Mathematics and numerics for data assimilation and state estimation

Summer semester 2020

Overview

Course information

Course webpage with schedule (alternatively Moodle):

`https://haakonahmatata.github.io/courses/data_assimilation/main.html`

9 ECTS

Examination: Written 90-120 minutes (early to mid-August).

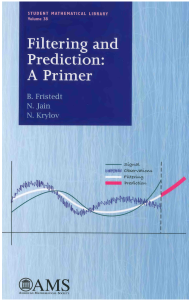
Student presentation in early July: Bonus points equivalent to 0.1*MaxEamScore is given those making a (roughly) 20 minutes presentation on important topic/paper in data assimilation.

Übungen: Almost every week. Will consist of sets of exercises to solve. You can work on the exercises and I/you will solve some of them in plenary.

Note: Please register also for übungen, LV 11.46000

Course literature

Main literature 1: "Filtering and Prediction: A Primer" by Fristedt, B, Jain, N. and Krylov, N., 1st ed. AMS (2007)
Main literature 2: "Data assimilation" by Law, K., Stuart, A., Zygalakis, K. 1st ed. Springer (2015).

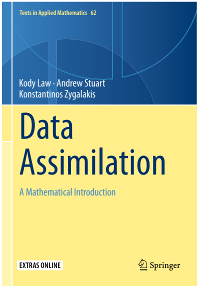


Supplementary literature:

- "Probability: Theory and Examples" Durrett, R, Version 5 January 11, 2019. Downloadable from:

  https://services.math.duke.edu/∼rtd/PTE/PTE5_011119.pdf

- "Probabilistic forecasting and Bayesian data assimilation" by Reich, S. and Cotter, C., 1st ed. Cambridge University Press (2015).



Who, where and when

Name and position: Jr. Prof. Håkon Hoel at the Chair of Numerics for Uncertainty Quantification.

Research interests: numerical analysis of stochastic differential equations, nonlinear filtering and Monte Carlo methods

Address: Room C364, Kackertstr. 9,

email: hoel AT uq.rwth-aachen.de

Office hours: Monday 14-15.

Overview

1. Examples of data assimilation
2. Course content
3. Lecture 1
   - Probability space and random variables
4. Other courses and seminars at our chair

What is data assimilation?

Definition: The combination of dynamical models with measurement data to estimate the past/current/future state of a system.

Examples

- Weather prediction.
- Source location of natural resource, contaminant, earthquake etc.
- Automated navigation systems.

Example: weather prediction (E. Kalnay 2003, V. Bjerknes 1904)

Observations 6h window

Background or first guess

Inverse problem: Assimilate observations into current state estimate

Initial condition

6h forecast

Forward model: Forecast solving system of PDE.

$$\frac{d\mathbf{v}}{dt} = -\alpha \nabla p - \nabla \phi + \mathbf{F} - 2\Omega \times \mathbf{v}$$ Cons. momentum

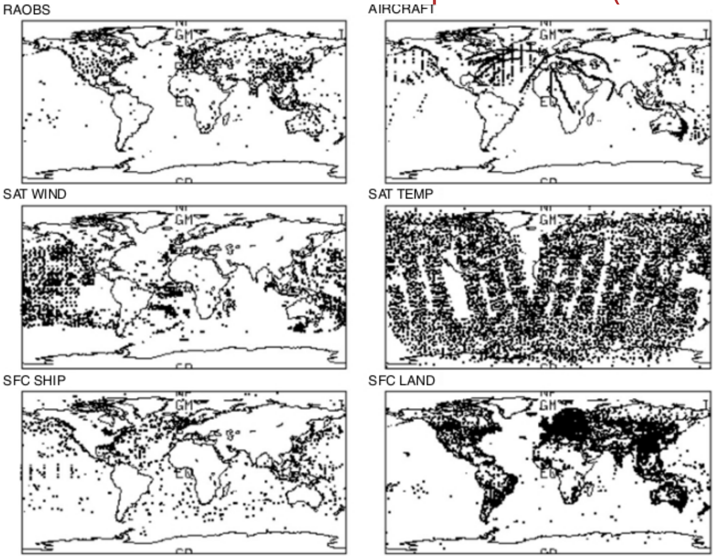
$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v})$$ Cons. mass

$$p\alpha = RT$$ Eq. of state

$$Q = C_p \frac{dT}{dt} - \alpha \frac{dp}{dt}$$ Cons. energy

$$\frac{\partial \rho q}{\partial t} = -\nabla \cdot (\rho \mathbf{v} q) + \rho(E - C)$$ Cons. water vapor mixing ratio

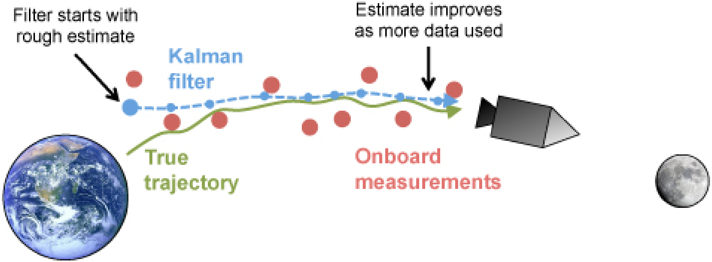
6h observations for weather predictions (E. Kalnay 2003)



How to incorporate/assimilate observation data into present state predictions?

Data assimilation is often used in combination with a control to make decisions:

- Space travel, autonomous cars: 1. Self-localization 2. Drive/use rocket fuel to navigate
- Weather: predict wind/solar power production tomorrow (and make actions)
- Oil exploration: 1. drill for oil 2. estimate most likely location for oil pocket given new info 3. drill for oil there ...

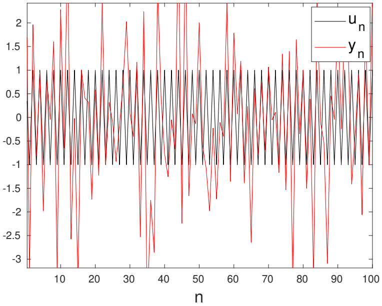


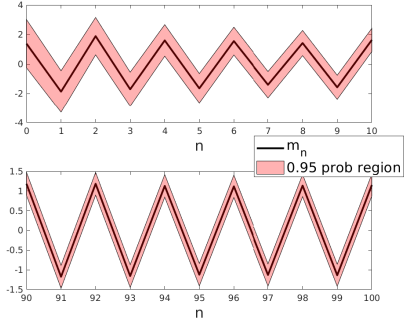
Down to earth example (Kalman filtering)

Unobserved dynamics: $u_{n+1} = -u_n, \quad \text{and } u_0 \sim N(0,1)$

Noisy observations: $y_n = u_n + \gamma_n, \quad \gamma_n \sim N(0,2).$

Problem: Determine $u_n | y_{0:n}$. (Sequence (y_n) is here generated from a sample of u_n with $u_0 = 1$.)





Overview

1. Examples of data assimilation
2. Course content
3. Lecture 1
   - Probability space and random variables
4. Other courses and seminars at our chair

Course content

- Bayesian inference
- Bayesian filtering for discrete time and space Markov chains (random walks . . .)
- Stochastic processes (Markov processes and stochastic differential equations).
- Linear and nonlinear discrete-time filtering algorithms and smoothing (Kalman filtering, Ensemble Kalman filtering, Particle filtering)
- The Fokker-Planck equation and the Bayes filter for discrete time, infinite state-space filtering
- Filtering in high-dimensional state space
- Continuous time filtering methods
- Model fitting/parameter fitting and model validation
- Student presentations on applications of filtering
- Tentative: Multilevel Monte Carlo methods and applications of control with data assimilation

Bayesian inference

- Given two events C, D, with $\mathbb{P}(D) > 0$, Bayes' theorem yields

$$\mathbb{P}(C \mid D) = \frac{\mathbb{P}(C \cap D)}{\mathbb{P}(D)}$$

  where

$$\mathbb{P}(C \mid D) := \text{Probability of C given D.}$$

- Is useful in filtering

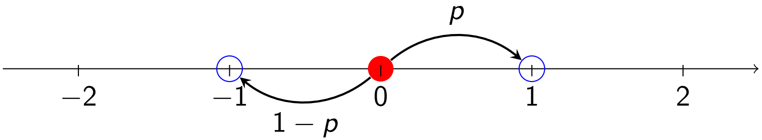
$$\mathbb{P}(X_1 = a | Y_0 = b_0, Y_1 = b_1) = \frac{\mathbb{P}(X_1 = a, Y_1 = b_1 | Y_0 = b_0)}{\mathbb{P}(Y_1 = b_1 | Y_0 = b_0)} = \dots$$

  ((X_n, Y_n) signal-observation pair).
- Subtlety: How to treat $\mathbb{P}(C|D)$ and conditional expectations when $\mathbb{P}(D) = 0$?

Bayesian filtering for discrete time and space Markov chains

- A random walk on $\mathbb{Z}^d$ is a sequence $X_0, X_1, \ldots$ with independent and identically distributed (iid) increments.
- Example below

$$\mathbb{P}\left(X_{n+1} - X_n = 1\right) = p \quad \text{and} \quad \mathbb{P}\left(X_{n+1} - X_n = -1\right) = 1 - p.$$

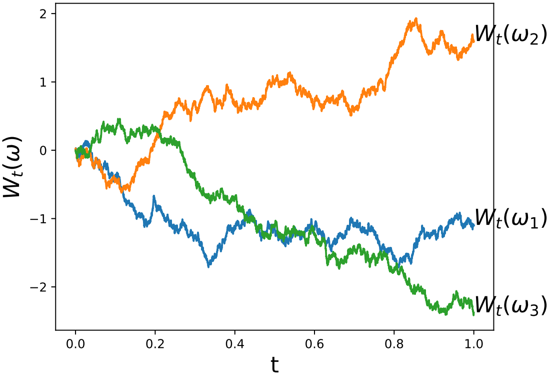


- Filtering problem: Given $X_0 = 0$ and observations $Y_k = X_k + W_k$ where $(W_1, W_2, \ldots)$ are iid random variables and also independent from $(X_1, X_2, \ldots)$, determine

$$\mathbb{P}\left(X_n | Y_0 = b_0, Y_1 = b_1, \ldots, Y_n = b_n\right).$$

Stochastic processes

- A stochastic process on $\mathbb{R}$ is a family of random variables $\{u(t)\}_{t\in[0,T]}$ such that $u(t) \in \mathbb{R}$ is a random variable for each $t \in [0, T]$.
- Examples: Wiener processes:



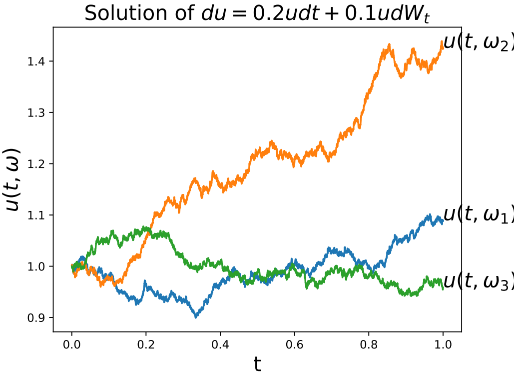
Stochastic differential equations

The solution of

$$du(t) = a(u(t))\,dt + b(u(t))\,dW(t),$$
$$u(0) = u_0,$$

is a stochastic process.

Example: Geometric Brownian Motion



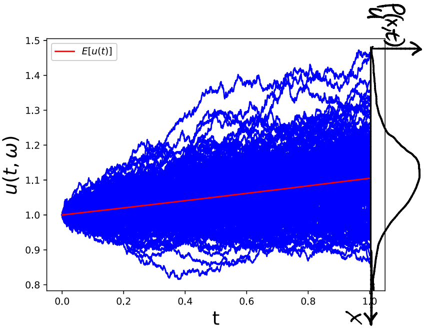
Stochastic differential equations

The solution of

$$
\begin{aligned}
du(t) &= a(u(t))\, dt + b(u(t))\, dW(t), \\
u(0) &= u_0,
\end{aligned}
$$

is a stochastic process.

Example: Density of Geometric Brownian Motion



Topics we will treat on stochastic processes

- Theory on Markov processes (Poisson, Wiener and Itô stochastic differential equations)
- Numerical methods for sampling realizations of stochastic processes
- Discrete time filtering problem: For continuous time process u and discrete time observations

$$y(k) = Q(u(k)) + \text{"noise"}$$

  determine

$$\mathbb{P}\left(u(n) \mid Y(0) = b_0, Y(1) = b_1, \dots, Y(n) = b_n\right).$$

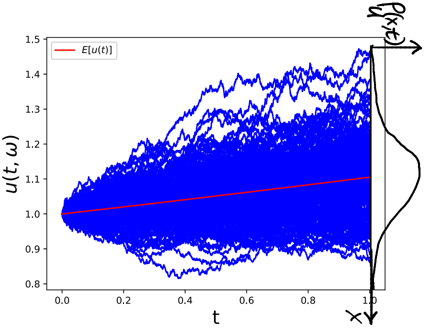
- Continuous time filtering: Given

$$\begin{aligned} u(t) &= u_0 + \int_0^t a(s)u(s)ds + W_1(t) \\ y(t) &= H(t)u(t) + W_2(t) \end{aligned}$$

  estimate $\mathbb{P}\left(u(t) | \{Y(s) = b(s)\}_{s\in[0,t]}\right)$.

Fokker-Planck equation

Density $\rho_u(x,t)$ for the SDE is the solution of a parabolic partial differential equation called the Fokker-Planck equation.



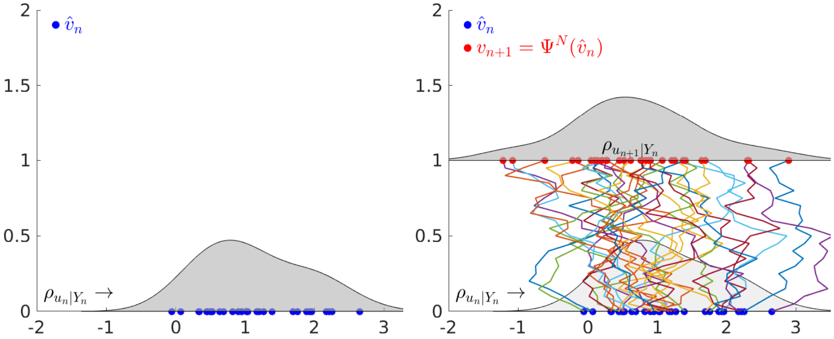
In many cases ρ_u can be used to derive the exact filters (called the Bayes filter).

Kalman filtering and nonlinear methods

In linear settings with additive Gaussian noise, Kalman filtering is an exact filtering method. However, for nonlinear settings:

$$u_{n+1} = \Psi(u_n),$$
$$y_{n+1} = Q(u_{n+1}) + \gamma_{n+1},$$

alternatives are needed. EnKF and particle filters are ensemble/particle based methods that approximate the filter density by empirical measures:

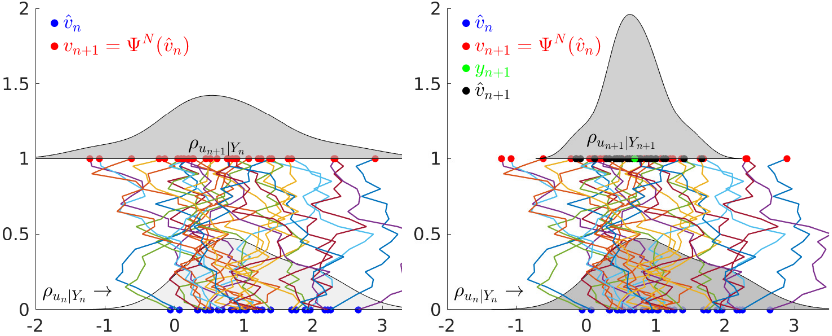


Kalman filtering and nonlinear methods

In linear settings with additive Gaussian noise, Kalman filtering is an exact filtering method. However, for nonlinear settings:

$$u_{n+1} = \Psi(u_n),$$
$$y_{n+1} = Q(u_{n+1}) + \gamma_{n+1},$$

alternatives are needed. EnKF and particle filters are ensemble/particle based methods that approximate the filter density by empirical measures:



Model fitting and validation

The basic filtering setup consists of these fundamental assumptions:

- Underlying dynamics $X_{n+1} = \Psi(X_n, t_n)$ where the mapping Ψ is *known* but $X_0, X_1, \ldots$ is only partially observed
- by $Y_k = Q(X_k) + "noise(k)"$ for $k = 0, 1, \ldots$ where
- the mapping Q and the distribution of $"noise(k)"$ are *assumed known*.

For real problems, Ψ and $"noise(k)"$ are of course often not known!

Fitting problem: for a parametrized class of mappings $\{\Psi_p\}_{p \in \mathcal{P}}$ find the "best" model given a collection of possibly different kinds of observations $Y_1, Y_2, \ldots$.

Student presentations, example topics

- Numerical weather prediction: "Atmospheric modeling, data assimilation and predictability" Kalnay.
- Oil reservoir state estimation: "Data Assimilation" Evensen and "An Iterative Ensemble Kalman Filter for Multiphase Fluid Flow Data Assimilation" Gu and Oliver.
- "Ensemble Kalman methods for inverse problems" Iglesias, Law and Stuart
- Data assimilation for the cardiovascular system (Sections 10 and 11 in): "The cardiovascular system: Mathematical modelling, numerical algorithms and clinical applications" Quateroni, Manzoni and Vergara.
- "On the convergence of the ensemble Kalman filter" Mandel, Cobb and Beezley.
- "On sequential Monte Carlo sampling methods for Bayesian filtering" Doucet, Godsill and Andrieu.
- "Multilevel Ensemble Kalman filtering" Hoel, Law and Tempone.
- Infinite dimensional Bayesian inference: "Inverse problems, a Bayesian perspective" Stuart.
- Data assimilation for virus pandemics.

Overview

Probability space

Definition 1 (Probability space)

A probability space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ consisting of

- the sample space Ω (set of outcomes),
- the set of events $\mathcal{F}$ which is a $\sigma-$algebra on Ω,
- a probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$.

Definition 2 ($\sigma-$algebra on Ω)

$\mathcal{F}$ consists of a collection of subsets of Ω such that

1. $\Omega \in \mathcal{F}$ **[contains the full set]**
2. if $D \in \mathcal{F}$, then $D^C \in \mathcal{F}$ also **[closed under complements]**
3. if $D_i \in \mathcal{F}$ for $i = 1, 2, \ldots$, then $\cup_{i=1}^{\infty} D_i \in \mathcal{F}$ **[closed under countable unions]**.

Exercise: show that $\emptyset \in \mathcal{F}$ and that $\mathcal{F}$ is closed under countable intersections.

Measurable spaces and probability measures

Definition 3

The pair $(\Omega, \mathcal{F})$ is called a **measurable space**, and $\mathbb{P} : \mathcal{F} \to [0,1]$ is a probability measure on $\mathcal{F}$ provided

1. $\mathbb{P}(D) \geq 0$ for all $D \in \mathcal{F}$ **[measures are non-negative-valued (obvious from the image I wrote)]**
2. if $D_i \in \mathcal{F}$ for $i = 1, 2, \ldots$ and the sequence is pairwise disjoint (meaning that $D_i \cap D_j = \emptyset$ for all $i \neq j$), then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} D_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(D_i)$$ **[countable additivity of disjoint sets]**

3. $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\emptyset) = 0$ **[measure of full space is 1!]**.

Example 4 (Measure on finite-state space)

$\Omega = \{-1, 0, 1\}$, $\mathcal{F} = \{\emptyset, \{-1\}, \{0,1\}, \Omega\}$ and

$$\mathbb{P}(\{-1\}) = 1/4, \quad \mathbb{P}(\{0,1\}) = 3/4.$$

Discrete random variables/vectors

Definition 5

A discrete random variable X defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is a mapping $X : \Omega \to \{a_1, a_2, \dots, \}$ where

1. $A = \{a_1, a_2, \dots, \} \subset \mathbb{R}^d$ is a finite or at most countable set of **distinct** outcomes
2. and it must hold that $X^{-1}(a_k) = \{\omega \in \Omega \mid X(\omega) = a_k\} \in \mathcal{F}$ for all a_k.

- X is described by the events and their probabilities

$$X^{-1}(a_k) = \{X = a_k\}, \qquad \mathbb{P}_X(a_k) := \mathbb{P}(X = a_k) = \mathbb{P}(X^{-1}(a_k)) \quad \forall a_k \in A.$$

- This is because X can be represented by a simple function

$$X(\omega) = \sum_{k=1} a_k \mathbb{1}_{X=a_k}(\omega). \quad \text{where } \mathbb{1}_{X=a_k}(\omega) \begin{cases} 1 & \text{if } X(\omega) = a_k \\ 0 & \text{otherwise} \end{cases}$$

- The measure $\mathbb{P}_X$ is called the **distribution** of X (it is a probability measure on the image space of X).

Discrete random variables 2

- Any function $f : \mathbb{R}^d \to \mathbb{R}^k$ also is a discrete rv, and can be represented
$$f(X)(\omega) = \sum_{k=1} f(a_k)\mathbb{1}_{X=a_k}(\omega).$$
- Note! The defintion for continuous random variables is more subtle for continuous rv, and (image-space) outcomes $\{a_1, a_2, \ldots, \}$ may not be associated uniquely to (probability-space) outcomes in Ω.

Example 6 (Coin toss, $X \sim$ Bernoulli(p))

- image-space outcomes $A = \{0, 1\}$,
- $$\Omega = \{Heads, Tails\}, \qquad \mathcal{F} = \{\emptyset, \{Heads\}, \{Tails\}, \Omega\}$$
- $X(Heads) = 1$ and $X(Tails) = 0$ and

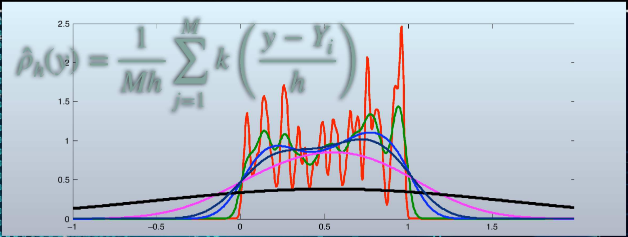
$$\mathbb{P}(X = 1) = \mathbb{P}(X^{-1}(1)) = \mathbb{P}(Heads) = p, \quad \mathbb{P}(X = 0) = \mathbb{P}(Tails) = 1 - p.$$

Next lecture

- Joint distributions
- Independence
- Expectations and variance
- Law of large numbers
- Conditional probability and expectation

Overview

Stochastic Numerics with applications in Simulation and Data Science



For whom?

This course is for CES, Mathematics, and Simulation Sciences master-students as well as everyone interested.

Content:

- Random variable generation
- Monte Carlo method: Error analysis
- Variance reduction techniques (antithetic variables, control variables, importance sampling)
- Large deviations and Rare events simulations
- Kernel density estimators
- Resampling techniques
- Simulation of stochastic processes: Gaussian fields and Kriging.
- Markov Chains
- Markov Chain Monte Carlo methods (Metropolis-Hasting, Gibbs sampler, Tempering)
- Bayesian Filters, Kalman Filters and data assimilation

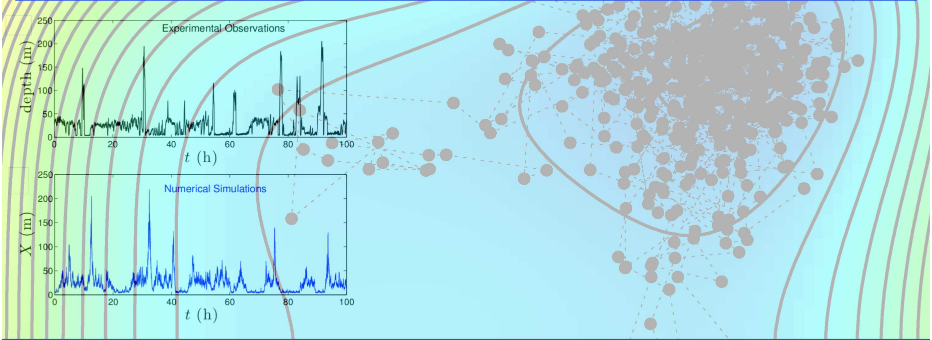
Lecturers: Prof. Dr. Raul Tempone
Dr. Nadhir Ben Rached

Lectures: Tue, Wed & Thurs, 10:30 - 12:00
Tutorials: Thurs, 14:30 - 16:00

Where: C301 (3010 | 301) | Kackertstr 9, 3rd floor

Web: www.uq.rwth-aachen.de

Data Science under Uncertainty



For whom?

This seminar is for Data Science, Mathematics, Simulation Sciences, and CES master students as well as everyone interested.

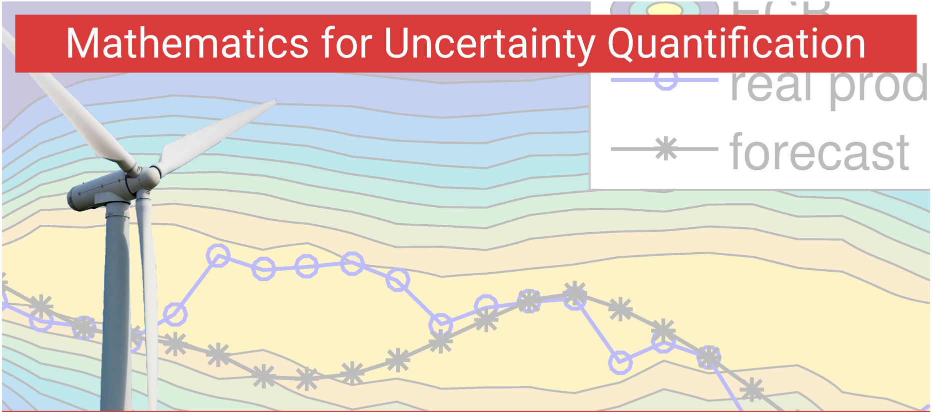
Mathematics for Uncertainty Quantification | RWTH AACHEN UNIVERSITY

Content

In this seminar we will cover data science applications subject to uncertainties. The focus will be on the mathematical and numerical analysis of stochastic tools used to treat these problems. For example, these tools include Markov chain Monte Carlo sampling methods, Data Assimilation techniques, optimal experimental design, model selection and validation, and statistical learning techniques such as clustering and support vector machines.

Lecturers:	Prof. Dr. R. Tempone Prof. S. Krumscheid, Ph.D.
	 
First Meeting:	Mon. 6 Apr 2020 14:30—16:00
Where:	C301 Kackertstr. 9 (3010 I 301)
Web:	www.uq.rwth-aachen.de

Mathematics for Uncertainty Quantification



For whom?

This seminar is for CES, Mathematics, and Simulation Sciences master-students as well as everyone interested.

Content

Mathematical modeling and numerical simulation are central components of modern scientific research. A key challenge is to quantify uncertainty in model predictions. This seminar will explore research topics in the context of mathematical models and analysis for simulation techniques used in uncertainty quantification.

Lecturers: Prof. Dr. Raul Tempone, Dr. Eric Hall

First Meeting: Tue 7 Apr 2020, 14:30—16:00

Where: C301 Kackertstr. 9 (3010 | 301)

Web: www.uq.rwth-aachen.de

UQ Mathematics for Uncertainty Quantification | RWTH AACHEN UNIVERSITY