comes with a closed-form solution of the global optimum, the estimates are poor. This is a result of the used ordinary least squares criterion that results in an biased estimator for the given problem formulation where the measurement uncertainty enters the regression matrix. The newly derived maximum likelihood estimators takes this into account resulting in better estimates. Moreover, in combination with the GPR as state estimator, spatial resolution and the quality of the estimated coefficients is improved. Although we used the GPs to apply regularization, there are still small oscillations in the estimated coefficients. This is a result of the small mismatch between the simulated and estimated state and the many degrees of freedom of the coefficients. For comparison, estimates based on noiseless observations are shown along the other estimates in Figure 6.2. The coefficients are estimated accurately except at the boundaries of the domain. Especially when using GPR. This is the result of small errors made in the state estimation using the Gaussian process as the amplitude of the signal quickly decreases due to the Dirichlet boundary condition and there is less information available at the boundaries for the GP.

6.7 Conclusion and discussion

This chapter presents a novel method to estimate the unknown space-dependent transport coefficients for 1D generic scalar transport from noisy measurements by first determining the probability density function of the state and then using this information to estimate the unknown parameters with a maximum likelihood estimator. By separating the state and parameter estimation problem, we avoid the artificial regularization of the unknown transport coefficients, but apply regularization on the state via GPR. This approach has shown to be successful, although we must note that it has a reduced performance at the boundaries of the domain. However, we expect that this can be improved by including more (prior) knowledge, e.g., by forcing the solution of the GP to satisfy the model (Raissi et al. 2017; Jidling et al. 2017), considering multiple frequencies simultaneously and embedding the stability of the transfer function into the kernel (Lataire and Chen 2016; Hallemans et al. 2021), and simultaneously estimating the state and parameters by considering the coefficients as hyperparameters of the GP (Raissi et al. 2017). Nonetheless, in the current state, the proposed methodology shows to be a significant improvement over the earlier linear least squares method that

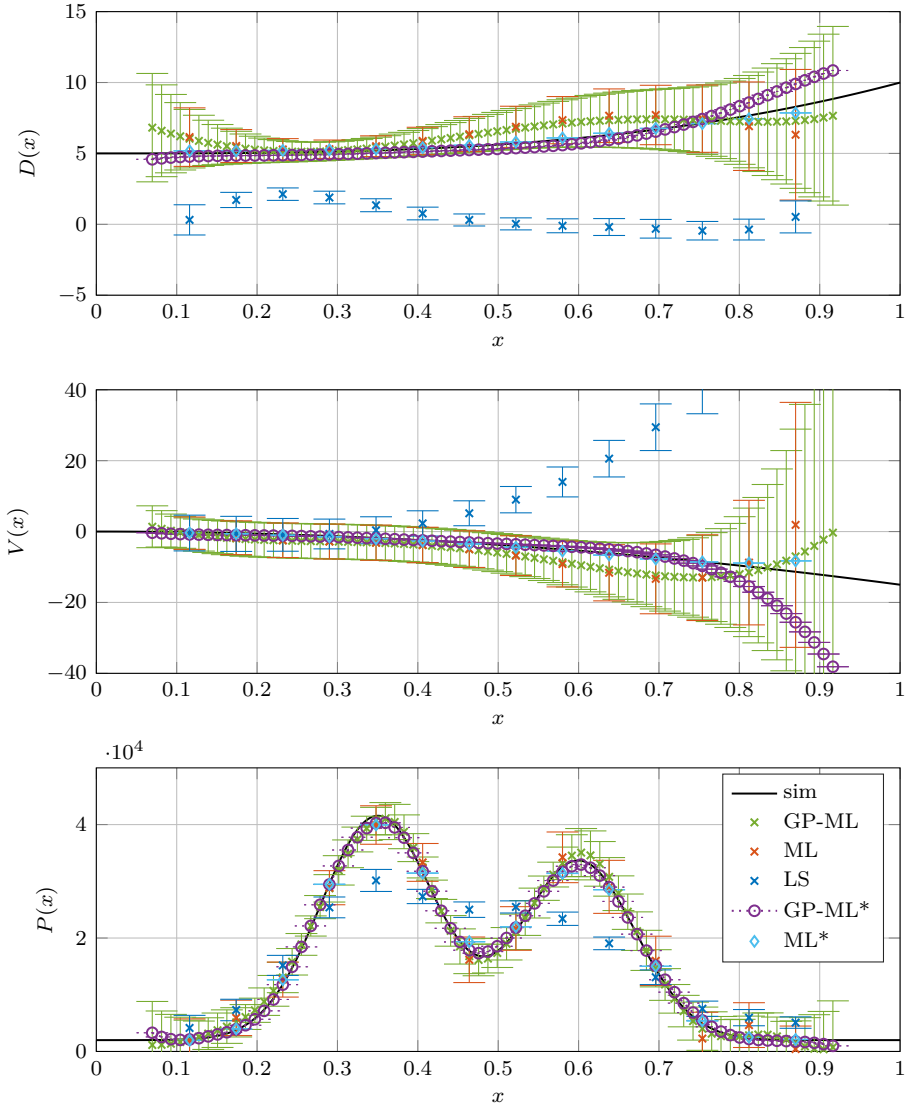


Figure 6.2. The estimated transport parameters: diffusion (top), convection (middle), source location (bottom) and their $\mathfrak{p} = 0.95$ confidence interval, estimated by the earlier presented ordinary least squares method (LS), the novel maximum likelihood method (ML) and the maximum likelihood method where the state and its uncertainty is estimated using Gaussian process regression (GP-ML) where the * indicates the noiseless scenario.

results in biased estimates when considering noise. Moreover, the novel derived maximum likelihood estimator can also work with only measurement data in case the GPR is distrusted and help verifying the correctness or influence of the GPR results as we showed for the noiseless scenario.

7 Conclusions, discussion and recommendations

7.1 Conclusions

The aim of this thesis is to the answer the research questions as stated in Section 1.4:

> *How can we accurately estimate the physical spatially varying coefficients in a generic scalar transport model in the presence of stochastic uncertainty?*

The three main contributions that helped answering this research question are as follows:

- A frequency domain approach has been investigated to estimate the three spatially varying transport coefficients: diffusivity, convectivity, reactivity/damping. In addition, the approach allowed to estimate the spatial distribution of the source/sink locations from data.
- The equation error criterion has been used instead of the commonly used output error criterion resulting in faster algorithms. Additionally, this allows for better analysis of the applied regularization and how it affects the coefficient estimates.
- It has been investigated how complex Gaussian process regression can be used to estimate the state in the frequency domain and how this state estimate can be used for regularization. This has been demonstrated by using a maximum likelihood estimator to estimate the spatially varying coefficients using the

state estimates. In this way, regularization is applied via the state without making assumptions on (the spatial variability of) the transport coefficients.

The research question constitutes two modeling choices, namely a deterministic part with the challenge of estimating the physical transport coefficients, and a stochastic part with the challenge how to deal with stochastic uncertainty. The conclusions for both parts are given below. Finally, the implementations of the proposed methodologies are discussed.

7.1.1 Deterministic challenges

The challenges regarding the deterministic modeling have led to answers to the following sub-research questions:

- **How does data processing and modeling choices affect the estimated coefficient?** This question is answered in two parts: i) the required data processing for working with frequency domain models and ii) the effect of boundary conditions on coefficient estimates.

  i) For the coefficient estimation process it is important that the model and experimental data are compatible. As this thesis takes a frequency domain approach to estimate the transport coefficients, the used models only consider the forced response of the system while the measurement data can contain transients, e.g., due to initial conditions. Transients have most of their energy in the low frequency range. Therefore, it is standard practice to compensate the Fourier coefficients in the low frequency range for transients. However, for distributed systems, the forced response often decays quicker with (spatial) depth for high frequencies than the transient decays with depth, significantly impacting the Fourier coefficients at high frequencies. For distributed systems, the transients have a negative impact over the complete frequency and spatial domain and need to be removed. This thesis uses the local polynomial method to remove transients and shows that this significantly improves the frequency response function and the corresponding estimated coefficients. More specifically, this is demonstrated on experimental data from the field of hydrology where they normally only use one strong frequency component that is the least affected by transients. After processing the data, it is demonstrated that all frequencies can be used as they are now compatible with the model (that only contains the forced response).

  ii) Boundary conditions play a key role as they determine the solution to the model, and accordingly impact the coefficient estimation. For hydrology, two approaches have been proposed. Firstly, the semi-infinite domain model, which is the standard model assumption. Secondly, the finite domain model where measurements are regarded as Dirichlet boundary conditions to create a locally bounded domain. This finite

domain model avoids errors related to semi-infinite domains (Berkel 2015). Application of the two models to the same dataset often results in different coefficient estimates, with extremes of flows in opposite direction.

To compare and discern the two models, existing methods for the semi-infinite and finite domain model are extended. For the finite-domain model that uses multiple frequencies to estimate constant transport coefficients, this extension includes the use of multiple arbitrary sensor locations. Extension of the existing methods allows for both models (semi-infinite and finite) to process exactly the same data. Application to synthetic datasets demonstrates that a small violation of the semi-infinite domain model assumption can result in coefficient estimates with a wrong sign. Application on experimental data shows that if the estimates obtained with both models agree, the model fit and cost function value are very similar. If the estimated coefficients do not agree, the bounded domain model fits the data better and has a much lower cost function value, indicating more trustworthy estimates. Therefore, in case three or more sensors are available, the semi-infinite domain should be abandoned.

- **When does the estimation of spatially varying coefficients require regularization and how does this affect the estimates?** The standard method to estimate the transport coefficients is by minimizing the output error criterion, i.e., taking the (weighted) sum of the squared error between the measurements of the state (transport variable) and the model predictions at the corresponding sample points. This approach suffers from two problems:
  - The methods used for finding the optimal parameters for the output error criterion are typically solved iteratively. This is time consuming and in many cases the problem is non-convex, i.e., numerical solvers do not guarantee convergence to the optimal solution.
  - As only the output error is optimized at the sampled locations, this approach does not impose constraints on the state in-between data points. Therefore, the coefficient estimates start oscillating. This can even happen when the observed (output) error is zero due to spatial aliasing of the state.

  The latter problem is generally resolved via regularization. This can be on the state level, by providing the state value in-between sensor locations, e.g., by interpolation, or on the level of the unknown coefficients by restricting the (variance of) spatial distribution of the unknown coefficients. Restricting the unknown coefficients is the most common method, but also artificial as the goal is to estimate the physical transport coefficients and there is usually no prior information on how the unknown parameters change as function of the spatial variable.

This thesis takes a different approach by separating the two problems. First, the solution of the state as a function of the spatial variable is estimated based on the (limited) measurement data, e.g., via inter- and extrapolation. Instead of the commonly used output error criterion, the equation error criterion is used. The equation error criterion takes the (weighted) sum of the squared error between the measured or estimated state and the model values at all points where the model is considered. Secondly, by using the equation error criterion in combination with a linear parameterization of the unknown coefficients, a closed-form solution for the unknown coefficients in the frequency domain has been derived. This approach brings the following advantages:

- The equation error clearly shows the relation between the (number of) unknown parameters that parameterize the coefficients, the (number of) spatial measurement locations, and the assumptions made on the state in-between spatial measurement locations. In this way it directly shows where regularization is applied and how it affects coefficient estimates.
    - Regularization via the state can be applied by estimating the state in-between measurement locations. This has the advantage of easy visual inspection and can be validated by taking additional spatial measurements.
    - Regularization via the coefficients can be applied by limiting the number of basis functions. Prior knowledge on the unknown coefficients can be incorporated in the selection of the basis functions. In principle, this regularization is not required for the noiseless case as the closed form solution is exact on the considered data points. For the noisy case regularization is applied via the state estimate. Hence, using a parametrization with more degrees of freedom than required avoids unwanted regularization.
- A closed-form solution can be used which significantly speeds up the process and avoids convergence to a local minimum.
    - This allows for the estimation of transport coefficients in n spatial dimensions as iterative methods suffer harder from the curse of dimensionality due to the required number of evaluations compared to this closed-form solution.
    - The low computational cost allows testing different levels of regularization in a machine-learning like manner, e.g., selecting the best parameterizations of the unknown coefficients via a model selection criterion such as minimum description length or Akaike information criterion.

7.1.2 Stochastic challenges

The presented equation error framework contributes to the deterministic part of the transport coefficients estimation but not the stochastic challenge: **How to deal with stochastic uncertainty when estimating spatially varying coefficients?** This thesis takes the viewpoint that only the measurement data is available without any other prior knowledge. Hence, the stochastic uncertainty on the measurements is unknown and needs to be estimated from the data itself. This has been incorporated via the local polynomial method that is also used to remove transients from the data. The local polynomial method is build upon the standard assumption of circular complex normal distributed noise on the Fourier coefficients. This is a weak assumption on the associated time domain noise distribution as several different additive noise distributions result in circular complex normal distribution of the Fourier coefficients. Provided that the local polynomial method or equivalent techniques are used, the probability density function of the (processed) Fourier coefficients is assumed to be known.

The transport coefficients that are most likely to have generated these Fourier coefficient distributions are determined by maximizing the likelihood function of the equation error criterion. Earlier maximum likelihood estimators are use a scalar expression as they only consider one input (semi-infinite domain) or one output (finite domain model). To deal with the multiple uncertain in- and outputs, required to estimate spatially varying coefficients, a new maximum likelihood estimator has been derived. This maximum-likelihood cost function is expressed in terms of vectors and matrices. Moreover, as the likelihood function is non-convex, the optimization problem is solved using iterative optimization scheme where the bi-linear structure of problem formulation is exploited. Furthermore, due to the linear parametrization of the unknown coefficients, the confidence interval of the estimated coefficients is determined via propagation of uncertainty. This results in spatially varying coefficients where the true physical coefficient lies within the boundaries of the estimated coefficients.

Often, regularization on the state level is required. Most methods that estimate spatially varying coefficients and that use regularization on the state level, do not consider stochastic uncertainties. However, it is important to determine the confidence bounds of the coefficient estimates. Therefore, it is also important to have access to the probability distribution of the estimated state as this information is required for the applied maximum likelihood estimator. This thesis estimates the state by considering the state for each frequency as an individual complex valued Gaussian process. In this way the (estimated) uncertainty of the measurements is used to determine the probability density function of the state. Then, the required samples for the equation error used in the maximum likelihood estimator can be drawn from this distribution. Prior information about the underlying PDE, such as smoothness and the (noise) distributions are embedded in the Gaussian process. This results in significantly improved estimates compared to the commonly used output-error criterion or methods that do not consider the uncertainties.

7.1.3 Implementations of the methodologies

This thesis delivered various MATLAB® implementations of the discussed methodologies that are applied and verified on both simulated and experimental data. The LPMLEn, the methodology to estimate constant transport coefficients using multiple frequencies, and multiple outputs, is published as open source software on HydroShare (Kampen et al. 2022c). The code is robust and reliable in terms of user interaction, but not in terms of the data verification and the corresponding estimation results. For example, for very noisy datasets it is more likely to end up in local minima or even with infeasible solutions such as negative diffusion coefficients. Therefore, the user should still verify if the estimates make sense or try different starting values for the iterative optimization routine.

The MATLAB implementation to estimate spatially varying coefficients in 1D is internally available at DIFFER and upon request for other interested researchers. The code is able to deal with non-uniform grids, contains various basis-functions for the parametrization such as B-splines and polynomials, has implementation for slab and cylindrical geometry, can deal with known source functions, known spatially varying coefficients such as the electron density for estimating heat transport in a nuclear fusion reactor, and allows the estimation of different coefficients combinations. When using the (linear) least square estimator, the closed-form solution is efficiently determined as it only requires solving a system of linear equations. Furthermore, the iterative optimization algorithm for the maximum likelihood estimator uses the pseudo Jacobian for faster convergence. Therefore, the most computational heavy steps are often the generation of basis functions and regression matrices instead of solving the system of equations. The computational time of the implemented algorithms is hard to evaluate as it depends on the number of data points, degrees of freedom of the coefficient parameterization and the sparsity of the problem, i.e., if the used basis functions have local support or have support over the complete domain. For the used synthetic datasets in this thesis the complete estimation time (including basis function generation) for the closed-form solution has been below 100 ms, where for the maximum likelihood solution it is a few seconds. These are the times on a regular laptop computer.

The implementation of the closed-form solution for nD transport coefficients is currently the most limited as it can only deal with square uniformly distributed measurement grids. For the example given in Section 5.3.2, generating the 1612 basis functions for the 249×249 grid (73206 spatial data points) takes a few seconds, but the estimation procedure for the corresponding 1612 unknown parameters takes approximately 100 ms on a regular laptop computer.

On a general note, the methods in this thesis have already been implemented and used in the fields of hydrology to quantifying vertical streambed fluxes around woody structures (Schneidewind et al. 2021) and nuclear fusion to quantify the electron cyclotron power deposition broadening (Slief et al. 2023). In principle, the methodology and the implementation are applicable to every transport problem described by (1.13). However, we have observed that significant systematic errors

in measurements or input parameters, for instance originating from the inversion of line integrated measurements to spatially distributed measurements or a source function that is assumed to be known, result in spatial distributed parameter estimates which are inconclusive. Hence, the quality of the data and especially to what extend systematic errors (bias) are present in the measurements or input parameters determines the applicability of the proposed (if not all) methodologies to produce reliable estimates.

7.2 Discussion and recommendations

This section describes possible future research and is subdivided into two parts. One focuses on the improvement of the presented methodology whereas the other focuses on the required extensions or applicability of the developed method in case the underlying assumptions do not hold.

7.2.1 Improvement of the estimation methodology

As shown in this thesis, the ability to estimate spatially varying transport coefficients is closely related to the number of spatial measurement points and the assumptions made on the state in-between spatial measurement points. Hence, avoiding (spatial) aliasing is key in the estimation of spatially varying transport coefficients. This thesis tackles the spatial aliasing problem by only considering the coefficient values on the spatial sample points. We use basis functions as spatial interpolation of the estimated coefficients. For regularization or artificially increased spatial resolution, the measured transport variable is interpolated via Gaussian process regression. The following two recommendations will help to further improve the current methodology:

- For a better fundamental understanding of the spatial aliasing problem and how it affects the estimation of spatially varying coefficients, it is recommend to use a spectral decomposition of the state. By writing the state as a summation of products of spatial functions and frequency (or time) depending coefficients, it clearly separates the spatial and/or temporal domain. Moreover, the selected basis function for the spectral decomposition directly implies the behavior of the state in-between spatial measurement points, i.e., interpolation. Furthermore, the error integrals over the spatial domain can be determined analytically. In this way, the basis functions for the spectral decomposition are directly related to the parametrization of the unknown coefficients over the complete spatial domain. Studying how these errors change by testing different combinations of basis functions and coefficient parametrizations can give more insight in the underlying spatial aliasing problem. Moreover, testing different combinations of basis functions and coefficient parameterization allows for a machine learning like approach to learn the best combinations.

- For better coefficient estimates, the state and coefficient estimation process should be done simultaneously, as the state depends on the coefficients. The recommended approach is to modify the kernel functions of the Gaussian process such that the state estimates satisfies the transport equations. Here, the modified kernels contain the unknown coefficients that can be treated as hyperparameters. Consequently, the state and coefficient estimation are coupled as the coefficients can be determined by optimizing the likelihood function of the Gaussian process while the state is used to provide the required regularization. Further improvements can be achieved by taking a Bayesian approach and assigning priors to the unknown coefficients or learning the model (structure) form the data.

7.2.2 Extensions of the estimation methodology

The methods presented in this thesis have been developed with specific assumptions in mind that simplify the analysis. However, in practice, these assumptions will not always be satisfied. Furthermore, by removing or reducing some assumptions it may make the developed methods useful for a larger group of users. Therefore, in the following paragraphs the most important assumptions are described and possible extensions are suggested.

Higher order partial differential equations

The methods presented in this thesis are focused around the transport equation, i.e. a second order parabolic partial differential equation. As the developed methods are generic, they can easily be extended to partial differential equations with higher order temporal and spatial derivatives (Vries 2021). However, a practical limitation is that higher order derivatives are more subject to noise, and therefore will limit the order of differential equations that be considered with this method.

Coupled transport equations

In many applications, the transport is not described by a single PDE but by a system of coupled PDEs, e.g., density and temperature. Solving the forward problem, i.e. determining the state, requires solving both PDEs simultaneously, whereas solving the inverse problem, i.e. estimating the unknown coefficients, allows decoupling of the PDEs. By treating the coupled transport variable (state) as an additional input, the coefficient estimation for each PDE can be solved individually. Hence, with a small extension of an additional input, the developed methods are applicable for coupled systems. However, it should be noted that uniqueness of the solution to the inverse problem is a concern for coupled PDEs thus should be investigated in the future (Vries 2021).

Nonlinear partial differential equations

The considered (linear) transport equation can be the result of a linearization, which only holds if the perturbations are sufficiently small. If this is not the case, e.g., larger amplitudes are required to rise above the noise level, then it will result in nonlinear dynamics. The considered linear model with the current methodology results in the best linear approximation. Extending the current methodology to nonlinear models is a big challenge as this can result in intermodulation or harmonic distortions that can significantly impact the excited harmonics. To start studying nonlinear transport models in the frequency domain, it is recommended to start with models where a periodic input results in the same period output (Berkel et al. 2017b). Then, for each frequency, the corresponding equation can be determined and estimating the coefficients can be done by following a similar methodology as presented in this thesis. However, this approach will result in many different coefficients as all combinations of harmonics should be considered. Hence, performing multiple experiments with different realizations or having prior knowledge about the model structure and parameters is a necessity to solve the estimation problem.

Bibliography

Abbott, B. W. et al. (2016). "Using multi-tracer inference to move beyond single-catchment ecohydrology". In: *Earth-Science Reviews* 160 (Sept. 2016), pp. 19–42. DOI: 10.1016/j.earscirev.2016.06.014.

Acworth, R. I., G. C. Rau, A. M. McCallum, M. S. Andersen, and M. O. Cuthbert (2014). "Understanding connected surface-water/groundwater systems using Fourier analysis of daily and sub-daily head fluctuations". In: *Hydrogeology Journal* 23.1 (Sept. 5, 2014), pp. 143–159. DOI: 10.1007/s10040-014-1182-5.

Acworth, R. I., L. J. S. Halloran, G. C. Rau, M. O. Cuthbert, and T. L. Bernardi (2016). "An objective frequency domain method for quantifying confined aquifer compressible storage using Earth and atmospheric tides". In: *Geophysical Research Letters* 43.22 (Nov. 2016). DOI: 10.1002/2016gl071328.

Aihara, S.-I. and A. Bagchi (1988). "Parameter identification for stochastic diffusion equations with unknown boundary conditions". In: *Applied Mathematics & Optimization* 17.1 (Jan. 1988), pp. 15–36. DOI: 10.1007/bf01448357.

Anderson, M. P. (2005). "Heat as a Ground Water Tracer". In: *Ground Water* 43.6 (Aug. 2005), pp. 951–968. DOI: 10.1111/j.1745-6584.2005.00052.x.

Anibas, C. et al. (2009). "Transient or steady-state? Using vertical temperature profiles to quantify groundwater-surface water exchange". In: *Hydrological Processes* 23.15 (July 2009), pp. 2165–2177. DOI: 10.1002/hyp.7289.

Anibas, C. et al. (2016). "From streambed temperature measurements to spatial-temporal flux quantification: using the LPML method to study groundwater-surface water interaction". In: *Hydrological Processes* 30.2 (Jan. 2016), pp. 203–216. DOI: 10.1002/hyp.10588.

Anibas, C. et al. (2018). "Delineation of spatial-temporal patterns of groundwater/surface-water interaction along a river reach (Aa River, Belgium) with transient thermal modeling". In: *Hydrogeology Journal* 26.3 (Dec. 2018), pp. 819–835. DOI: 10.1007/s10040-017-1695-9.

Antsaklis, P. J. and A. N. Michel (1997). *Linear Systems*. Boston, MA, USA: Birkhäuser, p. 670. ISBN: 978-0-8176-4434-5. DOI: 10.1007/0-8176-4435-0.

Baar, M. R. de, G. M. D. Hogeweij, N. J. Lopes Cardozo, A. A. M. Oomens, and F. C. Schüller (1997). "Electron Thermal Transport Barrier and Magnetohydrodynamic Activity Observed in Tokamak Plasmas with Negative Central

Shear". In: *Physical Review Letters* 78.24 (June 1997), pp. 4573–4576. DOI: 10.1103/physrevlett.78.4573.

Baker, B. (2007). "Delta-sigma ADCs in a nutshell". In: *EDN*. Aspencore Network, Dec. 2007. URL: https://www.edn.com/delta-sigma-adcs-in-a-nutshell/.

Baker, B. (2008). "Delta-sigma ADCs in a nutshell, part 2: the modulator". In: *EDN*. Aspencore Network, Jan. 2008. URL: https://www.edn.com/delta-sigma-adcs-in-a-nutshell-part-2-the-modulator/.

Banks, E. W. et al. (2018). "Active heat pulse sensing of 3-D-flow fields in streambeds". In: *Hydrology and Earth System Sciences* 22.3 (Mar. 20, 2018), pp. 1917–1929. DOI: 10.5194/hess-22-1917-2018.

Banks, H. T. and K. Kunisch (1989). *Estimation Techniques for Distributed Parameter Systems*. Systems & Control: Foundations & Applications. Boston, MA, USA: Birkhäuser. ISBN: 978-0-8176-3433-9. DOI: 10.1007/978-1-4612-3700-6.

Banks, H. T. and P. Lamm (1985). "Estimation of variable coefficients in parabolic distributed systems". In: *IEEE Transactions on Automatic Control* 30.4 (Apr. 1985), pp. 386–398. DOI: 10.1109/tac.1985.1103955.

Baukal, C. E., V. Y. Gershtein, and X. Li (2000). *Computational Fluid Dynamics in Industrial Combustion*. Ed. by C. E. Baukal, V. Gershtein, and X. J. Li. Boca Raton, FL, USA: CRC Press, Oct. 2000. DOI: 10.1201/9781482274363.

Baxter, C. V. and F. R. Hauer (2000). "Geomorphology, hyporheic exchange, and selection of spawning habitat by bull trout (Salvelinus confluentus)". In: *Canadian Journal of Fisheries and Aquatic Sciences* 57.7 (July 2000), pp. 1470–1481. DOI: 10.1139/f00-056.

Belov, Y. Y. (2002). *Inverse Problems for Partial Differential Equations*. Vol. 32. Inverse and Ill-Posed Problems Series. De Gruyter, Dec. 2002, p. 211. ISBN: 9789067643580. DOI: 10.1515/9783110944631.

Benosman, M. and A.-m. Farahmand (2017). "Gaussian Processes-based Parametric Identification for Dynamical Systems". In: *IFAC-PapersOnLine* 50.1 (July 2017), pp. 14034–14039. DOI: 10.1016/j.ifacol.2017.08.2431.

Benson, D. A., S. W. Wheatcraft, and M. M. Meerschaert (2000). "Application of a fractional advection-dispersion equation". In: *Water Resources Research* 36.6 (Feb. 2000), pp. 1403–1412. DOI: 10.1029/2000wr900031.

Berkel, M. van, G. W. Oosterwegel, M. J. H. Anthonissen, H. J. Zwart, and G. Vandersteen (2019). "A novel frequency domain maximum likelihood approach for estimating transport coefficients in cylindrical geometry for nuclear fusion devices". In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. Nice, France, Dec. 2019. DOI: 10.1109/cdc40024.2019.9029992.

Berkel, M. van, G. Vandersteen, H. J. Zwart, G. M. D. Hogeweij, and M. R. de Baar (2013). "Maximum Likelihood Estimation of diffusion and convection in tokamaks using infinite domains". In: *2013 IEEE International Conference on Control Applications (CCA)*. Hyderabad, India, Aug. 2013. DOI: 10.1109/cca.2013.6662920.

Berkel, M. van, H. J. Zwart, G. M. D. Hogeweij, and M. R. de Baar (2017a). "Technical note on the linearity and power dependence of the diffusion coefficient

in W7-AS". In: *Plasma Physics and Controlled Fusion* 59.6 (May 3, 2017), p. 062001. DOI: 10.1088/1361-6587/aa6a2b.

Berkel, M. van et al. (2014a). "Estimation of the thermal diffusion coefficient in fusion plasmas taking frequency measurement uncertainties into account". In: *Plasma Physics and Controlled Fusion* 56.10 (Sept. 2014), p. 105004. DOI: 10.1088/0741-3335/56/10/105004.

Berkel, M. van et al. (2014b). "Explicit approximations to estimate the perturbative diffusivity in the presence of convectivity and damping. I. Semi-infinite slab approximations". In: *Physics of Plasmas* 21.11 (Nov. 24, 2014), p. 112507. DOI: 10.1063/1.4901309.

Berkel, M. van et al. (2014c). "Frequency domain sample maximum likelihood estimation for spatially dependent parameter estimation in PDEs". In: *Automatica* 50.8 (Aug. 2014), pp. 2113–2119. DOI: 10.1016/j.automatica.2014.05.027.

Berkel, M. van et al. (2017b). "New evidence and impact of electron transport nonlinearities based on new perturbative inter-modulation analysis". In: *Nuclear Fusion* 57.12 (Sept. 28, 2017), p. 126036. DOI: 10.1088/1741-4326/aa827a.

Berkel, M. van et al. (2018a). "A systematic approach to optimize excitations for perturbative transport experiments". In: *Physics of Plasmas* 25.8 (Aug. 2018), p. 082510. DOI: 10.1063/1.5010325.

Berkel, M. van et al. (2018b). "Heat flux reconstruction and effective diffusion estimation from perturbative experiments using advanced filtering and confidence analysis". In: *Nuclear Fusion* 58.9 (July 2018), p. 096036. DOI: 10.1088/1741-4326/aad13e.

Berkel, M. van et al. (2020). "Correcting for non-periodic behaviour in perturbative experiments: application to heat pulse propagation and modulated gas-puff experiments". In: *Plasma Physics and Controlled Fusion* 62.9 (July 22, 2020), p. 094001. DOI: 10.1088/1361-6587/ab9eaa.

Berkel, M. van et al. (2023). "A framework to improve the reliability of temperature-based estimates of flow and diffusion in streambeds with frequency domain examples". In: *in preparation*.

Berkel, M. van (2015). "Estimation of heat transport coefficients in fusion plasmas". PhD thesis. Eindhoven University of Technology, June 2015.

Billingsley, P. (2012). *Probability and measure*. Hoboken, NJ, USA: Wiley, p. 624. ISBN: 9781118122372.

Bishop, C. M. and J. W. Connor (1990). "Heat-pulse propagation in tokamaks and the role of density perturbations". In: *Plasma Physics and Controlled Fusion* 32.3 (Mar. 1990), pp. 203–211. DOI: 10.1088/0741-3335/32/3/005.

Boano, F. et al. (2014). "Hyporheic flow and transport processes: Mechanisms, models, and biogeochemical implications". In: *Reviews of Geophysics* 52.4 (Oct. 2014), pp. 603–679. DOI: 10.1002/2012rg000417.

Bock, H., W. Jäger, and O. Venjakob (2013). *Model Based Parameter Estimation*. Ed. by H. G. Bock et al. Berlin, Germany: Springer Berlin Heidelberg. ISBN: 978-3-642-30366-1. DOI: 10.1007/978-3-642-30367-8.

Boor, C. de (1978). *A Practical Guide to Splines*. New York, NY, USA: Springer. ISBN: 978-0-387-95366-3.

Boorn, B. F. H. van den (2021). "Sensitivity analysis of the heat transport coefficient estimation in fusion reactors and their identifiability under noisy conditions". MA thesis. Eindhoven University of Technology, Feb. 12, 2021.

Boulton, A. J. (2007). "Hyporheic rehabilitation in rivers: restoring vertical connectivity". In: *Freshwater Biology* 52.4 (Apr. 2007), pp. 632–650. DOI: 10.1111/j.1365-2427.2006.01710.x.

Boulton, A. J., T. Datry, T. Kasahara, M. Mutz, and J. A. Stanford (2010). "Ecology and management of the hyporheic zone: stream-groundwater interactions of running waters and their floodplains". In: *Journal of the North American Benthological Society* 29.1 (Mar. 2010), pp. 26–40. DOI: 10.1899/08-017.1.

Boulton, A. J., S. Findlay, P. Marmonier, E. H. Stanley, and H. M. Valett (1998). "The functional significance of the hyporheic zone in streams and rivers". In: *Annual Review of Ecology and Systematics* 29.1 (Nov. 1998), pp. 59–81. DOI: 10.1146/annurev.ecolsys.29.1.59.

Bourdelle, C. (2005). "Turbulent particle transport in magnetized fusion plasma". In: *Plasma Physics and Controlled Fusion* 47.5A (Apr. 2005), A317–A326. DOI: 10.1088/0741-3335/47/5A/023.

Bracewell, R. N. (1986). *The Fourier transform and its applications*. 2nd. Ney York, NY, USA: McGraw-Hill. ISBN: 0070070156.

Briggs, M. A., L. K. Lautz, S. F. Buckley, and J. W. Lane (2014). "Practical limitations on the use of diurnal temperature signals to quantify groundwater upwelling". In: *Journal of Hydrology* 519 (Nov. 2014), pp. 1739–1751. DOI: 10.1016/j.jhydrol.2014.09.030.

Briggs, M. A., L. K. Lautz, J. M. McKenzie, R. P. Gordon, and D. K. Hare (2012). "Using high-resolution distributed temperature sensing to quantify spatial and temporal variability in vertical hyporheic flux". In: *Water Resources Research* 48.2 (Feb. 23, 2012). DOI: 10.1029/2011wr011227.

Brookman, M. W. et al. (2021). "Resolving ECRH deposition broadening due to edge turbulence in DIII-D". In: *Physics of Plasmas* 28.4 (Apr. 21, 2021), p. 042507. DOI: 10.1063/1.5140992.

Brunke, M. and T. Gonser (1997). "The ecological significance of exchange processes between rivers and groundwater". In: *Freshwater Biology* 37.1 (Feb. 1997), pp. 1–33. DOI: 10.1046/j.1365-2427.1997.00143.x.

Brunner, P. and C. T. Simmons (2011). "HydroGeoSphere: A Fully Integrated, Physically Based Hydrological Model". In: *Ground Water* 50.2 (Dec. 2011), pp. 170–176. DOI: 10.1111/j.1745-6584.2011.00882.x.

Bryant, S. R. et al. (2020). "Seasonal manganese transport in the hyporheic zone of a snowmelt-dominated river (East River, Colorado, USA)". In: *Hydrogeology Journal* 28.4 (Apr. 2020), pp. 1323–1341. DOI: 10.1007/s10040-020-02146-6.

Buffington, J. M. and D. Tonina (2009). "Hyporheic Exchange in Mountain Rivers II: Effects of Channel Morphology on Mechanics, Scales, and Rates of Exchange".

In: *Geography Compass* 3.3 (Mar. 2009), pp. 1038–1062. DOI: 10.1111/j.1749-8198.2009.00225.x.

Buss, S. et al. (2009). *The Hyporheic Handbook: A Handbook on the Groundwater-Surface Water Interface and Hyporheic Zone for Environment Managers*. Bristol, UK: Environment Agency. ISBN: 978-1-84911-131-7.

Caissie, D. and C. H. Luce (2017). "Quantifying streambed advection and conduction heat fluxes". In: *Water Resources Research* 53.2 (Feb. 2017), pp. 1595–1624. DOI: 10.1002/2016wr019813.

Carslaw, H. S. and J. C. Jaeger (1959). *Conduction of heat in solids*. 2nd. Oxford, UK: Oxford Clarendon Press.

Chellaï, O. et al. (2018). "Millimeter-Wave Beam Scattering by Field-Aligned Blobs in Simple Magnetized Toroidal Plasmas". In: *Physical Review Letters* 120.10 (Mar. 8, 2018), p. 105001. DOI: 10.1103/physrevlett.120.105001.

Chellaï, O. et al. (2021). "Millimeter-wave beam scattering and induced broadening by plasma turbulence in the TCV tokamak". In: *Nuclear Fusion* 61.6 (Apr. 29, 2021), p. 066011. DOI: 10.1088/1741-4326/abf43f.

Chen, K., H. Zhan, and Q. Wang (2018). "An innovative solution of diurnal heat transport in streambeds with arbitrary initial condition and implications to the estimation of water flux and thermal diffusivity under transient condition". In: *Journal of Hydrology* 567 (Dec. 2018), pp. 361–369. DOI: 10.1016/j.jhydrol.2018.10.008.

Citrin, J. et al. (2017). "Tractable flux-driven temperature, density, and rotation profile evolution with the quasilinear gyrokinetic transport model QuaLiKiz". In: *Plasma Physics and Controlled Fusion* 59.12 (Nov. 2017), p. 124005. DOI: 10.1088/1361-6587/aa8aeb.

Clark, M. P. et al. (2015). "A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies". In: *Water Resources Research* 51.4 (Apr. 2015), pp. 2515–2542. DOI: 10.1002/2015WR017200.

Constantz, J. (2008). "Heat as a tracer to determine streambed water exchanges". In: *Water Resources Research* 44.4 (Apr. 2008). DOI: 10.1029/2008wr006996.

Cranswick, R. H., P. G. Cook, M. Shanafield, and S. Lamontagne (2014). "The vertical variability of hyporheic fluxes inferred from riverbed temperature data". In: *Water Resources Research* 50.5 (May 2014), pp. 3994–4010. DOI: 10.1002/2013wr014410.

Curtain, R. F. and H. J. Zwart (1995). *An Introduction to Infinite-Dimensional Linear Systems Theory*. Vol. 50. 5. New York, NY, USA: Springer-Verlag, May 1995. ISBN: 978-0-387-94475-3. DOI: 10.1007/978-1-4612-4224-6.

Cuthbert, M. O. and R. Mackay (2013). "Impacts of nonuniform flow on estimates of vertical streambed flux". In: *Water Resources Research* 49.1 (Jan. 10, 2013), pp. 19–28. DOI: 10.1029/2011wr011587.

Das, A., S. Weiland, and M. van Berkel (2019). "Frequency Domain Estimation of Spatially Varying Parameters in Heat and Mass Transport". In: *2019 American Control Conference (ACC)*. IEEE, July 2019. DOI: 10.23919/acc.2019.8814465.

Deuflhard, P., A. Schiela, and M. Weiser (2012). "Mathematical cancer therapy planning in deep regional hyperthermia". In: *Acta Numerica* 21 (Apr. 2012), pp. 307–378. DOI: 10.1017/s0962492912000049.

DeWeese, T., D. Tonina, and C. H. Luce (2017). "Monitoring Streambed Scour/Deposition Under Nonideal Temperature Signal and Flood Conditions". In: *Water Resources Research* 53.12 (Dec. 2017), pp. 10257–10273. DOI: 10.1002/2017wr020632.

Domenico, P. A. and F. W. Schwartz (1997). *Physical and Chemical Hydrogeology*. 2nd. Wiley, p. 528. ISBN: 978-0-471-59762-9.

Eason, E. D. (1976). "A review of least-squares methods for solving partial differential equations". In: *International Journal for Numerical Methods in Engineering* 10.5, pp. 1021–1046. DOI: 10.1002/nme.1620100505.

Ebrahimian, M., R. Pourgholi, M. Emamjome, and P. Reihani (2007). "A numerical solution of an inverse parabolic problem with unknown boundary conditions". In: *Applied Mathematics and Computation* 189.1 (June 2007), pp. 228–234. DOI: 10.1016/j.amc.2006.11.062.

Escande, D. F. and F. Sattin (2012). "Calculation of Transport Coefficient Profiles in Modulation Experiments as an Inverse Problem". In: *Physical Review Letters* 108.12 (Mar. 2012), p. 125007. DOI: 10.1103/physrevlett.108.125007.

Ewing, R. E. and T. Lin (1988). "A direct method for parameter estimation in a hyperbolic partial differential equation". In: *Proceedings of the 27th IEEE Conference on Decision and Control*. Vol. 2. IEEE, pp. 1662–1667. DOI: 10.1109/cdc.1988.194611.

Fanelli, R. M. and L. K. Lautz (2008). "Patterns of Water, Heat, and Solute Flux through Streambeds around Small Dams". In: *Ground Water* 46.5 (Sept. 2008), pp. 671–687. DOI: 10.1111/j.1745-6584.2008.00461.x.

Fenstermacher, M. E. et al. (2022). "DIII-D research advancing the physics basis for optimizing the tokamak approach to fusion energy". In: *Nuclear Fusion* 62.4 (Apr. 2022), p. 042024. DOI: 10.1088/1741-4326/ac2ff2.

Fletcher, R. (1980). *Practical Methods of Optimization. Volume 1: Unconstrained Optimization*. New York, NY, USA: John Wiley and Sons Ltd, May 1980. ISBN: 9780471277118.

Folegot, S., S. Krause, R. Mons, D. M. Hannah, and T. Datry (2018). "Mesocosm experiments reveal the direction of groundwater-surface water exchange alters the hyporheic refuge capacity under warming scenarios". In: *Freshwater Biology* 63.2 (Nov. 2018), pp. 165–177. DOI: 10.1111/fwb.13049.

Ford, R. G., B. K. Lien, S. D. Acree, and R. R. Ross (2021). "Spreadsheet Tools for Quantifying Seepage Flux Across the GW-SW Interface". In: *Water Resources Research* 57.1 (Jan. 2021). DOI: 10.1029/2019wr026232.

Franklin, G. F., J. D. Powell, and A. Emami-Naeini (2015). *Feedback Control of Dynamic Systems*. 7th. Boston, MA, USA: Pearson. ISBN: 9780133496598.

Garbet, X. et al. (2004). "Physics of transport in tokamaks". In: *Plasma Physics and Controlled Fusion* 46.12B (Nov. 2004), B557–B574. DOI: 10.1088/0741-3335/46/12B/045.

Gariglio, F. P., D. Tonina, and C. H. Luce (2013). "Spatiotemporal variability of hyporheic exchange through a pool-riffle-pool sequence". In: *Water Resources Research* 49.11 (Nov. 2013), pp. 7185–7204. DOI: 10.1002/wrcr.20419.

Gentle, K. W. (1988). "Dependence of heat pulse propagation on transport mechanisms: Consequences of nonconstant transport coefficients". In: *Physics of Fluids* 31.5, p. 1105. DOI: 10.1063/1.866790.

Gevers, M. (2005). "Identification for Control: From the Early Achievements to the Revival of Experiment Design". In: *European Journal of Control* 11.4-5 (Jan. 2005), pp. 335–352. DOI: 10.3166/ejc.11.335-352.

Ghysels, G. et al. (2021). "The Significance of Vertical and Lateral GroundwaterSurface Water Exchange Fluxes in Riverbeds and Riverbanks: Comparing 1D Analytical Flux Estimates with 3D Groundwater Modelling". In: *Water* 13.3 (Jan. 27, 2021). Ed. by C. Allan, p. 306. DOI: 10.3390/w13030306.

Gils, R. W. van, M. F. M. Speetjens, H. J. Zwart, and H. Nijmeijer (2014). "Output-based modal control of three-dimensional pool-boiling systems". In: *International Journal of Thermal Sciences* 82 (Aug. 2014), pp. 34–46. DOI: 10.1016/j.ijthermalsci.2014.03.012.

Glose, T. J., C. S. Lowry, and M. B. Hausner (2019). "Limits on Groundwater-Surface Water Fluxes Derived from Temperature Time Series: Defining Resolution-Based Thresholds". In: *Water Resources Research* 55.12 (Dec. 2019), pp. 10678–10689. DOI: 10.1029/2019wr025643.

Glose, T. J., C. S. Lowry, and M. B. Hausner (2021). "Examining the utility of continuously quantified Darcy fluxes through the use of periodic temperature time series". In: *Journal of Hydrology* 595 (Apr. 2021), p. 125675. DOI: 10.1016/j.jhydrol.2020.125675.

Goodman, N. R. (1963). "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)". In: *The Annals of Mathematical Statistics* 34.1 (Mar. 1963), pp. 152–177. DOI: 10.1214/aoms/1177704250.

Goto, S., M. Yamano, and M. Kinoshita (2005). "Thermal response of sediment with vertical fluid flow to periodic temperature variation at the surface". In: *Journal of Geophysical Research: Solid Earth* 110.B1 (Jan. 29, 2005). DOI: 10.1029/2004jb003419.

Graf, A. et al. (2014). "Spatiotemporal relations between water budget components and soil water content in a forested tributary catchment". In: *Water Resources Research* 50.6 (June 2014), pp. 4837–4857. DOI: 10.1002/2013WR014516.

Greenwald, M. (2002). "Density limits in toroidal plasmas". In: *Plasma Physics and Controlled Fusion* 44.8 (July 22, 2002), R27–R53. DOI: 10.1088/0741-3335/44/8/201.

Guillaume, P. and R. Pintelon (1996). "A Gauss-Newton-like optimization algorithm for "weighted" nonlinear least-squares problems". In: *IEEE Transactions on Signal Processing* 44.9, pp. 2222–2228. DOI: 10.1109/78.536679.

Haas, J. de, J. O'Rourke, A. Sips, and N. Lopes Cardozo (1991). "Interpretation of heat and density pulse measurements in JET in terms of coupled transport".

In: *Nuclear Fusion* 31.7 (July 1991), pp. 1261–1274. DOI: 10.1088/0029-5515/31/7/003.

Hallemans, N., R. Pintelon, D. Peumans, and J. Lataire (2021). "Improved frequency response function estimation by Gaussian process regression with prior knowledge". In: *IFAC-PapersOnLine* 54.7, pp. 559–564. DOI: 10.1016/j.ifacol.2021.08.419.

Harvey, J. and M. Gooseff (2015). "River corridor science: Hydrologic exchange and ecological consequences from bedforms to basins". In: *Water Resources Research* 51.9 (Sept. 2015), pp. 6893–6922. DOI: 10.1002/2015WR017617.

Hatch, C. E., A. T. Fisher, J. S. Revenaugh, J. Constantz, and C. Ruehl (2006). "Quantifying surface water-groundwater interactions using time series analysis of streambed thermal records: Method development". In: *Water Resources Research* 42.10 (Oct. 11, 2006). DOI: 10.1029/2005wr004787.

Hatch, C. E., A. T. Fisher, C. R. Ruehl, and G. Stemler (2010). "Spatial and temporal variations in streambed hydraulic conductivity quantified with time-series thermal methods". In: *Journal of Hydrology* 389.3-4 (Aug. 2010), pp. 276–288. DOI: 10.1016/j.jhydrol.2010.05.046.

Havinga, L. and H. Schellen (2019). "The impact of convective vapour transport on the hygrothermal risk of the internal insulation of post-war lightweight prefab housing". In: *Energy and Buildings* 204 (Dec. 2019), p. 109418. DOI: 10.1016/j.enbuild.2019.109418.

Hespanha, J. P. (2018). *Linear Systems Theory Second Edition. Second Edition*. 2nd. Princeton, NJ, USA: Princeton University Press, Dec. 2018. ISBN: 9781400890088. DOI: 10.23943/9781400890088.

Ho, A. et al. (2019). "Application of Gaussian process regression to plasma turbulent transport model validation via integrated modelling". In: *Nuclear Fusion* 59.5 (Mar. 2019), p. 056007. DOI: 10.1088/1741-4326/ab065a.

Hogeweij, G. M. D., N. J. Lopes Cardozo, M. R. de Baar, and A. M. R. Schilham (1998). "A model for electron transport barriers in tokamaks, tested against experimental data from RTP". In: *Nuclear Fusion* 38.12 (Dec. 1998), pp. 1881–1891. DOI: 10.1088/0029-5515/38/12/312.

Hutchinson, I. H. (2002). *Principles of Plasma Diagnostics*. Cambridge University Press, July 2002. DOI: 10.1017/cbo9780511613630.

Ida, K. (1998). "Experimental studies of the physical mechanism determining the radial electric field and its radial structure in a toroidal plasma". In: *Plasma Physics and Controlled Fusion* 40.8 (Aug. 1998), pp. 1429–1488. DOI: 10.1088/0741-3335/40/8/002.

Incropera, F. P., D. P. Dewitt, T. L. Bergman, and A. S. Lavine (2006). *Fundamentals of Heat and Mass Transfer*. 6th. Hoboken, NJ, USA: Wiley. ISBN: 9780471457282.

Irsyad, M., A. D. Pasek, Y. S. Indartono, and A. W. Pratomo (2017). "Heat transfer characteristics of building walls using phase change material". In: *IOP Conference Series: Earth and Environmental Science*. Vol. 60. IOP Publishing, Mar. 2017, p. 012028. DOI: 10.1088/1755-1315/60/1/012028.

Irvine, D. J., B. L. Kurylyk, and M. A. Briggs (2020). "Quantitative guidance for efficient vertical flow measurements at the sedimentwater interface using temperaturedepth profiles". In: *Hydrological Processes* 34.3 (Nov. 2020), pp. 649–661. DOI: 10.1002/hyp.13614.

Irvine, D. J., L. K. Lautz, M. A. Briggs, R. P. Gordon, and J. M. McKenzie (2015). "Experimental evaluation of the applicability of phase, amplitude, and combined methods to determine water flux and thermal diffusivity from temperature time series using VFLUX 2". In: *Journal of Hydrology* 531 (Dec. 2015), pp. 728–737. DOI: 10.1016/j.jhydrol.2015.10.054.

Irvine, D. J. et al. (2017). "Using Diurnal Temperature Signals to Infer Vertical Groundwater-Surface Water Exchange". In: *Groundwater* 55.1 (Oct. 2017), pp. 10–26. DOI: 10.1111/gwat.12459.

Ito, K. and K. Kunisch (2008). *Lagrange Multiplier Approach to Variational Problems and Applications*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, Jan. 2008. DOI: 10.1137/1.9780898718614.

Jensen, J. K. and P. Engesgaard (2011). "Nonuniform Groundwater Discharge across a Streambed: Heat as a Tracer". In: *Vadose Zone Journal* 10.1 (Feb. 2011), pp. 98–109. DOI: 10.2136/vzj2010.0005.

Jidling, C., N. Wahlström, A. Wills, and T. B. Schön (2017). "Linearly constrained Gaussian processes". In: *31st Conference on Neural Information Processing Systems*.

Kalbus, E., F. Reinstorf, and M. Schirmer (2006). "Measuring methods for groundwater surface water interactions: a review". In: *Hydrology and Earth System Sciences* 10.6 (Nov. 2006), pp. 873–887. DOI: 10.5194/hess-10-873-2006.

Kampen, R. J. R. van, M. van Berkel, and H. J. Zwart (2022a). "Estimating Space-Dependent Coefficients for 1D transport using Gaussian Processes as State Estimator in the Frequency Domain". In: *61st IEEE Conference on Decision and Control (CDC)*. Dec. 2022, p. 1717.

Kampen, R. J. R. van, M. van Berkel, and H. J. Zwart (2023a). "Estimating Space-Dependent Coefficients for 1D transport using Gaussian Processes as State Estimator in the Frequency Domain". In: *IEEE Control Systems Letters* 7 (2023), pp. 247–252. DOI: 10.1109/LCSYS.2022.3186626.

Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2020). "Complex Gaussian Process Regression for Estimating Spatially Varying Coefficients in Thermal Transport". In: *39th Benelux Meeting on Systems and Control*. Elspeet, Netherlands, Mar. 2020, p. 38.

Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021a). "A closed-form solution to estimate space-dependent parameters in heat and mass transport". In: *Physics@Veldhoven*. Veldhoven, Netherlands, Jan. 2021.

Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021b). "A Closed-Form Solution to Estimate Spatially Varying Parameters in Heat and Mass Transport". In: *2021 American Control Conference (ACC)*. May 2021, pp. 286–291. DOI: 10.23919/ACC50511.2021.9482757.

Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021c). "A Closed-Form Solution to Estimate Spatially Varying Parameters in Heat and Mass Transport". In: *IEEE Control Systems Letters* 5.5 (Nov. 2021), pp. 1681–1686. DOI: 10.1109/lcsys.2020.3042933.

Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021d). "A Linear Least Squares Approach to Estimate Space-Dependent Parameters in Heat and Mass Transport". In: *Benelux Workshop on Systems and Control*. Rotterdam, Netherlands, June 2021, p. 68.

Kampen, R. J. R. van, J. de Vries, S. Weiland, M. R. de Baar, and M. van Berkel (2023b). "Fast simultaneous estimation of nD transport coefficients and source function in perturbation experiments". In: *Scientific Reports* 13.1 (Feb. 2023), p. 3241. DOI: 10.1038/s41598-023-30337-0.

Kampen, R. J. R. van, S. Weiland, H. J. Zwart, and M. van Berkel (2022b). "A frequency domain maximum likelihood Approach to Estimate Space-Dependent Parameters in heat and mass transport". In: *41st Benelux Meeting on Systems and Control*. Brussels, Belgium, July 2022, p. 200.

Kampen, R. J. R. van et al. (2021e). "On the Validity of Flux Estimates using Semi-Infinite Domains.Comparing Flux Estimates from Semi-Infinite and Bounded Domains using the LPMLEn - a Multi Frequency, Multi Sensor Method to Estimate Vertical Streambed Fluxes and Sediment Thermal Properties". In: *AGU Fall Meeting 2021*. Hybrid, New Orleans, LA, USA, Dec. 2021.

Kampen, R. J. R. van et al. (2022c). *LPMLEn - A code for estimating heat transport parameters in 1D*. online code and dataset. 2022. DOI: 10.4211/hs.3b13760174174c31988120baeb84e2e8.

Kampen, R. J. R. van et al. (2022d). "LPMLEn A Frequency Domain Method to Estimate Vertical Streambed Fluxes and Sediment Thermal Properties in Semi-Infinite and Bounded Domains". In: *Water Resources Research* 58.3 (Feb. 2022), e2021WR030886. DOI: 10.1029/2021wr030886.

Karan, S., P. Engesgaard, and J. Rasmussen (2014). "Dynamic streambed fluxes during rainfall-runoff events". In: *Water Resources Research* 50.3 (Mar. 2014), pp. 2293–2311. DOI: 10.1002/2013wr014155.

Keery, J., A. Binley, N. Crook, and J. W. Smith (2007). "Temporal and spatial variability of groundwatersurface water fluxes: Development and application of an analytical method using temperature time series". In: *Journal of Hydrology* 336.1-2 (Mar. 2007), pp. 1–16. DOI: 10.1016/j.jhydrol.2006.12.003.

Keesman, K. J. (2011). *System Identification. An Introduction*. London, UK: Springer-Verlag London Limited. ISBN: 978-0-85729-522-4. DOI: 10.1007/978-0-85729-522-4.

Khalil, H. K. (2002). *Nonlinear systems*. Upper Saddle River, NJ, USA: Prentice Hall, p. 750. ISBN: 0-13-067389-7.

Kirov, K. K. et al. (2002). "ECRH power deposition studies in ASDEX Upgrade". In: *Plasma Physics and Controlled Fusion* 44.12 (Nov. 22, 2002), pp. 2583–2602. DOI: 10.1088/0741-3335/44/12/307.

Kirsch, A. (2011). *An Introduction to the Mathematical Theory of Inverse Problems*. 2nd. New York, NY, USA: Springer New York. ISBN: 978-1-4419-8473-9. DOI: 10.1007/978-1-4419-8474-6.

Koch, F. W. et al. (2016). "1DTempPro V2: New Features for Inferring Groundwater/Surface-Water Exchange". In: *Groundwater* 54.3 (Sept. 2016), pp. 434–439. DOI: 10.1111/gwat.12369.

Köchl, F. (2009). "Modelling of Transport and Pellet Fuelling in ITER-relevant Plasmas". PhD thesis. Technische Universität Wien, Oct. 2009.

Krause, S., D. M. Hannah, and T. Blume (2011). "Interstitial pore-water temperature dynamics across a pool-riffle-pool sequence". In: *Ecohydrology* 4.4 (Feb. 2011), pp. 549–563. DOI: 10.1002/Eco.199.

Krause, S. et al. (2017). "Ecohydrological interfaces as hot spots of ecosystem processes". In: *Water Resources Research* 53.8 (Aug. 2017), pp. 6359–6376. DOI: 10.1002/2016WR019516.

Kravaris, C. and J. H. Seinfeld (1985). "Identification of Parameters in Distributed Parameter Systems by Regularization". In: *SIAM Journal on Control and Optimization* 23.2 (Mar. 1985), pp. 217–241. DOI: 10.1137/0323017.

Krieger, K., G. Fussman, and the ASDEX Team (1990). "Determination of impurity transport coefficients by harmonic analysis". In: *Nuclear Fusion* 30.11 (Nov. 1990), pp. 2392–2396. DOI: 10.1088/0029-5515/30/11/015.

Kunisch, K. and G. Peichl (1991). "Estimation of a temporally and spatially varying diffusion coefficient in a parabolic system by an augmented Lagrangian technique". In: *Numerische Mathematik* 59.1 (Dec. 1991), pp. 473–509. DOI: 10.1007/bf01385792.

Kuzmin, D., R. Löhner, and S. Turek (2012). *Flux-Corrected Transport*. 2 nd. Springer Netherlands. ISBN: 978-94-007-4037-2. DOI: 10.1007/978-94-007-4038-9.

Lapham, W. W. (1989). *Use of temperature profiles beneath streams to determine rates of vertical ground-water flow and vertical hydraulic conductivity*. Tech. rep. Department of the Interior, U.S. Geological Survey. DOI: 10.3133/wsp2337.

Lataire, J. and R. Pintelon (2011). "Frequency-domain weighted non-linear least-squares estimation of continuous-time, time-varying systems". In: *IET Control Theory & Applications* 5.7 (May 2011), pp. 923–933. DOI: 10.1049/iet-cta.2010.0223.

Lataire, J. and T. Chen (2016). "Transfer function and transient estimation by Gaussian process regression in the frequency domain". In: *Automatica* 72 (Oct. 2016), pp. 217–229. DOI: 10.1016/j.automatica.2016.06.009.

Lautz, L. K. (2010). "Impacts of nonideal field conditions on vertical water velocity estimates from streambed temperature time series". In: *Water Resources Research* 46.1 (Jan. 2010). DOI: 10.1029/2009WR007917.

Lautz, L. K. (2012). "Observing temporal patterns of vertical flux through streambed sediments using time-series analysis of temperature records". In: *Journal of Hydrology* 464-465 (Sept. 2012), pp. 199–215. DOI: 10.1016/j.jhydrol.2012.07.006.

Leibundgut, C., P. Maloszewski, and C. Ülls (2009). *Tracers in Hydrology*. Wiley, Sept. 2009. DOI: 10.1002/9780470747148.

Levenberg, K. (1944). "A method for the solution of certain non-linear problems in least squares". In: *Quarterly of Applied Mathematics* 2.2, pp. 164–168. DOI: 10.1090/qam/10666.

Liao, W., M. Dehghan, and A. Mohebbi (2009). "Direct numerical method for an inverse problem of a parabolic partial differential equation". In: *Journal of Computational and Applied Mathematics* 232.2 (Oct. 2009), pp. 351–360. DOI: 10.1016/j.cam.2009.06.017.

Ljung, L. (1999). *System Identification: theory for the user. Theory for the User (2nd Edition)*. 2nd. Upper Saddle River, NJ, USA: Prentice Hall PTR. ISBN: 9780136566953.

Lopes Cardozo, N. J. (1995). "Perturbative transport studies in fusion plasmas". In: *Plasma Physics and Controlled Fusion* 37.8 (Aug. 1995), pp. 799–852. DOI: 10.1088/0741-3335/37/8/001.

Luce, C. H., D. Tonina, R. Applebee, and T. DeWeese (2017). "Was That Assumption Necessary? Reconsidering Boundary Conditions for Analytical Solutions to Estimate Streambed Fluxes". In: *Water Resources Research* 53.11 (Nov. 2017), pp. 9771–9790. DOI: 10.1002/2017wr020618.

Luce, C. H., D. Tonina, F. Gariglio, and R. Applebee (2013). "Solutions for the diurnally forced advection-diffusion equation to estimate bulk fluid velocity and diffusivity in streambeds from temperature time series". In: *Water Resources Research* 49.1 (Jan. 2013), pp. 488–506. DOI: 10.1029/2012WR012380.

Magin, R. L. (2004). "Fractional Calculus in Bioengineering, Part 1". In: *Critical Reviews in Biomedical Engineering* 32.1, pp. 1–104. DOI: 10.1615/critrevbiomedeng.v32.10.

Mantica, P. et al. (2008). "Chapter 10: Core Transport Studies in JET". In: *Fusion Science and Technology* 53.4 (May 2008), pp. 1152–1216. DOI: 10.13182/FST08-A1750.

Markovsky, I. and S. V. Huffel (2007). "Overview of total least-squares methods". In: *Signal Processing* 87.10 (Oct. 2007), pp. 2283–2302. DOI: 10.1016/j.sigpro.2007.04.004.

Marquardt, D. W. (1963). "An Algorithm for Least-Squares Estimation of Nonlinear Parameters". In: *Journal of the Society for Industrial and Applied Mathematics* 11.2 (June 1963), pp. 431–441. DOI: 10.1137/0111030.

Marsily, G. de (1986). *Quantitative hydrogeology. groundwater hydrology for engineers*. Orlando, FL, USA: Academic Press, p. 440. ISBN: 9780122089152.

Mauri, R. (2015). *Transport Phenomena in Multiphase Flows*. Fluid Mechanics and Its Applications. Springer International Publishing. ISBN: 978-3-319-15792-4. DOI: 10.1007/978-3-319-15793-1.

McAliley, W. A. et al. (2022). "Application of Recursive Estimation to Heat Tracing for Groundwater/Surface-Water Exchange". In: *Water Resources Research* 58.6 (June 2022). DOI: 10.1029/2021WR030443.

McKelvey, T. and G. Guérin (2012). "Non-parametric frequency response estimation using a local rational model". In: *IFAC Proceedings Volumes* 45.16 (July 2012), pp. 49–54. DOI: 10.3182/20120711-3-be-2027.00299.

Mechhoud, S., E. Witrant, L. Dugard, and D. Moreau (2015). "Estimation of Heat Source Term and Thermal Diffusion in Tokamak Plasmas Using a Kalman Filtering Method in the Early Lumping Approach". In: *IEEE Transactions on Control Systems Technology* 23.2 (Mar. 2015), pp. 449–463. DOI: 10.1109/TCST.2014.2342760.

Mochi, M., G. Pacelli, M. C. Recchioni, and F. Zirilli (1999). "Inverse Problem for a Class of Two-Dimensional Diffusion Equations with Piecewise Constant Coefficients". In: *Journal of Optimization Theory and Applications* 100.1 (Jan. 1999), pp. 29–57. DOI: 10.1023/a:1021712830465.

Monteyne, G., G. Vandersteen, R. Pintelon, and D. Ugryumova (2013). "Transient suppression in FRF measurement: Comparison and analysis of four state-of-the-art methods". In: *Measurement* 46.7 (Aug. 2013), pp. 2210–2222. DOI: 10.1016/j.measurement.2013.03.032.

Moret, J.-M., T. Dudok de Wit, B. Joye, and J. Lister (1993). "Investigation of plasma transport processes using the dynamical response of soft X-ray emission". In: *Nuclear Fusion* 33.8 (Aug. 1993), pp. 1185–1200. DOI: 10.1088/0029-5515/33/8/i07.

Munz, M. and C. Schmidt (2017). "Estimation of vertical water fluxes from temperature time series by the inverse numerical computer program FLUX-BOT". In: *Hydrological Processes* 31.15 (June 2017), pp. 2713–2724. DOI: 10.1002/hyp.11198.

Nocedal, J. and S. J. Wright (1999). *Numerical Optimization*. New York, NY, USA: Springer-Verlag. ISBN: 978-0-387-98793-4. DOI: 10.1007/b98874.

Nuij, P., M. Steinbuch, and O. Bosgra (2008). "Measuring the higher order sinusoidal input describing functions of a non-linear plant operating in feedback". In: *Control Engineering Practice* 16.1 (Jan. 2008), pp. 101–113. DOI: 10.1016/j.conengprac.2007.04.003.

Onderka, M., S. Banzhaf, T. Scheytt, and A. Krein (2013). "Seepage velocities derived from thermal records using wavelet analysis". In: *Journal of Hydrology* 479 (Feb. 2013), pp. 64–74. DOI: 10.1016/j.jhydrol.2012.11.022.

Oppenheim, A. V., A. S. Willsky, and S. H. Nawab. (1997). *Signals and Systems*. 2nd. Upper Saddle River, NJ, USA: Prentice Hall, p. 957. ISBN: 9780136511755.

Paul, E. L., V. Atiemo-Obeng, S. M. Kresta, and North American Mixing Forum (2003). *Handbook of Industrial Mixing. Science and Practice*. Wiley-Interscience. ISBN: 9780471269199.

Pintelon, R., J. Schoukens, G. Vandersteen, and K. Barbé (2010a). "Estimation of nonparametric noise and FRF models for multivariable systemsPart I: Theory". In: *Mechanical Systems and Signal Processing* 24.3 (Apr. 2010), pp. 573–595. DOI: 10.1016/j.ymssp.2009.08.009.

Pintelon, R., J. Schoukens, G. Vandersteen, and K. Barbé (2010b). "Estimation of nonparametric noise and FRF models for multivariable systemsPart II:

Extensions, applications". In: *Mechanical Systems and Signal Processing* 24.3 (Apr. 2010), pp. 596–616. DOI: 10.1016/j.ymssp.2009.08.010.

Pintelon, R. and J. Schoukens (2012). *System Identification*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Mar. 2012. DOI: 10.1002/9781118287422.

Quarteroni, A. and A. Valli (1994). *Numerical Approximation of Partial Differential Equations*. Hoboken, NJ, USA: Springer Berlin Heidelberg, Mar. 1994. DOI: 10.1007/978-3-540-85268-1.

Rai, P. K. and S. Tripathi (2019). "Gaussian process for estimating parameters of partial differential equations and its application to the Richards equation". In: *Stochastic Environmental Research and Risk Assessment* 33.8-9 (July 2019), pp. 1629–1649. DOI: 10.1007/s00477-019-01709-8.

Raissi, M., P. Perdikaris, and G. E. Karniadakis (2017). "Machine learning of linear differential equations using Gaussian processes". In: *Journal of Computational Physics* 348 (Nov. 2017), pp. 683–693. DOI: 10.1016/j.jcp.2017.07.050.

Raissi, M., A. Yazdani, and G. E. Karniadakis (2020). "Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations". In: *Science* 367.6481 (Feb. 2020), pp. 1026–1030. DOI: 10.1126/science.aaw4741.

Rasmussen, C. E. and C. K. I. Williams (2005). *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: The MIT Press. DOI: 10.7551/mitpress/3206.001.0001.

Rau, G. C., M. S. Andersen, and R. I. Acworth (2012). "Experimental investigation of the thermal dispersivity term and its significance in the heat transport equation for flow in sediments". In: *Water Resources Research* 48.3 (Mar. 2012). DOI: 10.1029/2011WR011038.

Rau, G. C., M. S. Andersen, A. M. McCallum, and R. I. Acworth (2010). "Analytical methods that use natural heat as a tracer to quantify surface watergroundwater exchange, evaluated using field temperature records". In: *Hydrogeology Journal* 18.5 (Mar. 2010), pp. 1093–1110. DOI: 10.1007/s10040-010-0586-0.

Rau, G. C., M. S. Andersen, A. M. McCallum, H. Roshan, and R. I. Acworth (2014). "Heat as a tracer to quantify water flow in near-surface sediments". In: *Earth-Science Reviews* 129 (Feb. 2014), pp. 40–58. DOI: 10.1016/j.earscirev.2013.10.015.

Ravensbergen, T. et al. (2021). "Real-time feedback control of the impurity emission front in tokamak divertor plasmas". In: *Nature Communications* 12.1 (Feb. 2021), p. 1105. DOI: 10.1038/s41467-021-21268-3.

Reeves, J. and C. E. Hatch (2016). "Impacts of three-dimensional nonuniform flow on quantification of groundwater-surface water interactions using heat as a tracer". In: *Water Resources Research* 52.9 (Sept. 2016), pp. 6851–6866. DOI: 10.1002/2016wr018841.

Reimerdes, H. et al. (2022). "Overview of the TCV tokamak experimental programme". In: *Nuclear Fusion* 62.4 (Mar. 2022), p. 042018. DOI: 10.1088/1741-4326/ac369b.

Ren, T., G. J. Kluitenberg, and R. Horton (2000). "Determining Soil Water Flux and Pore Water Velocity by a Heat Pulse Technique". In: *Soil Science Society*

of America Journal 64.2 (Mar. 2000), pp. 552–560. DOI: 10.2136/sssaj2000.642552x.

Rijlaarsdam, D., P. Nuij, J. Schoukens, and M. Steinbuch (2017). "A comparative overview of frequency domain methods for nonlinear systems". In: *Mechatronics* 42 (Apr. 2017), pp. 11–24. DOI: 10.1016/j.mechatronics.2016.12.008.

Roques, L., M. D. Chekroun, M. Cristofol, S. Soubeyrand, and M. Ghil (2014). "Parameter estimation for energy balance models with memory". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 470.2169 (Sept. 2014), p. 20140349. DOI: 10.1098/rspa.2014.0349.

Roshan, H., M. Cuthbert, M. Andersen, and R. Acworth (2014). "Local thermal non-equilibrium in sediments: Implications for temperature dynamics and the use of heat as a tracer". In: *Advances in Water Resources* 73 (Nov. 2014), pp. 176–184. DOI: 10.1016/j.advwatres.2014.08.002.

Roshan, H., G. C. Rau, M. S. Andersen, and I. R. Acworth (2012). "Use of heat as tracer to quantify vertical streambed flow in a two-dimensional flow field". In: *Water Resources Research* 48.10 (Oct. 2012). DOI: 10.1029/2012wr011918.

Rossi, L., B. Krishnamachari, and C.-C. Kuo (2004). "Distributed parameter estimation for monitoring diffusion phenomena using physical models". In: *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004.* IEEE. DOI: 10.1109/sahcn.2004.1381948.

Ryter, F., R. Dux, P. Mantica, and T. Tala (2010). "Perturbative studies of transport phenomena in fusion devices". In: *Plasma Physics and Controlled Fusion* 52.12 (Nov. 15, 2010), p. 124043. DOI: 10.1088/0741-3335/52/12/124043.

Ryter, F. et al. (2001). "Experimental Evidence for Gradient Length-Driven Electron Transport in Tokamaks". In: *Physical Review Letters* 86.11 (Mar. 2001), pp. 2325–2328. DOI: 10.1103/physrevlett.86.2325.

Sattin, F. et al. (2012). "Estimate of convectiondiffusion coefficients from modulated perturbative experiments as an inverse problem". In: *Plasma Physics and Controlled Fusion* 54.12 (Nov. 21, 2012), p. 124025. DOI: 10.1088/0741-3335/54/12/124025.

Schmidt, C., O. Büttner, A. Musolff, and J. H. Fleckenstein (2014). "A method for automated, daily, temperature-based vertical streambed water-fluxes". In: *Fundamental and Applied Limnology* 184.3 (June 2014), pp. 173–181. DOI: 10.1127/1863-9135/2014/0548.

Schneidewind, U. et al. (2016). "LPMLE3: A novel 1-D approach to study water flow in streambeds using heat as a tracer". In: *Water Resources Research* 52.8 (Aug. 2016), pp. 6596–6610. DOI: 10.1002/2015wr017453.

Schneidewind, U. (2016). "Water Flow and Contaminant Transformation in the Hyporheic Zone of Lowland Rivers". PhD thesis. Ghent University. ISBN: 978-90-5989-864-6.

Schneidewind, U. et al. (2021). "Quantifying vertical streambed fluxes around woody structures using high-resolution streambed temperature measurements".

In: *EGU General Assembly 2021*. online, Mar. 2021. DOI: 10.5194/egusphere-egu21-9228.

Schumer, R., D. A. Benson, M. M. Meerschaert, and B. Baeumer (2003a). "Fractal mobile/immobile solute transport". In: *Water Resources Research* 39.10 (Oct. 2003). DOI: 10.1029/2003WR002141.

Schumer, R., D. A. Benson, M. M. Meerschaert, and B. Baeumer (2003b). "Multiscaling fractional advection-dispersion equations and their solutions". In: *Water Resources Research* 39.1 (Jan. 2003). DOI: 10.1029/2001WR001229.

Schumer, R., D. A. Benson, M. M. Meerschaert, and S. W. Wheatcraft (2001). "Eulerian derivation of the fractional advectiondispersion equation". In: *Journal of Contaminant Hydrology* 48.1-2 (Mar. 2001), pp. 69–88. DOI: 10.1016/S0169-7722(00)00170-4.

Schweizer, D., V. Ried, G. C. Rau, J. E. Tuck, and P. Stoica (2021). "Comparing Methods and Defining Practical Requirements for Extracting Harmonic Tidal Components from Groundwater Level Measurements". In: *Mathematical Geosciences* 53.6 (Feb. 2021), pp. 1147–1169. DOI: 10.1007/s11004-020-09915-9.

Selker, J. S. et al. (2006). "Distributed fiber-optic temperature sensing for hydrologic systems". In: *Water Resources Research* 42.12 (Dec. 2006), pp. 6596–6610. DOI: 10.1029/2006WR005326.

Shanafield, M., G. Pohll, and R. Susfalk (2010). "Use of heat-based vertical fluxes to approximate total flux in simple channels". In: *Water Resources Research* 46.3 (Mar. 2010). DOI: 10.1029/2009WR007956.

Slief, J. H., R. J. R. van Kampen, M. W. Brookman, and M. van Berkel (2022). "Extension of the flux fit method for estimating power deposition profiles". In: *Physics of Plasmas* 29.1 (Jan. 20, 2022), p. 010703. DOI: 10.1063/5.0069869.

Slief, J. H. et al. (2023). "Quantifying electron cyclotron power deposition broadening in DIII-D and the potential consequences for the ITER EC system". In: *Nuclear Fusion* 63.2 (Jan. 2023), p. 026029. DOI: 10.1088/1741-4326/acaedc.

Söderström, T. and P. Stoica (1989). *System identification*. New York, NY, USA: Prentice Hall, p. 612. ISBN: 9780131276062.

Sohn, R. A. and R. N. Harris (2021). "Spectral Analysis of Vertical Temperature Profile Time-Series Data in Yellowstone Lake Sediments". In: *Water Resources Research* 57.4 (Apr. 2021). DOI: 10.1029/2020WR028430.

Stallman, R. W. (1965). "Steady one-dimensional fluid flow in a semi-infinite porous medium with sinusoidal surface temperature". In: *Journal of Geophysical Research* 70.12 (June 1965), pp. 2821–2827. DOI: 10.1029/JZ070i012p02821.

Stonestrom, D. A. and J. Constantz (2003). *Heat as a Tool for Studying the Movement of Ground Water Near Streams*. Tech. rep. Department of the Interior, U.S. Geological Survey.

Sundqvist, H. and G. Veronis (1970). "A simple finite-difference grid with non-constant intervals". In: *Tellus* 22.1 (Jan. 1970), pp. 26–31. DOI: 10.3402/tellusa.v22i1.10155.

Takenaga, H., K. Nagashima, A. Sakasai, T. Oikawa, and T. Fujita (1998). "Determination of particle transport coefficients in reversed shear plasma of JT-60U".

In: *Plasma Physics and Controlled Fusion* 40.2 (Feb. 1998), pp. 183–190. DOI: 10.1088/0741-3335/40/2/002.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, Jan. 2005. DOI: 10.1137/1.9780898717921.

Tartakovsky, A., C. Ortiz Marrero, P. Perdikaris, G. Tartakovsky, and D. Barajas-Solano (2018). "Learning Parameters and Constitutive Relationships with Physics Informed Deep Neural Networks". In: Aug. 21, 2018. eprint: arXiv:1808.03398v2[math.AP].

Thije Boonkkamp, J. H. M. ten and M. J. H. Anthonissen (2011). "The Finite Volume-Complete Flux Scheme for Advection-Diffusion-Reaction Equations". In: *Journal of Scientific Computing* 46.1 (June 2011), pp. 47–70. DOI: 10.1007/s10915-010-9388-8.

Vandersteen, G. et al. (2015). "Determining groundwater-surface water exchange from temperature-time series: Combining a local polynomial method with a maximum likelihood estimator". In: *Water Resources Research* 51.2 (Feb. 2015), pp. 922–939. DOI: 10.1002/2014WR015994.

Vereecken, H. et al. (2016). "On the role of patterns in understanding the functioning of soil-vegetation-atmosphere systems". In: *Journal of Hydrology* 542 (Nov. 2016), pp. 63–86. DOI: 10.1016/j.jhydrol.2016.08.053.

Vogel, C. R. (2002). *Computational Methods for Inverse Problems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, Jan. 2002, p. 199. ISBN: 9780898715507. DOI: 10.1137/1.9780898717570.

Vogt, T., P. Schneider, L. Hahn-Woernle, and O. A. Cirpka (2010). "Estimation of seepage rates in a losing stream by means of fiber-optic high-resolution vertical temperature profiling". In: *Journal of Hydrology* 380.1-2 (Jan. 2010), pp. 154–164. DOI: 10.1016/j.jhydrol.2009.10.033.

Vries, J. de (2021). "Closed-form solutions of inverse 2D-and coupled transport problems". MA thesis. Eindhoven University of Technology, May 12, 2021.

Wesson, J. (2004). *Tokamaks*. 3rd. Clarendon Press. ISBN: 0198509227.

Wörman, A. et al. (2012). "Spectral scaling of heat fluxes in streambed sediments". In: *Geophysical Research Letters* 39.23 (Dec. 2012), n/a–n/a. DOI: 10.1029/2012GL053922.

Wörman, A., A. I. Packman, L. Marklund, J. W. Harvey, and S. H. Stone (2006). "Exact three-dimensional spectral solution to surface-groundwater interactions with arbitrary surface topography". In: *Geophysical Research Letters* 33.7. DOI: 10.1029/2006gl025747.

Wu, L. et al. (2020). "Impact of Flow Alteration and Temperature Variability on Hyporheic Exchange". In: *Water Resources Research* 56.3 (Mar. 2020). DOI: 10.1029/2019wr026225.

Xun, X., J. Cao, B. Mallick, A. Maity, and R. J. Carroll (2013). "Parameter Estimation of Partial Differential Equation Models". In: *Journal of the American Statistical Association* 108.503 (Sept. 2013), pp. 1009–1020. DOI: 10.1080/01621459.2013.794730.

Young, P. C., D. J. Pedregal, and W. Tych (1999). "Dynamic harmonic regression". In: *Journal of Forecasting* 18.6 (Nov. 1999), pp. 369–394. DOI: 10.1002/(SICI)1099-131X(199911)18:6<369::AID-FOR748>3.0.CO;2-K.

Zhang, Y. et al. (2021). "Upscaling Heat Flow in Porous Media With Periodic Surface Temperature Fluctuation Using a One-Dimensional Subordinated Heat Transfer Equation". In: *Water Resources Research* 57.7 (July 2021). DOI: 10.1029/2020WR027266.

Zhao, Z.-X., M. K. Banda, and B.-Z. Guo (2016). "Simultaneous identification of diffusion coefficient, spacewise dependent source and initial value for one-dimensional heat equation". In: *Mathematical Methods in the Applied Sciences* 40.10 (Nov. 18, 2016), pp. 3552–3565. DOI: 10.1002/mma.4245.

Zhou, T. et al. (2017). "A New Approach to Quantify Shallow Water Hydrologic Exchanges in a Large Regulated River Reach". In: *Water* 9.9 (Sept. 2017), p. 703. DOI: 10.3390/w9090703.

Samenvatting

Schatten van ruimtelijk variërende transportcoëfficiënten in het frequentiedomein

Het kwantificeren van transportverschijnselen die de uitwisseling van energie, lading, massa, of momentum beschrijven, speelt een belangrijke rol in de natuurwetenschappen en techniek. Het gegeneraliseerde wiskundig model dat deze transportverschijnselen beschrijft wordt het generieke scalaire transportmodel genoemd en behoort tot de klasse van parabolische partiële differentiaalvergelijkingen (PDVs). Hoewel transportmodellen vaak zijn afgeleid vanuit de bekende onderliggende fysica, zijn de bijbehorende fysische parameters die de transportmechanismen beschrijven onzeker of onbekend. Dit resulteert in een schattingsprobleem waarvoor er in de literatuur verschillende methoden zijn ontwikkeld.

Schattingsmethoden variëren, afhankelijk van het toepassingsgebied, van het schatten van constante coëfficiënten met behulp van analytische oplossingen van de PDV tot het schatten van ruimtelijk variërende coëfficiëntfuncties met behulp van numerieke PDV-benaderingstechnieken. Dit proefschrift gaat verder op beide strategieën door verschillende frequentiedomeinmethodieken uit te breiden en te vergelijken. De keuze voor het frequentiedomein is een bewuste keuze door het tal van voordelen ten opzichte van tijddomeinmethodieken. Een van de belangrijkste voordelen is de reductie van de PDV tot een gewone differentiaalvergelijking met complexe waarden. Daarnaast is de inhoud van het informatie dragende spectrum van de gebruikte verstoringsexperimenten voor PDVs doorgaans geconcentreerd, wat ten volste wordt benut.

Het eerste deel van dit proefschrift richt zich op het uitbreiden van de maximum likelihood (meest aannemelijke) schatter voor het transportmodel met één ruimtelijke dimensie (1D) en constante parameters. De uitbreiding maakt het mogelijk om het effect van de verschillende randvoorwaarden voor het transportmodel op de geschatte coëfficiënten te bestuderen. Om precies te zijn, dit proefschrift breidt de maximum likelihood schatter uit zodat zowel het lokaal begrensde domein als het semi-oneindige domein als randvoorwaarden gebruikt kunnen worden voor schattingen met meerdere met ruis vervuilde metingen. Daardoor kan de schatter

met beide randvoorwaarden hetzelfde aantal metingen gebruiken waardoor er een eerlijke vergelijking mogelijk is. Deze nieuwe methode is toegepast op zowel simulatie resultaten als experimentele gegevens uit de hydrologie. Dit laat duidelijk de verdiensten van het schatten met meerdere metingen op een lokaal begrensd domein ten opzichte van het semi-oneindige domein.

Het tweede en grootste deel van dit proefschrift richt zich op het ontwikkelen van methodologieën voor het schatten van ruimtelijk variërende coëfficiënten voor parabolische PDVs. De standaardbenadering is het minimaliseren van de uitgangsfout, d.w.z. het verschil tussen de metingen en de modelwaarde op de meetlocaties. Als de ruimtelijke afhankelijkheid van de coëfficiënten niet beperkt wordt, bevatten de geschatte coëfficiënten vaak grote (niet-fysieke) oscillerende artefacten als gevolg van ruisversterkingen. De klassieke aanpak van dit probleem is door middel van regularisatie, d.w.z. het beperken van de complexiteit van de functies die de onbekende coëfficiënten beschrijven. Dit proefschrift vermijdt deze (vaak) willekeurige regularisatie op de onbekende coëfficiënten door de vergelijkingsfout te minimaliseren, d.w.z. het verschil tussen de gemeten of geschatte toestand en de toestand van het model worden geminimaliseerd. De aanpak is als volgt: 1) er wordt een discretisatie toegepast op de ruimtelijke coördinaat; 2) een lineaire parametrisering wordt toegepast op de onbekende coëfficiënten. Het verschil tussen de vrijheidsgraden van deze parametrisering en het beschouwde aantal ruimtelijke punten weerspiegelt direct de mate van toegepaste regularisatie op het schattingsprobleem. Bovendien resulteert de lineaire parametrisering, in combinatie met de vergelijkingsfout, in een lineair kleinste-kwadratenprobleem met een bekende oplossing in een gesloten vorm. Dit maakt 3) directe berekening van het globale optimum mogelijk door een lineaire matrixvergelijking op te lossen, hetgeen dat een lage rekenkracht vereist. Dit is een belangrijke verbetering aangezien de schattingen van ruimtelijk variërende coëfficiënten normaal veel rekenkracht vereist vanwege de noodzakelijke iteratieve optimalisatiemethoden. Daarom maakt deze nieuwe methodologie het mogelijk om snel verschillende parametrisaties van coëfficiënten te schatten en te vergelijken op een manier die lijkt op machinaal leren. Bovendien maakt dit het mogelijk om snel ruimtelijke variërende coëfficiënten in meerdere ruimtelijke dimensies te schatten, iets waar normaliter veel rekenkracht voor nodig is.

Als verdere verbetering van de 1D-schattingsmethode wordt de casus beschouwd waarin alle metingen ruis bevatten. In dit geval bevat de gesloten-vorm oplossing een systematische fout. Daarom wordt de methodologie uitgebreid door eerst de waarschijnlijkheidsdichtheidsfunctie van de toestand te bepalen door middel van Gaussiaanse proces regressie en vervolgens met de maximum likelihood schatter de transportcoëfficiënten te bepalen.

De voorgestelde methodologieën zijn toegepast binnen twee verschillende domeinen. In de hydrologie voor het schatten van de vermenging van grondwater en oppervlaktewater. Hetgeen dat verantwoordelijk is voor de uitwisseling van verontreinigingen en zodoende de waterkwaliteit beïnvloed. Ten tweede, op het gebied van kernfusie voor het schatten van warmte en deeltjestransport die de

reactorefficiëntie bepalen.

Het is te verwachten dat deze methoden hun weg zullen vinden naar meer toepassingen en gebruikers zullen helpen hun transportmodellen te identificeren, en de aanleiding zijn voor een nieuw perspectief op het schatten transportcoëfficiënten.

List of publications

Peer-reviewed journal articles

- Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021c). "A Closed-Form Solution to Estimate Spatially Varying Parameters in Heat and Mass Transport". In: *IEEE Control Systems Letters* 5.5 (Nov. 2021), pp. 1681–1686. DOI: 10.1109/lcsys.2020.3042933.

- Kampen, R. J. R. van, U. Schneidewind, C. Anibas, et al. (2022d). "LPMLEn A Frequency Domain Method to Estimate Vertical Streambed Fluxes and Sediment Thermal Properties in Semi-Infinite and Bounded Domains". In: *Water Resources Research* 58.3 (Feb. 2022), e2021WR030886. DOI: 10.1029/2021wr030886.

- Kampen, R. J. R. van, M. van Berkel, and H. J. Zwart (2023a). "Estimating Space-Dependent Coefficients for 1D transport using Gaussian Processes as State Estimator in the Frequency Domain". In: *IEEE Control Systems Letters* 7 (2023), pp. 247–252. DOI: 10.1109/LCSYS.2022.3186626.

- Kampen, R. J. R. van, J. de Vries, S. Weiland, et al. (2023b). "Fast simultaneous estimation of nD transport coefficients and source function in perturbation experiments". In: *Scientific Reports* 13.1 (Feb. 2023), p. 3241. DOI: 10.1038/s41598-023-30337-0.

Co-authored peer-reviewed journal articles

- Berkel, M. van, R. J. R. van Kampen, G. Vandersteen, et al. (2020). "Correcting for non-periodic behaviour in perturbative experiments: application to heat pulse propagation and modulated gas-puff experiments". In: *Plasma Physics and Controlled Fusion* 62.9 (July 22, 2020), p. 094001. DOI: 10.1088/1361-6587/ab9eaa.

- Ravensbergen, T., M. van Berkel, A. Perek, et al. (2021). "Real-time feedback control of the impurity emission front in tokamak divertor plasmas". In: *Nature Communications* 12.1 (Feb. 2021), p. 1105. DOI: 10.1038/s41467-021-21268-3.
- Slief, J. H., R. J. R. van Kampen, M. W. Brookman, and M. van Berkel (2022). "Extension of the flux fit method for estimating power deposition profiles". In: *Physics of Plasmas* 29.1 (Jan. 20, 2022), p. 010703. DOI: 10.1063/5.0069869.
- Reimerdes, H., M. Agostini, E. Alessi, et al. (2022). "Overview of the TCV tokamak experimental programme". In: *Nuclear Fusion* 62.4 (Mar. 2022), p. 042018. DOI: 10.1088/1741-4326/ac369b.
- Fenstermacher, M. E., J. Abbate, S. Abe, et al. (2022). "DIII-D research advancing the physics basis for optimizing the tokamak approach to fusion energy". In: *Nuclear Fusion* 62.4 (Apr. 2022), p. 042024. DOI: 10.1088/1741-4326/ac2ff2.
- Slief, J. H., R. J. R. van Kampen, M. W. Brookman, et al. (2023). "Quantifying electron cyclotron power deposition broadening in DIII-D and the potential consequences for the ITER EC system". In: *Nuclear Fusion* 63.2 (Jan. 2023), p. 026029. DOI: 10.1088/1741-4326/acaedc.
- Berkel, M. van, C. H. Luce, R. J. R. van Kampen, et al. (2023). "A framework to improve the reliability of temperature-based estimates of flow and diffusion in streambeds with frequency domain examples". In: *in preparation*.

Peer-reviewed conference proceedings

- Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021b). "A Closed-Form Solution to Estimate Spatially Varying Parameters in Heat and Mass Transport". In: *2021 American Control Conference (ACC)*. May 2021, pp. 286–291. DOI: 10.23919/ACC50511.2021.9482757.
- Kampen, R. J. R. van, M. van Berkel, and H. J. Zwart (2022a). "Estimating Space-Dependent Coefficients for 1D transport using Gaussian Processes as State Estimator in the Frequency Domain". In: *61st IEEE Conference on Decision and Control (CDC)*. Dec. 2022, p. 1717.

Non peer-reviewed abstracts in conference proceedings

- Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2020). "Complex Gaussian Process Regression for Estimating Spatially Varying Coefficients in Thermal Transport". In: *39th Benelux Meeting on Systems and Control*. Elspeet, Netherlands, Mar. 2020, p. 38.
- Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021a). "A closed-form solution to estimate space-dependent parameters in heat and mass transport". In: *Physics@Veldhoven*. Veldhoven, Netherlands, Jan. 2021.
- Kampen, R. J. R. van, A. Das, S. Weiland, and M. van Berkel (2021d). "A Linear Least Squares Approach to Estimate Space-Dependent Parameters in Heat and Mass Transport". In: *Benelux Workshop on Systems and Control*. Rotterdam, Netherlands, June 2021, p. 68.
- Kampen, R. J. R. van, U. Schneidewind, C. Anibas, et al. (2021e). "On the Validity of Flux Estimates using Semi-Infinite Domains.Comparing Flux Estimates from Semi-Infinite and Bounded Domains using the LPMLEn - a Multi Frequency, Multi Sensor Method to Estimate Vertical Streambed Fluxes and Sediment Thermal Properties". In: *AGU Fall Meeting 2021*. Hybrid, New Orleans, LA, USA, Dec. 2021.
- Kampen, R. J. R. van, S. Weiland, H. J. Zwart, and M. van Berkel (2022b). "A frequency domain maximum likelihood Approach to Estimate Space-Dependent Parameters in heat and mass transport". In: *41st Benelux Meeting on Systems and Control*. Brussels, Belgium, July 2022, p. 200.

Dankwoord

Net zoals bij de partiële differentiaalvergelijkingen die ik de afgelopen vier jaar heb bestudeerd, zijn ook bij het promotietraject de randvoorwaarden zeer bepalend voor de uitkomst. Eenieder die, op welke manier dan ook, een bijdrage heeft geleverd aan de totstandkoming van dit proefschrift, behoort tot die randvoorwaarden. Ik wil jullie allen hartelijk bedanken en een aantal van jullie benoem ik graag in het bijzonder.

Allereerst mijn promotoren. Matthijs, jouw ongeremde enthousiasme, gedrevenheid en high-gain feedback zijn onmiskenbaar en onmisbaar geweest in de afgelopen vier jaar. Jouw altijd kritische input en intensieve begeleiding heb ik erg gewaardeerd. Daarnaast ben ik je dankbaar voor de prachtige start van mijn promotietraject door mij direct naar Brussel en Lausanne te laten reizen. Dit waren de meest leuke en leerzame weken in de afgelopen vier jaar. De thuiswerkperiode was onvermijdelijk, maar jij zorgde ervoor dat de motivatie op peil bleef door allerlei mooie samenwerkingen op te zetten. Onze schattingsmethoden toepassen in de hydrologie was een leuke onverwachte wending. Alle moeite die je voor mij hebt gedaan is niet ongezien gebleven en hiervoor wil ik je dan ook graag bedanken.

Hans en Siep, jullie scherpe commentaren, positieve houding en fascinatie voor alles wat maar met wiskunde te maken heeft werkt aanstekelijk. Na elke bijeenkomst vertrok ik in een betere stemming dan waarmee ik arriveerde. Fijnere promotoren dan jullie kan ik me niet voorstellen.

Furthermore, I would like to express my gratitude towards the chair Patrick Anderson and the committee members Marijke Huysmans, Paul van den Hof, Karel Keesman, and Tuomas Tala for their time and effort spent reading my dissertation and participating in my defense.

This thesis has been realized through indispensable collaboration with various inspiring people. Artur, Joost, Matthijs and Timo, although it is not included in this thesis, the two experimental campaigns at TCV were a great way to get some hands-on experience with system identification and nuclear fusion. This experience laid the foundation of my fusion knowledge and my further research. Our achievements, all the amazing places we had dinner, and the victory beers at the lake, made these trips “unaniem een belachelijk groot success”. Amritam, thank you for the fruitful collaboration, it really helped me get up to speed with my own research. Charlie, Daniele, and Stefan, many thanks for all the inspiring

discussions regarding heat transport in hydrology. This really opened up my eyes to how widely applicable and important my research is. Uwe, I really want to thank you for the close collaboration, interesting discussions, and most of all, taking care of all the references in this horrible program called 'Word'. Verder wil ik nog Marco de Baar, Koen Tiels, Gerd Vandersteen en Egbert Westerhof bedanken voor jullie advies gedurende de afgelopen vier jaar, dit heeft erg geholpen bij het scherp krijgen van verschillende details.

Daarnaast wil ik ook alle studenten die ik heb begeleid bedanken. Andreas, Bart, Jelle, andere Jelle en Jochem, de samenwerking met jullie heeft mij veel belangrijke inzichten opgeleverd die de basis hebben gevormd voor een aantal van mijn publicaties.

A big thanks to all the people at DIFFER, Bart, Beatrix, Chris, Fabio, Jos, Jonathan, Lennard, Michele, Marek and Paul for making it such an enjoyable workplace. I really appreciate all the interesting talks and fun moments we had during and after work. Special thanks to Laura and Maria, for pushing everybody (including me) to participate in social activities outside work and often bringing (home-made) food to lift the spirit during the coffee breaks. A special shout out to Aaron, Laura and Jesse for making so many great memes and WhatsApp stickers of all the DIFFER Ph.D. students.

Daarnaast wil ik ook mijn (voormalige) kantoorgenoten bedanken, Artur, Bob, Garud, Gerard, Gijs, Jesse, Joost en Timo. Dit kantoor was de geboorteplaats voor veel legendarische dingen: happy tunes Friday, quote van de week, computer-puntenscorelijst, MATLAB memes en zelfs de Energy Systems and Control groep is daar uiteindelijk opgericht. Het was altijd een plek waar ik met plezier naar toe ging en terecht kon met mijn 'domme' natuurkunde vragen, ongenuanceerde opmerkingen en frustraties. Garud, Jesse en Timo, ik ben jullie dankbaar voor het delen van alle nieuwsberichten en columns die we vervolgens bediscussieerden tijdens de ochtendkoffie en/of lunch. Then I moved to a new office, with new officemates: Aaron, Jelle, and Thomas. With the three of you it was not only our research that connected us, but also our taste in music and sense of humor. Opening the door in the morning and hearing great music blasting from the speakers made the final stretch of the Ph.D. a bit easier. Jelle, jou wil ik toch nog even extra bedanken voor alle zeer scherpe opmerkingen en het corrigeren van mijn Nederlands taalgebruik. Al blijf ik erbij, sommige (spreek)taalfouten zijn cultureel erfgoed en dienen behouden te worden.

Verder wil ik nog iedereen uit de CST en D&C groep van de TU/e bedanken, in het bijzonder, Brandon, Camiel, Fahim, Hao, Jilles, Johan, Koen, Koen, Lars, Leontine, Menno, Roeland, Roy, Sebastiaan en Wouter. Met alle plezier liet ik de goede koffie en het mooie kantoor bij DIFFER achter om één dag per week met jullie in de betonnen kelder te spenderen. Bedankt voor alle leuke koffiepauzes, lunch wandelingen, vrijdagmiddag borrels en de onvergetelijke tijden tijdens conferenties en borrels. Sven, dankjewel voor de mooie discussies over het potentieel toepassen van mijn werk op hyperthermie voor kankerbehandelingen. Bardia and Tomas, it was a real pleasure to be your office mate in the final stretch of the Ph.D. Especially

during our trip to Cancún where we had to write our thesis/paper while having the tempting view of the beach and the pools. Fortunately, after submitting, we were still able to enjoy it for a bit.

Rinus, Ruud en Timo, ik ben blij dat ik deel mocht uitmaken van dit 'boze witte mannen' gezelschap. De diepgaande gesprekken onder het genot van koffie, bier of whisky waarbij alles in termen van wiskunde en regeltechniek werd geanalyseerd, deden mij goed. Ik waardeer het enorm dat ik altijd op jullie kan rekenen als vaste gesprekspartners voor advies, inspiratie en de nodige afleiding.

Ook wil ik de docenten werktuigbouwkunde aan de Hogeschool Rotterdam bedanken, in het bijzonder Andrea, Jac en Jos. Jullie hebben mij niet alleen als student een goede basis meegegeven, maar daarna heb ik als collega nog veel meer van jullie geleerd. Zonder jullie hulp en ondersteuning had ik nooit aan mijn Ph.D. kunnen beginnen. Andrea, de appjes met updates over de opleiding en de sporadische vragen over de vakken die ik gaf waren een welkome afleiding.

Daarnaast wil ik al mijn vrienden bedanken voor de leuke tijden buiten werk. In het bijzonder, de 'spelletjes groep', Joris, Leontine, Menno, Roeland, Thomas en Yi Ying. Fabian, voor alle leuke gesprekken en updates omtrent je toffe hobbyprojecten. Die gaven mij vaak weer hernieuwde energie voor mijn eigen hobbyprojecten, waarvan ik er nu eindelijk eentje ga afronden, namelijk mijn Ph.D. Mike en Dirk-Jan, dankjewel voor de jarenlange vriendschap en alle onvergetelijke momenten die we samen hebben beleefd. Ik kan altijd waarderen hoe ik met jullie gewoon even alles achter me kan laten en kan genieten van de altijd bijzondere momenten samen.

Als volgende wil ik Els en Peter bedanken voor de gezelligheid, ondersteuning, altijd aanwezige interesse in mijn werk en natuurlijk de frietjes op vrijdag.

De laatste woorden zijn voor mijn familie. Opa en oma, dankjewel voor de gezelligheid en het begrip voor het 'oneindige' studeren, maar na vele jaren is dit dan uiteindelijk het sluitstuk.

Wesley en Kimberley, jullie ook bedankt voor het creëren van een fijne thuissituatie waarin ik altijd al mijn verhalen kwijt kon en tegenwoordig het enthousiasme als ik weer met een nieuw bordspel kom opdagen.

Pa en ma, bedankt voor het creëren van een stabiele basis. De levenslessen die jullie mij van kinds af aan probeerde bij te brengen hebben mij in staat gesteld om dit te bereiken. Vooral de zin "even diep ademhalen en gewoon nog een keer proberen" heeft mij door de afgelopen vier jaar geholpen.

Tot slot, Shari, al meer dan 10 jaar ben jij mijn steun en toeverlaat. Bedankt voor alles dat je voor mij hebt gedaan, van 'het streepje' tot het bieden van een luisterend oor en van alle relativerende opmerkingen tot het verbreden van mijn horizon met leuke activiteiten en nieuwe hobby projecten. Mijn proefschrift maakt mij trots, maar jij maakt mij gelukkig.

Ricky van Kampen
Eindhoven, Maart 2023

About the author

Ricky van Kampen was born on the eighteenth of September, 1993 in Spijkenisse, the Netherlands. After finishing his secondary education in 2010 at PENTA college CSG Blaise Pascal in Spijkenisse, he studied Mechanical Engineering at the Rotterdam University of Applied Sciences. After receiving his B.Eng degree in 2014, he started his pre-master in Mechanical Engineering at the Eindhoven University of Technology, while simultaneously starting his career as a lecturer in mechanical engineering at the Rotterdam University of Applied Sciences. In 2016 he received the teaching qualification for universities of applied sciences from the VU Amsterdam. In 2018 he received the M.Sc. degree (*with great appreciation*) in Mechanical Engineering from the Eindhoven University of Technology. The subject of his Master's Thesis was on geometric social force interaction models for autonomous driving with mobile robots. In 2019 he quit his job as a lecturer at the Rotterdam University of Applied Sciences to start his Ph.D. research on the estimation of transport coefficients in the newly founded Energy Systems and Control group at DIFFER - Dutch Institute for Fundamental Energy Research in Eindhoven and the Control Systems Technology Group at the Eindhoven University of Technology. The main results of this research are presented in this thesis.