Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Øyvind Stormark Auestad

Identifying Soil Heat Dynamics

Master's thesis in Industrial mathematics
Supervisor: Henning Omre
August 2020

NTNU
Norwegian University of
Science and Technology

Øyvind Stormark Auestad

Identifying Soil Heat Dynamics

Master's thesis in Industrial mathematics
Supervisor: Henning Omre
August 2020

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Summary

We have studied a Gaussian process for modelling soil heat flow. It is the stationary solution to a linear stochastic system based on the stochastic heat equation with additive noise. With temperature measurements at different locations in the soil, filtering, and computing the likelihood of the observations are efficiently performed using the Kalman recursions. The maximum likelihood estimates may in turn be found using some numerical optimization routine with quick gradient computations by automatic differentiation. Finally, the proposed model is applied to real temperature measurements in a problem related to the load capacity of buried electric cables.

Sammendrag

Vi har studert en Gaussisk prosess for å modellere varmeflyt i jord. Prosessen er den stasjonære løsningen på et lineært stokstisk system basert på varmeledningsligningen med additiv støy. Med temperaturmålinger rundt omkring i jorden, løses filtreringsproblemet, og man kan beregne observasjonssannsynligheten, effektivt ved hjelp av Kalmanrekursjonene. Videre kan man finne sannsynlighetsmaksimeringsestimatene med en gradientbasert optimeringsrutine, hvor gradienten beregnes hurtig ved hjelp av automatisk derivasjon. Avslutningsvis anvender vi modellen på faktiske temperaturmålinger fra jorden i et problem relatert til lastkapasiteten til strømkabler.

Table of Contents

Chapter 1

Introduction

The bottleneck for the amount of current which can be passed through buried electric cables is the temperatures of the cable components. The problem of determining the greatest amount, or cable load capacity, is an active topic of research among electric engineers. Even though the physics governing the cable temperature alone is well understood by the electric engineer, the soil heat dynamics, when exposed to external variables such as varying weather and precipitation, is much less understood.

Every buried cable has a predetermined static load capacity, as per the industrial standard IEC 60287. This is the constant current which can be applied for an infinite amount of time while keeping the temperature of the components below their maximum tolerable temperature. As it is power consumption and energy production that mainly determine the current passing through cables, the cable load is rarely constant. Therefore, cables are also rated according to the dynamic current rating standard, IEC 60853. This standard determine the cable load capacity when subject to periodic load profiles, and short lasting, high, loads. Due to the varying thermal properties of the surroundings, conservative values for the thermal properties must be used in the computations. It follows that the computed load capacities usually end up becoming conservative, and electric cables generally go underutilized.

1.1 Problem formulation

In order to better understand and quantify the uncertainties introduced into the heat dynamics by varying conditions, different research projects have been initiated, and thermal measurements from the soil surrounding buried electric cables in different actual cable installations are being made. Among them, we have the Tronsholen Skeiane cable field, displayed in Figure 1.1. This motivates us to consider the problem of identifying parameter driven stochastic heat flow. In this study, we aim to propose a simple, stochastic model for soil heat dynamics, and verify its validity on measurements from an actual cable installation. The goal of the study is to be able to produce reliable forecasts of the cable temperature and its uncertainty.

Figure 1.1: Tronsholen Skeiane. Source: Lyse Elnett.

1.2 Thesis structure

Chapter 2 presents the underlying theoretical framework of linear stochastic systems, in addition to filtering, smoothing and forecasting for linear Gaussian state space models. At the end of the chapter, the nonlinear optimization problem of maximising the observation likelihood is addressed. In Chapter 3 we study the particular stochastic heat flow problem related to cable soil systems, and justify our choice of model. In the fourth and final chapter, parameter inference on actual temperature measurements are performed.

1.3 Main definitions and notation

A random variable, $X : \Omega \to \mathbf{R}^n$, is a measurable function on a probability space $(\Omega, \mathcal{F}, P)$, and a stochastic process is an indexed collection of random variables,

$$\{X_t : t \in T \subseteq \mathbf{R}^d\}. \tag{1.1}$$

The probability distribution of a random variable X is the probability measure, μ_X, on $(\mathbf{R}^n, \mathcal{B})$, satisfying $\mu_X(A) = P(X^{-1}(A))$, $A \in \mathcal{B}$. The density of the distribution, p_X, is the $\mathcal{B}$-measurable function, satisfying $\mu_X(A) = \int_A p_X(x)dx$ for all $A \in \mathcal{B}$. A Gaussian process is a stochastic process $\{X_t\}_{t\in T}$ such that for every subset $\{t_1, \ldots, t_k\} \subseteq T$, the random variable $(X_{t_1}, \ldots, X_{t_k}) : \Omega^k \to (\mathbf{R}^n)^k$ is multivariate Gaussian distributed.

Chapter 2

Framework

We are interested in the two dimensional stochastic heat equation with additive noise,

$$\partial_t U(t,x) = k\Delta U(t,x) + q(t,x) + W(t,x); \quad (t,x) \in [0,\infty) \times \mathbf{R}^2, \tag{2.1}$$

where ∂_t, Δ denotes time differentiation and the Laplacian operator respectively, with $U_{0,x} : \mathbf{R}^2 \times \Omega \to \mathbf{R}$, $W_{t,x} : [0,\infty) \times \mathbf{R}^2 \times \Omega \to \mathbf{R}$ independent, and $q : [0,\infty) \times \mathbf{R}^2 \to \mathbf{R}$. The solution to (2.1) is,

$$U_{t,x} = \int_{\mathbf{R}^2} \phi(0, x-y) U_{0,y} dy + \int_0^t \int_{\mathbf{R}^2} \phi(t-s, x-y)(q(s,y) + W_{s,y}) dy ds, \tag{2.2}$$

where $\phi(t,x) = (4\pi kt)^{-d/2} \exp(-\|x\|^2/(4kt))$ is the heat kernel, provided it makes sense for the choice of $W_{t,x}$. A detailed description of the solution form (2.2) is given in Hairer (2004), and it turns out there are a lot of distributions $W_{t,x}$ for which it is meaningful. We consider the case when $W_{t,x}$ takes values according to some twice differentiable potential, $Z_{t,x} : [0,\infty) \times \mathbf{R}^2 \times \Omega \to \mathbf{R}$, such that $W(t,x) = k\Delta Z(t,x)$. This corresponds to heat flow subject to an additional uncertain heat flux,

$$J = -k\nabla Z_{t,x}. \tag{2.3}$$

In the remainder of this study, we consider the discrete version of the conservation law (2.1), where the spatial domain is divided in a finite number of volumes, and has Dirichlet boundaries. In this case, the conservation equation may be expressed by a linear stochastic system. We are interested in inferring the parameters driving the solution to this system. That is, the parameters determining the stochastic process, $Z_{t,x}$, the thermal diffusion coefficient, k, in addition to parameters related to boundary conditions and source terms. In the remainder of this chapter we establish necessary theory on identifying linear stochastic systems, and start by introducing state space processes and filtering.

2.1 State space process

A stochastic process, $\{(X_t, Y_t) : t = 1, \ldots, T\}$, determined through

$$\begin{aligned} \text{(system)} \quad X_{t+1} &= a_t(X_t, V_t), \\ \text{(measurement)} \quad Y_t &= b_t(X_t, W_t), \end{aligned} \tag{2.4}$$

with $a_t : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$, $b_t : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m$, and X_1, V_t, W_t independent random variables, is referred to as a state space process, or hidden Markov process. The equation defining the development of the state, $\{X_t\}$, is commonly referred to as the state, system or dynamic equation. The equation defining the observations, $\{Y_t\}$, is commonly referred to as the measurement or observation equation. That is, the process naturally arises when modelling the development of some unobserved stochastic quantity, where we have available some, possibly noisy, observations related to the unobserved state.

The state space process above satisfies the Markov property,

$$p(x_t|x_{1:t-1}) = p(x_t|x_{t-1}), \quad p(x_1|x_0) := p(x_1), \tag{2.5}$$

where we have omitted the subscript of the probability density as it is clear from the argument, and denoted the collections $(x_1, \ldots, x_t)$ by $x_{1:t}$. It also satisfies the conditional independence properties,

$$p(y_t|y_{1:t-1}, y_{t+1:T}, x_{1:T}) = p(y_t|x_t), \tag{2.6}$$

and,

$$p(x_t|x_{t+1}, y_{1:T}) = p(x_t|x_{t+1}, y_{1:t}). \tag{2.7}$$

Finally, the joint probability distribution of the state and observations may be expressed,

$$p(x_{1:T}, y_{1:T}) = p(x_1)p(y_1|x_1) \prod_{t=2}^{T} p(x_t|x_{t-1})p(y_t|x_t). \tag{2.8}$$

This study is concerned with linear state space processes. When the time dynamics are discrete, this process may be expressed,

$$\begin{aligned} X_{t+1} &= A_t X_t + q_t + D_t V_t, \\ Y_t &= B_t X_t + r_t + H_t W_t, \end{aligned} \tag{2.9}$$

for matrices $A_t \in \mathbf{R}^{n \times n}$, $B_t \in \mathbf{R}^{m \times n}$, $D_t \in \mathbf{R}^{n \times k}$ and $H_t \in \mathbf{R}^{m \times l}$, and $q_t \in \mathbf{R}^n$, $r_t \in \mathbf{R}^m$, and noise terms $V_t \in \mathbf{R}^k, W_t \in \mathbf{R}^l$. In the particular case when X_1, V_t and W_t are all Gaussian and independent, we refer to the process as a linear Gaussian state space process.

2.2 Filtering, smoothing and forecasting

Many engineering problems are concerned with estimating or monitoring some unknown quantity that varies continuously or discretely through time, based on some, possibly noisy,

observations. Rephrased, we are interested in estimating the state, given the observations, preferably the conditional probability distribution, $p(x_t|y_{1:T})$, if possible. The filtering problem is the problem of obtaining the best estimate of X_t based on the observations $Y_s = y_s$ for $s = 1, \dots, t$. The smoothing problem concerns obtaining the best estimate using all observations, while the forecasting problem entail estimating the state $X_{t'}$, at time $t' > t$ given observations up to time t.

From Bayes rule and (2.8) we may deduce

$$p(x_{1:t}|y_{1:t}) \propto p(x_1)p(y_1|x_1)\prod_{t=2}^{T} p(x_t|x_{t-1})p(y_t|x_t), \tag{2.10}$$

moreover, the marginal distributions, $p(x_t|y_{1:t})$, may be computed recursively as

$$\begin{aligned} p(x_t|y_{1:t}) &\propto p(y_t|x_t)p(x_t|y_{1:t-1}), \\ p(x_{t+1}|y_{1:t}) &= \int p(x_{t+1}|x_t)p(x_t|y_{1:t})dx_t. \end{aligned} \tag{2.11}$$

Even though these conditional distributions are rarely easily evaluated, the state inference problems of the linear state space process of (2.9) has an efficient and elegant solution. The derivation of which relies on elementary Hilbert space theory and the notion that for random variables taking values in $\mathbf{R}$, $L^2(\Omega, P)$, with inner product $E[\,\cdot\,\cdot]$, is a Hilbert space (see Brockwell and Davis (1991)). For random variables X taking values in $\mathbf{R}^n$ and Y in $\mathbf{R}^m$, we may define $\hat{X}$ as the $\mathbf{R}^n$ valued random variable, having as entries the affine transformations of the components of Y, minimizing the mean squared errors, $E[X^{(i)} - \hat{X}^{(i)}]^2,\ i = 1, \dots, n$. From the projection theorem, we know that this projection exists, and satisfies the sufficient and necessary orthogonality conditions,

$$E[X^{(i)} - \hat{X}^{(i)}] = 0,\ E[\big(X^{(i)} - \hat{X}^{(i)}\big)Y^{(j)}] = 0,\ j = 1, \dots, m, \tag{2.12}$$

for $i = 1, \dots n$. This condition may be expressed compactly as

$$E[X - MY - \mu] = 0,\ E[\big(X - MY - \mu\big)Y^T] = 0, \tag{2.13}$$

with $\hat{X} = MY + \mu$ for some matrix $M \in \mathbf{R}^{n\times m}$ and $\mu \in \mathbf{R}^n$. The Kalman recursions, originally introduced in Kalman (1960), compute the projection of X_t, denoted $\hat{X}_{t|t'}$, into the space $\{\hat{X} : \hat{X} = \mu + M_1Y_1 + \cdots + M_{t'}Y_{t'},\ \mu \in \mathbf{R}^n,\ M_i \in \mathbf{R}^{n\times m}\}$, and its error, $S_{t|t'} := E[\big(X_t - \hat{X}_{t|t'}\big)\big(X_t - \hat{X}_{t|t'}\big)^T]$, efficiently recursively, in a Gram-Schmidt like manner.

2.2.1 Kalman recursions

We now derive the recursions for the linear state space process (2.9) where the entries of X_1, V_t, W_t are all in $L^2(\Omega, P)$ and orthogonal. The covariance matrices of V_t and W_t are identity matrices in $\mathbf{R}^{k\times k}$ and $\mathbf{R}^{l\times l}$, respectively. Initially, set $\hat{X}_{1|0} = \hat{X}_1, S_{1|0} = S_1$. Note that since W_t is orthogonal to $Y_1, \dots, Y_{t-1}$, the best linear predictor of Y_t given $Y_1, \dots, Y_{t-1}$ is $B_t\hat{X}_{t|t-1} + r_t$. Let $I_t := Y_t - B_t\hat{X}_{t|t-1} - r_t$, be the t'th innovation, and

note that this sequence of random variables is per construction orthogonal. Moreover, the span of the innovations up to time t coincides with the span of the observations up to time t. The orthogonality of the innovations implies that we have

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + M_t I_t, \tag{2.14}$$

where the $n \times m$-matrix M_t may be found from the orthogonality condition,

$$E[\big(X_t - M_t I_t\big) I_t^T] = 0, \tag{2.15}$$

giving, $M_t = S_{t|t-1} B_t^T \Delta_t^{-1}$, with Δ_t^{-1} being any generalized inverse of $B_t S_{t|t-1} B_t^T + H_t H_t^T$. Due to the orthogonality of V_t and $Y_1, \ldots, Y_t$ we have that $\hat{X}_{t+1|t} = A_t \hat{X}_{t|t} + q_t$. The errors of the projections can be found from these expressions and algebraic manipulations to be

$$\begin{aligned} S_{t|t-1} &= A_t S_{t-1|t-1} A_t^T + D_t D_t^T, \\ S_{t|t} &= (I - M_t B_t) S_{t|t-1}. \end{aligned} \tag{2.16}$$

The forecasting problem is solved simply by noting that $\hat{X}_{t+1|t'} = A_t \hat{X}_{t|t'} + q_t$, and $S_{t+1|t'} = A_t S_{t|t'} A_t^T + D_t D_t^T$ for $t' < t$. In his original paper, Kalman does not treat the smoothing problem. Although, the same idea applies when computing $\hat{X}_{t|t'}, t' > t$. The orthogonality of the innovations allows us to write,

$$\hat{X}_{t|t'} = \hat{X}_{t|t'-1} + M_{t,t'} I_{t'}, \tag{2.17}$$

where the matrix $M_{t,t'}$ is found from the orthogonality condition $E[\big(X_t - M_{t,t'} I_{t'}\big) I_{t'}^T] = 0$, giving $M_{t,t'} = E[X_t I_{t'}^T] \Delta_{t'}^{-1}$. We may write $E[X_t I_{t'}^T] = S_{t,t'} B_{t'}^T$, where $S_{t,t'} = E[X_t (X_{t'} - \hat{X}_{t'|t'-1})^T]$ is computed from the recurrence relation

$$S_{t,t'} = A_{t'-1}(I - M_{t'-1} B_{t'-1}) S_{t,t'-1}, \tag{2.18}$$

with $S_{t,t} = S_{t|t-1}$. The error of the projection then becomes

$$S_{t|t'} = S_{t|t'-1} - M_{t,t'} \Delta_t M_{t,t'}^T. \tag{2.19}$$

Rausch-Tung-Streibel smoother

The smoother implementation above is not particularly efficient if the intention is to compute the smoothed estimate of the state at all times. Originally introduced in Rauch et al. (1965), we may obtain a faster implementation by noting that the condition $\hat{X}_{t|t,X_{t+1}} = E[X_t | I_{1:t}, X_{t+1}]$ (i.e. the conditional expectation is linear), in addition to the conditional independence condition (2.7), is sufficient for $\hat{X}_{t|T,X_{t+1}} = \hat{X}_{t|t,X_{t+1}}$ to hold, where the subscript notation $\hat{X}_{t|T,X_{t+1}}$ denotes the projection of X_t into the space $\{\hat{X} : \hat{X} = \mu + M_0 X_{t+1} + \sum_{t=1}^{T} M_t I_t\}$. That is,

$$\hat{X}_{t|t,X_{t+1}} = E[X_t | I_{1:t}, X_{t+1}] = E[X_t | I_{1:T}, X_{t+1}], \tag{2.20}$$

where the latter equality is due to (2.7). This implies that we have,

$$E[(X_t - \hat{X}_{t|t,X_{t+1}})I_k^T] = 0, \quad k > t, \tag{2.21}$$

which in turn implies, $\hat{X}_{t|T,X_{t+1}} = \hat{X}_{t,X_{t+1}}$. The orthogonality condition yields,

$$\hat{X}_{t|t,X_{t+1}} = \hat{X}_{t|t} + M_t(X_{t+1} - \hat{X}_{t+1|t}), \tag{2.22}$$

with $M_t = S_{t|t}A_t^T S_{t+1|t}^{-1}$. It follows that

$$\hat{X}_{t|T} = \hat{X}_{t|t} + M_t(\hat{X}_{t+1|T} - \hat{X}_{t+1|t}), \tag{2.23}$$

with smoothing errors $S_{t|T} = S_{t|t} + M_t(S_{t+1|T} - S_{t+1|t})M_t^T$. The equality (2.23) follows noting that the projection $\hat{X}_{t|T}$ is equal to the projection of $\hat{X}_{t|T,X_{t+1}}$ into the span of $\{I_1, \ldots, I_T\}$ (see Brockwell and Davis (1991)). Performing the smoothing computations is in practice performed by first running the filter recursions once, storing the predictors $\hat{X}_{t|t-1}, \hat{X}_{t|t}$ and the errors, $S_{t|t-1}, S_{t|t}$, and then, starting at $\hat{X}_{T|T}, S_{T|T}$, compute $\hat{X}_{t|T}, S_{t|T},\ t < T$, recursively.

Best linear predictor vs. conditional expectation

Up until now, we have not defined the conditional expectation of a random variable $X : \Omega \to \mathbf{R}$, $E[|X|] < \infty$. In elementary statistics, this is defined in terms of the conditional probability density. That is, "conditioning X on the observation $Y = y$", where Y is another random variable, we have,

$$E[X|Y] = \int_{\mathbf{R}} xp(x|y)dx, \tag{2.24}$$

and in turn, we are left with a function of y. When X is also in $L^2(\Omega, P)$, Brockwell and Davis (1991) defines the conditional expectation of X given some random variable $Y : \Omega \to \mathbf{R}^n$, as the projection of X into the space of all random variables in $L^2(\Omega, P)$ which can be written of the form $\phi(Y) : \Omega \to \mathbf{R}$, for some measureable $\phi : \mathbf{R}^n \to \mathbf{R}$. In other words, this is the function $E[X|Y] : \Omega \to \mathbf{R}$ of Y minimizing the distance $E[(X - E[X|Y])^2]$. The best linear predictor we defined above therefore differs by the conditional expectation in that we constrain $E[X|Y]$ to be an affine function of the components of Y. For this study, the definition in Brockwell and Davis (1991) suffices, since we are only working with random variables in $L^2(\Omega, P)$, and only conditioning on observations of Y. It does not hold in general for random variables in $L^1(\Omega, P)$, and for conditioning on more general events. The general definition of conditional expectation is given in Øksendal (2000).

When X_1, V_t, W_t in the state space model (2.9) are independent and Gaussian, the best linear predictor and conditional mean coincides, and so the Kalman recursions compute the first and second moments of the exact conditional distributions recursively. More generally, Brockwell and Davis (2002) notes that the best linear predictor and conditional expectation coincides in the case when X_1, V_t, W_t are uncorrelated, provided $\{V_t\}$ is a Martingale difference sequence with respect to $\{X_t\}$; that is, $E[V_t|\{X_s\}_{s \le t}] = 0, t = 1, \ldots, T$.

We note that this condition is implied by independence of X_1, V_t, W_t, but not by orthogonality. For all practical reasons, we only concern ourselves with state space processes where X_1, V_t, W_t are independent from now on.

2.3 Linear stochastic systems

The way state space models were introduced in this study were as a discrete time stochastic process, defined for integer time steps, $t = 1, \ldots, T$. Although, when modelling some quantities, it may be more natural to have a continuous time development. This is commonly done through a stochastic differential equation. That is, it is meaningful to define

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \tag{2.25}$$

in the sense that,

$$X_t - X_0 = \int_0^t (b(s, X_s)ds + \sigma(s, X_s)dB_s), \tag{2.26}$$

is the limit in mean square of the Riemann sum,

$$\sum_{j=1}^{k-1} \big(b(s_j, X_{s_j})\Delta s_j + \sigma(s_j, X_{s_j})\Delta B_{s_j}\big), \quad \Delta s_j = s_{j+1} - s_j,\ s_1 = 0,\ s_k = t, \tag{2.27}$$

with $\Delta B_{s_j} = B_{s_{j+1}} - B_{s_j}$, for a Brownian motion $\{B_t\}$, in addition to the technical conditions described in Øksendal (2000). That is, the limit of (2.27) exists in the sense of X_t being a well defined random variable, and is referred to as the Itô-integral, and the resulting stochastic process, an Itô diffusion.

Generally, depending on the construction of the Riemann sum, different limits are obtainable. In fact, in our particular application of linear stochastic system where $\sigma(.)$ does not depend on X_t, the two most popular such limits, the one corresponding to Itô calculus, and that of Stratonovich calculus where the integrand in the Riemann sum is evaluated as the average of the interval endpoints, actually coincide.

In the case when the stochastic process, $\{B_t\}$, in the Riemann sum defining the Itô-integral above is a standard Brownian motion, satisfying

$$\begin{gathered} B_t - B_{t'},\ t' < t \text{ are independent of all } B_s,\ s \le t', \\ (B_{t_1}, \ldots, B_{t_k}), \text{ is multivariate Gaussian for all } \{t_1, \ldots, t_k\} \subseteq [0, T], \\ E[B_t] = 0, \\ E[B_t B_t^T] = tI; \quad B_0 = 0, \end{gathered} \tag{2.28}$$

$b(.)$ affine in X_t, $\sigma(.)$ does not depend on X_t, the solution to the resulting system of linear stochastic differential equations,

$$dX_t = AX_t dt + q_t dt + C_t dB_t; \quad X_0 \sim N(x_0, S_0), \tag{2.29}$$

with $X_t \in \mathbf{R}^n$, $B_t \in \mathbf{R}^k$, $q_t \in \mathbf{R}^n$, $A \in \mathbf{R}^{n\times n}$, $C_t \in \mathbf{R}^{n\times k}$, may be expressed,

$$X_t = \exp(At)X_0 + \int_0^t \exp(A(t-s))(q_s ds + C_s dB_s). \tag{2.30}$$

Or, equivalently, as the stochastic process, Gaussian distributed at each time t, with mean and variance $\hat{X}(t)$, $S(t)$, given by the systems of linear ordinary differential equations,

$$\begin{aligned} \frac{d}{dt}\hat{X} &= A\hat{X} + q_t, \\ \frac{d}{dt}S &= AS + SA^T + C_t C_t^T, \end{aligned} \tag{2.31}$$

with $\hat{X}(0) = x_0 \in \mathbf{R}^n$, and $S(0) = S_0 \in \mathbf{R}^{n\times n}$. Note that the solution (2.30) may be expressed by the discrete dynamics equation,

$$X_t = FX_0 + r + V, \tag{2.32}$$

with

$$\begin{aligned} F = \exp(At), \quad r = \int_0^t \exp(A(t-s))q_s ds, \\ V \sim N(0, Q_t), \text{ with } Q_t = \int_0^t \exp\big(A(t-s)\big)C_s C_s^T \exp\big(A^T(t-s)\big)ds, \end{aligned} \tag{2.33}$$

where $Q_t \in \mathbf{R}^{n\times n}$ solves the latter equation in (2.31) (referred to as the Lyapunov equation) with initial condition $S(0) = 0$, and is assumed to be positive definite. The expression for Q_t may also be found using the expression (2.30) and the Itô isometry (see Øksendal (2000)). With the linear dynamics equation above, and with observations $\{Y_{t_j} : j = 1, \ldots, k\}$ affine in X_t and with added noise, we have full consistency with the linear Gaussian state space process as introduced in (2.9). However, the Kalman recursion still hold in the more general case where A varies with time, provided we replace the discrete mean and variance update equations with the solution to the differential equations (2.31). The filtering problem arising when also the observation process is continuous was originally solved in Kalman and Bucy (1961), and gives rise to the well studied Ricatti differential equation.

An important example, which will be applied a lot in Chapter 3 and 4, is the Ornstein-Uhlenbeck process, defined through,

$$dZ_t = -\phi Z_t dt + D dB_t; \quad Z_0 \sim N(z_0, S_0), \tag{2.34}$$

where $Z_t \in \mathbf{R}^n, D \in \mathbf{R}^{n\times k}, B_t \in \mathbf{R}^k$, which entails that the distribution of Z_t is Gaussian with mean $\exp(-\phi t)z_0$ and variance $\exp(-2\phi t)S_0 + \frac{DD^T}{2\phi}(1 - \exp(-2\phi t))$. A Technical notion: If we for instance take $S_0 = 0$, Z_t is technically only multivariate Gaussian if DD^T is positive definite; this is generally not true. However, we usually abuse the multivariate normal notation, since it does not matter in the filtering recursions.

Stability of linear stochastic systems

In this and the next section, we assume q and C in (2.29) are fixed in time as well. When the eigenvalues of A have strictly negative real part, the system is globally asymptotically stable, and tends to the same random variable independent of initial state. The moments of the steady state solution is given by the steady states of the linear systems (2.31). In particular, the steady state variance may be found by solving (the continuous Lyapunov equation),

$$AS + SA^T + CC^T = 0. \tag{2.35}$$

This equation has a simple solution whenever A commutes with CC^T and its own transpose, namely $S^* = -(A + A^T)^{-1}CC^T$. It follows that we may then solve the Lyapunov differential equation by the integral shifting trick, $S(t) = S^* - \exp(At)S^* \exp(A^T t)$, with $S(0) = 0$. In the general case when the matrices do not commute, one could in principle find the solution by collecting the linear system of $\frac{n}{2}(n+1)$ equations and solving them, although, there exists method designed specifically for this problem. In a practical setting, the added noise in (2.33) can just be approximated by the trapezoidal rule, $S(t) \approx \frac{t}{2}(CC^T + \exp(At)CC^T \exp(A^T t))$, with $S(0) = 0$, for a small time step t.

For the linear state space processes (2.9), there may also be a steady state for the conditional distribution of the the state at uniform discrete time steps. Collecting the variance update equations of the discrete time Kalman recursions, we find that,

$$S_{j+1|j} = A(I - M_j B)S_{j|j-1}A^T + Q, \tag{2.36}$$

and so the steady state variance of the one step predictions is given by the equation,

$$S^* = A(I - S^*B^T(BS^*B^T + R)^{-1}B)S^*)A^T + Q, \tag{2.37}$$

known as the algebraic Ricatti equation (see e.g. Davis and Vinter (1985)).

Stationary solution

Closely related to stability of linear stochastic systems are stationarity. A stationary stochastic process $\{X_t\}_{t\in[0,T]}$ is a stochastic process such that for all collections $\{t_1, \ldots, t_k\} \subseteq [0, T]$, the joint probability density of $(X_{t_1+h}, \ldots, X_{t_k+h})$, is independent of h. In the Gaussian case, this is equivalent to the mean and covariance of the process being time shift invariant. Assuming the system (2.29) is stable and time invariant as described in the previous section, it admits a unique stationary solution (see Brockwell and Davis (1991)). To see this, note that we may represent the stationary solution (where we have set $q = 0$ for simplicity),

$$X_t = \int_{-\infty}^{t} \exp(A(t-s))dB_s, \quad E[X_t X_t^T] = S^*, \quad E[X_t] = 0, \tag{2.38}$$

where S^* is the steady state variance. It follows that for $h > 0$,

$$E[X_{t+h}X_t^T] = E[\big(\exp(Ah)X_t + \int_t^{t+h} \exp(A(t-s))dB_s\big)X_t] = \exp(Ah)S^*, \tag{2.39}$$

owing to the independence of X_t and $\{B_s\}_{s>t}$.

In Chapter 4, we employ the zero mean stationary Ornstein-Uhlenbeck process (2.34). This is the solution of the system (2.34) with initial condition $Z_0 \sim N(0, \frac{1}{2\phi}DD^T)$, and is the centered Gaussian process, Z_t, with covariance,

$$E[Z_t {Z_s}^T] = \exp(-\phi|t-s|)\frac{DD^T}{2\phi}. \tag{2.40}$$

2.4 Parameter inference

The remainder of this chapter is concerned with parameter inference for linear Gaussian state space models. Formally, we are interested in inferring the parameters driving the process (2.9), by computing the maximum likelihood estimates given the observations, $\{(Y_t = y_t) : t = 1, \ldots, T\}$,

$$\hat{\theta} = \text{argmax}_\theta L(\theta), \tag{2.41}$$

where $L(\theta) = p(y_{1:T}; \theta)$ is the likelihood of the observations.

We may avoid the cumbersome task of integrating out the unknown state from (2.8), by noting that the innovations, I_t, are independent zero mean multivariate Gaussian with variance Δ_t. The likelihood of the observations may thus be expressed,

$$L(\theta) = \prod_{t=1}^{T}(2\pi)^{-m/2}\det(\Delta_t)^{-1/2}\exp(-\frac{1}{2}I_t^T\Delta_t^{-1}I_t). \tag{2.42}$$

In the general case, there are no closed form expression for the maximum likelihood estimates, and they must be found by maximizing the likelihood numerically. Note that maximising the likelihood is equivalent to minimizing the negative log likelihood,

$$\ell(\theta) = \frac{1}{2}\sum_{t=1}^{T}\big(\log\det(\Delta_t) + I_t^T\Delta_t^{-1}I_t\big), \tag{2.43}$$

which is efficiently computed using the filtering recursions. Minimizing (2.43) is a well studied problem in continuous nonlinear optimization. In order to do so, we search for a stationary point by numerically solving $\nabla\ell(\theta) = 0$. This is commonly done using a quasi-Newton method. That is, we hope for convergence of the scheme,

$$\hat{\theta}_{k+1} = \hat{\theta}_k - H_k^{-1}\nabla\ell(\hat{\theta}_k), \tag{2.44}$$

where H_k is some approximation of the Hessian of ℓ at $\hat{\theta}_k$.

Gradient and Hessian

In our application, we are concerned with time invariant state space processes of the form (2.9) where B does not depend on θ. In this case, the derivative of (2.43) with respect to

$\theta \in \mathbf{R}$ is,

$$\partial_\theta \ell = \frac{1}{2}\sum_{t=1}^{T}\left(2I_t^T\Delta_t^{-1}(\partial_\theta I_t) - I_t^T\Delta_t^{-1}(\partial_\theta\Delta_t)\Delta_t^{-1}I_t + \mathrm{tr}(\Delta_t^{-1}(\partial_\theta\Delta_t))\right) \tag{2.45}$$

where,

$$\partial_\theta I_t = -B\partial_\theta\hat{X}_{t|t-1} - \partial_\theta s_t, \tag{2.46}$$

and,

$$\partial_\theta\Delta_t = B(\partial_\theta S_{t|t-1})B^T + \partial_\theta R. \tag{2.47}$$

The derivatives of $\hat{X}_{t|t-1}, S_{t|t-1}$ may be computed recursively; collecting the Kalman recursion equations, we obtain,

$$\begin{aligned}\hat{X}_{t+1|t} &= F(\hat{X}_{t|t-1} + M_t(Y_t - B\hat{X}_{t|t-1})) + r_t,\\ S_{t+1|t} &= F(I - M_tB)S_{t|t-1}F^T + Q.\end{aligned} \tag{2.48}$$

so that,

$$\begin{aligned}\partial_\theta\hat{X}_{t+1|t} = {}&(\partial_\theta F)(\hat{X}_{t|t-1} + M_t(Y_t - B\hat{X}_{t|t-1}))\\ &+ F(I - M_tB)\partial_\theta\hat{X}_{t|t-1} + F(\partial_\theta M_t)(Y_t - B\hat{X}_{t|t-1})\\ &+ \partial_\theta r_t,\end{aligned} \tag{2.49}$$

and,

$$\begin{aligned}\partial_\theta S_{t+1|t} = {}&(\partial_\theta F)(I - M_t)S_{t|t-1}F^T\\ &- F(\partial_\theta M_t)S_{t|t-1}F^T\\ &+ F(I - M_t)(\partial_\theta S_{t|t-1})F^T\\ &+ F(I - M_t)S_{t|t-1}(\partial_\theta F^T) + \partial_\theta Q,\end{aligned} \tag{2.50}$$

where,

$$\begin{aligned}\partial_\theta M_t = {}&(\partial_\theta S_{t|t-1})B^T\Delta_t^{-1}\\ &- S_{t|t-1}B^T\Delta_t^{-1}(\partial_\theta\Delta_t)\Delta_t^{-1},\end{aligned} \tag{2.51}$$

With initial values $\partial_\theta S_{1|0} = \partial_\theta S_1,\ \partial_\theta\hat{X}_{1|0} = \partial_\theta\hat{X}_1$. All matrix derivatives are componentwise, such that $(\partial_\theta M)_{i,j} = \partial_\theta(M)_{i,j}$. Differentiating once more we obtain recursions for the second derivatives of (2.48). However, an approximation to the information matrix (see Gupta and Mehra (1974), Goodrich and Caines (1979) for details),

$$M := E[\nabla^2\ell(\theta)], \tag{2.52}$$

is popularly used instead. That is,

$$\begin{aligned}M_{i,j} = E[(\partial_{\theta_i}\ell(\theta))(\partial_{\theta_j}\ell(\theta))] \approx \sum_{t=1}^{T}\big(&2(\partial_{\theta_i}I_t)^T\Delta_t^{-1}(\partial_{\theta_j}I_t)\\ &+ \mathrm{tr}(\Delta_t^{-1}(\partial_{\theta_i}\Delta_t)\Delta_t^{-1}(\partial_{\theta_j}\Delta_t)) + \frac{1}{2}\mathrm{tr}(\Delta_t^{-1}(\partial_{\theta_i}\Delta_t))\mathrm{tr}(\Delta_t^{-1}(\partial_{\theta_j}\Delta_t))\big).\end{aligned} \tag{2.53}$$

The derivative of the likelihood may be computed by hand, as described above, but since we are going to be experimenting with a lot of different parameterizations when fitting models to real data, manually computing the gradient can become tedious and error prone. Instead, we use automatic differentiation when computing the gradient.

The Stan Math Library and automatic differentiation

Automatic differentiation has found wide application in many engineering problems, perhaps most notably in the domain of continuous optimization, particularly parameter estimation for statistical models with large numbers of parameters. The Stan Math Library is a C++ library implementing reverse mode automatic differentiation using operator overloading. It contains a wide selection of supported matrix operations, including those required to perform parameter inference using the approach described in this chapter. A detailed description of automatic differentiation and the Stan math library may be found in the library documentation, Carpenter et al. (2015). Most relevant for our application are specialized log determinant and inverse functions for symmetric positive definite matrices, and matrix exponentials.

Direct maximization

Computing the gradient by automatic differentiation, and using a quasi-Newton method which approximates the Hessian by recent gradient computations, makes the optimization very simple implementation wise. In this study we use BFGS with a simple backtracking linesearch, as outlined in Nocedal and Wright (2006). Due to the nonlinear and complicated nature of the problem, and since we do not expect global convergence, we use a very relaxed, if any, curvature condition in the line search. Furthermore, we also reset the inverse Hessian approximation regularly to avoid it becoming ill conditioned. In turn, the resulting method varies between steepest descent, setting $H_k = \alpha I$ in (2.44), where $\alpha > 0$ is found from the line search, and ordinary BFGS for convex problems.

Since the likelihood generally has multiple stationary points, it is important to verify that the stationary point we find is in fact a minimizer. This is done by checking that the Hessian is positive definite at the candidate minimizer. It is worth mentioning that Gupta and Mehra (1974) has some valuable notions on the choice of optimization scheme for this problem. Both Gupta and Mehra (1974) and Goodrich and Caines (1979) advocate the use of the Fisher Scoring method, using (2.53) as the Hessian approximation. However, the general purpose quasi-Newton method has a clear advantage implementation wise.

Expected maximisation

An alternative approach to maximizing the likelihood directly, popular among statisticians, is the expected maximization approach, originally outlined in Shumway and Stoffer (1982) for this problem. In the proceeding we assume that the noise terms in (2.9) are such that $Q_t = D_t D_t^T, R_t = H_t H_t^T$ are positive definite. Taking the logarithm of the joint likelihood (2.8), scaling by a factor of -2 and subtracting constant terms, we get,

$$\begin{aligned}\ell'(\theta) &= \log\det(S_1) + (x_1-\mu_1)^T S_1^{-1}(x_1-\mu_1) \\ &+ \sum_{t=2}^{T}\big(\log\det(Q_{t-1}) + (x_t - A_{t-1}x_{t-1} - q_{t-1})^T Q_{t-1}^{-1}(x_t - A_{t-1}x_{t-1} - q_{t-1})\big) \\ &+ \sum_{t=1}^{T}\big(\log\det(R_t) + (y_t - B_t x_t - r_t)^T R_t^{-1}(y_t - B_t x_t - r_t)\big).\end{aligned} \tag{2.54}$$

Using the trace product property, treating (2.54) as a random variable and interchanging the order of expectation and trace, we obtain,

$$\begin{aligned}E[\ell'|Y_{1:T};\theta] &= \log\det(S_1) + \operatorname{tr}\big(E[(X_1-\mu_1)(X_1-\mu_1)^T|Y_{1:T}]S_1^{-1}\big) \\ &+ \sum_{t=2}^{T}\Bigg(\log\det(Q_{t-1}) \\ &+ \operatorname{tr}\big(E[(X_t - A_{t-1}X_{t-1} - q_{t-1})(X_t - A_{t-1}X_{t-1} - q_{t-1})^T|Y_{1:T}]Q_{t-1}^{-1}\big)\Bigg) \\ &+ \sum_{t=1}^{T}\Bigg(\log\det(R_t) + \operatorname{tr}\big(E[(Y_t - B_t X_t - r_t)(Y_t - B_t X_t - r_t)^T|Y_{1:T}]R_t^{-1}\big)\Bigg).\end{aligned} \tag{2.55}$$

From the Rauch-Trung-Streibel smoothing recursions, we can compute the expected value of ℓ' conditioned on the observations efficiently. Using the best linear predictor notation, and the property, $E[XY|Z] = E[(X - E[X|Z])(Y - E[Y|Z])^T] + E[X|Z]E[Y|Z]^T$, we note that,

$$\begin{aligned}E[X_t|Y_{1:T}] &= \hat{X}_{t|T}, \\ E[X_t X_t^T|Y_{1:T}] &= S_{t|T} + \hat{X}_{t|T}\hat{X}_{t|T}^T, \\ E[X_t X_{t-1}^T|Y_{1:T}] &= A_{t-1}S_{t-1|T} + \hat{X}_{t|T}\hat{X}_{t-1|T}^T,\end{aligned} \tag{2.56}$$

from which the conditional expectation may be computed. The EM iteration scheme becomes

$$\hat{\theta}_{k+1} = \operatorname{argmax}_\theta E[\ell'|Y_{1:T};\hat{\theta}_k](\theta), \tag{2.57}$$

where intermediate optimization steps may be performed using for example quasi-Newton methods. In the particular case when all matrices are constant in time and $q_t, r_t = 0$,

Shumway and Stoffer (1982) notes that the EM iteration scheme takes the form,

$$\begin{aligned}
\mu_1^{(s+1)} &= \hat{X}_{1|T} \\
S_1^{(s+1)} &= S_{1|T} \\
A^{(s+1)} &= \Big(\sum_{t=2}^{T} E[X_t X_{t-1}^T | Y_{1:T}]\Big)\Big(\sum_{t=2}^{T} E[X_{t-1} X_{t-1}^T | Y_{1:T}]\Big)^{-1}, \\
B^{(s+1)} &= \Big(\sum_{t=1}^{T} y_t E[X_t^T | Y_{1:T}]\Big)\Big(\sum_{t=1}^{T} E[X_t X_t^T | Y_{1:T}]\Big)^{-1}, \\
Q^{(s+1)} &= \frac{1}{T-1}\Big(\sum_{t=2}^{T} E[X_t X_t^T | Y_{1:T}] - A^{(s+1)} \sum_{t=2}^{T} E[X_{t-1} X_t^T | Y_{1:T}]\Big), \\
R^{(s+1)} &= \frac{1}{T}\Big(\sum_{t=1}^{T} y_t y_t^T - B^{(s+1)} \sum_{t=1}^{T} E[X_t | Y_{1:T}] y_t^T\Big),
\end{aligned} \tag{2.58}$$

where the expectations are computed using the current parameter estimates. The iteration scheme for A and B has a natural interpretation, namely as average projection matrices of X_t onto X_{t-1}, and Y_t onto X_t, using the full conditional distribution with the current parameter estimates.

The expected maximization method is appealing for a number of reasons when closed form updates for the parameters exists. However, for most state space processes, it might be hard to find those closed form updates. A quick informal comparison of ML estimation for the state space process (2.9) using direct maximization and EM, verifies that the EM approach is superior to direct maximization when closed form updates are available. When closed form updates are not available, and quasi Newton methods are used for both approaches, direct maximization seems superior.

2.4.1 Model selection

In the final part of this chapter we briefly note some important aspects of model selection related to linear Gaussian state space models.

Asymptotic properties of ML estimates

Under the hypothesis that the data is generated from the proposed model, in addition to some regularity conditions (see Hamilton (1994)), ML estimates are asymptotically multivariate normal, such that

$$\hat{\theta} \sim N(\theta, M^{-1}(\theta)), \text{ as } T \to \infty, \tag{2.59}$$

with T the number of observations, θ the true parameter value, and M the Fisher information matrix as defined in (2.52). The covariance may be approximated by evaluating the negative inverse Hessian of the log likelihood at the ML estimate, and in turn be used to estimate parameter uncertainty. It is important to point out that inferring parameter significance using this asymptotic distribution is not necessarily valid if the null hypothesis

lies on the boundary of the parameter space. However, for finite sample sizes, simulation based methods may be employed to study the finite sample distribution of the parameter estimates under any hypothesis on the true parameter values.

Akaike information criterion

A much applied criteria in model selection is the Akaike information criterion,

$$\mathrm{AIC} = 2k + 2\ell(\hat{\theta}), \tag{2.60}$$

with k the number of estimated model parameters, and $\ell(\hat{\theta})$ the negative log likelihood evaluated at the ML estimate. Note that the criteria decreases with increasing likelihood, and increases with model complexity. Hence, we seek a model minimizing the AIC criteria. A thorough motivation for minimizing the criteria may be found in Akaike (1974).

Diagnostics

The Gaussian assumption may be wrong, and in order to verify if it is reasonable, we note that under the assumption that the proposed model generated the data, the scaled innovations,

$$\Delta_t^{-1/2} I_t \sim N(0, I), \tag{2.61}$$

where $\Delta_t^{-1/2}$ is the inverse square root innovation variance matrix. Hence, for a given set of data, we expect the collection of the mT in total entries of scaled innovations to be independent and standard Gaussian distributed. The scaled innovations may be approximated by the innovations and variances computed by the model when using the ML parameter estimates. The distribution of the resulting sample may be studied using test of normality (e.g. the Anderson Normality test, Q-Q plots), and direct inspection. However, these approaches may not reveal possible time dependencies of the residuals; to check for this, the residuals should be plotted against time, and their autocorrelation function inspected.

Chapter 3

Soil Cable System

Before we develop a stochastic model which suits the problem, we present the underlying deterministic heat flow problem, and note some of its characteristic properties.

3.1 Model

We are concerned with the 2-dimensional heat problem,

$$\begin{cases} u_t - \nabla \cdot (k \nabla u) = f, & x \in \mathbf{R} \times (-\infty, 0),\ t > 0, \\ u(t,x)|_{x_2=0} = h(t), & \text{BC}, \\ u(0,x) = u_0(x), & \text{IC}, \end{cases} \tag{3.1}$$

where $u(t,x)$ is the temperature of the soil, and $f(t,x)$ is the source term, due to cables passing through this cross section. The temperature of the ground surface is $h(t)$, and is located at $\{x \in \mathbf{R}^2,\ x_2 = 0\}$. The diffusion coefficient, k, may be expressed κ/c with κ being the thermal conductivity, and c the volumetric thermal capacity. The source term, f, may be expressed g/c with g being the actual heat loss per volume per time unit. The problem (3.1) may readily be solved by finite difference or finite volume methods. However, in capturing the radial heat flow around the sources, and in incorporating arbitrary measurement locations, a high resolution discretization might be required. We may obtain reasonable results by making some further simplifications.

We initially assume that the cables may be modelled as point sources, and hence that $f = \sum_i f_i$, $f_i = a_i(t)\delta(x - x_i)$, with $\delta(.)$ the Dirac delta distribution and x_i the location of source i. In the simplified case when the thermal diffusion coefficient is constant in space, the problem simplifies to the linear inhomogeneous heat equation,

$$\begin{cases} u_t - k\Delta u = f, & x \in \mathbf{R} \times (-\infty, 0), t > 0, \\ u(x,t)|_{x_2=0} = h(t), & \text{BC}, \\ u(x,0) = u_0(x), & \text{IC}. \end{cases} \tag{3.2}$$

A solution to (3.2) may be found by summing solutions of the problems,

$$\begin{cases} u_t - k\Delta u = f_i, & x \in \mathbf{R}^2, t > 0, \\ u(0,x) = u_0(\|x - x_i\|_2), & \text{IC}, \end{cases} \tag{3.3}$$

and,

$$\begin{cases} u_t - k\Delta u = 0, & x \in \mathbf{R} \times (-\infty, 0), t > 0, \\ u(t,x)|_{x_2=0} = h(t), & \text{BC}, \\ u(0,x) = u_0(x_2), & \text{IC}, \end{cases} \tag{3.4}$$

provided $u(0,x)$ is expressable as a linear combination of the initial conditions of (3.3) and (3.4). Moreover, note that the solution to (3.3) varies only in the radial direction. That is, we may reduce the problem to the radial heat equation,

$$\begin{cases} u_t - \frac{k}{r}u_r - ku_{rr} = a_i(t)\delta(r), & r > 0, t > 0, \\ u(0,r) = u_0(r), & \text{IC}, \end{cases} \tag{3.5}$$

while the problem (3.4) may be reduced to the one dimensional problem,

$$\begin{cases} u_t - ku_{ll} = 0, & l > 0, t > 0, \\ u(0,t) = h(t), & \text{BC}, \\ u(l,0) = u_0(l), & \text{IC}. \end{cases} \tag{3.6}$$

Suppose we have n sources and that the solution to (3.3) may be expressed $u_i^{(s)}(t,x)$, while the solution to (3.4) is expressed $u^{(b)}(t,x)$. Furthermore, we denote the solution of (3.3) with a source at $\tilde{x_i}$, by $\tilde{u}_i^{(s)}(t,x)$, where $\tilde{x}_i$ is equal to x_i but with opposite sign of the second coordinate. Then, the solution to (3.2) with certain restrictions on the initial condition, may be expressed,

$$u = u^{(b)}(t,x) + \sum_i u_i^{(s)}(t,x) - \sum_i \tilde{u}_i^{(s)}(t,x). \tag{3.7}$$

Note that the $\tilde{u}_i^{(s)}$ terms cancel out the contribution of the radial problems at the boundary, so that only the vertical problem, $u^{(b)}$, contributes to the boundary condition.

We may approximate the solutions the 1-dimensional problems (3.3) and (3.4) by solving the system of ODEs obtained by either, discretizing the derivatives in space using finite difference approximations, or, use finite volume methods with the original conservation laws on integral form. In the particular case when finite difference methods are used for the derivatives in space, suppose we discretize the domain around a source uniformly radially, with increments d. For the radial problem (3.5), we then obtain,

$$\begin{aligned} \dot{u}_1 &= 2\frac{k}{d^2}(u_2 - u_1) + \frac{a(t)}{d^2}, \\ \dot{u}_i &= \frac{k}{d^2}\left(\frac{r_i - d/2}{r_i}u_{i-1} - 2u_i + \frac{r_i + d/2}{r_i}u_{i+1}\right), \quad i = 2, \ldots, n_s, \end{aligned} \tag{3.8}$$

where the inverse factor of d^2 in the source term reflects the thermal capacity of the inner soil volume. For the vertical problem, (3.4), discretizing uniformly with increments d yields,

$$\begin{aligned} u_1 &= h(t), \\ \dot{u}_i &= \frac{k}{d^2}\Big(u_{i-1} - 2u_i + u_{i+1}\Big), \quad i = 2, \dots, n_b. \end{aligned} \tag{3.9}$$

Note that we have obtained two additional boundary conditions, namely u_{n_s+1} for the radial problem, and u_{n_b+1} for the vertical (the former may always taken to be zero).

Boundary conditions

The boundary condition at the soil surface, $l = 0$ in problem (3.6), is the interface between air and soil. It follows that $h(t)$, the soil temperature at the boundary, is not known to us, although varies with for example the air temperature and radiation just above the surface. In order to model it we employ a finite volume approximation based on the original conservation law in integral form,

$$\partial_t \int_{(-\epsilon,\epsilon)} u dl = j(t,-\epsilon) - j(t,\epsilon) + \int_{(-\epsilon,\epsilon)} q(l) dl \tag{3.10}$$

for some $\epsilon > 0$ where $j(.)$ is the heat flux. The last term in (3.10) represents heat inflow due to radiation. We assume that the source density, q, may be expressed $r(t)\delta(l)$, so that the integral is always equal to $r(t)$. In our case, $j(-\epsilon) = -\rho u_l(-\epsilon)$, and $j(\epsilon) = -k u_l(\epsilon)$. Setting $\epsilon = d/2$ and using central differences for the space derivatives yields the approximate relation, where $g(t)$ is the air temperature,

$$\dot{u}_1 = \frac{\rho}{d^2}(g(t) - u_1) + \frac{k}{d^2}(u_2 - u_1) + \frac{r(t)}{d}, \tag{3.11}$$

which replaces the equation for u_1 in the scheme (3.9).

A natural form of the radiation term, $r(t)$, where t is given in hours, is,

$$r(t) = -\mu_1 + \gamma^{\chi(\text{cloudy})}\big(\mu_2 + \mu_3 \cos^2\left(\pi \frac{t+\delta_3}{365 \cdot 24}\right)\big) \cos^2\left(\pi \frac{t+\delta_2}{24}\right), \tag{3.12}$$

with $\mu_1, \mu_2, \mu_3 > 0$, and $\chi(.)$ the indicator function. The first term is due to radiation out from the soil. It is kept constant for simplicity, although it depends on the surface temperature in reality. The term μ_2 is the strongest radiation at the time of year when radiation is weakest, while the term μ_3 is the difference between μ_2 and the radiation at its overall strongest. In the next chapter, we will be working with measurements from Tronsholen-Skeiane. Here, radiation is strongest some time late in June. The hourly measurements we have available start at 02-07-2015, 18:00, and we assume a maximum of the radiation onto the ground surface at 22-06, 13:30, so that $\delta_2 = 4.5, \delta_3 = 10 \cdot 24$. It is important to recognize that the effect of radiation depends highly on the presence of clouds. Whenever clouds are present, we should scale the radiation inflow by a factor of γ.

However, this does not really give an accurate description of the actual heat flow at the boundary. The soil surface is exposed to wind, rain and snow, which complicates the above relationships. For example, on rainy days, we have significant heat contributions due to convection. That is, water with a certain temperature rains down and enters the soil. These complicated relationships are only included in the sense that we model the uncertainties they introduce into the simpler soil temperature model. This will be discussed further in Section 3.1.2.

Ideally, we would like to have the boundary condition, u_{n_b+1} of (3.9), as far away from the soil surface as possible, and with constant temperature, s, as this is most reasonable physically. However, this will possibly require a very high number of grid points in order to maintain a reasonable resolution scheme. This can in turn become computationally demanding. A solution to this problem is to use a variable resolution scheme, finer by the measurement devices and in their immediate vicinity, while coarser far below the soil surface, and keep s constant. Another solution is to keep the Dirichlet boundary condition at a relatively shallow depth, but allow it to change slightly with time, with yearly periods. In the proceeding we use a slightly varying Dirichlet boundary condition at lm below the ground surface, such that the soil temperature at the boundary becomes,

$$u_{n_b+1} = s(t) = \eta_1 + \eta_2 \cos^2\left(\pi \frac{t+\delta}{24 \cdot 365}\right), \tag{3.13}$$

for some auxiliary parameters, η_1, η_2, δ.

3.1.1 Source term, and extension to non-point sources

We initially made the simplifying assumption that the cables could be treated as points sources. In this case, all heat losses are in essence placed at the cable conductor. In reality, different losses occur at different part of the cable components, while the cable conductor usually ends up being the hottest. Although, not considering the distribution of heat losses in the cable by treating cables as point sources, could give inaccurate and conservative results. Cable heat dynamics are in fact very complicated, and we limit ourselves to just noting the most relevant heat losses. There are three types of heat losses we concerns ourselves with, and they are displayed in Figure 3.1.

Conductor losses

The first and most important type of loss, is the conductor loss, q_c. They are due to the electrical resistance of the cable conductor, and the applied current. This loss may be expressed at the electric resistance times the squared current. In the case of direct current, the electrical resistance is approximately a linear increasing function of the conductor temperature, while in the case of alternating currents, the resistance increases, and its relationship with temperature generally becomes very involved.

Sheath losses

As the cable sheath is also usually made of metal, the conductor current induces a current in these parts of the cable as well, due to magnetic forces. It follows that we get a heat loss

in the cable sheath, q_s. It turns out that this loss may be modelled as being proportional to the conductor losses, with proportionality constant dependent on the temperature of the sheath.

Dielectric losses

The last type of heat loss we present is the dielectric loss, q_d, which occur over the cable dielectric. It is due the electric field trying to rotate the dipoles of the dielectric material. A detailed description may be found in Ilstad (2009). For all practical reasons, it is constant, and usually negligible.

Nonlinear, non point source, cable heat dynamics

In the simple temperature model displayed in Figure 3.1, the radial finite volume representation gives rise to a nonlinear system of ODEs for the temperatures in the volume, $u = (u_1, u_2)$. This may be expressed,

$$\begin{gathered}\dot{u} = Au + B(u)v(t) + b(t),\\ A = \begin{pmatrix} -a_{1,2} & a_{1,2} \\ a_{2,1} & -a_{2,1} - a_{2,3} \end{pmatrix}, B(u)v(t) = \begin{pmatrix} q_c/C_1 \\ q_s/C_2 \end{pmatrix}, \; b(t) = \begin{pmatrix} 0.5q_d/C_1 \\ 0.5q_d/C_2 + a_{2,3}u_3 \end{pmatrix},\end{gathered} \tag{3.14}$$

with $v(t) \in \mathbf{R}$ the square of the applied current at time t, $a_{i,j} = (T_{i,j}C_i)^{-1}$, with $T_{i,j}$ the thermal resistance between volume i and j, and C_i the thermal capacity of layer i. The losses, depending on u, may be expressed,

$$\begin{aligned} q_c &= R_{AC}(u_1)v(t),\\ q_s &= \lambda(u_2)q_c, \end{aligned} \tag{3.15}$$

where R_{AC} is the alternating current resistance of the cable conductor. We will not go into details on the alternating current resistance and its dependence on temperature, as this is not that relevant for this study.

The nonlinear cable temperature model described above (or a higher resolution system constructed in the same way), may be concatenated with a numerical solution of the linear heat flow problem,

$$\begin{cases} u_t - \frac{k}{r}u_r - ku_{rr} = 0, & r > r_{\text{cable}}, t > 0,\\ u(t, r_{\text{cable}}) = u_{\text{cable}}(t), & \text{BC},\\ u(0, r) = u_0(r), & \text{IC}, \end{cases} \tag{3.16}$$

where r_{cable} is the radius of the cable, which in turn may replace the problem (3.5) in order to obtain a more realistic model.

From the notions above, it is apparent that the cable heat dynamics are non linear in reality. However, when the temperature is limited to a small range of values, keeping the thermal properties of the system constant may be justified. This is the case for the data collected at Tronsholen Skeiane, which we will be studying in the next chapter. External

variables are much more relevant than the nonlinear dynamics in this case (in fact, the entire heat contribution from the cables are almost negligible). When the cable temperature varies greatly, and the linear assumption cannot be justified, we may instead locally linearize the solution. That is, set, $B(u(t)) \approx B(u(t_0)) + \partial_u B(u(t_0))(u(t) - u(t_0))$, such that

$$\begin{aligned} \dot{u}(t) = &\big(A + v(t)\partial_u B(u(t_0))\big)u(t) \\ &+ \big(B(u(t_0)) - \partial_u B(u(t_0))u(t_0)\big)v(t) + b(t), \end{aligned} \tag{3.17}$$

for $t > t_0$ and $\partial_u B \in \mathbf{R}^{n\times n}$ the Jacobian of B with respect to u.

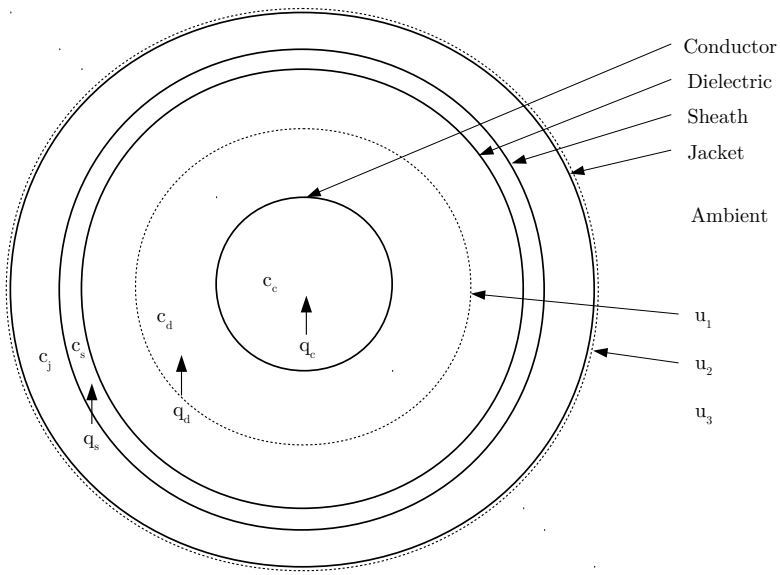


Figure 3.1: A simple temperature model for the cable. Finite volume layers are separated by dotted lines.

Dynamical system

The scheme illustrated above may be collected in the system of ODEs,

$$\dot{u} = M(t)u + b(t, u), \tag{3.18}$$

where $M \in \mathbf{R}^{(s\cdot n_s + n_b)\times(s\cdot n_s + n_b)}$ is tridiagonal, $b \in \mathbf{R}^{s\cdot n_s + n_b}$ contains the boundary conditions and the source terms, and s is the number of different sources. The matrix M may be constructed either by finite difference or finite volume methods, but in either case

it takes the form,

$$M_t = \begin{pmatrix} -m_{1,0} - m_{1,2} & m_{1,2} & 0 & \dots & 0 \\ m_{2,1} & -m_{2,1} - m_{2,3} & m_{2,3} & \dots & 0 \\ & & \vdots & & \\ 0 & & & & 0 \\ 0 & \dots & 0 & m_{n,n-1} & -m_{n,n-1} - m_{n,n+1} \end{pmatrix}, \tag{3.19}$$

with $m_{i,j} > 0$.

It follows from (3.7) that the modelled temperature is defined at all points in some domain, $[-l, l] \times [-2l, 0]$, where the contributions in the sum of the different solutions is the interpolated value of adjacent finite difference nodes' or volume's temperatures. In the case when we have three equal point sources, and a decaying vertical temperature profile, we obtain the contours displayed in Figure 3.2.

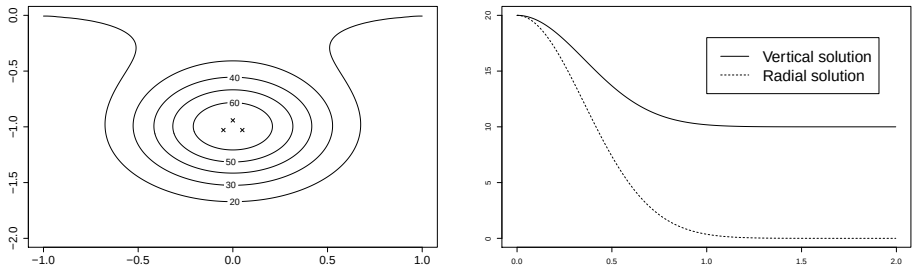


Figure 3.2: Contours of the solution, constructed by the sums of the solutions to the 1-dimensional problems. The sources are the crosses in the leftmost figure.

3.1.2 Linear stochastic system

In the proceeding, we study the system only for a small range of different temperatures, and therefore ignore the nonlinear dynamics of the heat losses. We may extend the dynamic model to be a stochastic linear system, by including noise in a natural way,

$$dU_t = M_t U_t dt + b_t dt + K_t Z_t dt, \tag{3.20}$$

for some noise process Z_t. When we are dealing with stochastic heat flow, we want to include the noise such that we do not break the underlying conservation laws, but instead represent heat fluxes which the deterministic model cannot explain. We may include some uncertainty in the flow of heat between adjacent volumes by letting K_t be a Laplacian matrix (also called Kirchoff matrix, owing to the discrete conservation law), such that,

$$(K)_{i,j} \propto -|\text{adj}(i)|\chi(i = j) + \chi(j \in \text{adj}(i)), \tag{3.21}$$

where $\text{adj}(i)$ is the set of nodes adjacent to node i, and the proportionality constant is positive. The discretized Laplacian, M, and the Laplacian matrix K, are not equal, since

the columns of M need not sum to zero due to the boundary conditions. Note that, heat flow is "contained within the system" provided all columns of K sum to 0.

An important property of the included noise is that it is correlated with its past values. That is, events which affect the heat flow in ways our model cannot explain, for example local soil moisture differences, lingers in time. The perhaps most clear suggestion is using Ornstein–Uhlenbeck processes as noise. Note that this may be implemented by expanding the state space,

$$\begin{gathered} dX_t = A_t X_t dt + q_t dt + C_t dB_t, \\ X_t = \begin{pmatrix} U_t \\ Z_t \end{pmatrix}, \quad A_t = \begin{pmatrix} M & K \\ 0 & H \end{pmatrix}, \quad q_t = \begin{pmatrix} b \\ 0 \end{pmatrix}, \quad C_t = \begin{pmatrix} 0 \\ D \end{pmatrix}, \end{gathered} \tag{3.22}$$

and setting $H = -\phi I, \phi > 0$, where B_t is a standard Brownian motion. The same effects also have dependencies in space. Including these dependencies may be achieved by constructing DD^T appropriately using a suitable covariance function.

Inflow boundaries

In order to allow for some random heat flow through the top boundary in an appropriate way, we may add another, independent and with its own set of parameters, Ornstein-Uhlenbeck process, L_t, to (3.11), scaled by an inverse factor of d to reflect the soil layer thermal capacity. This may be achieved with simple modifications to the matrices above. There should not be any random heat flow from far beneath the ground surface. There could however be an inflow boundary at the cable sources, since we do not have any knowledge of the true cable current between measurements. For simplicity, we ignore this effect.

3.1.3 Properties of the solution

The mean and variance of the system (3.22) may be computed by integrating the systems of ODEs (2.31) as described in Chapter 2. Although, we point out some of the properties of the solution under certain simplifying conditions. If we assume that all matrices and source terms are constant between time steps $t_{j+1} = t_j + \Delta t,\ j = 1, \ldots, k-1$, we may express $X_{t_{j+1}}$, given X_{t_j}, through the discrete dynamics equation,

$$X_{t_{j+1}} = \exp(A_{t_j}\Delta t) X_{t_j} + (\exp(A_{t_j}\Delta t) - I) A_{t_j}^{-1} q_{t_j} + V_{t_j}, \tag{3.23}$$

where $X_{t_1} \sim N(x_1, S_1)$, and $V_{t_j} \sim N(0, Q_{t_j})$. Q_{t_j} is found from the Lyapunov differential equation (2.31) with initial condition $S_{t_j} = 0$,

$$Q_{t_j} = \int_0^{\Delta t} \exp(A_{t_j} s) C_{t_j} C_{t_j}^T \exp(A_{t_j}^T s) ds, \tag{3.24}$$

which we may evaluate using either of the two methods proposed in Chapter 2. Without loss of generality, we may decouple the deterministic model for the temperature mean, $u(t)$,

$$\dot{u} = Mu + b; \quad u(0) = E[U_0], \tag{3.25}$$

and the linear stochastic system describing the uncertain heat flow,

$$\begin{aligned} d(U_t - u_t) &= M(U_t - u_t)dt + KZ_t dt, \\ dZ_t &= HZ_t + DdB_t, \end{aligned} \tag{3.26}$$

and only study the stochastic heat flow problem (3.26) in the following.

Stationary solution

When the system (3.26) is time invariant and stable, it has a unique stationary solution. The system matrix is negative definite provided M and H are negative definite, since A shares the eigenvalues of M and H (note that $\det(A) = \det(M)\det(H)$). Clearly $H = -\phi I$ is negative definite. Note that M, constructed as (3.19), is always negative semi definite owing to Greshgorin's disk theorem. In fact, if we view the discretized domain as a graph, M must be negative definite if every node is connected with a boundary node/condition. This is the case for our M (K on the other hand is negative semi definite, and has at least one zero eigenvalue). It follows that the covariance of $X_t = (U_t, Z_t)$ in (3.26) may be expressed,

$$E[X_{t+h}X_t^T] = \exp(Ah)S^*, \tag{3.27}$$

with $h > 0$ and $S^* = \int_0^\infty \exp(A\tau)CC^T \exp(A^T\tau)d\tau$, the steady state variance. The properties of the covariance function above is not easily interpreted, and we may instead study the properties of the one dimensional problem, $U_t \in \mathbf{R}$,

$$dU = mUdt + kZdt; \quad U(0) = 0, \tag{3.28}$$

where $m < 0$, and Z_t is the scalar stationary Ornstein-Uhlenbeck process with variance $\frac{d^2}{2\phi}$. U_t may in this case be expressed,

$$U_t = \int_0^t \exp(m(t-s))kZ_s ds, \tag{3.29}$$

so that, owing to the stationary covariance function of Z_t, as shown in Chapter 2,

$$E[U_{t+h}U_t] = \frac{d^2k^2}{2\phi}\int_0^{t+h}\int_0^t \exp(m(2t+h-s_1-s_2) - \phi|s_1 - s_2|)ds_2 ds_1, \tag{3.30}$$

with $h > 0$. Evaluating the integral and simplifying, we eventually find that,

$$\begin{aligned} E[U_{t+h}U_t] = e^{mh}\frac{d^2k^2}{2\phi}\Bigg(&\frac{1}{(m-\phi)m} + \frac{1}{(\phi-m)m}e^{2mt} \\ &+ \frac{2}{\phi - m}(e^{-(\phi+m)t} - 1)\frac{1}{\phi+m}e^{2mt}\Bigg) \\ &+ \frac{d^2k^2}{2\phi}\frac{1}{m^2-\phi^2}\Big(e^{-\phi h} - e^{mh} - e^{mt-\phi(t+h)} + e^{m(t+h)-\phi t}\Big), \end{aligned} \tag{3.31}$$

which it may be verified converges as $m \to -\phi$. Letting $t \to \infty$, we get,

$$E[U_{t+h}U_t] = \frac{d^2k^2}{2\phi}\frac{1}{(m^2-\phi^2)m}\big(\phi e^{mh} + me^{-\phi h}\big), \tag{3.32}$$

which is the stationary covariance function of U_t. Setting $h = 0$, $k = m$, the expression for the stationary variance becomes,

$$\frac{d^2}{2}\frac{m}{(m-\phi)\phi}, \tag{3.33}$$

which is monotonically increasing in $-m$ and bounded above by $\frac{d^2}{2\phi}$. In Figure 3.3 the stationary variance is plotted against $-m$ using $d^2 = 2, \phi = 1$. If we on the other hand set

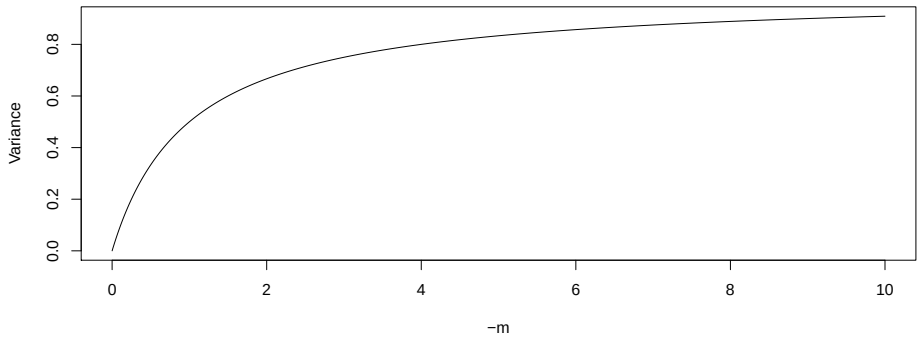


Figure 3.3: Variance of steady state.

$d = \phi$, and let $\phi \to \infty$, we get,

$$E[U_{t+h}U_t] = -\frac{k^2}{2m}e^{mh}, \tag{3.34}$$

which is a stationary Ornstein-Uhlenbeck process with parameter $m < 0$, and noise coefficient k, as we would expect.

In Figure 3.4, 3.5, 3.6 and 3.7, four realizations using different parameters of the stationary solution to (3.26) are displayed at some time t, along with the potential Z_t at the same time. U_t and Z_t are dicretized in the domain $[0,1] \times [0,1]$, and the potential Z_t is taken to be a zero mean Gaussian process with covariance function,

$$E[Z_t Z_s^T] = \frac{\sigma^2}{2\phi}\exp(-\phi|t-s|)DD^T, \tag{3.35}$$

where DD^T is constructed using a squared exponential covariance function, such that,

$$(DD^T)_{i,j} = \exp(-\omega\|x_i - x_j\|_2^2). \tag{3.36}$$

The parameters $\phi, \omega > 0$ are varied in the different figures, while the thermal diffusion coefficient and variance, k, σ^2, are fixed equal to 1. At each time t, the true heat potential in the domain is equal to $U_{t,x}+Z_{t,x}$. Intuitively, we expect $U_{t,x} \to -Z_{t,x}+$some constant, as $k \gg \phi$. The constant must ensure that $\int_{[0,1]\times[0,1]} U_{t,x} dV = 0$, as we have allowed for no inflow boundaries, and is therefore equal to $\int_{[0,1]\times[0,1]} Z_{t,x} dV$. This is apparent in Figure 3.6 and 3.7. Simulating from the stationary solution may be achieved by running the linear system (3.22) for some time.

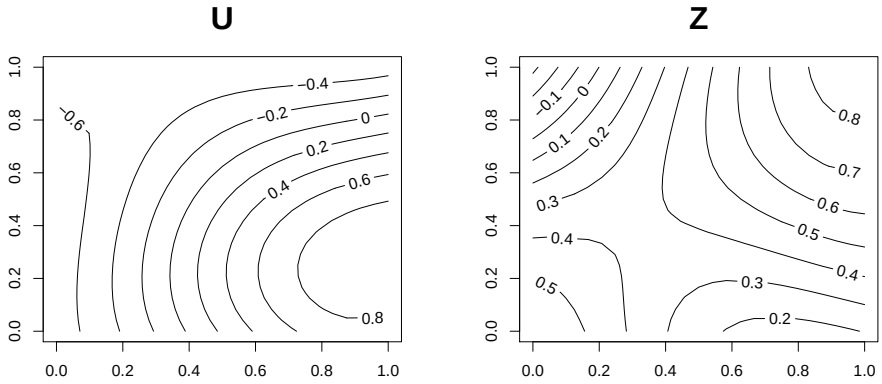


Figure 3.4: $\phi = 1, \omega = 1$: Stationary $U_{t,x}$, with $Z_{t,x}$ also displayed.

Reference model

To compare the proposed model, we use a different Gaussian process. We model the difference between the true temperature, $U_{t,x}$, and the deterministic modelled temperature, $u(t,x)$, by (3.2), as a stationary Ornstein-Uhlenbeck process. If we let $U_t, u_t \in \mathbf{R}^m$ be the true temperature and modelled temperature at the measurement locations $x_1, \ldots, x_m$, respectively, we can express $U_t - u_t$,

$$d(U_t - u(t)) = -\phi(U_t - u(t))dt + DdB_t; \quad U_0 - u_0 \sim N\left(0, \frac{DD^T}{2\phi}\right), \tag{3.37}$$

where $(DD^T)_{i,j} = \sigma^2\kappa(\|x_i - x_j\|_2)$, for some covariance function $\kappa(.,.)$.

3.2 Previous related work and the industrial standard

At the time of hand computations, Neher and McGrath (1957), outlined an iterative scheme for computing the load capacity of buried electric cables. This has more or less been the industrial standard used in determining the load capacity in steady state operation. The scheme corresponds to fixing the conductor temperature at its maximum tolerable value

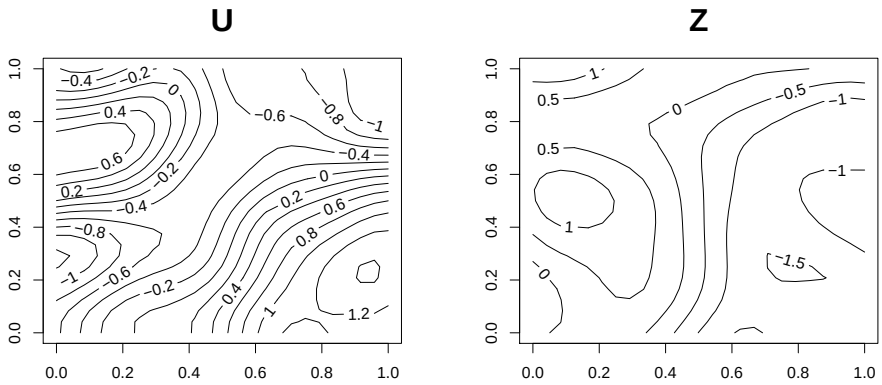


Figure 3.5: $\phi = 1, \omega = 10$: Stationary $U_{t,x}$, with $Z_{t,x}$ also displayed.

and the thermal properties of the surroundings at their worst case, and then compute the steady state solution of the cable soil system. This is the same as computing the steady state of the simplified temperature model (3.14), when the appropriate thermal capacities and resistances are used. The solution involves the constant applied current and the temperature of the other cable components; provided the steady state temperature of the other cable components are below their maximum value, the computed current is the cable load capacity. Further details may be found in IEC 60287.

When dealing with varying conditions, the industrial standard for dynamic current ratings IEC 60853 use the solution of the decoupled heat problem given in (3.7), when a varying load profile is applied. In this context, the model we have proposed is a stochastic extension of this method, where we model the error between the true temperature and the model by a Gaussian process. Some important cables are also equipped with temperature measurement devices, which fit very well with the model we propose.

Some further comments on model choice

Initially, it was attempted to include uncertainty in the thermal parameters directly. State and parameter inference were in this case performed using sequential Monte Carlo methods. It was recognized that, compared to parameter inference for linear state space processes, this was very inconvenient. Similarly to when including uncertainties in the thermal parameters (e.g. say that the diffusion coefficient is some stochastic process) the linear formulation (3.22) also allows modelling uncertain heat flow which do not break the conservation laws. Note that, in this problem, there is also travel of heat due to travel of moisture (convection), and so a model including uncertainty only in the thermal diffusion coefficient is not fully representative anyways, while the convection diffusion equation would possibly complicate the model unnecessarily.

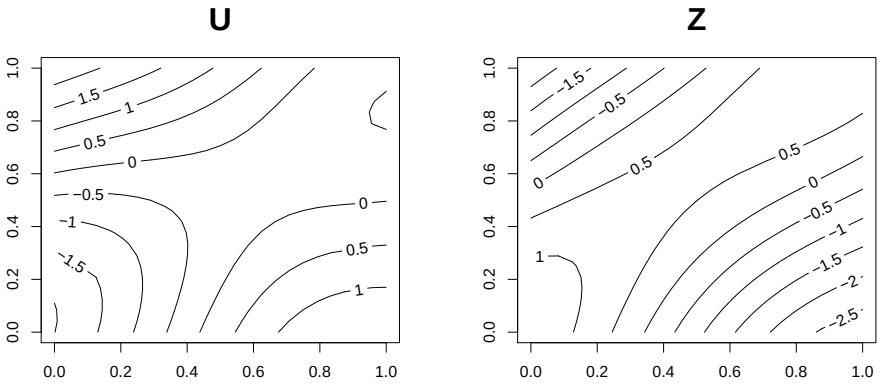


Figure 3.6: $\phi = 0.1, \omega = 1$: Stationary $U_{t,x}$, with $Z_{t,x}$ also displayed.

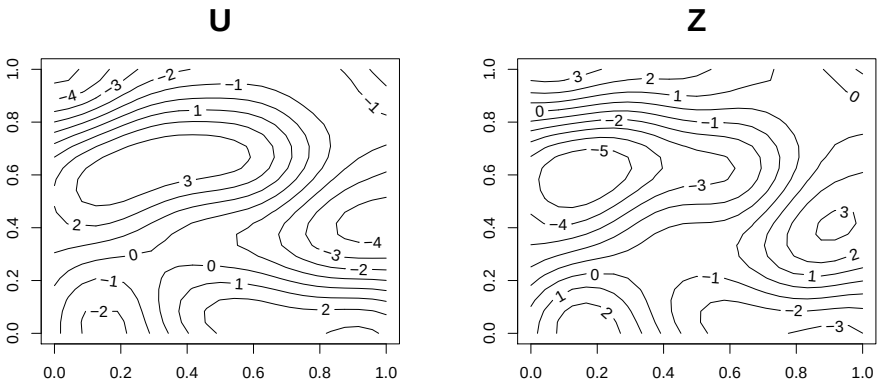


Figure 3.7: $\phi = 0.1, \omega = 10$: Stationary $U_{t,x}$, with $Z_{t,x}$ also displayed.

Chapter 4

Application to Tronsholen-Skeiane Measurements

In this chapter, we apply the proposed model to real soil heat measurements. We start by presenting the measurements and the measurement setup.

4.1 Tronsholen-Skeiane measurements

This section presents some of the data collected at Tronsholen-Skeiane. In Figure 4.1 the measurement setup at the location of Figure 1.1 is displayed. Along three probes, temperature and humidity measurements are made evenly distributed, at an hourly rate. The measurements are displayed in Figure 4.1 and 4.2. Note the clear correlation with air temperature and the temperature measurements, particularly those along the vertical probe. The correlation between the temperature of the measurement devices closest to the ground surface and the air temperature is 0.84, decreasing to 0.77, at the bottom of the vertical probe. The top vertical soil temperatures has a daily periodic component which decays as we move further down. For the locations closest to the cables (7,8), the daily periodicity is again apparent. This is due to the periodicity of the cable load; power consumption has a 24 hour period, and peaks during popular household dinner times.

It is worth pointing out that there is a clustering in the vertical temperatures. This is perhaps due to the fiber sheath or some other installation feature. Furthermore, the temperature of the top soil layers seem very stable for a longer period of time around January. This period was subject to a lot of snowfall, which we note from the soil moisture measurements in Figure 4.2, and the fact that the top soil temperature is fixed and equal to zero. Even though the air temperature varies greatly, the soil temperature remains seemingly unchanged. It is likely isolated from heat contribution from the ground surface by snow. Studying the horizontal temperature measurements, we find that there is very little variation in temperature in the horizontal direction on the other hand.

In Figure 4.2 the current in the middle cable during this period is displayed. There were

no current in the rightmost cable during this period, while the leftmost cable had a similar current profile. In the same figure, thermal conductivity and soil humidity measured just above the middle cables are displayed, in addition to the rainfall during the period. There is a clear relationship between thermal conductivity and soil moisture content. The soil's ability to transport heat increases with its moisture content, and the effect lingers. This is due to the soil generally becoming more compressed after being exposed to moisture. Note that, even though humidity increases thermal conductivity, it also increases the soil's thermal capacity, and so the effect of humidity on the thermal diffusion coefficient is a bit more involved.

Measurement setup

In the proceeding we assume that the center of the trefoil cable installation in Figure 4.1 is $1m$ below the ground surface. Setting the vertical coordinate to 0 at the ground surface, the horizontal coordinate to 0 at the center of the trefoil installation, it follows that the bottom of the cable trench has vertical coordinate $-1-(1+\frac{1}{\sqrt{3}})r \approx 1-1.58r$, where $r = 0.05m$ is the radius of the cable. In Figure 4.3 the middle part of the measurement setup is displayed. The location of the measurement devices along the vertical probe is found by matching the 15'th measurement location with that of the 8'th; the correlation between measurement location 8 and 15 are 0.9993. We obtain,

$$x_{\text{vertical}}^{(8-i)} = (-2r, 0.0865 + 0.1i - 1 - 1.58r), \quad i = 0, \dots, 7, \tag{4.1}$$

while the second coordinate of the horizontal devices are $-1+(1+\frac{2}{\sqrt{3}})r \approx -1+2.15r$, and so the locations are approximately,

$$x_{\text{horizontal}}^{(8-i)} = (0.1i, 2.15r - 1), \quad i = 0, \dots, 7. \tag{4.2}$$

The cable centers of the middle are located at

$$\begin{aligned} x_{\text{cable}}^{(1)} &= (-r, -1 - 0.58r), \\ x_{\text{cable}}^{(2)} &= (0, -1 + 1.15r), \\ x_{\text{cable}}^{(3)} &= (r, -1 - 0.58r). \end{aligned} \tag{4.3}$$

There are, however, uncertainties associated with these coordinates. They are guesses based on no knowledge other than Figure 4.1.

Model setup

For the sake of not introducing too much systematic error into our results, and since the horizontal heat flow is seemingly not very large, we limit ourselves only to consider the vertical measurement devices. Furthermore, we ignore the leftmost and rightmost cables, as they are too far from the measurement devices to have any significant impact on the measured temperature when the cable load has been this low. That leaves the middle trefoil installation to be included, which consists of three equal cables with identical load profiles. It follows that we get a system consisting of one vertical and one radial solution.

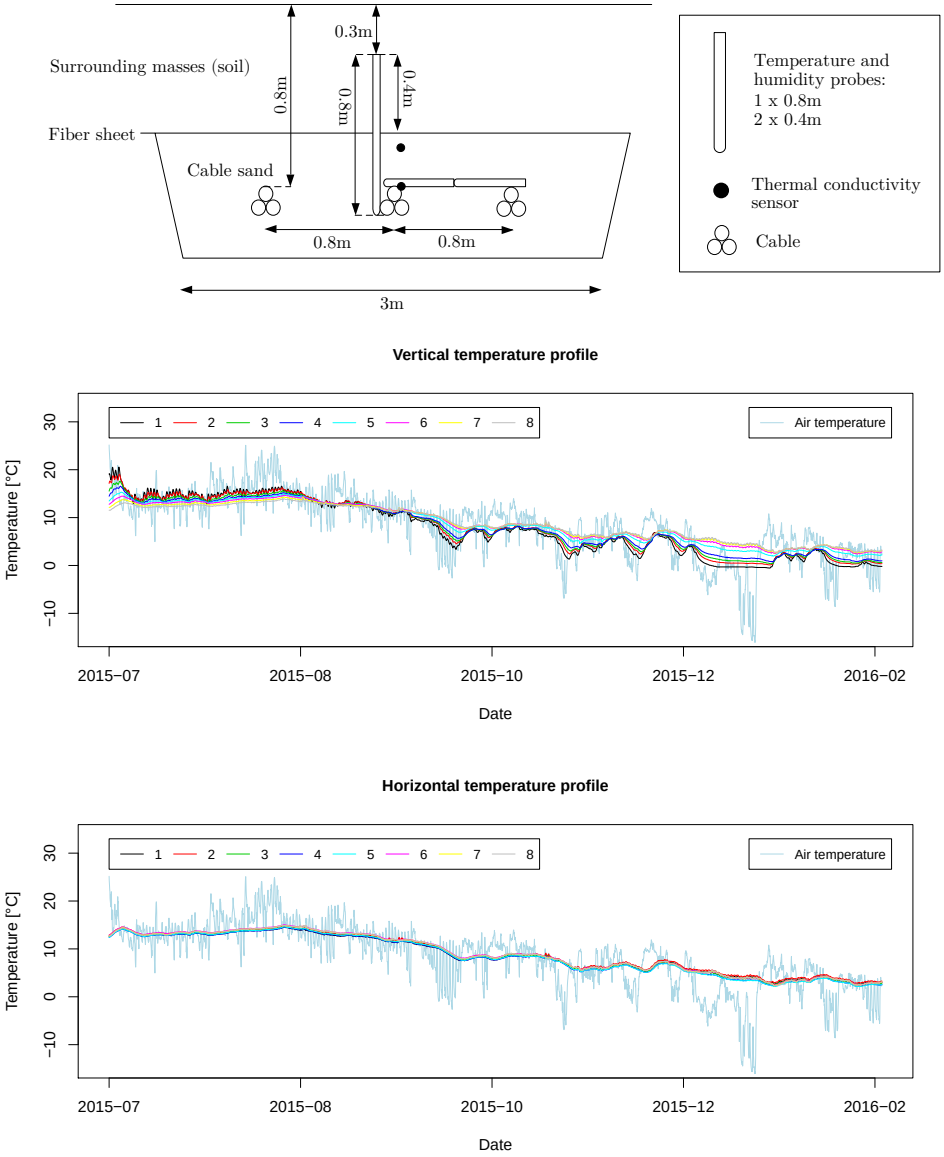


Figure 4.1: Measurement setup with some of the measurements collected at Tronsholen-Skeiane. The measurements devices are numbered from top to bottom and right to left, respectively. Source: SINTEF Energi.

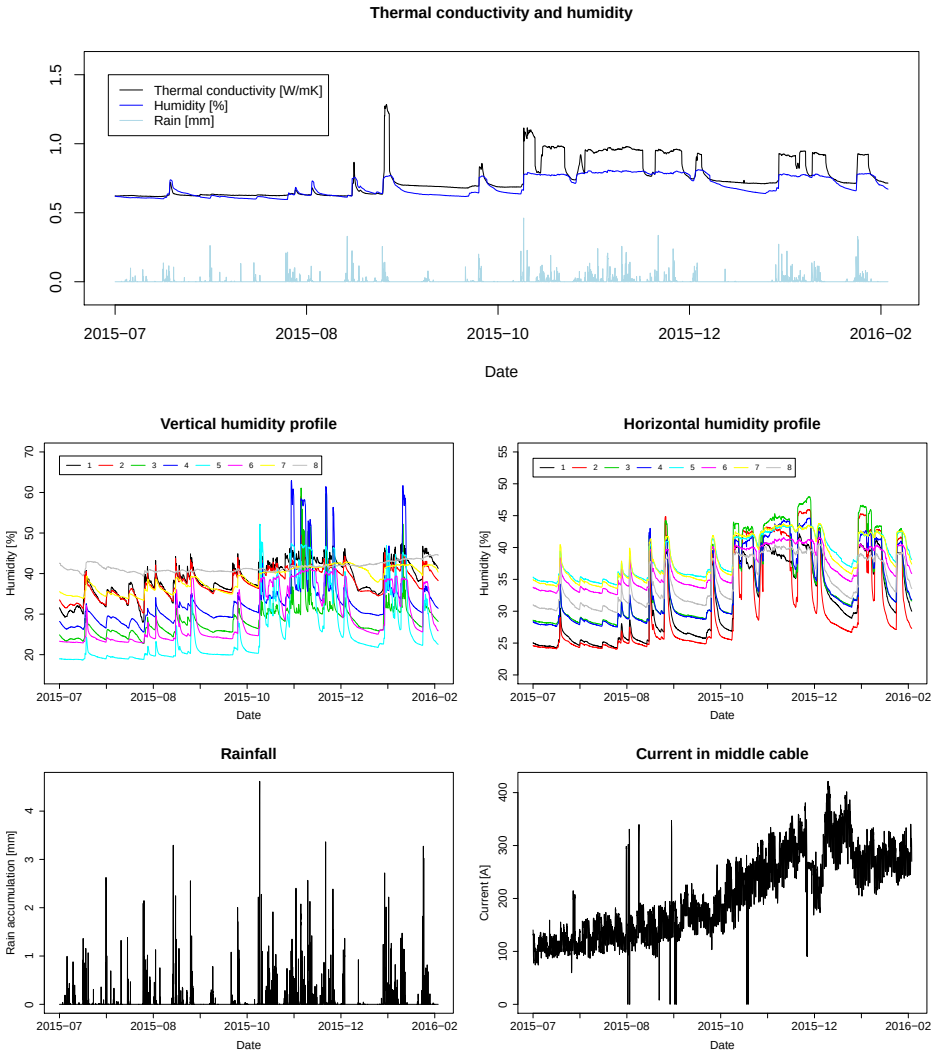


Figure 4.2: Thermal conductivity, rainfall and applied current at the Tronsholsen-Skeiane cable field.

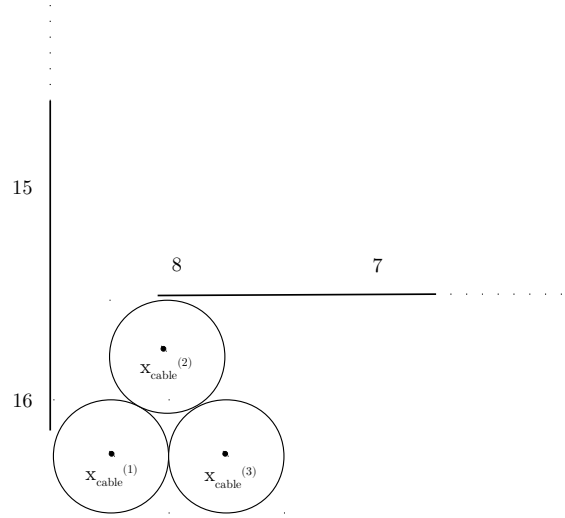


Figure 4.3: The middle part of the trefoil installation. The measurement locations along the probes are numbered from 1 to 16.

For simplicity, we ignore random heat flow in the radial direction, so that the stochastic part of the system (3.26) is only due to vertical heat flow, while the deterministic part (3.25) takes into account both the vertical heat flow, and radial.

It follows that we have a mean parameterized by the thermal diffusion coefficient, β (we have replaced k to consistently use greek letters for parameters), and the soil air boundary flow parameter ρ, in addition to the cable heat loss per squared kilo Ampere, α. That is, we assume for simplicity that the cable loss term is proportional to the squared current, with proportionality constant α. In terms of the notation used in (3.22), we set,

$$d^2(M)_{i,j} = \begin{cases} -\beta - \rho, & i, j = 1, \\ -2\beta, & i = j, \\ \beta, & j = i + 1, \text{ or } j = i - 1, \end{cases} \tag{4.4}$$

where d is the distance between grid nodes. The Kirchoff matrix $K \in \mathbf{R}^{n \times n}$ is equal to M except at the top left and bottom right entries; to avoid introducing random heat flow into the system, we must set,

$$\begin{aligned} d^2(K)_{1,1} &= -\beta, \\ d^2(K)_{n,n} &= -\beta. \end{aligned} \tag{4.5}$$

This choice of K maintain the interpretation of $Z_{t,x}$ as the error in the heat potential.

Furthermore, we must concatenate a column vector to the right side of the system matrix, and a row vector to the bottom, in order to allow an independent Ornstein-Uhlenbeck process determine the random heat flow into the system. The column vector has top entry $1/d$, and bottom entry $-\phi_2$, while all other entries of the vectors are zero.

In the following we construct the spatial covariance of $Z_t \in \mathbf{R}^n$ using two different covariance functions. A squared exponential, such that

$$(DD^T)_{i,j} = \sigma_1^2 \exp(-\omega d^2 (i-j)^2), \tag{4.6}$$

and the Ornstein-Uhlenbeck covariance,

$$(DD^T)_{i,j} = \sigma_1^2 \exp(-\omega d |i-j|), \tag{4.7}$$

for $i, j = 1, \ldots, n$, and parameters $\sigma_1^2, \omega > 0$. Gaussian processes with squared exponential covariance function are infinitely mean square differentiable, and so the interpretation of $Z_{t,x}$ as the error in the heat potential should be meaningful also as the scheme resolution becomes very fine (see Abrahamsen (1997)). The Ornstein-Uhlenbeck covariance function does not have this property however, but we include it for the sake of comparison (the linear stochastic system is nevertheless always well defined). Similarly to the system matrix, we must add a row and column to the matrix CC^T in (3.22) in order to include the independent Ornstein-Uhlenbeck process at the soil air boundary. The bottom right entry of which then becomes σ_2^2. Finally, we set,

$$H = -\phi_1 I \in \mathbf{R}^{n \times n}, \quad R = \sigma_3^2 I \in \mathbf{R}^{8 \times 8}. \tag{4.8}$$

In summary, we have a state of dimension $2n+1$, containing $U_t \in \mathbf{R}^n$ the true vertical temperatures, $Z_t \in \mathbf{R}^n$ the error in the heat potential, and $L_t \in \mathbf{R}$, the independent Ornstein-Uhlenbeck process used as noise at the soil air boundary. The parameters of the latter process are $\sigma_2^2, \phi_2 > 0$. In comparing the model proposed above, we have a "general purpose" Gaussian process, as noted in Chapter 3. It has the same parameterized mean as the proposed model, but is just a simple Ornstein-Uhlenbeck process with different choices of spatial covariance.

4.1.1 Synthetic example

Before we start working on the real measurements we illustrate the properties of the proposed model with a synthetic example. It is relatively simple, ignoring any random heat flow from the top soil layers, and radiation terms. The parameter values used in the synthetic example are displayed in Table 4.1. In Figure 4.4 we have simulated a realization of the soil temperature using the two different spatial covariance functions. The temperatures are from the top soil layer (1), down to a depth of $1.5m$ (8), where we have used the actual air temperature and cable load measurements from Tronsholen-Skeiane. Even though the choice of parameter values may seem random at first glance, their values will become more clear as we move on to the real measurements in the next section. For the moment, it suffices to treat them without any frame of reference.

In Figure 4.5 the filtered (first 3000 observations) and forecasted (remaining) estimates at the vertical temperature measurement devices are displayed, using squared exponential

spatial covariance. As the forecasts run for some time, we note that the process tends to its stationary distribution. In which case, the variance is the steady state variance, and the forecasted mean is just the deterministic modelled mean by (3.25). From the forecasts at the 7. measurement device, it is possible to spot a slight heat contribution from the cables in the form of a small periodic component.

Synthetic test example parameters		
Parameter	Value	Unit
β	$3. \times 10^{-3}$	m^2h^{-1}
σ_1^2	$3. \times 10^{-3}$	$^\circ C^2h^{-2}$
ϕ	$3. \times 10^{-3}$	1
σ_3^2	$5. \times 10^{-5}$	$^\circ C^2$
ρ	$1. \times 10^{-3}$	m^2h^{-1}
ω	$10.$	m^{-1}
η_1	$5.$	$^\circ C$
η_2	$5.$	$^\circ C$
δ	-1000	h
α	$2. \times 10^{-2}$	$^\circ C(kA)^{-2}h^{-1}$

Table 4.1: Parameter values for the synthetic example.

Distribution of parameter estimates

In Figure 4.6, 2000 samples of ML estimates computed by repeated simulation and estimation using the parameters displayed in Table 4.1 and a squared exponential spatial covariance are displayed. The samples are plotted against their asymptotic distributions given by (2.59), where the Fisher information matrix has been approximated by the Hessian evaluated at an ML estimate using a finite difference approximation. The correspondence between the finite sample distribution and asymptotic distribution is good, owing to the fact that the simulated time series are long. We have set $T = 5814$, and used the external variable measurements from Tronsholen-Skeiane.

4.1.2 Real data

The aim of this section is to check if the proposed model is an appropriate model for soil heat dynamics. In order to do so, we study how different scheme resolutions and depths of boundary conditions affect the model when real soil temperature measurements are used. In addition, the results are compared with the reference model.

Model comparison, effect of scheme resolution, depth, and spatial covariance

In Table 4.2 we have estimated the parameters of the model with squared exponential spatial covariance using the hourly temperature measurements for two different choices of model depth (l), $1.5m$ and $3m$, and three different numbers of grid nodes (n), $20, 30, 40$ (only the latter two for the deepest model). The radial problem has fixed maximum radi

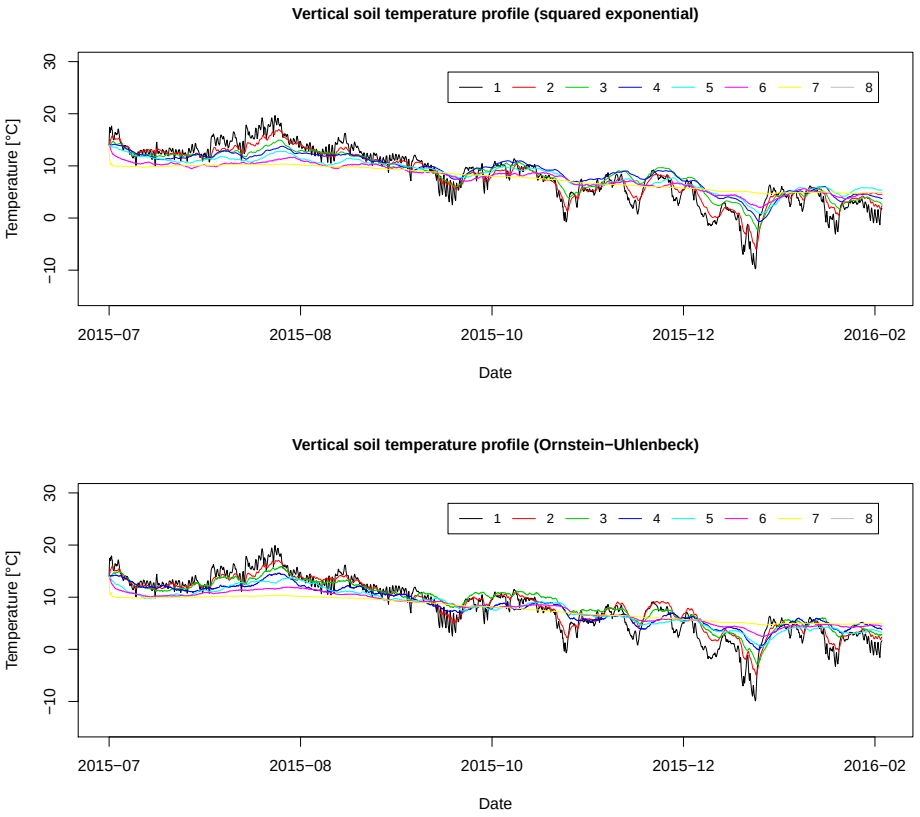


Figure 4.4: Synthetic example: A realization of the vertical soil temperature profile.

equal to $0.5m$, outside of which the temperature is set to zero, and uses the same number of grid nodes. Increasing the number of grid nodes should then give a more accurate model for the mean and covariance. Increasing the vertical depth should allow for a more accurate physical model. In Table 4.3, the same parameter estimates are displayed for the Ornstein-Uhlenbeck spatial covariance model. Table 4.4 displays the estimates of the reference model using $1.5m$ depth and 20 grid nodes. For all models, the initial condition has mean $14°C$ for the vertical problem, and $0°C$ for the radial. For simplicity, we have used a diagonal matrix with entries 4 as initial variance for the proposed model, while we have used the steady state variance as initial condition for the reference model.

Interpretation of estimated parameters

From the analysis of the measurements we pointed out that the thermal diffusion coefficient varies with soil humidity. It also depends on type of soil; if it is sandy, or coarse, all of which varies around the cables. In fact, typical values of soil thermal diffusion coeffi-

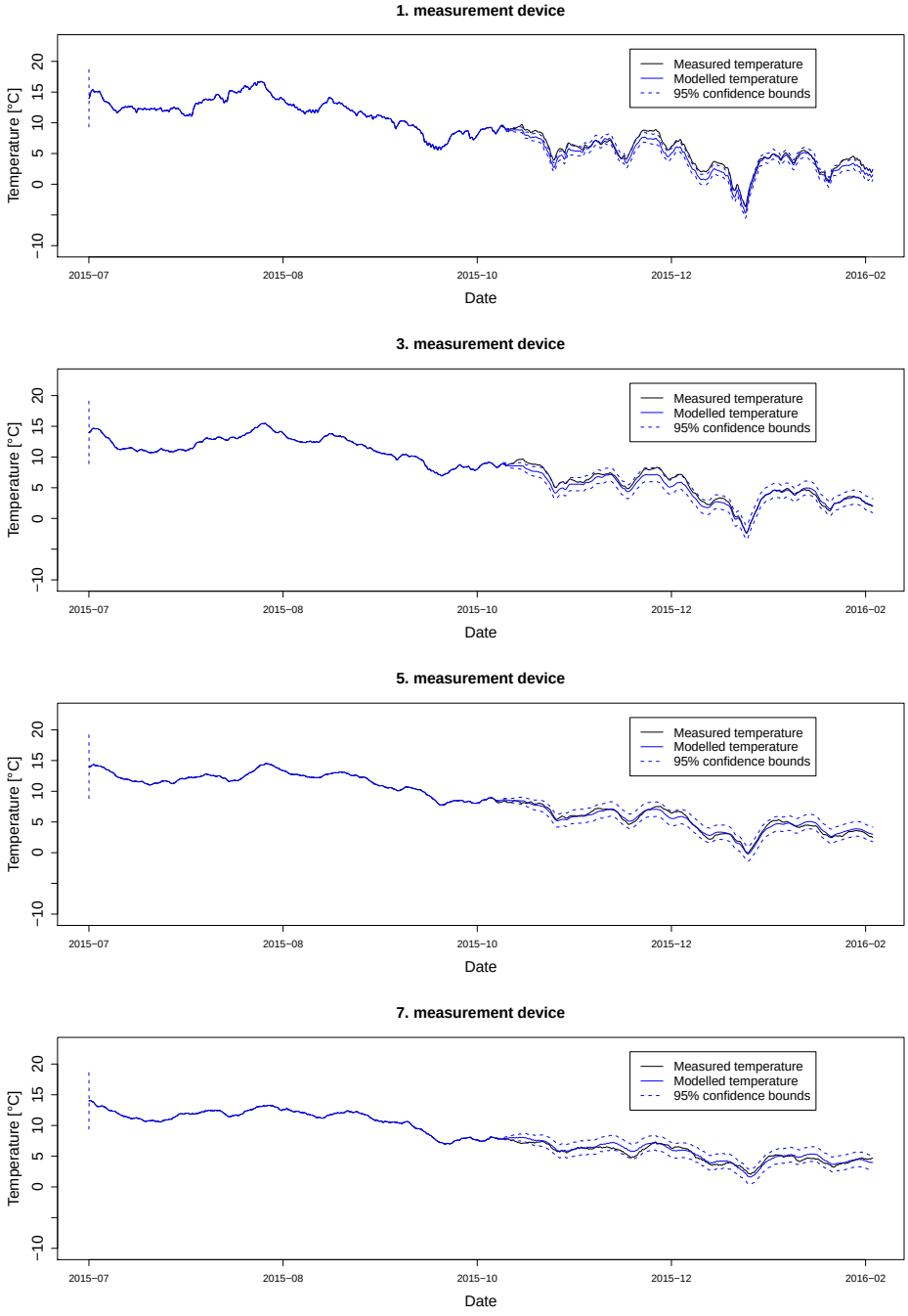


Figure 4.5: Synthetic example: Filtered and forecasted soil temperatures at the measurement locations.

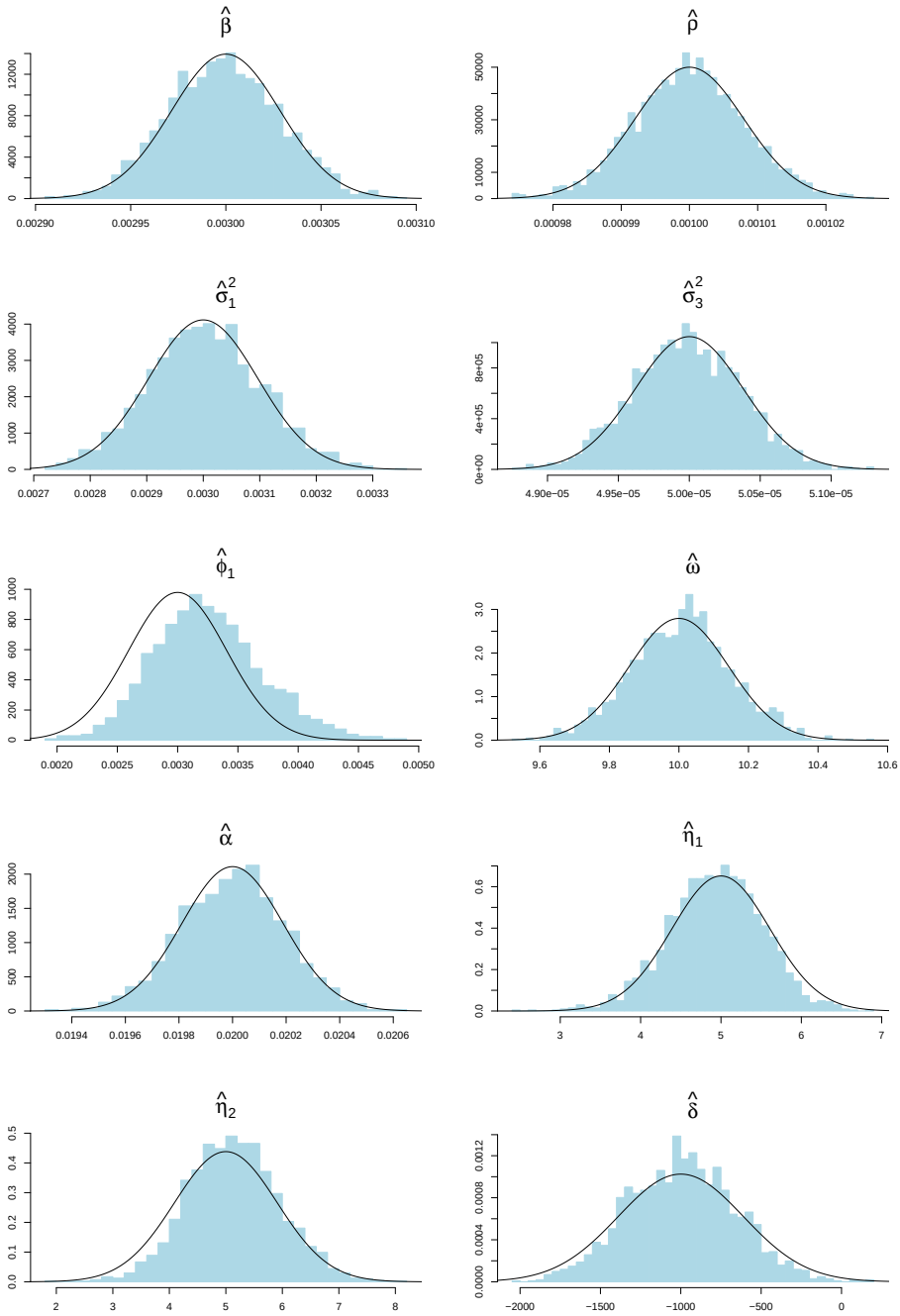


Figure 4.6: Synthetic example: Sample of 2000 ML estimates on simulated measurements, compared with their asymptotic distribution (black line).

Proposed model (squared exponential spatial covariance)					
l	$1.5m$			$3m$	
n	20	30	40	30	40
$\hat{\beta}$	2.63×10^{-3}	2.45×10^{-3}	2.49×10^{-3}	2.70×10^{-3}	2.53×10^{-3}
$\hat{\sigma}_1^2$	8.79×10^{-3}	8.35×10^{-3}	8.19×10^{-3}	8.61×10^{-3}	8.32×10^{-3}
$\hat{\phi}_1$	3.44×10^{-3}	4.63×10^{-3}	4.55×10^{-3}	6.69×10^{-3}	5.16×10^{-3}
$\hat{\sigma}_3^2$	8.19×10^{-5}	7.96×10^{-5}	8.00×10^{-5}	7.54×10^{-5}	8.02×10^{-5}
$\hat{\rho}$	4.36×10^{-3}	2.42×10^{-3}	1.21×10^{-3}	6.55×10^{-3}	5.60×10^{-3}
$\hat{\omega}$	22.91	23.13	21.68	30.35	23.55
$\hat{\eta}_1$	3.22	3.37	3.35	5.72	5.05
$\hat{\eta}_2$	9.13	8.88	9.02	2.32	3.20
$\hat{\delta}$	-739.75	-789.65	-766.64	-1101.06	-497.59
$\hat{\alpha}$	1.68×10^{-2}	1.70×10^{-2}	1.70×10^{-2}	1.70×10^{-2}	1.70×10^{-2}
$\hat{\sigma}_2^2$	1.21×10^{-2}	1.14×10^{-2}	7.27×10^{-3}	1.39×10^{-2}	1.81×10^{-2}
$\hat{\phi}_2$	1.74×10^{-1}	1.89×10^{-1}	2.10×10^{-1}	1.58×10^{-1}	1.58×10^{-1}
AIC	-237696	-237890	-237896	-238378	-238902

Table 4.2: Proposed model estimates using squared exponential spatial covariance.

Proposed model (Ornstein-Uhlenbeck spatial covariance)					
l	$1.5m$			$3m$	
n	20	30	40	30	40
$\hat{\beta}$	5.30×10^{-3}	4.64×10^{-3}	4.05×10^{-3}	4.60×10^{-3}	4.87×10^{-3}
$\hat{\sigma}_1^2$	3.59	3.59	3.91	4.92	3.91
$\hat{\phi}_1$	5.54×10^{-3}	6.11×10^{-3}	7.59×10^{-3}	9.20×10^{-3}	6.89×10^{-3}
$\hat{\sigma}_3^2$	4.39×10^{-5}	4.32×10^{-5}	3.24×10^{-5}	4.33×10^{-5}	4.47×10^{-5}
$\hat{\rho}$	1.17×10^{-3}	7.64×10^{-4}	6.10×10^{-4}	1.98×10^{-3}	1.20×10^{-3}
$\hat{\omega}$	2.50×10^{-3}	2.31×10^{-3}	2.00×10^{-3}	3.04×10^{-3}	2.21×10^{-3}
$\hat{\eta}_1$	4.19	4.29	4.46	6.47	6.94
$\hat{\eta}_2$	7.87	7.70	7.35	2.64	2.13
$\hat{\delta}$	-931.15	-975.74	-1093.09	-2342.39	-1583.01
$\hat{\alpha}$	2.11×10^{-2}	2.19×10^{-2}	2.14×10^{-2}	1.86×10^{-2}	2.19×10^{-2}
$\hat{\sigma}_2^2$	9.03×10^{-4}	8.81×10^{-4}	1.03×10^{-3}	1.37×10^{-3}	9.83×10^{-4}
$\hat{\phi}_2$	1.41×10^{-1}	1.31×10^{-1}	1.21×10^{-1}	1.29×10^{-1}	1.30×10^{-1}
AIC	-238196	-238886	-238571	-242737	-239434

Table 4.3: Proposed model estimates using Ornstein-Uhlenbeck spatial covariance.

Reference model		
Param.	Squared exponential	Ornstein-Uhlenbeck
$\hat{\beta}$	3.26×10^{-3}	4.83×10^{-3}
$\hat{\sigma}_1^2$	72.38	7.98×10^{-1}
$\hat{\phi}$	48.74	1.49×10^{-1}
$\hat{\sigma}_3^2$	3.94×10^{-3}	8.41×10^{-24}
$\hat{\rho}$	4.70×10^{-4}	1.88×10^{-3}
$\hat{\omega}$	18.71	2.39×10^{-2}
$\hat{\eta}_1$	1.07	4.68
$\hat{\eta}_2$	11.45	7.99
$\hat{\delta}$	-584.15	-2656.05
$\hat{\alpha}$	8.61×10^{-2}	1.45×10^{-2}
AIC	43978	-182182

Table 4.4: Reference model parameter estimates.

cients ranges between $1 \times 10^{-3} m^2 h^{-1}$ and $1 \times 10^{-2} m^2 h^{-1}$, depending on soil type and moisture content. Hence, there is no "true" thermal diffusion coefficient for the system. The estimate should instead be interpreted as some effective thermal diffusion coefficient. This estimate is in turn apparently very dependent on the choice of model.

It is worth pointing out that the parameter estimates of the proposed model do not seem to change very much and systematically with scheme resolution. We do not expect the estimate of ρ to remain unchanged when resolution changes. This is because ρ must incorporate all heat flow at the boundary, since the radiation term has been omitted for simplicity, in which case the scaling by $1/d^2$ in (3.11) is not entirely meaningful. Scaling by $1/d$ would possibly make more sense, in which case we expect ρ to decrease with n, as is the case. When model depth increases, we expect the constant part of the boundary condition at $-l$, η_1, to increase, while the periodic part, η_2, to decrease. This is apparent in the estimates of the proposed models displayed in Table 4.2 and 4.3. As η_2 decrease, the influence of δ on the model decreases, and there are large uncertainties associated with the estimate of this parameter, as we can see from the proposed model estimates using depth $3m$.

If we for example underestimate the distance between the top vertical measurement device and the soil air boundary, we could expect to find an estimate of ρ which is smaller than β, since it must represent the thermal resistance of the missing soil layers. If we on the other hand have overestimated the same distance, we might find a very large estimate of ρ, as the model wants the temperature of the top soil layer to effectively be set equal to the air temperature. In both of the scenarios above, the estimate of the noise process at the soil air boundary would be very different. Its variance, σ_2^2 would have to be large to cope with a large ρ, while it would not need to if ρ was small. From this discussion it could be argued that, using this parameterization, it would be reasonable to deliberately underestimate the distance between the ground surface and the top measurement device. However, doing so would underestimate the thermal capacity of the top soil layers, and so should be done with caution.

The model using Ornstein-Uhlenbeck spatial covariance achieves slightly higher AIC

values than the model using squared exponential spatial covariance. However, these values vary inconsistently among different scheme resolutions and depths. Furthermore, it seems to be much harder to find minimizers of the objective using this spatial covariance. It becomes very hard to address parameter uncertainty when we struggle finding global maximizers of the likelihood. Using squared exponential spatial covariance on the other hand, it seems we always converge quickly to a minimizer of the objective. It is clear that the reference models do not fit the measurements at all. Their AIC values are very low compared to the proposed model. In terms of model diagnostics, these models are very far off as well. In the next section, we make some necessary further improvements to the proposed model with squared exponential spatial covariance.

Forecasts and their reliability

In Table 4.5 we have estimated the parameters of two models ignoring the snowy period (last 2000 observations). Both models use a squared exponential spatial covariance, a depth of $1.5m$, and 20 grid nodes. After some trial and error, we also find that the results are generally better provided we move the ground surface $5cm$ closer to the measurement setup, so this has been done as well. The two models displayed in the table differ by the radiation term in (3.12).

In Figure 4.7 the filtered and forecasted estimates for the squared exponential model with no radiation term using all observations are displayed. It is clear that the snowy period in winter poses a bit of a challenge for the model. In Figure 4.8, the filtered and forecasted estimates are displayed for the first model in Table 4.5. They appear marginally better looking, with a smaller steady state variance than the model including the snowy period. The steady state standard deviation of the temperature at cable depth is $0.66°C$ for the model including all observations, and $0.44°C$ for the one ignoring the snowy observations.

In Figure 4.9, filtered and forecasted estimates are displayed for the model including radiation in Table 4.5. It is worth pointing out that we consistently get an estimated γ (parameter to scale the inflow radiation when it is cloudy) between 0.4 and 0.5 when fitting different models using the radiation term, which makes sense. We used the rainfall measurements to determine if it was cloudy, even though the converse is not true. The estimate of μ_2 on the other hand rarely seems significant, which is hard to justify from a physical perspective. This seems to be the case for the estimates displayed in Table 4.5 as well. The steady state standard deviation for the temperature at cable depth using this model is $0.35°C$. The scaled innovations for the model including radiation are plotted in Figure 4.10 and their distribution in Figure 4.11. For all models, these diagnostic plots are very similar. As we can see from the histogram and Q-Q plot, the Gaussian assumption on the innovations is not correct. We also note that the innovations are not completely uncorrelated either, as we move down the vertical measurement probe.

In practical application, parameter uncertainty must be included in the forecasted temperature. This can be achieved by sampling parameter values from the approximate asymptotic distribution (2.59) using the ML estimate, $\hat{\theta}$, as mean, and the Hessian of the negative log likelihood evaluated at the ML estimate as an approximation to the Fisher information matrix. For every sampled set of parameters, sample the future temperatures using the system (3.22). Then choose among the samples the 2.5% largest upper and lower future temperatures. The resulting bound is an approximation to the true 95% prediction inter-

val for the temperature. In this setting, the small confidence bounds of the temperature at the 7. measurement device in Figure 4.7, 4.8 and 4.9 are a bit deceiving; as we can see from the asymptotic standard deviations in Table 4.5, there are relatively large uncertainties associated with the parameters related to the boundary condition at $1.5m$ depth, η_1, η_2, which are $0.5m$ below this point.

Out of sample testing is naturally an important part of validating the reliability of our model. If we estimate the parameters ignoring the snowy part of the observations, and then attempt to forecast the snowy part of the observations, our model is far off. The pit falls of our model, when it comes to out of sample testing, are not necessarily owed to the uncertainty in the estimated parameters (provided we do not include a lot of parameters), but possibly due to the dynamics of the entire system changing with time. As for the snowy period, this can be solved appropriately using a time invariant system. That is, we "cut off" heat flow into the system from the soil air boundary when snow is present, and set the top boundary soil temperature equal to zero. This can be done by setting the air temperature equal to zero, and letting ρ take a very large value. The noise parameters of the boundary, σ_2^2, ϕ_2, must be changed appropriately.

Model	Without radiation		With radiation	
Param.	Estimate	Sd. (asym.)	Estimate	Sd. (asym.)
$\hat{\beta}$	2.13×10^{-3}	4.10×10^{-5}	2.12×10^{-3}	4.75×10^{-5}
$\hat{\sigma}_1^2$	1.17×10^{-2}	5.47×10^{-4}	1.16×10^{-2}	6.04×10^{-4}
$\hat{\phi}$	1.30×10^{-2}	1.37×10^{-3}	1.53×10^{-2}	1.28×10^{-3}
$\hat{\sigma}_3^2$	7.63×10^{-5}	1.31×10^{-6}	7.57×10^{-5}	1.30×10^{-6}
$\hat{\rho}$	4.73×10^{-3}	5.90×10^{-4}	4.74×10^{-3}	6.82×10^{-4}
$\hat{\omega}$	24.96	5.05×10^{-1}	25.29	5.18×10^{-1}
$\hat{\eta}_1$	4.40	1.39	4.93	1.26
$\hat{\eta}_2$	7.50	1.42	7.15	1.33
$\hat{\delta}$	-832.12	3.01	-983.92	2.48
$\hat{\alpha}$	9.40×10^{-3}	6.44×10^{-4}	9.34×10^{-3}	6.44×10^{-4}
$\hat{\sigma}_2^2$	9.69×10^{-3}	1.73×10^{-3}	9.52×10^{-3}	1.97×10^{-3}
$\hat{\phi}_2$	1.76×10^{-1}	1.82×10^{-2}	1.97×10^{-1}	1.83×10^{-2}
$\hat{\mu}_1$	–	–	1.15×10^{-1}	1.99×10^{-2}
$\hat{\mu}_2$	–	–	7.04×10^{-14}	6.46×10^{-13}
$\hat{\mu}_3$	–	–	2.31×10^{-1}	3.46×10^{-2}
$\hat{\gamma}$	–	–	4.14×10^{-1}	1.33×10^{-1}
AIC	-136385		-136504	

Table 4.5: Final model estimates with asymptotic standard deviation.

Discussion and further comments

The goal of the study was to suggest a simple model capable of making confident forecasts of the cable temperature and its uncertainty. As noted in the previous section, the model diagnostics are not perfect. However, they are not very far off. We are able to get partly uncorrelated innovations, in addition to reasonably valued parameter estimates, and very

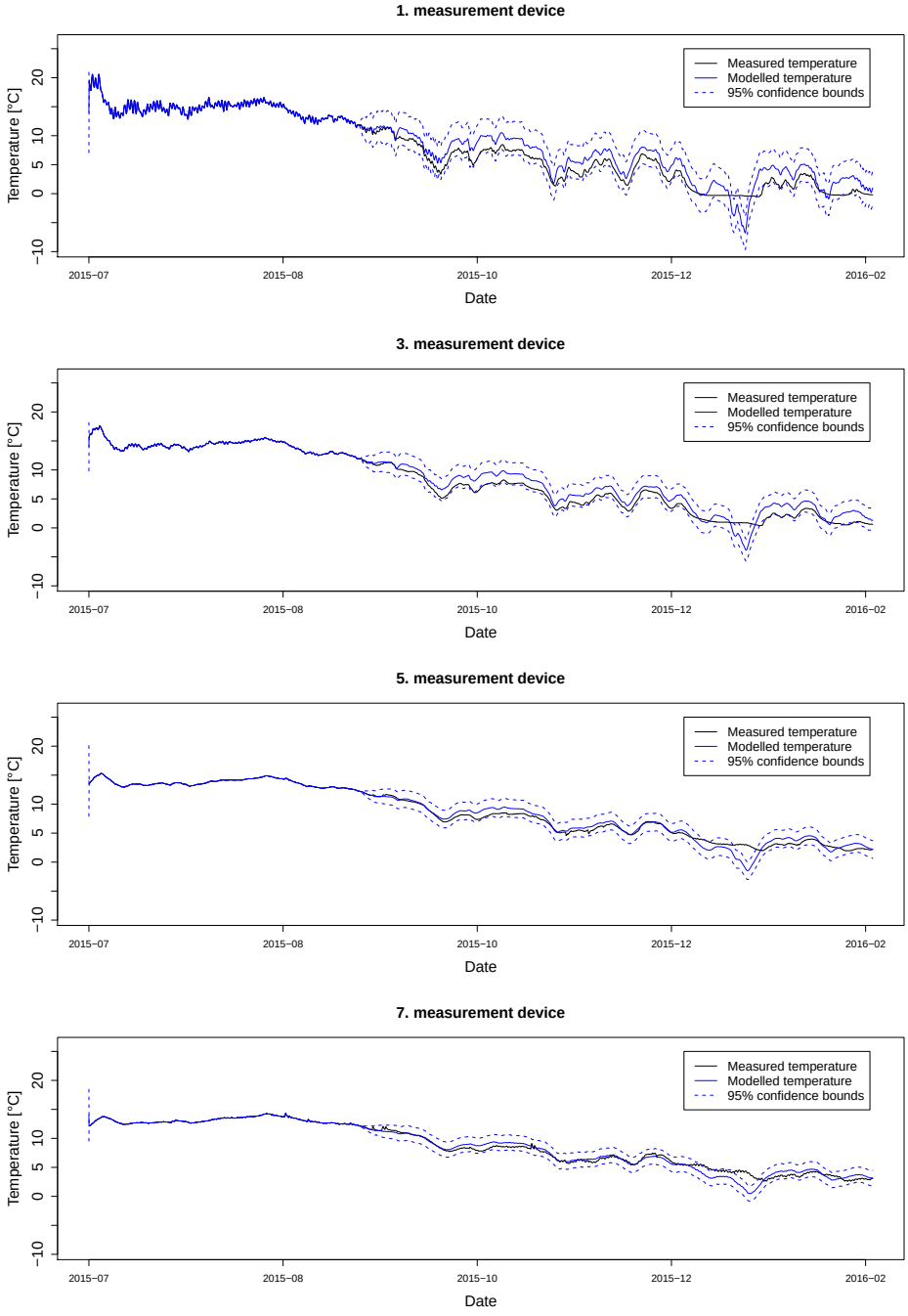


Figure 4.7: All observations: Filtered and forecasted soil temperatures at the measurement locations.

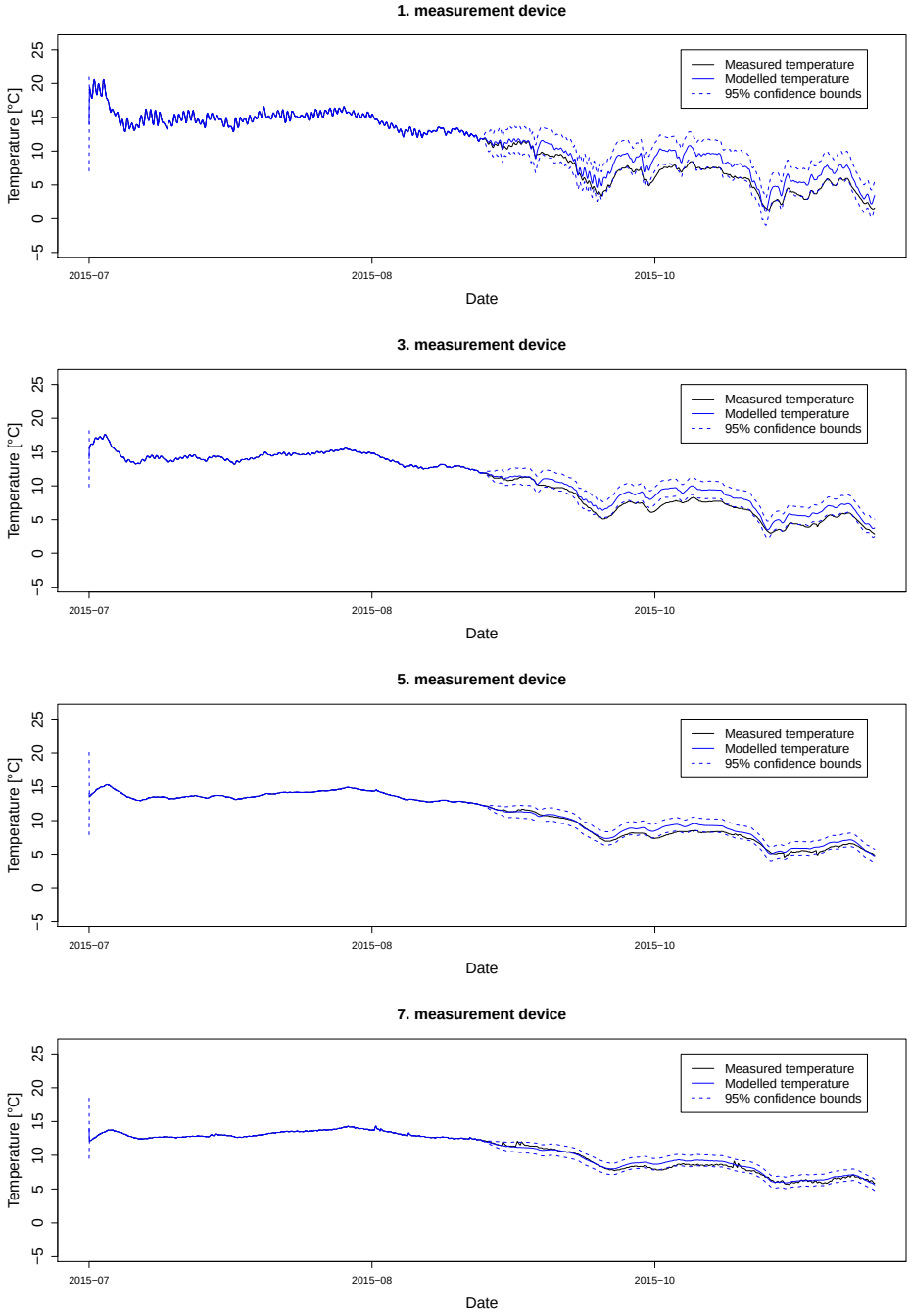


Figure 4.8: Ignoring the snowy period: Filtered and forecasted soil temperatures at the measurement locations.

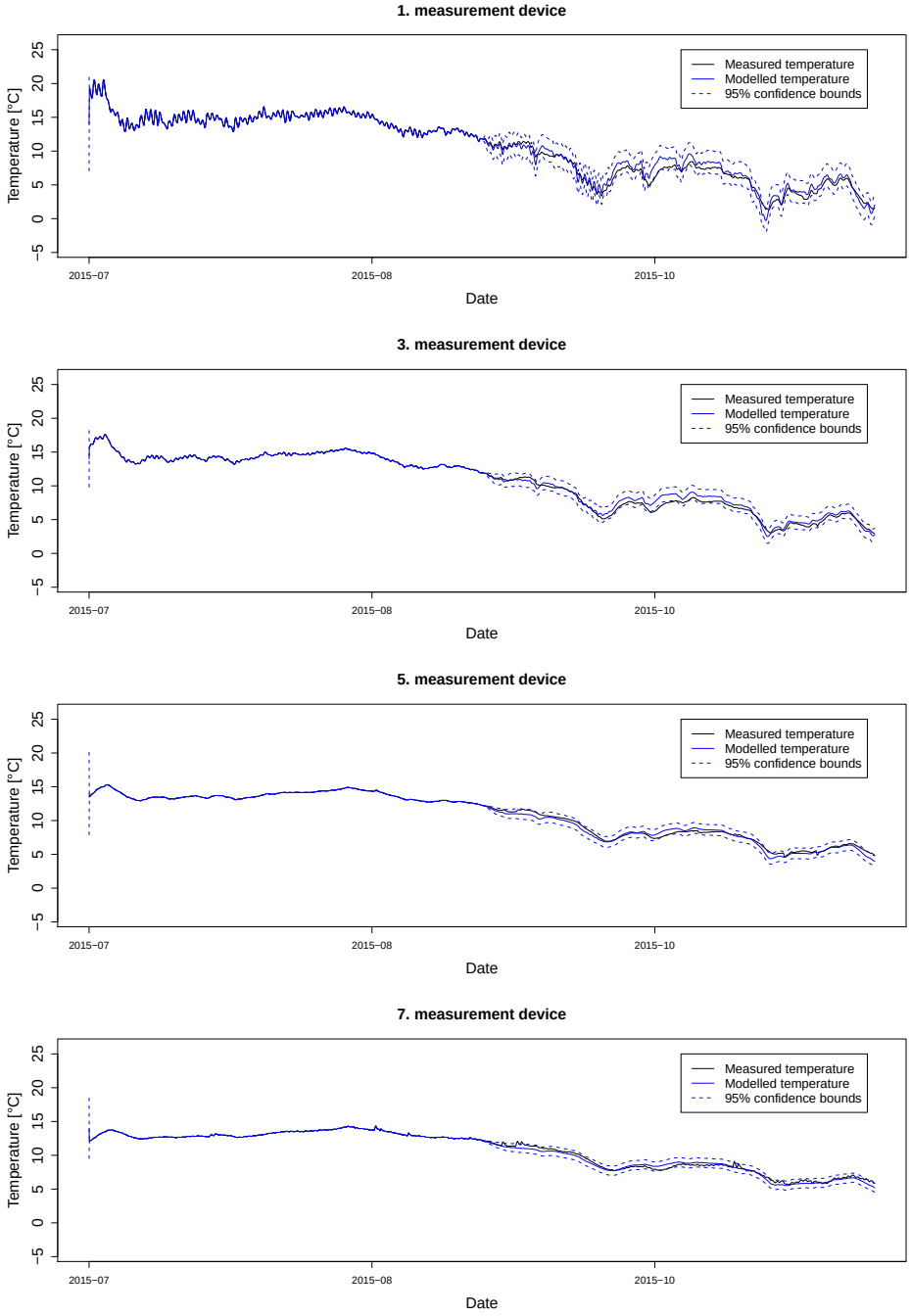


Figure 4.9: Ignoring the snowy period and including radiation: Filtered and forecasted soil temperatures at the measurement locations.

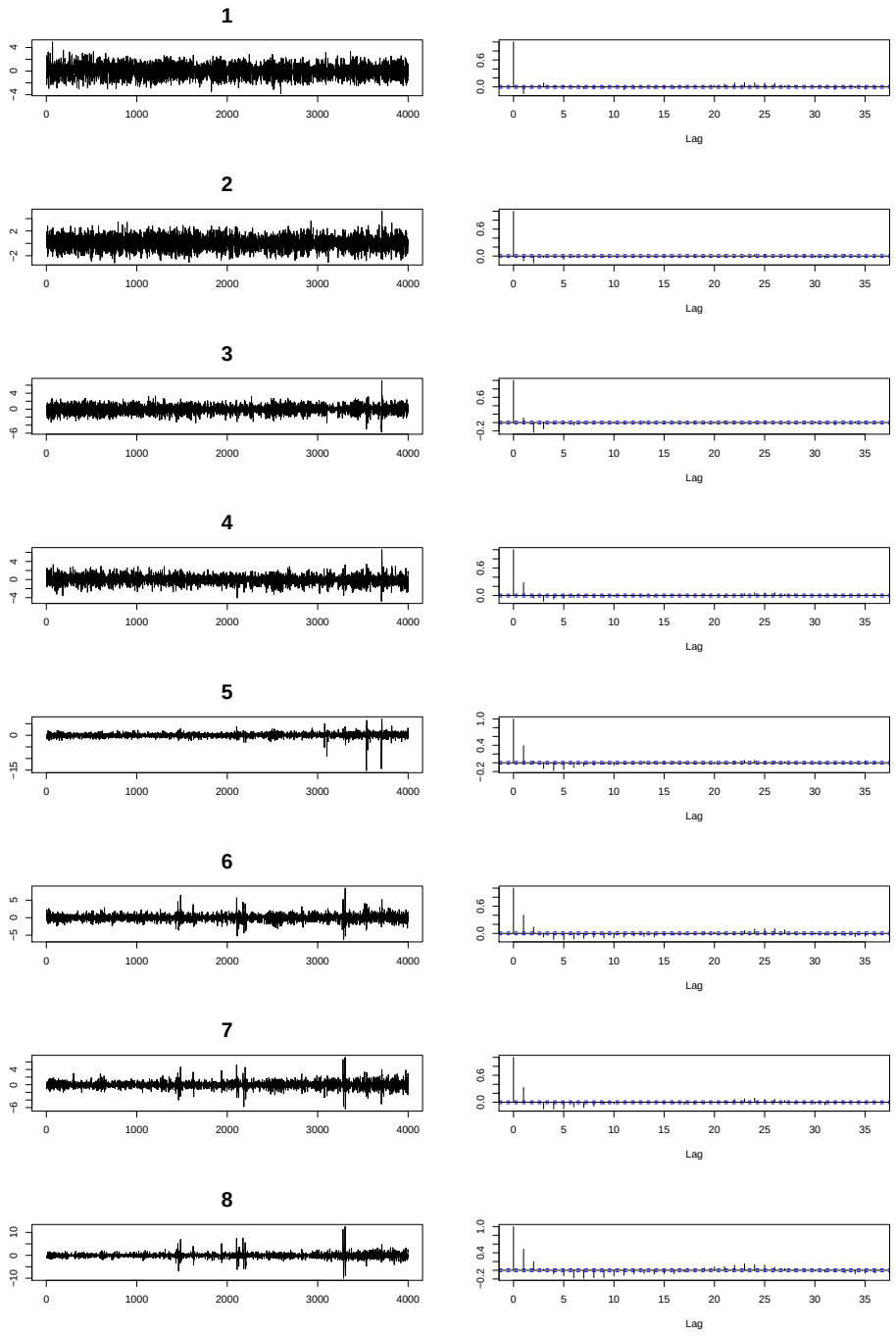


Figure 4.10: Real data: Scaled residuals at the vertical measurement devices.

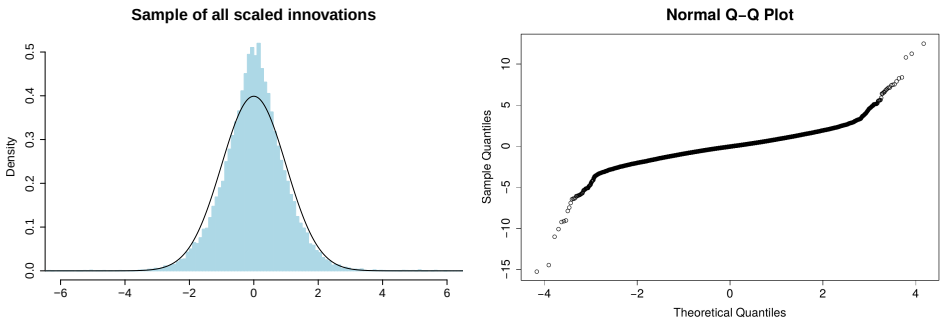


Figure 4.11: Real data: Scaled innovation sample, and it's Q-Q plot using the squared exponential model.

good looking forecasts. Comparing the simplicity of the model with the complexity of the problem, this is almost surprising. The estimates we find, being maximum likelihood estimates, also have appealing properties in terms of their uncertainty, which speaks to the reliability of the model.

Maximizing the likelihood is performed reasonably quickly. Using automatic differentiation, a gradient computation takes on average 2 seconds for the simplest model used above ($n = 20$, 12 parameters, $8 \cdot 5814$ observations). For the same model, using a finite difference approximation takes 18 seconds. The former also scales much better with increasing numbers of parameters. The time it takes for the optimization routine to converge depends on initial guess, and how well the proposed model matches the measurements. All implementations are in C++ where we have used the Stan Math Library.

It is important to point out that in order to appropriately address the problem of determining the cable load capacity, the model should be tested on data where the cable temperature has taken higher values, and exhibited larger variations. Since the heat flow in our data set was mainly in the vertical direction, a vertical one dimensional stochastic heat flow model sufficed. As radial heat flow becomes more significant, this might not be the case. The radial heat problems are nonlinear, and including random heat flow in the radial direction, we get a nonlinear stochastic system. In Chapter 3 we outlined how this can be handled.

4.2 Conclusion and final remarks

We have studied a Gaussian process for modelling soil heat flow. It is the stationary solution of a linear stochastic system based on the heat equation with additive noise. The noise was random heat fluxes in the soil, defined in terms of the gradient field of a stationary, spatially smooth, Gaussian Markov process. A fast and reliable framework for identifying the parameters driving this process has been established, and tested using simulated and real measurements. The real data example consisted of soil temperature measurements from an electric cable installation, and was related to the problem of determining cable load capacity.

The model appears to be a good contribution to extending the dynamic current rating standard. We found that it worked fine, especially compared to simpler models, even though the Gaussian assumption was not entirely correct. A natural extension of the work would be to test the model on data generated from cables with higher temperatures and temperature variations. This makes the heat dynamics nonlinear, in which case a nonlinear stochastic system or a linear system with some conservative parameter values would have to be used. Validating the model in these cases is necessary in order to properly address the problem of determining the load capacity.

We finally point out a key aspect of the proposed model, namely the fact that it is very simple, both mathematically and computationally. Filtering and maximum likelihood parameter estimation is done exactly and efficiently, and implementing the latter is relatively easy when automatic differentiation is used. In this context it is hard to justify using a more complicated model. Moreover, it is possibly an appropriate model for a lot of stochastic heat flow phenomena.

Bibliography

Abrahamsen, P., 1997. A Review of Gaussian Random Fields and Correlation Functions. Technical report. Norsk regnesentral.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716–723.

Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods. Springer.

Brockwell, P.J., Davis, R.A., 2002. Introduction to Time Series and Forecasting. Springer.

Carpenter, B., Hoffman, M.D., Brubaker, M., Lee, D., Li, P., Betancourt, M., 2015. The Stan Math Library: Reverse-Mode Automatic Differentiation in C++. Documentation of C++ library.

Davis, M.H.A., Vinter, R.A., 1985. Stochastic Modelling and Control. Chapman and Hall.

Goodrich, R.L., Caines, P.E., 1979. Linear system identification from cross sectional nonstationary data. IEEE AC-24, 403–411.

Gupta, N.K., Mehra, R.K., 1974. Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations. IEEE AC-19, 774–783.

Hairer, M., 2004. An introduction to stochastic PDEs. Lecture notes. University of Warwick.

Hamilton, J.D., 1994. State-space models. Handbook of Econometrics 4, 3041–3077.

IEC 60287, 2001. Calculation of the Current Ratings. Standard. International Electrotechnical Commission. Geneva, Switzerland.

IEC 60853, 1989. Calculation of the Cyclic and Emergency Rating of Cables. Standard. International Electrotechnical Commission. Geneva, Switzerland.

Ilstad, E., 2009. TET4195 High voltage equipment cable technology. NTNU Department of electric power engineering.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Transactions of the ASME - Journal of Basic Engineering 82, 35–45.

Kalman, R.E., Bucy, R.S., 1961. New results in linear filtering and prediction theory. Journal of Basic Engineering 83, 95–108.

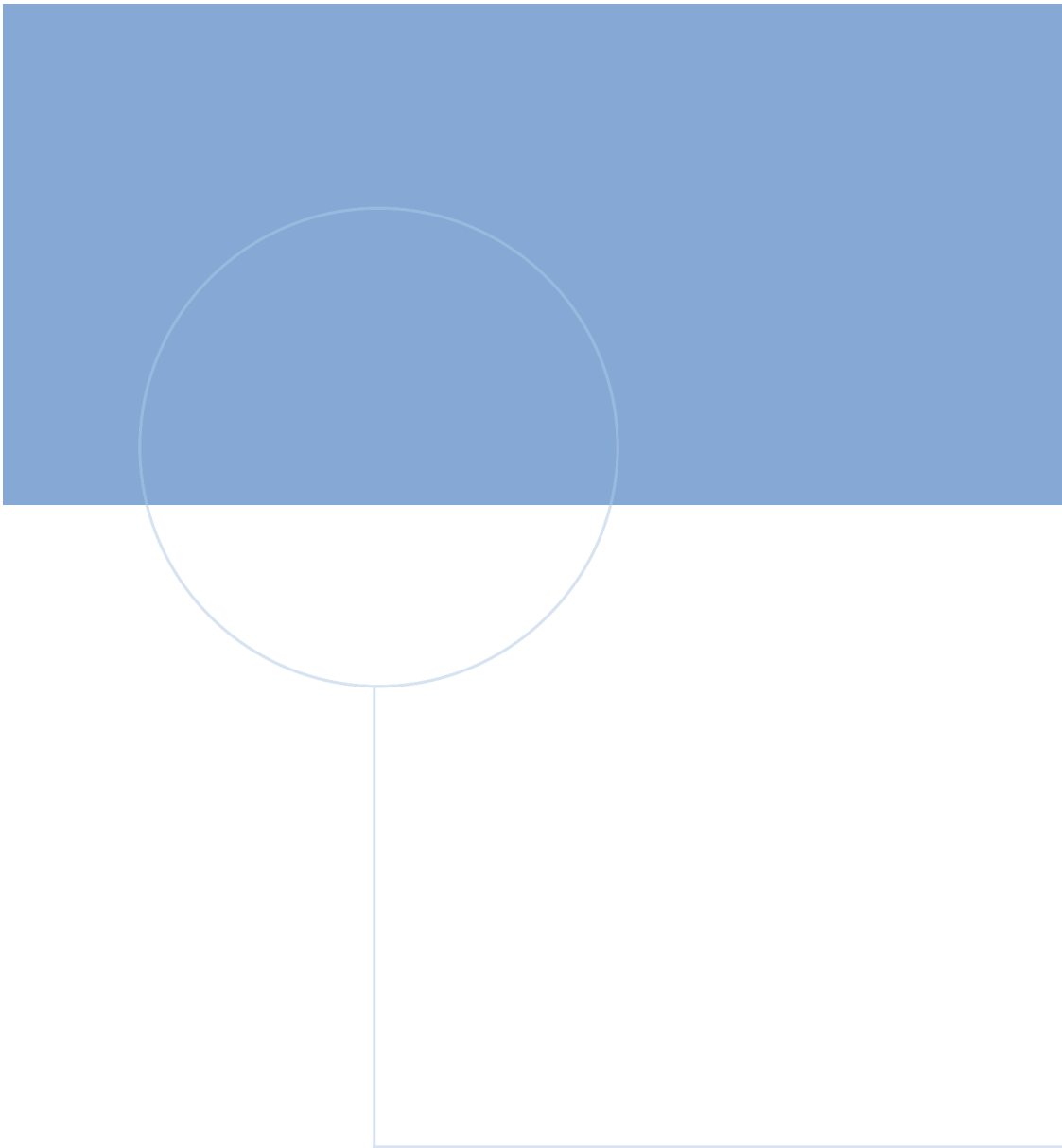
Neher, J.H., McGrath, M.H., 1957. The calculation of the temperature rise and load capability of cable systems. IEEE 73, 752–764.

Nocedal, J., Wright, S.J., 2006. Numerical Optimization. Springer.

Rauch, H.E., Tung, F., Striebel, C.T., 1965. Maximum likelihood estimates of linear dynamic systems. AIAA 3, 253–264.

Shumway, R.H., Stoffer, D.S., 1982. An approach to time series smoothing and forecasting using the EM algorithm. Jounal of Time Series Analysis 3, 253–264.

Øksendal, B., 2000. Stochastic Differential Equations. Springer-Verlag.



NTNU

Norwegian University of
Science and Technology