Modelling with Gaussian Processes

Juš Kocijan

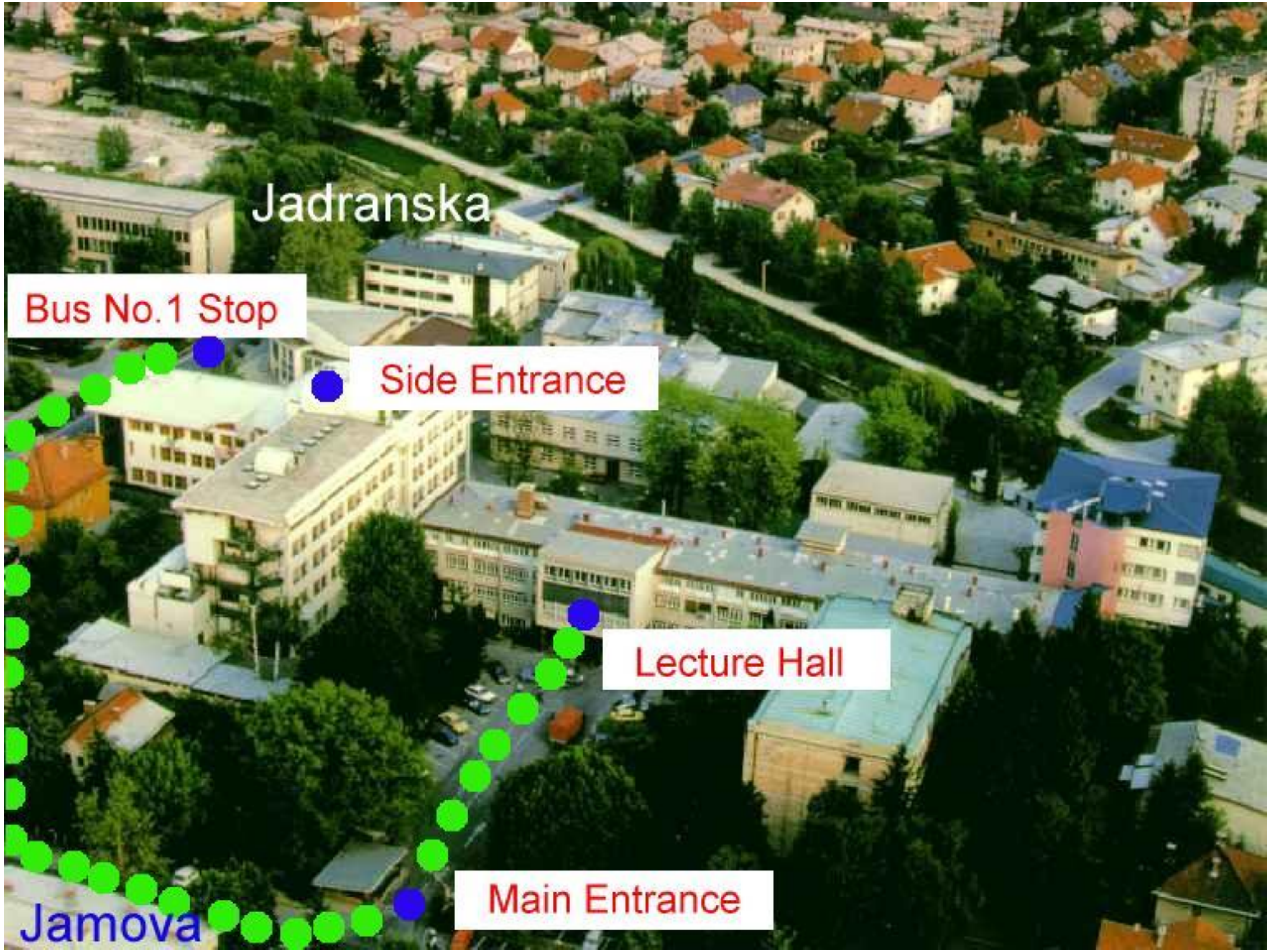
Jozef Stefan Institute, Ljubljana, Slovenia

University of Nova Gorica, Nova Gorica, Slovenia

Guest lecture @ UCM, Madrid, Spain, March 15, 2024

Sea
Guernsey(U.K.)
Jersey(U.K.)
Paris
FRANCE
Frankfurt
Luxembourg
CZECH REP.
SLOVAKIA
Bratislava
Vienna
Munich
LIECH.
AUSTRIA
SLOVENIA
Budapest
HUNGARY
ROMAN
Bern
SWITZ.
Lyon
Ljubljana
Milan
Turin
CROATIA
Belgrade
Bay of
Biscay
Bordeaux
Toulouse
MONACO
Genoa
SAN
MARINO
BOS. &
HER.
SERBIA
Sarajevo
ITALY
Bilbao
Marseille
Pristina
Sofia
BULG
ANDORRA
Corsica
(FR.)
VATICAN
CITY
Rome
Podgorica
KOS.
MONT.
Skopje
Tirana
MACEDONIA
ALB.
Porto
Madrid
Barcelona
Naples
PORTUGAL
SPAIN
Valencia
BALEARIC
ISLANDS
(SP.)
Sardinia
(IT.)
GREECE
Lisbon
Palermo
Sevilla
Málaga
Algiers
Sicily
(IT.)
Athens
Gibraltar(U.K.)
Constantine
Tunis
MALTA
Valletta
Ceuta(SP.)
Oran
TUNISIA
Melilla
(SP.)
Rabat
Casablanca
Fès
Mediterranean Sea
Tripoli

Jozef Stefan Institute



The Jožef Stefan Institute is the **leading Slovenian scientific research institute**, covering a broad spectrum of basic and applied research. The staff of more than 950 specializes in natural sciences, life sciences and engineering. The subjects concern production and control technologies, communication and computer technologies, knowledge technologies, biotechnologies, environmental technologies, nanotechnologies, nuclear engineering, material sciences, etc. **The mission** of the Jožef Stefan Institute is the accumulation - and dissemination - of knowledge at the highest international levels of **excellence**.

Department of systems and control

The domain of work

CONTROL OF SYSTEMS AND PROCESSES
(cybernetics, informatics, automation)

Staff

- 25 - 30 people (13-15 with PhD, 6-9 PhD students)

Financing

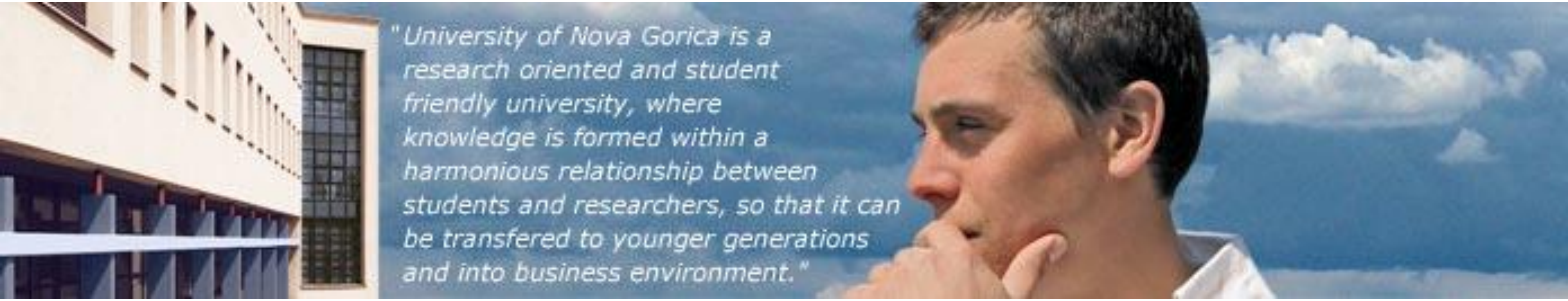
- 60% different ministries and state agencies
- 40% other sources (industry, international projects, etc.)

Current areas of work

- Modelling, control and optimization of (complex) systems
- Detection and localization of faults
- Smart factories
- Device and product development

UNIVERSITY OF NOVA GORICA

SCIENTIA 1995 VINCES



Management

- Student Council
- Governing Board
- Senate
- Int. Advisory Council
- Vice-rector for education
- Rector
- Vice-rector for research and arts

Education

- School of Engineering and Managment
- School for Viticulture and Enology
- School of Environmental Sciences
- School of Arts
- School of Humanities
- Graduate School
- School of Science

Research

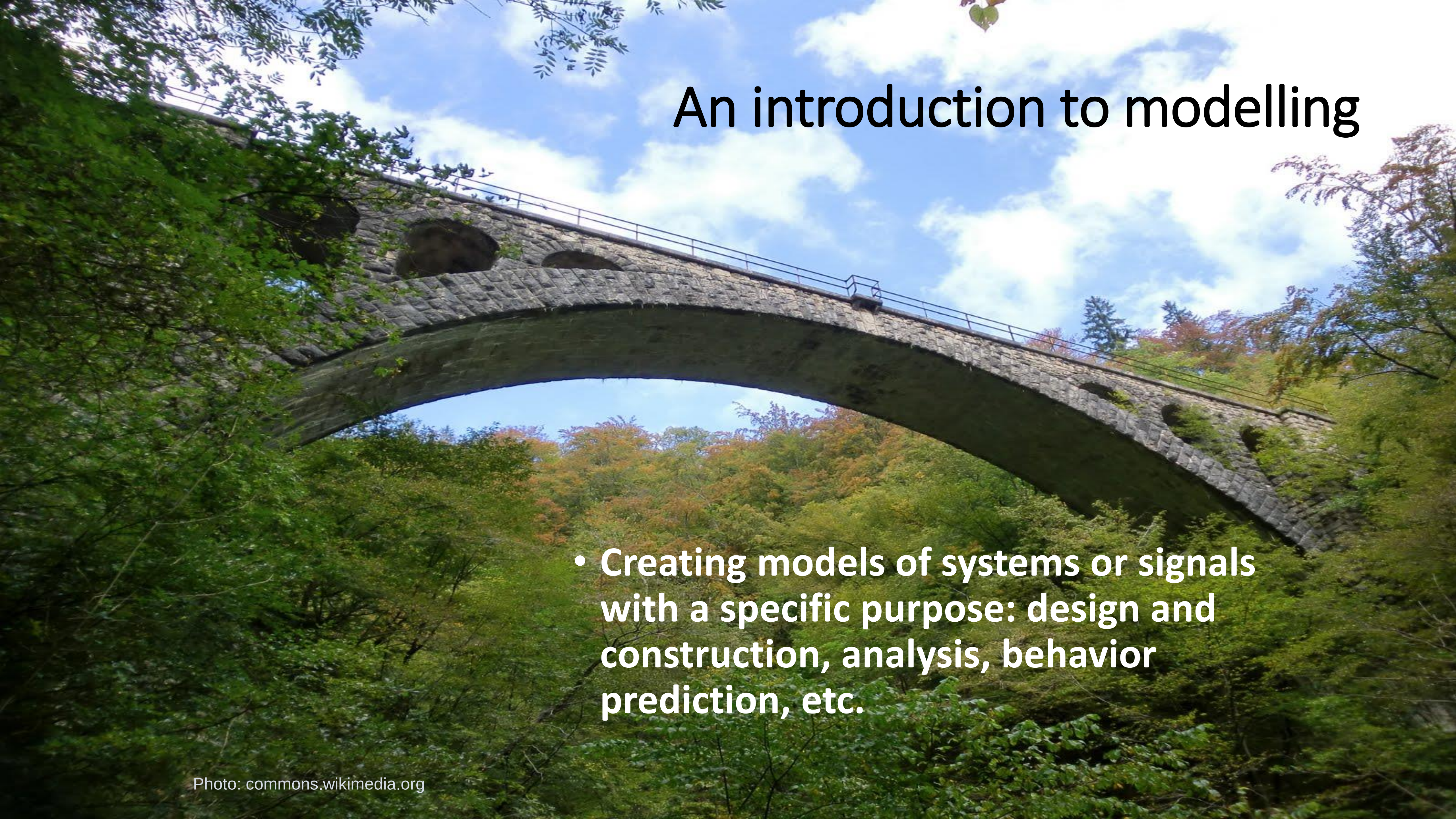
- Materials Research Laboratory
- Centre for Atmospheric Research
- Laboratory for Environmental and Life Sciences
- Center for Information Technologies and Applied Mathematics
- Laboratory of Organic Matter Physics
- Research Centre for Humanities
- Laboratory for Quantum Optics
- Wine Research Centre
- Center for Astrophysics and Cosmology
- Centre for Cognitive Science of Language

Support

- Library
- Administration
- Publisher

- University of Nova Gorica Foundation
- Primorska Technology Park

An introduction to modelling



- **Creating models of systems or signals with a specific purpose: design and construction, analysis, behavior prediction, etc.**

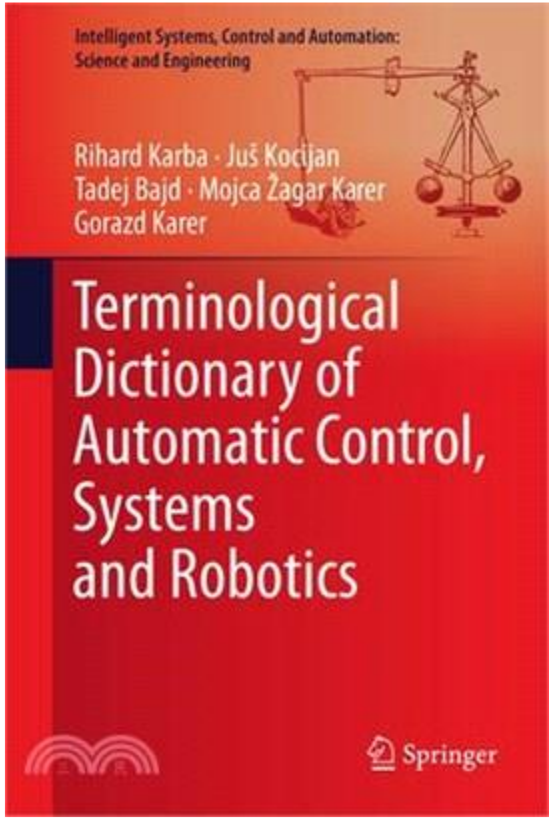
Photo: commons.wikimedia.org

Modelling – why and how

- **Modelling**

Various complex activities that make up the process of model development aiming to improve, e.g., the understanding of the functioning mechanisms of the modelled system, the prediction of its behaviour, the design and the evaluation of control systems, the estimation of the unmeasurable system states, the optimisation of the system behaviour, the development of simulators, as well as to enable the sensitivity analysis and the fault diagnosis.

- First-principles modelling (physics-based modelling, analytical modelling, etc.)
- System identification (experimental modelling , data-driven modelling, black-box modelling, statistical modelling, etc.)
- Hybrid modelling (integrated modelling, statistical postprocessing, explainable AI, etc.)



Linear regression

- Illustrative example,
- static input-output example

$$f(x) = x^3 + \epsilon$$

- 9 input data:

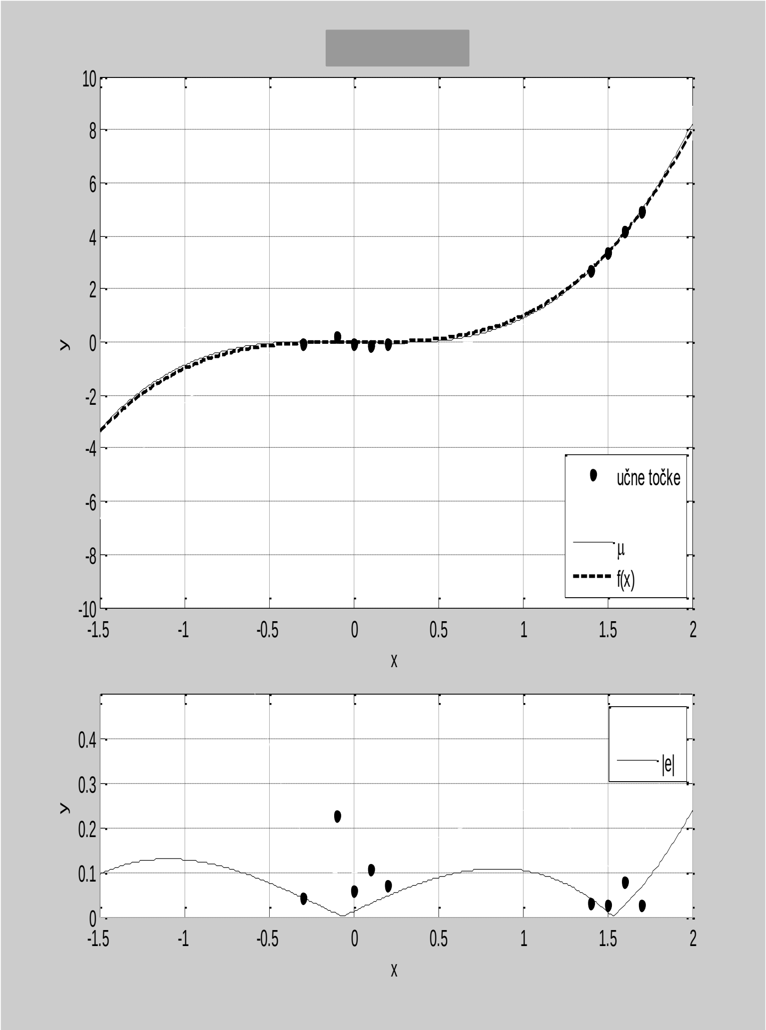
$x_i, y_i \quad \mathcal{N}(0, 0.01)$

Prediction Error Method (PEM):

$$\phi(\mathbf{x}) = [f_1(\mathbf{x}), \ldots, f_D(\mathbf{x})]^T$$

$$J = \sum_{i=1}^{N} (y_i - t_i)^2$$

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=0}^{M-1} w_j \phi_j(\mathbf{x}) = \mathbf{w}^{\mathrm{T}} \phi(\mathbf{x})$$



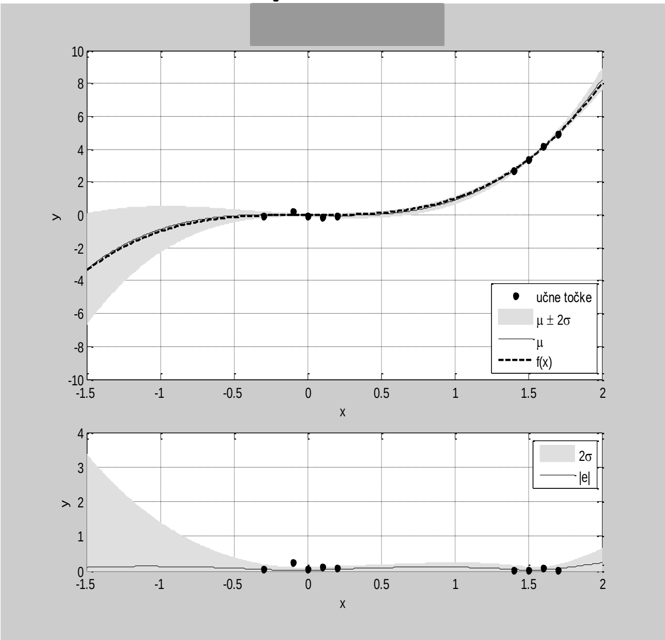
$y = f(\boldsymbol{x}) + \varepsilon =>$ $\quad cov\{\mathbf{w}\} = \sigma_n(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi})^{-1}$

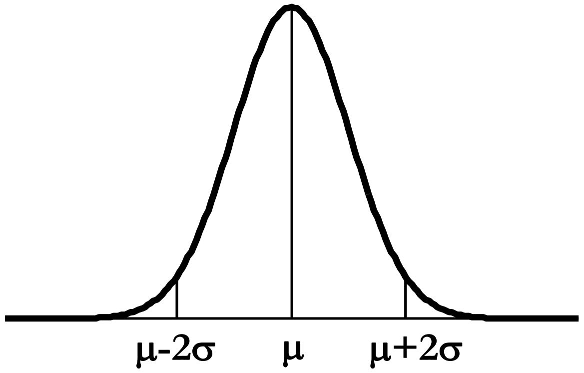
$[\mathbf{w} - \sigma_w \quad \mathbf{w} + \sigma_w]$

$$\mathbf{w} = \left(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^{\mathrm{T}}\mathbf{t}$$

Random variables with a normal distribution at the output

- Random variables with a normal or Gaussian distribution at the output of a linear model,





- uncertainty of parameters,

$$cov\{\mathbf{w}\} = \sigma_n(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi})^{-1}$$

- prediction uncertainty.

$$cov\{y\} = \mathbf{\Phi} cov\{\mathbf{w}\}\mathbf{\Phi}^{\mathrm{T}}$$

Kernel methods for nonlinear systems

- Alternative to modelling with basis functions

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=0}^{M-1} w_j \phi_j(\mathbf{x}) = \mathbf{w}^{\mathrm{T}} \boldsymbol{\phi}(\mathbf{x}) \qquad y(\mathbf{x}, \mathbf{w}) = f\left(\sum_{j=1}^{M} w_j \phi_j(\mathbf{x})\right)$$

  Linear and nonlinear models, e.g. MLP, fuzzy

- Substitute of large number of basis function with a kernel function

$$k(\mathbf{x}, \mathbf{x}') = \boldsymbol{\phi}(\mathbf{x})^{\mathrm{T}} \boldsymbol{\phi}(\mathbf{x}')$$

for linear system $k(\mathbf{x}, \mathbf{x}') = \mathbf{x}^{\mathrm{T}} \mathbf{x}'$

infinite number of basis functions $k(\mathbf{x}, \mathbf{x}') = \exp\left(-\|\mathbf{x} - \mathbf{x}'\|^2 / 2\sigma^2\right)$

radial basis functions $k(\mathbf{x}, \mathbf{x}') = k(\|\mathbf{x} - \mathbf{x}'\|)$

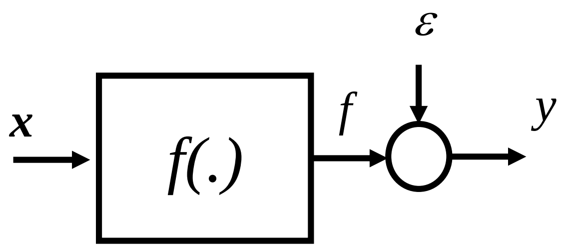
- Kernel functions => kernel methods (e.g., SVM-support vector machines)
- memory-based methods, where model contains training data => nonparametric model
- Parameters of kernel functions are called hyperparameters => fewer optimisation parameters

Gaussian process

- Gaussian process (GP) is a stochastic process: collection of random variables where each has a Gaussian distribution.
- Gaussian proces modelling is a nonparametric method that uses Bayesian learning for a model identification. Result is a Gaussian-process model (GP model).

Gaussian-process models

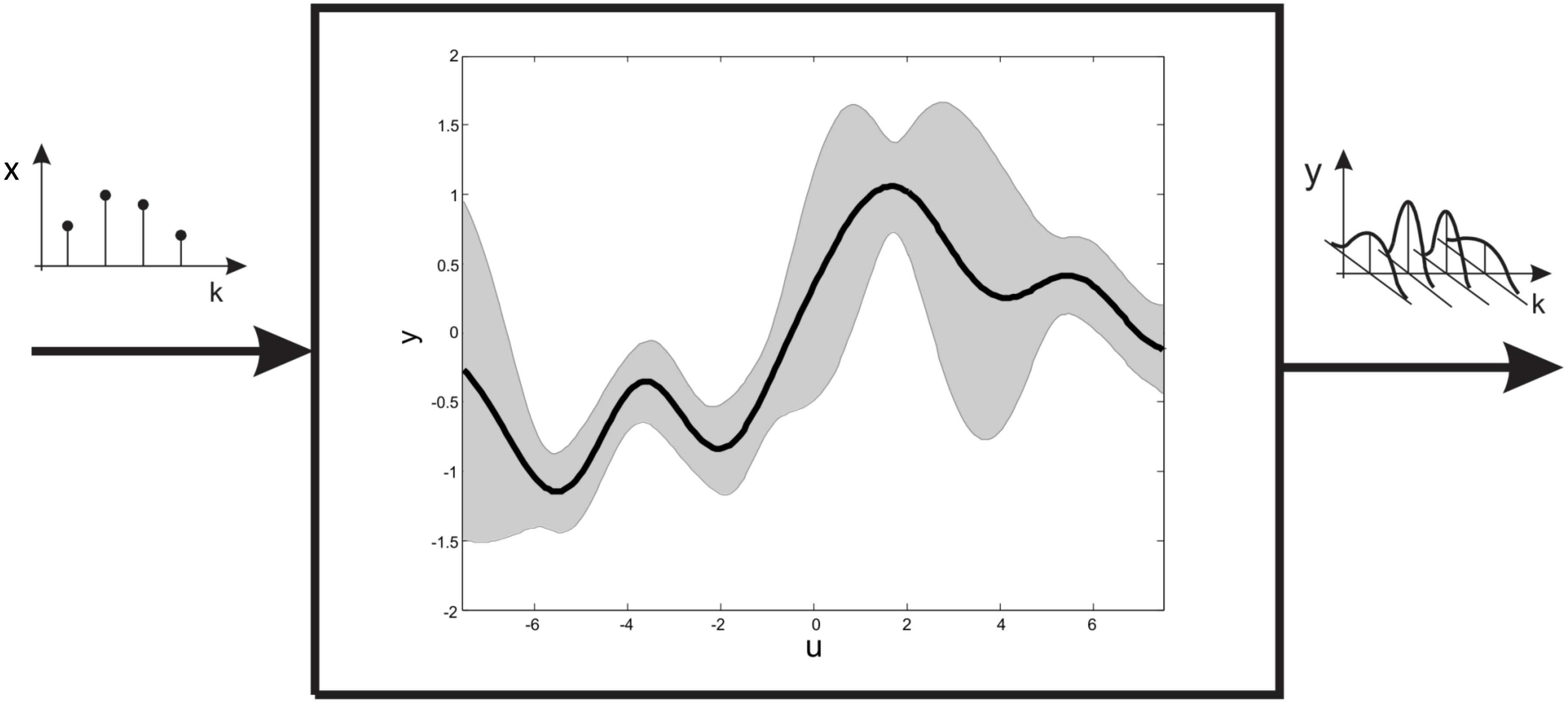
- Kernel model + Bayesian statistics = GP model



$$\mathbf{y} = f(\mathbf{X}) + \boldsymbol{\varepsilon}$$

- Output random variables, GP mapping function, kernel model
- Input is a regression vector of samples,
- Output is a random variable with normal distribution
- Mapping function is a Gaussian process
- Kernels in GP models are covariance functions
- Motives for GP use:
  - Optimisation of complex model structure can be avoided
  - Dealing with limited amount of data
  - Reducing the potential for overfitting

Gaussian-process model



Gaussian-process model

- Probabilistic (Bayes) model. Nonparametric model – no predetermined structure (basis functions) depending on system

$$\underbrace{p(w|y;\theta,\mathcal{M})}^{\text{posterior distribution}} = \frac{\int \overbrace{p(y|w,f;\mathcal{M})}^{\text{likelihood}} \overbrace{p(w,f;\theta,\mathcal{M})}^{\text{prior distribution}} df}{\underbrace{p(y;\theta,\mathcal{M})}_{\text{marginal likelihood}}},$$

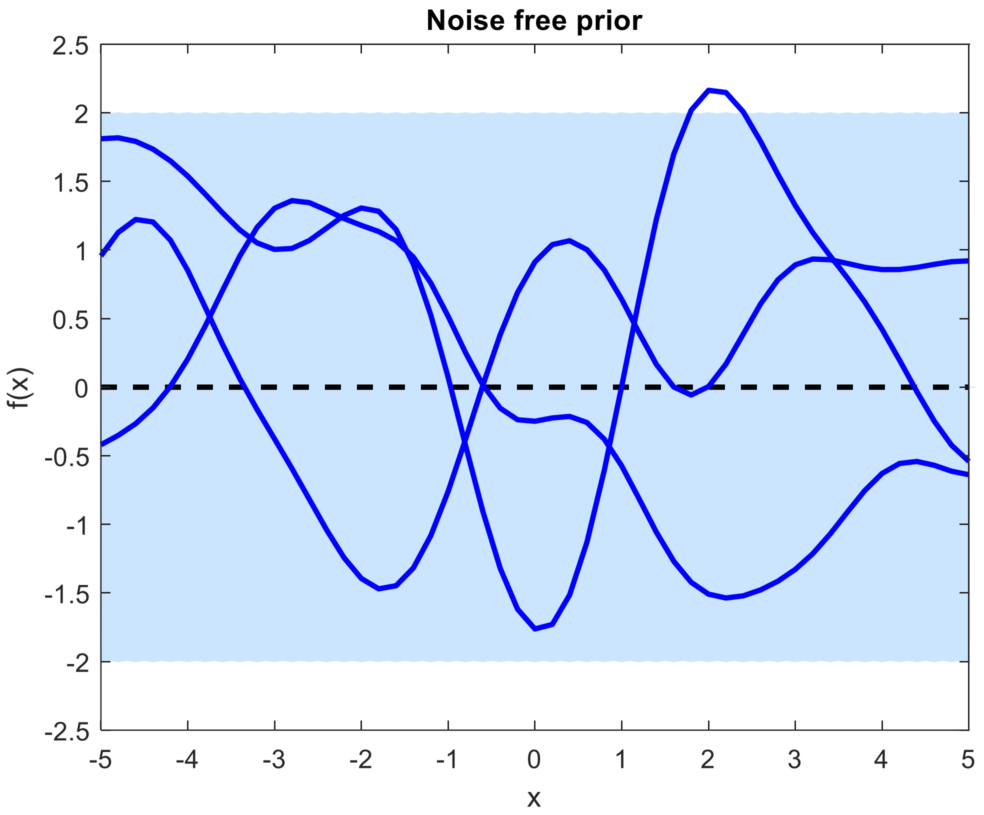
- Determined by:
  - Input/output data $\mathbf{x}_i, y_i$ (data points, not signals) $\mathcal{D} = \{\mathbf{X}, \mathbf{y}\}$ (training data – identification data):
  - Covariance matrix:

$$K_{ij} = \operatorname{cov}(y_i, y_j)$$

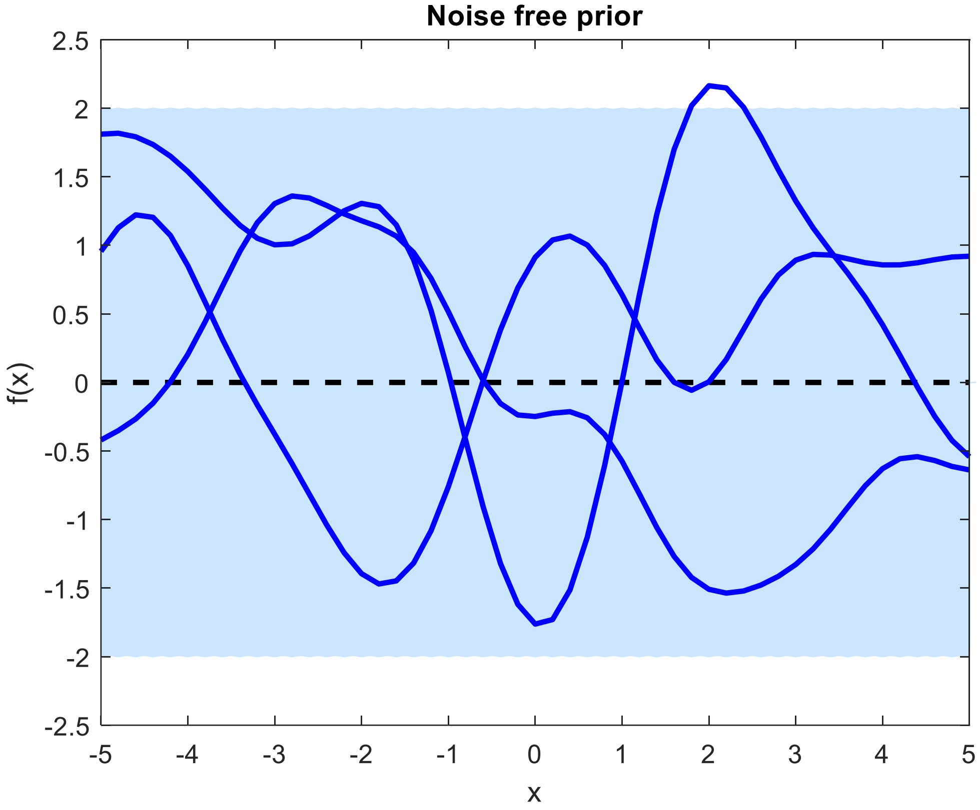
$$\mathbf{K} = \begin{bmatrix} K_{11} & \dots & K_{1N} \\ \vdots & \ddots & \vdots \\ K_{N1} & \dots & K_{NN} \end{bmatrix}$$

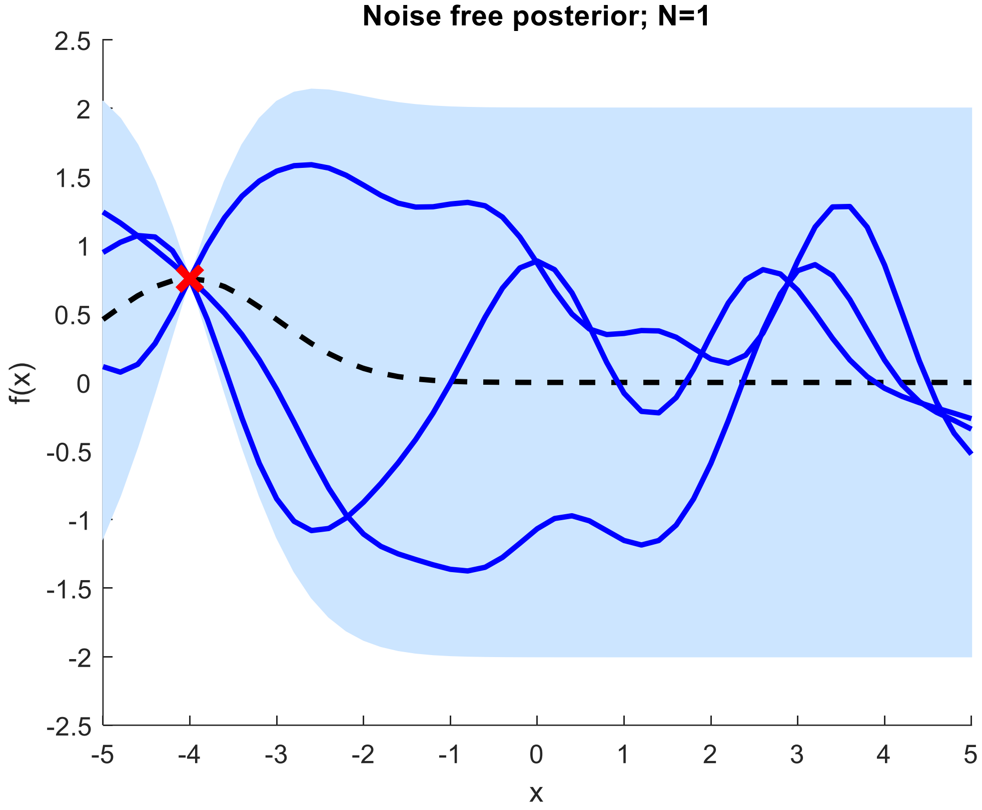
$$\operatorname{cov}(f(\mathbf{x}_i), f(\mathbf{x}_j)) = C(\mathbf{x}_i, \mathbf{x}_j)$$

GP model



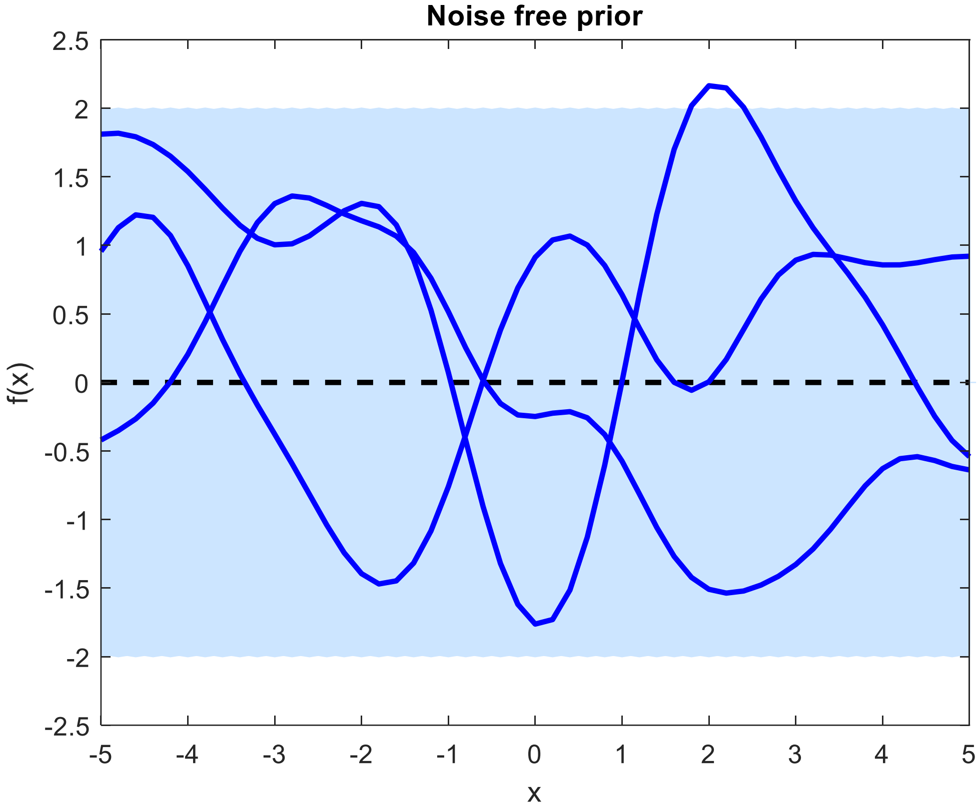
GP model

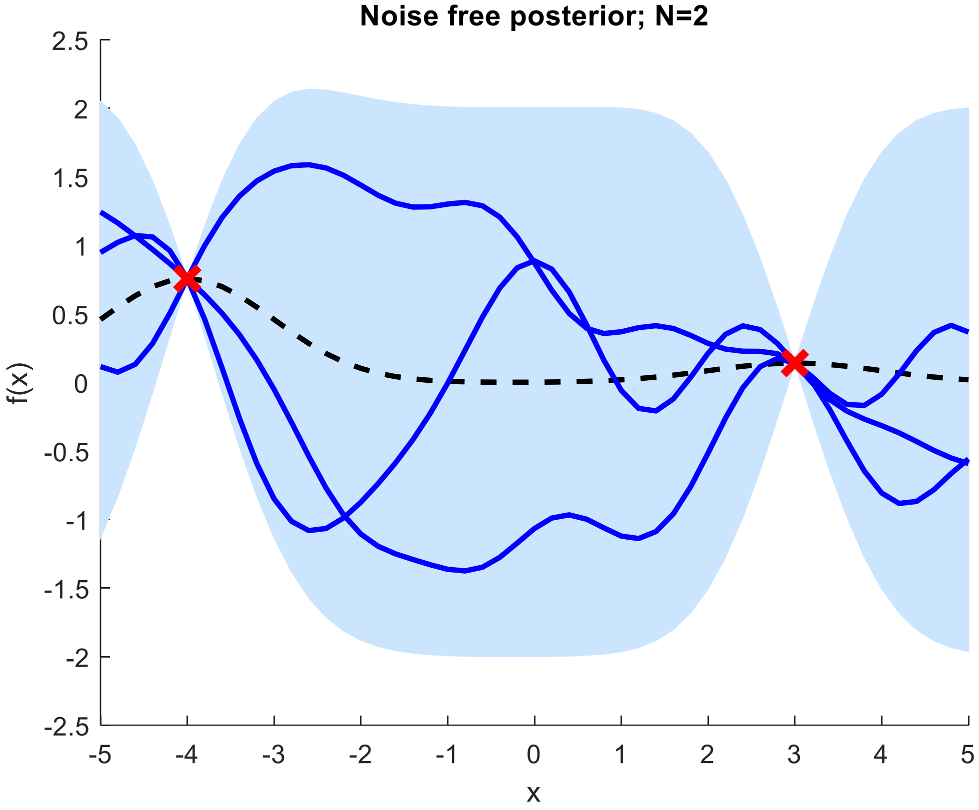




$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right)=\frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

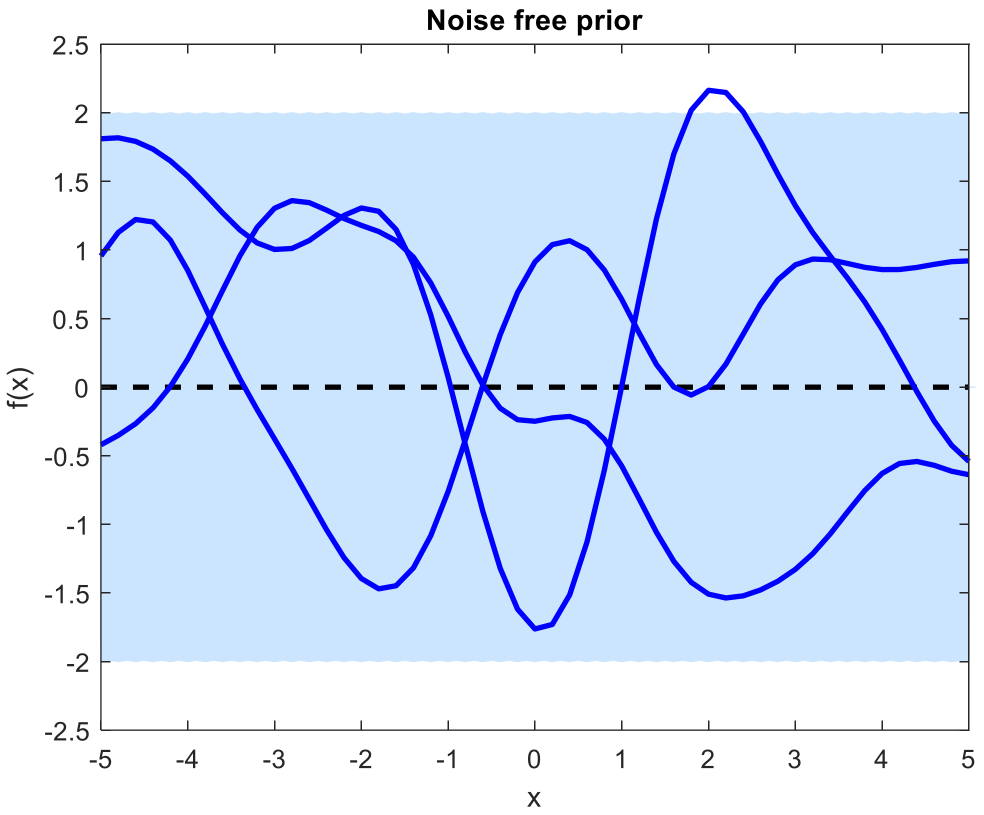
GP model

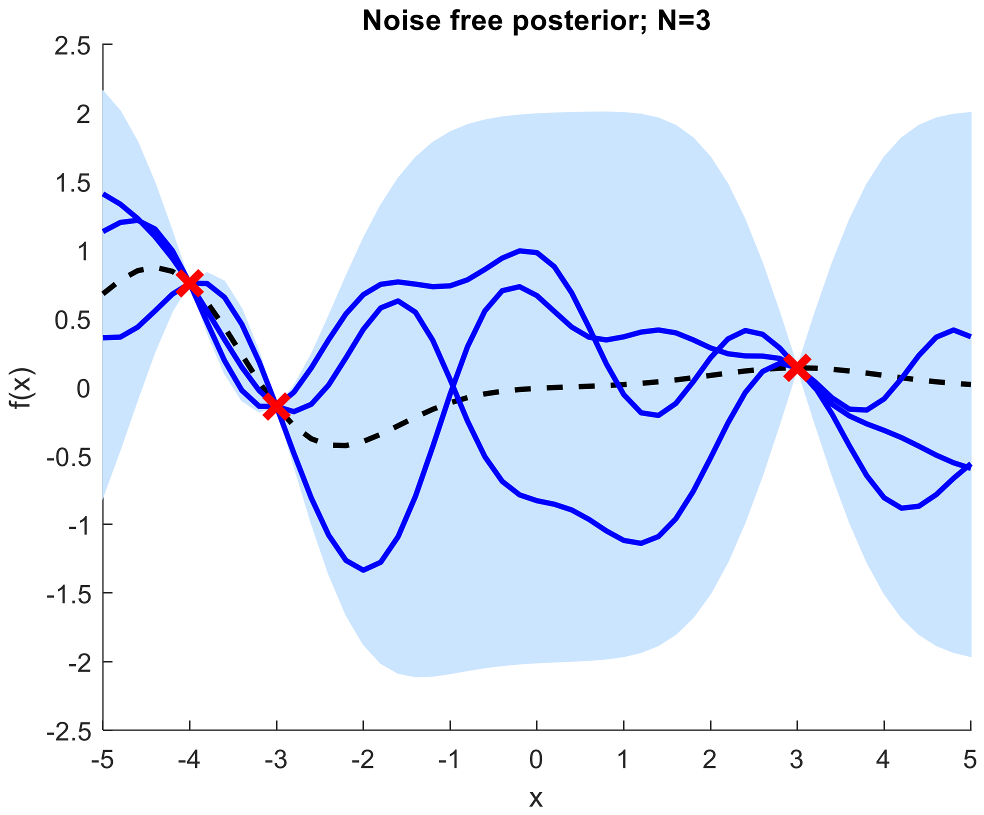




$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

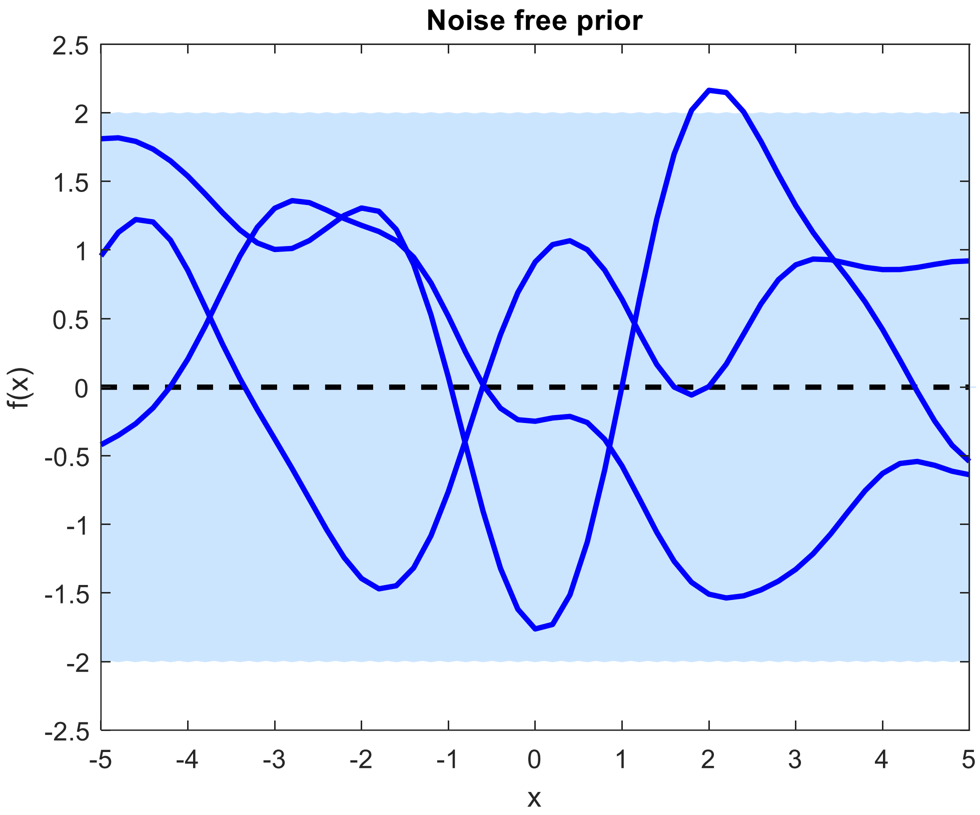
GP model

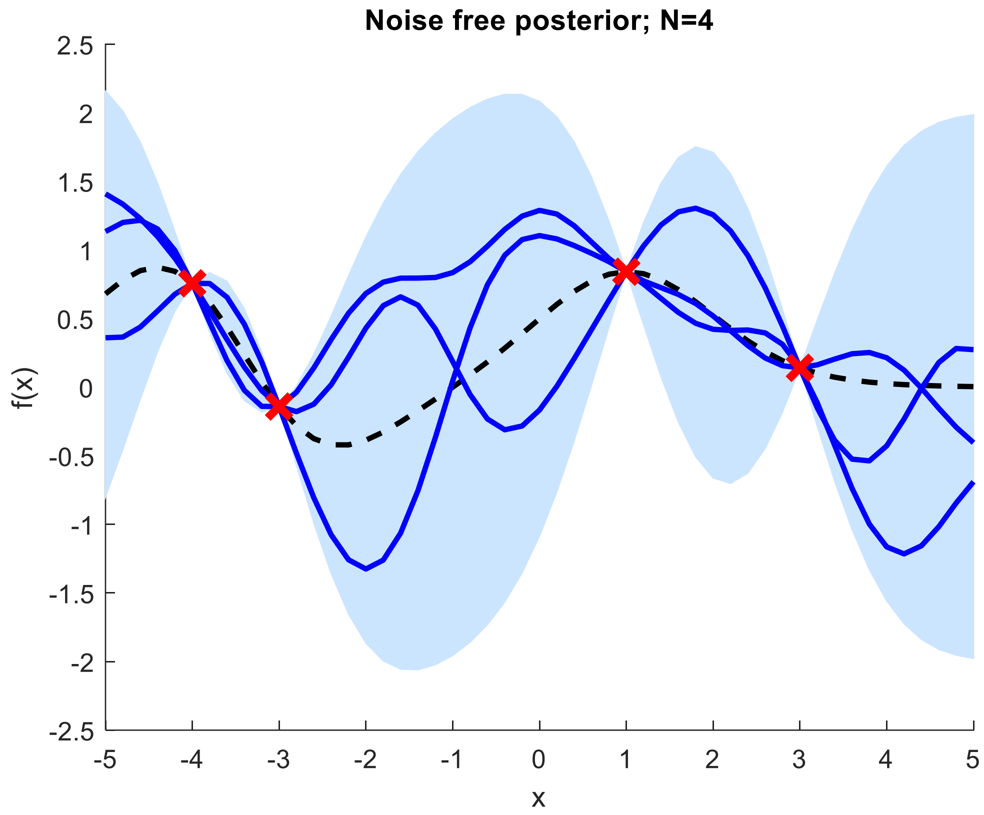




$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

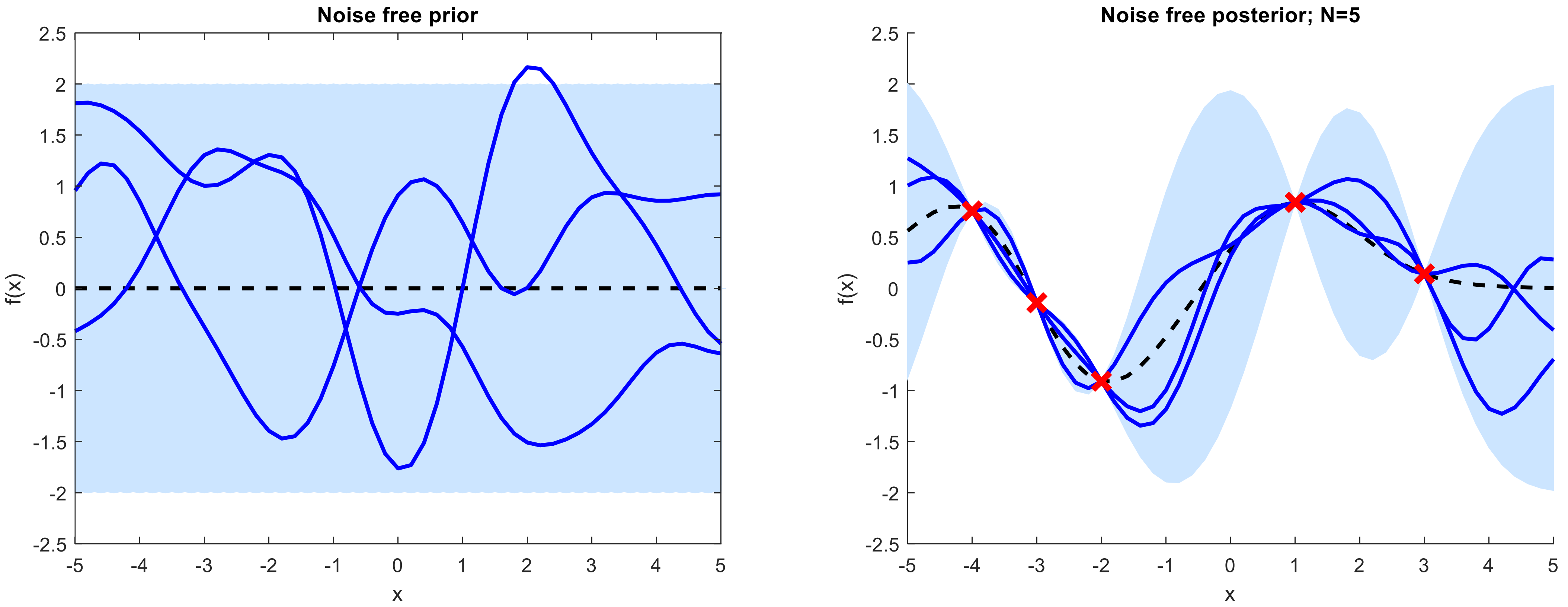
GP model





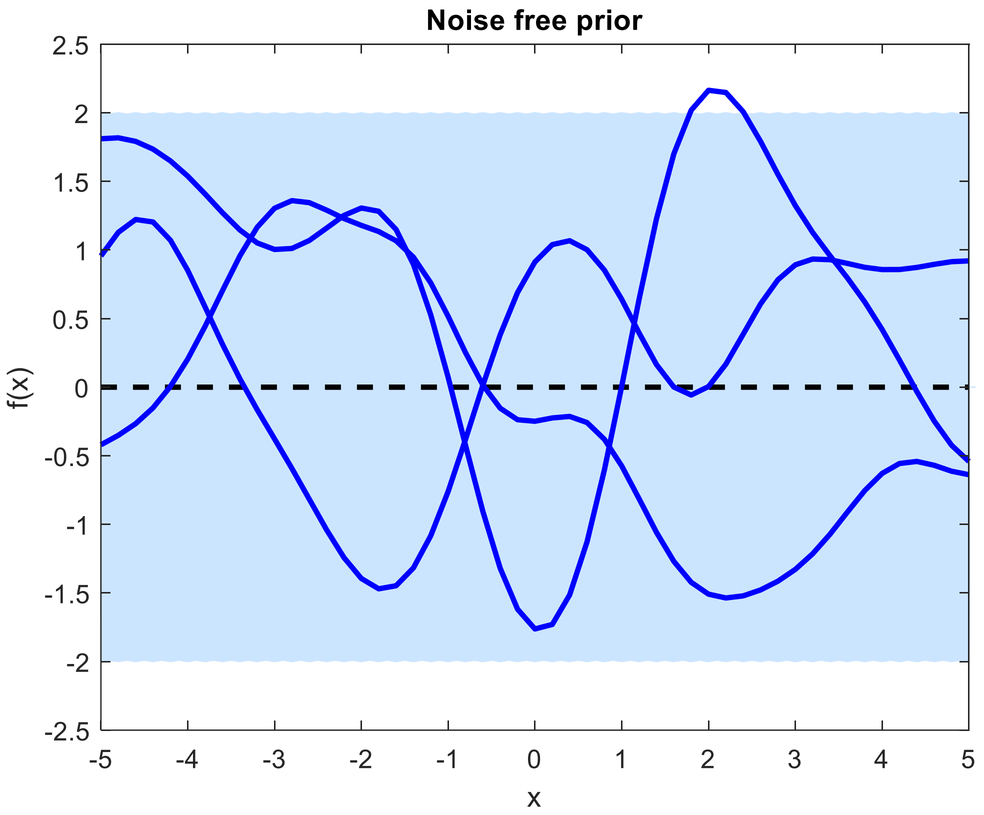
$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

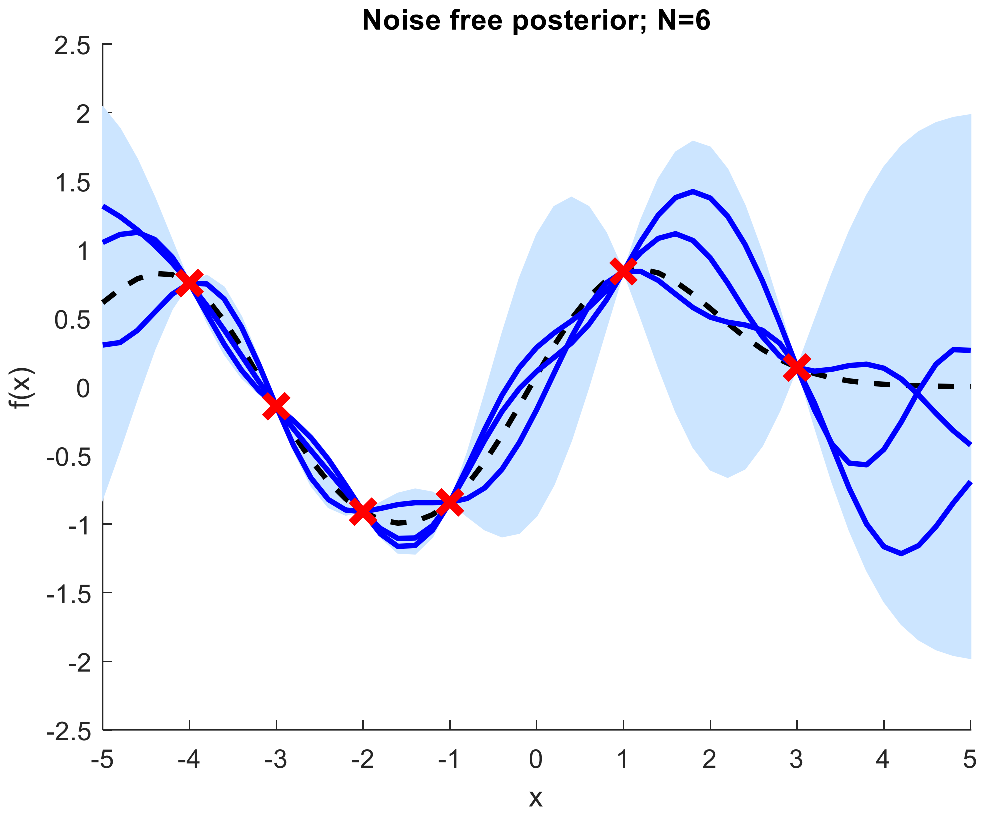
GP model



$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

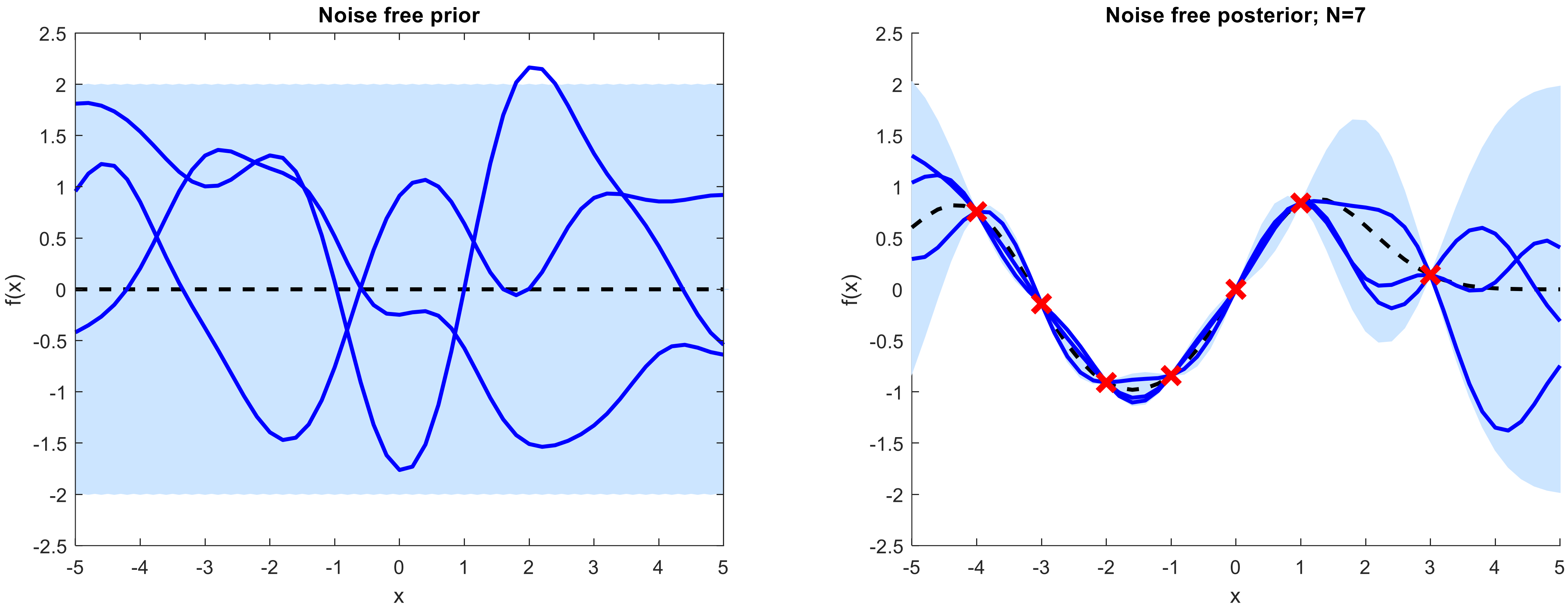
GP model





$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

GP model



$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

Training of hyperparameters

$$p\left(\boldsymbol{\theta}, \sigma_n^2 | \mathbf{y}\right) = \frac{p\left(\mathbf{y} | \boldsymbol{\theta}, \sigma_n^2\right) p\left(\boldsymbol{\theta}, \sigma_n^2\right)}{p\left(\mathbf{y}\right)}$$

- Optimisation:
  - Cost function: marginal likelihood

$$\hat{\boldsymbol{\theta}}, \hat{\sigma}_n^2 = \underset{\boldsymbol{\theta}, \sigma_n^2}{\arg\max} \left[\log\, p\left(\mathbf{y}; \boldsymbol{\theta}, \sigma_n^2\right)\right],$$

where

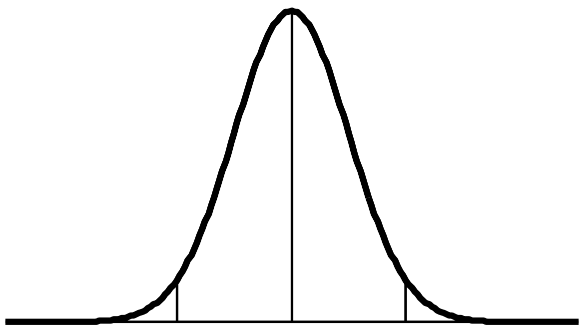
$$\log\, p\left(\mathbf{y}; \boldsymbol{\theta}, \sigma_n^2\right) = \underbrace{-\frac{1}{2}\mathbf{y}^T \left(\mathbf{K}_{f,f} + \mathbf{I}\sigma_n^2\right)^{-1} \mathbf{y}}_{\text{data fit term}} - \underbrace{\frac{1}{2}\log\left|\mathbf{K}_{f,f} + \mathbf{I}\sigma_n^2\right|}_{\text{regularization term}} - \underbrace{\frac{n}{2}\log 2\pi}_{\text{constant}}.$$

Prediction with GP model

- Prediction of the output based on similarity test input – training inputs

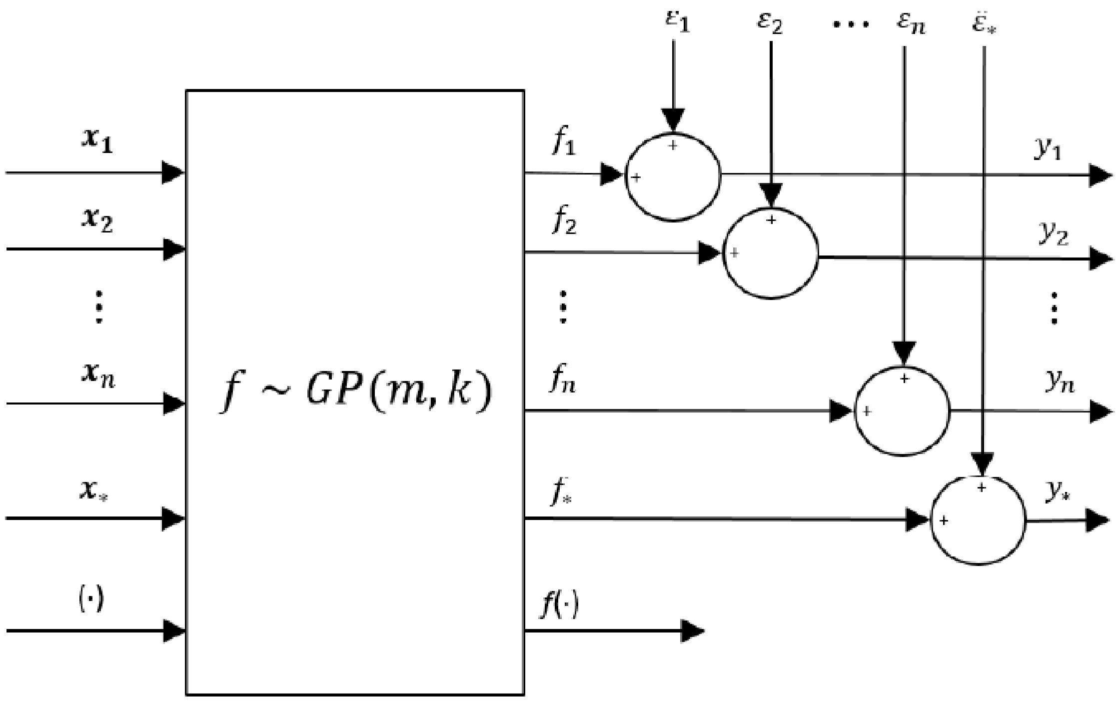
$$\mathcal{GP}\left(f\left(\cdot\right)|\mathbf{y}\right) = \frac{p\left(\mathbf{y}|\mathbf{f}\right)\mathcal{GP}\left(f\left(\cdot\right)\right)}{p\left(\mathbf{y}\right)}$$

- Output: normal distribution
  - Predicted mean $\mathbb{E}[\mathbf{f}_*|\mathbf{y}] = \mu(\mathbf{x})$
  - Prediction variance $\mathbb{V}[\mathbf{f}_*|\mathbf{y}] = \sigma^2(\mathbf{x})$

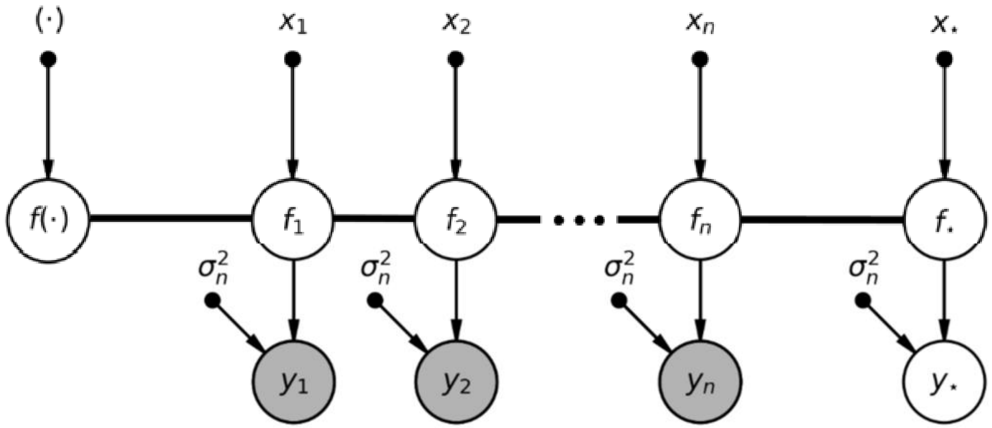


$$\mathbb{E}[\mathbf{f}_*|\mathbf{y}] = \mathbf{K}_{*,f}\left(\mathbf{K}_{f,f} + \mathbf{I}\sigma_n^2\right)^{-1}\mathbf{y},$$

$$\mathbb{V}[\mathbf{f}_*|\mathbf{y}] = \mathbf{K}_{*,*} - \mathbf{K}_{*,f}\left(\mathbf{K}_{f,f} + \mathbf{I}\sigma_n^2\right)^{-1}\mathbf{K}_{f,*}$$



(a) Block diagram of a GP regression model.

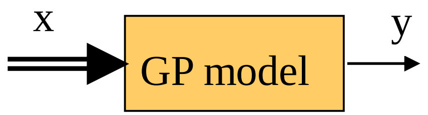


Probabilistic Graphical Model - PGM

(b) PGM of a GP regression model.

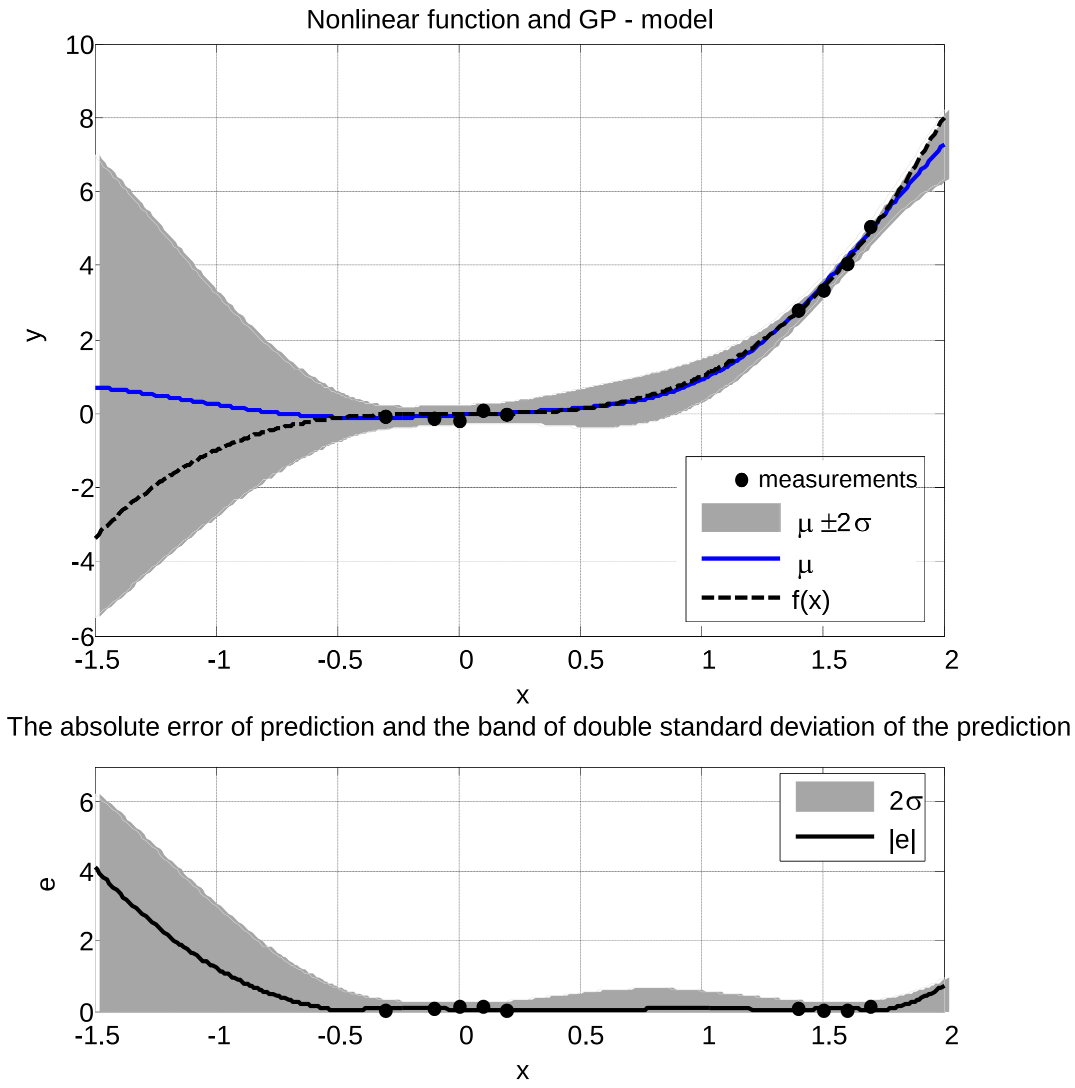
Static illustrative example

- Static example:



$$f(x) = x^3 + \epsilon$$

- x_i, y_i $\mathcal{N}(0, 0.01)$
- 9 measured data:
- Gray belt
- Sparse data → increased variance (higher uncertainty).



Properties of GP models (compared to a model based on basis functions) :

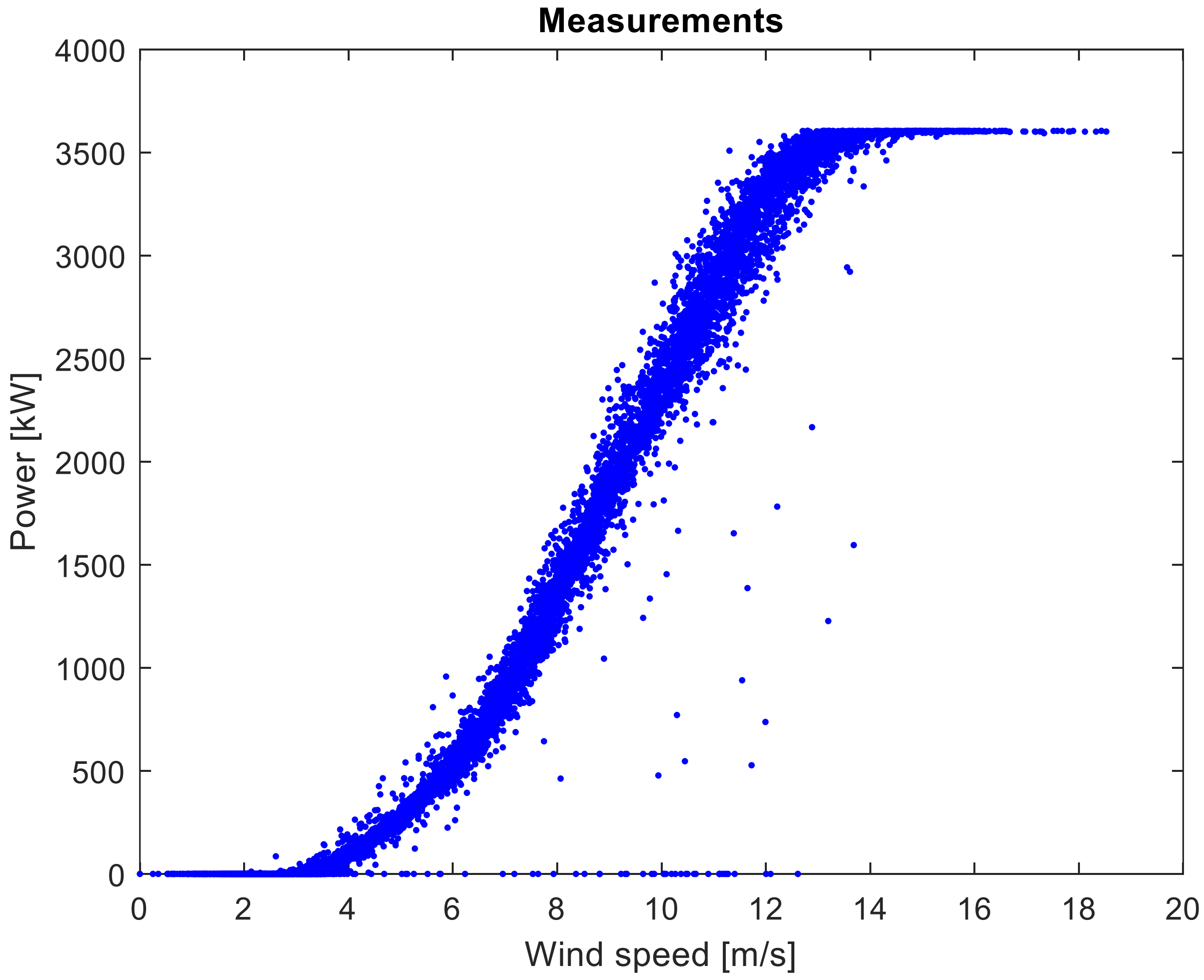
- a smaller number of optimization and design parameters,
- prediction uncertainty depends on the amount of identification data,
- overfitting is minimised,
- works satisfactorily even with a small number of identification data,
- smoothing data with noise,
- enables the use of prior knowledge, *
- easy to use (engineering practice),

- the computational load increases with the amount of identification data,
- a relatively young methodology, still in development
- nonparametric model.

* Also possible with some other models

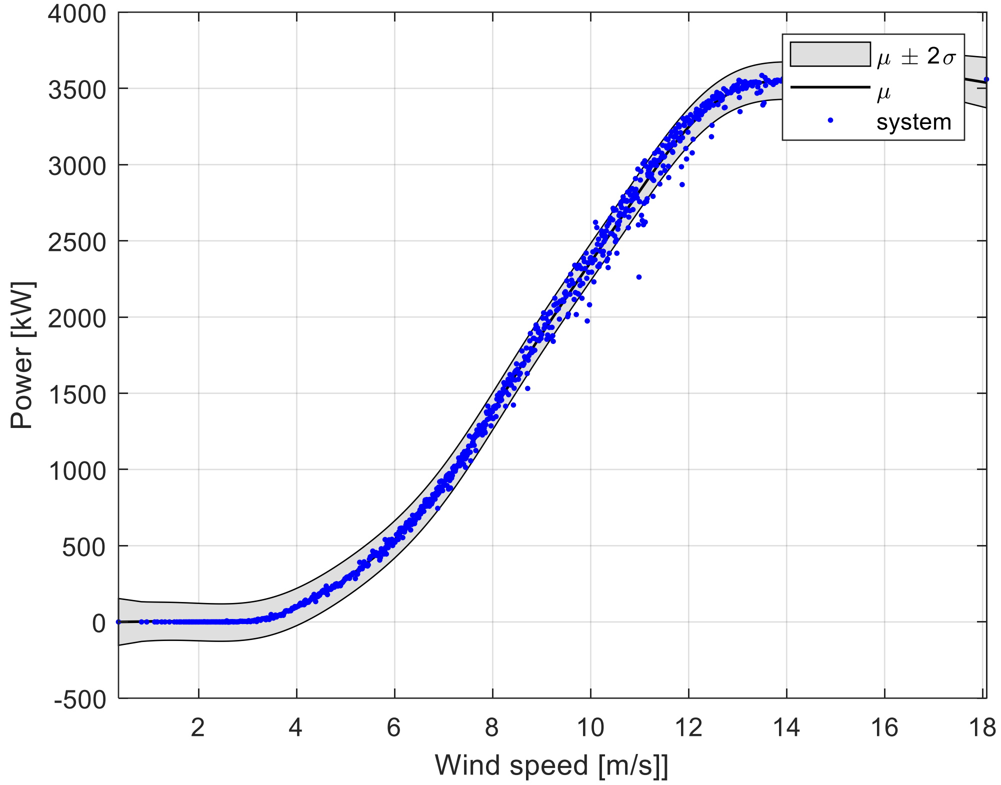
Example of smoothing - the power curve of wind turbine

- Static system

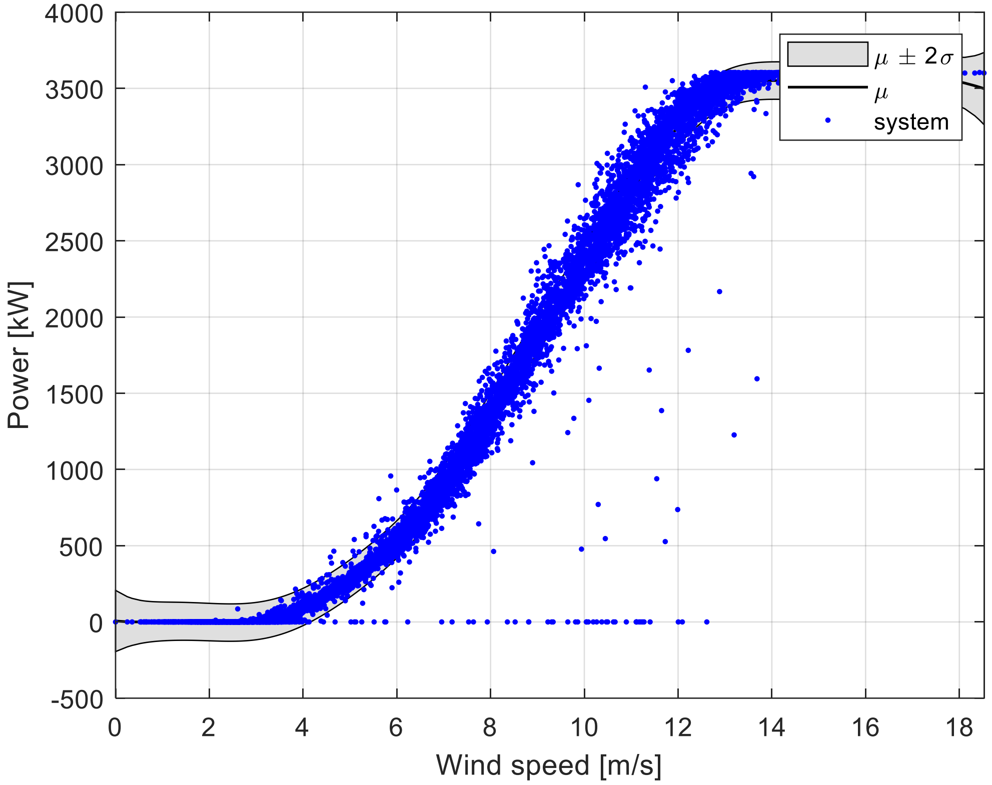


Power curve (2)

1/10 data for training

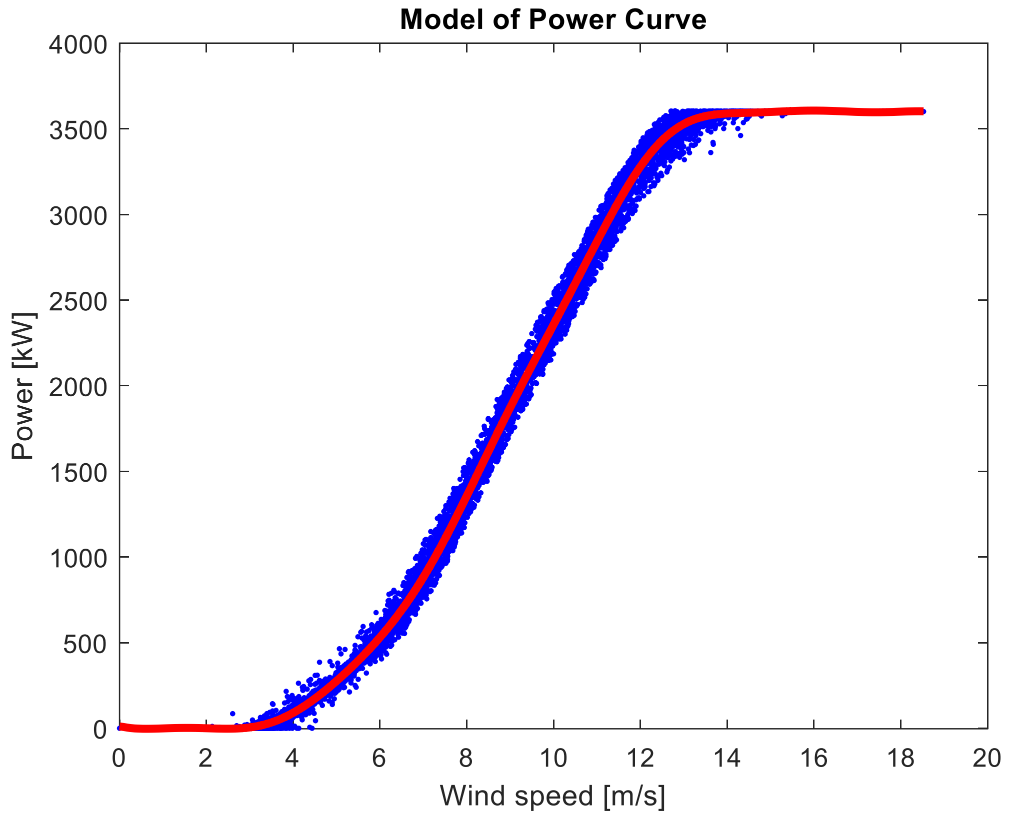


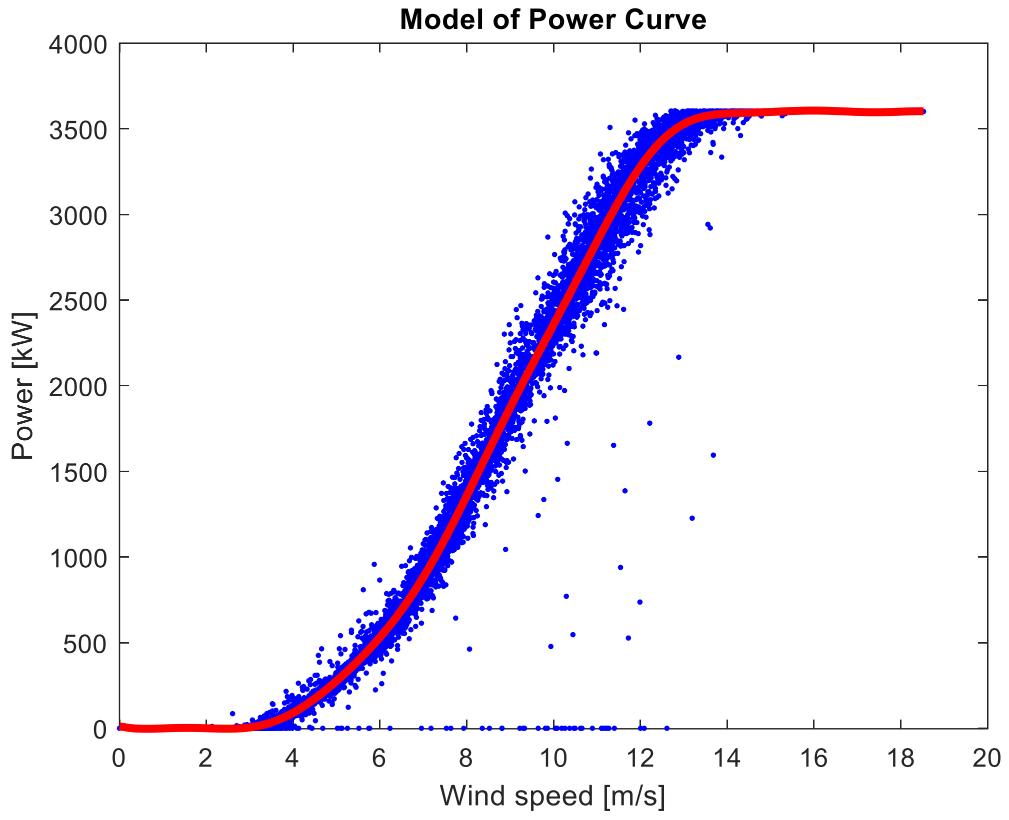
complete data



Next step: elimination of data
outside +/- 3σ band

Power curve (3)





NRMSE = 0.051

Static vs. dynamic

- **Dynamic system**

A system in which time-dependent qualitative or quantitative changes are taking place. The output of such a system depends not only on the current input value but also on the previous input values.

A dynamic system is a system or process in which motion occurs, or includes active forces, as opposed to static conditions with no motion.

- Dynamic models (system identification):

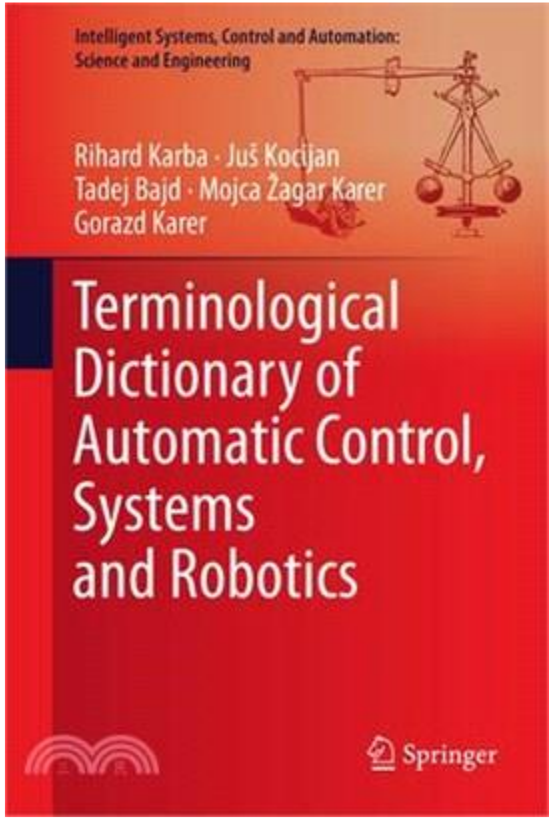
conventional approach (ANN, decision trees, fuzzy models, etc.) delayed inputs and outputs as regressors.

Modelling – why and how

- **Modelling**

Various complex activities that make up the process of model development aiming to improve, e.g., the understanding of the functioning mechanisms of the modelled system, the prediction of its behaviour, the design and the evaluation of control systems, the estimation of the unmeasurable system states, the optimisation of the system behaviour, the development of simulators, as well as to enable the sensitivity analysis and the fault diagnosis.

- First-principles modelling (physics-based modelling, analytical modelling, etc.)
- System identification (experimental modelling , data-driven modelling, black-box modelling, statistical modelling, etc.)
- Hybrid modelling (integrated modelling, statistical postprocessing, explainable AI, etc.)



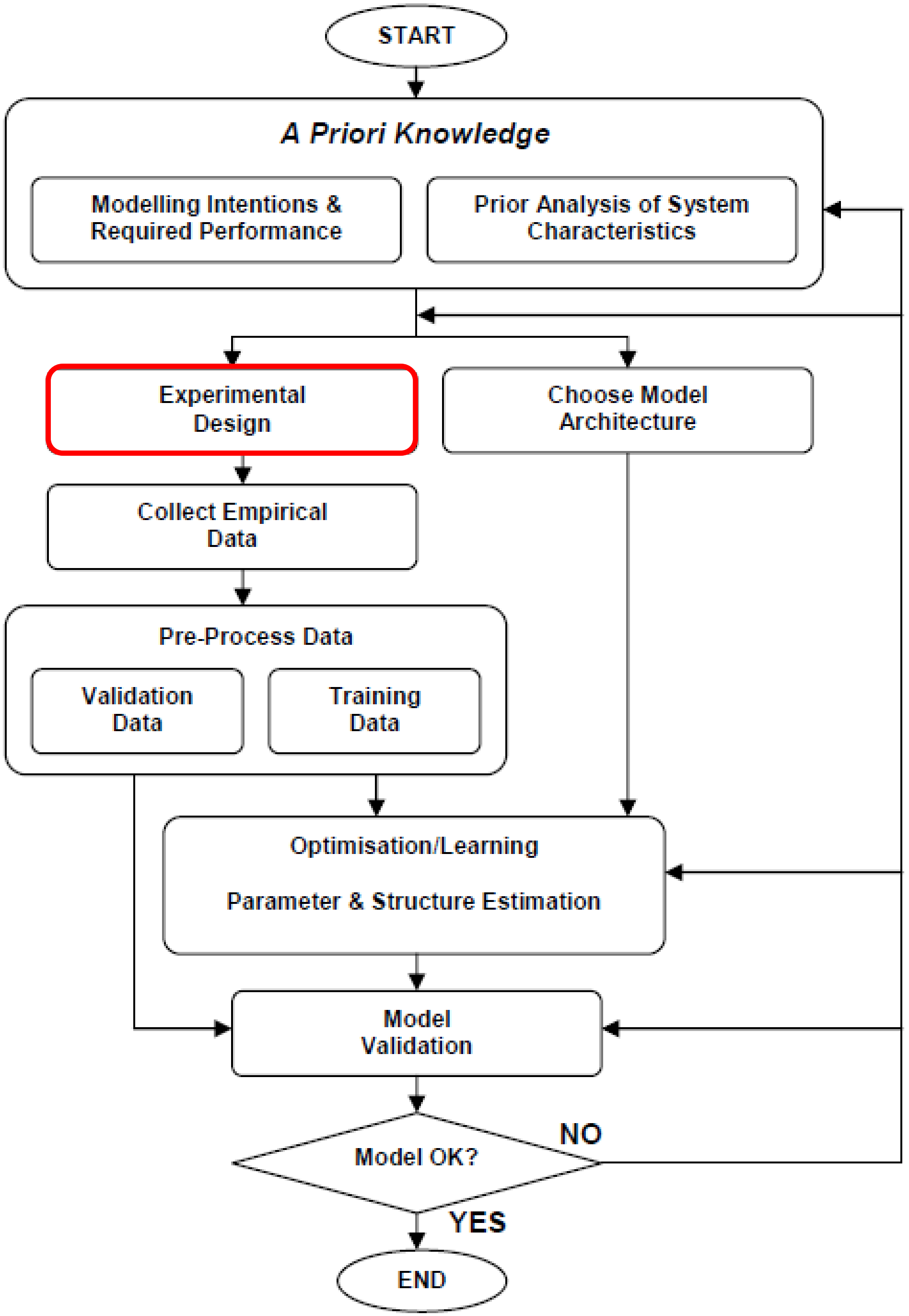
- What is **system identification**?

*Methods to build mathematical models of **dynamic systems** from measured data.*

- Purpose: prediction, forecasting, simulation, control design, systems analysis, etc.

Identification - why and how

- Application: unknown detailed physical background of the systems.
- Application: constraints of the engineering practice.



- ✓ Identification of **nonlinear** dynamic systems.
- ✓ Connection with statistics and machine learning.
  - Various methods → artificial neural networks (UNM), decision trees, fuzzy models, etc.
  - Difficult to use (structure selection, large number of optimization parameters and structure complexity, large amount of data required for optimization, bias-variance, inadequate fit, overfitting).
  - Planning in practice: the problem of confidence in the model.

- Dynamic models:

A common approach (UNM, fuzzy models, etc.) is that the regressor vector contains delayed samples of the input and output signal.

Input-output identification pairs $\mathbf{x}_i / y_i$:

$\mathbf{x}_i$ ... Regressor vector

[u(k-1),..,u(kL), y(k-1),..,y(kL)]

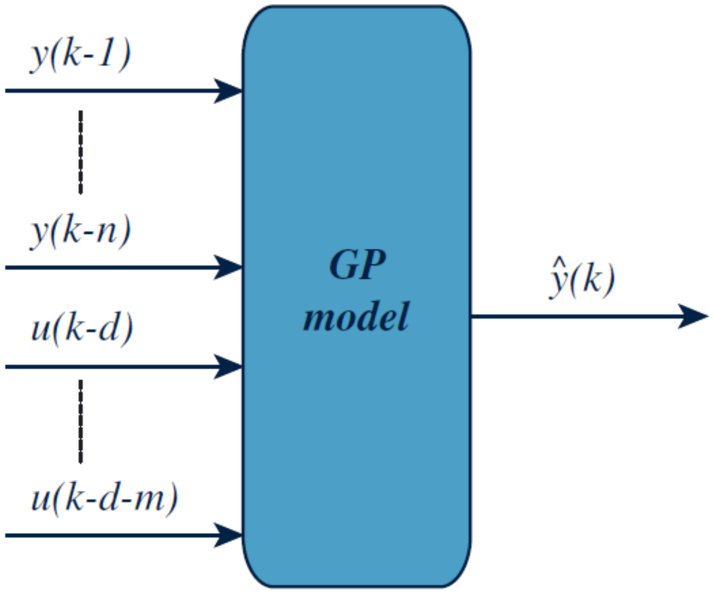
y_i ... Output value of the model

y(k)

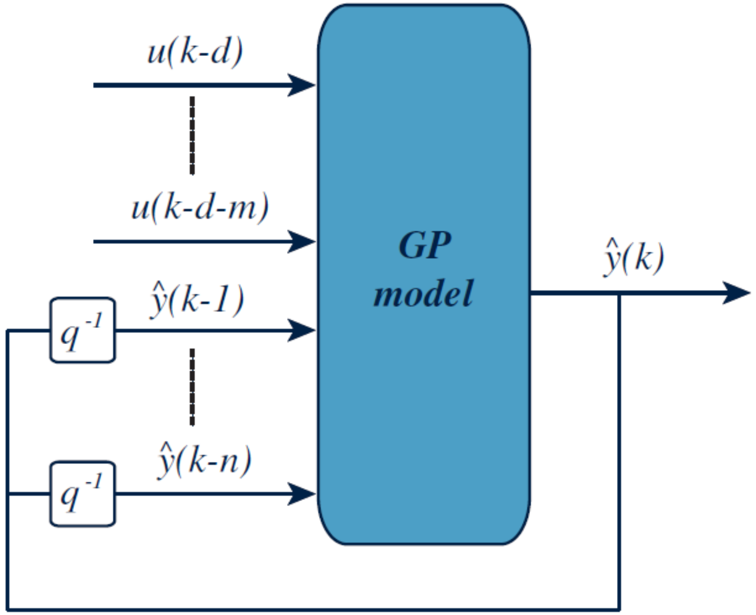
Why modeling dynamic systems with GP-models?

- Dynamic systems in engineering.
- Relatively large dimensions of the entrance space.
- Specific distribution of input data.
- Dynamic systems:
  - prediction, simulation,
  - simulation evaluation,
  - noise handling,
  - convergence.

Model prediction and simulation



Model with generalized error
- ARX model



Model with output error
- OE model

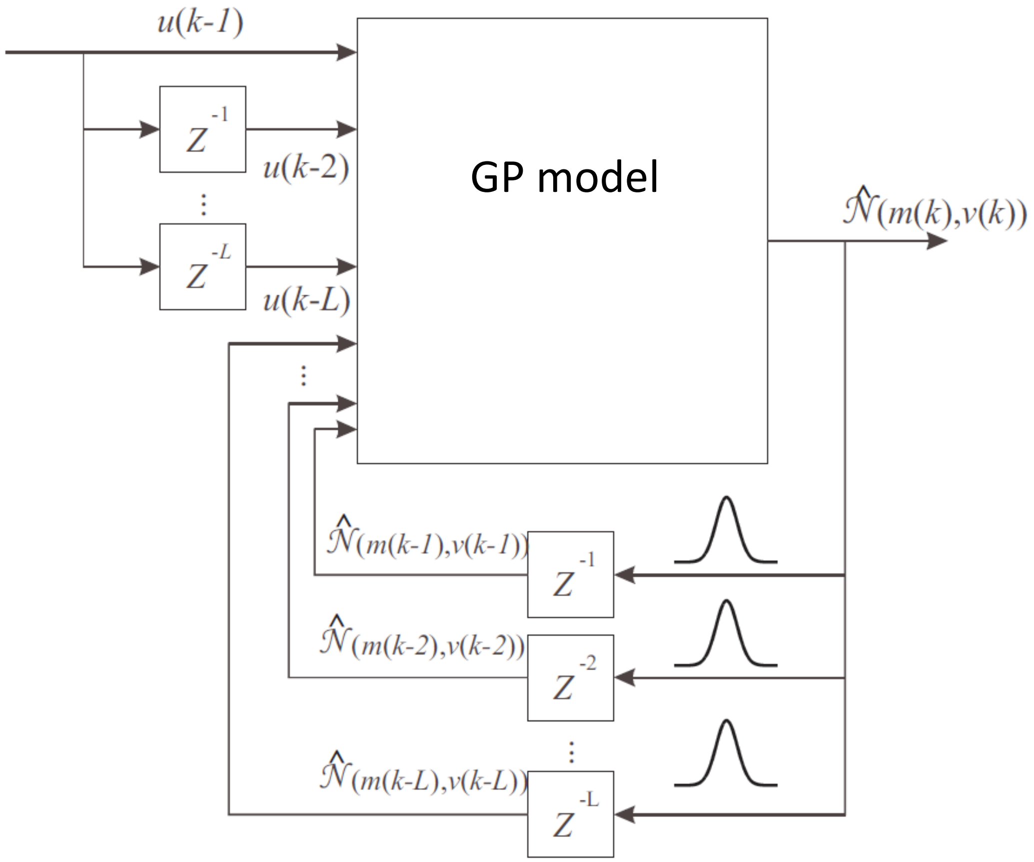
Dynamic model simulation

- A simulation
  - **"naive" ... m(k),**
  - by propagating

  m(k),v(k):

  - analytical approximation:

    - **The Taylor method,**
    - **Model matching,**
  - numerical approximation:
    - **Monte Carlo method.**



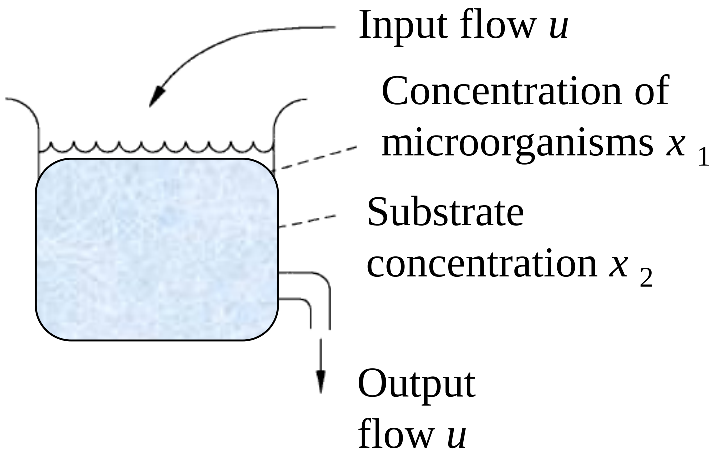
Example of identification – bioreactor

Identification process, properties of the identified GP-model →

Bioreactor - discrete dynamic nonlinear system of the 2nd order:

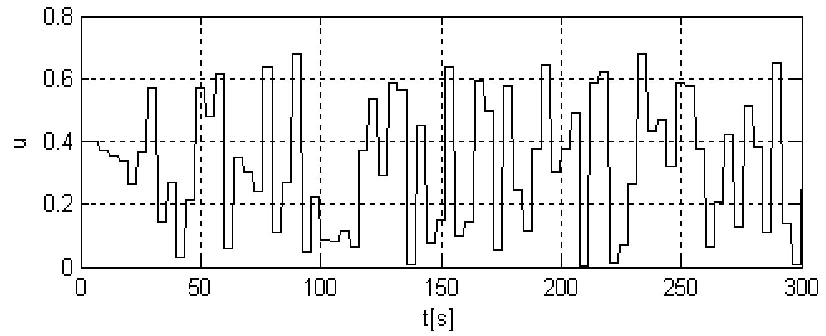
$$
\begin{aligned}
x_1(k+1) &= x_1(k) + 0.5\frac{x_1(k)x_2(k)}{x_1(k)+x_2(k)} - 0.5u(k)x_1(k) \\
x_2(k+1) &= x_2(k) - 0.5\frac{x_1(k)x_2(k)}{x_1(k)+x_2(k)} - 0.5u(k)x_2(k) + 0.05u(k) \\
y(k) &= x_1(k) + \epsilon(k)
\end{aligned}
$$

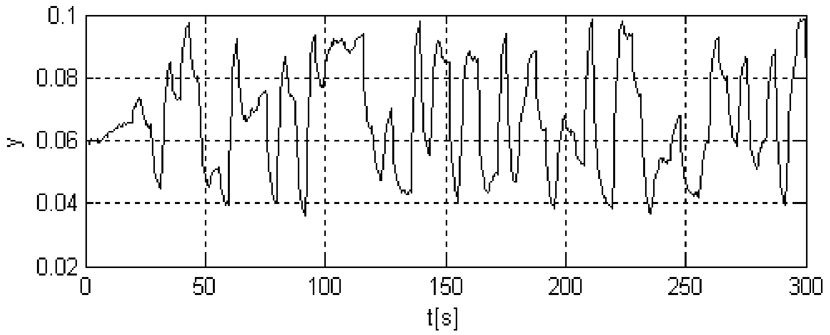
x_1 … concentration of microorganisms,
x_2 … substrate concentration,
u … output flow, $0 \leq u(k) \leq 0.7$,
ϵ. white Gaussian noise, $\sigma = 0.5\%$ (ymax-ymin).



Bioreactor (2)

1. Definition of the purpose of the model: response to an input signal.
2. Experiment design:
   - input and output signal →
   - ~ 600 points for identification.





3. Conducting an experiment,

data processing.

4. Selection of a model structure:
   - Gaussian covariance function (iterative)

   (stationarity, smoothness),
   - selection of regressors. [y(k-1)…y(kL) u(k-1)…u(kL)], y(k)

Bioreactor (3)

- validation → n=2,
- ARD → reduction of the number of regressors,

[u (k-1) u (k-2) ~~y (k-1)~~ y (k-2)]

[u (k-1) u (k-2) y (k-2)]

Table 2.3 Values of the validation performance measures and the hyperparameters of different bioreactor GP models with the best measure values in bold.

Model	Ident. data			Valid. data		Hyperparameters									noise‡
order	ℓ_1	$SMSE_2$	$MSLL_2$	SMSE	MSLL	$l_{y_4} = \frac{1}{\sqrt{w_{y_4}}}$	$l_{y_3} = \frac{1}{\sqrt{w_{y_3}}}$	$l_{y_2} = \frac{1}{\sqrt{w_{y_4}}}$	$l_{y_1} = \frac{1}{\sqrt{w_{y_1}}}$	$l_{u_4} = \frac{1}{\sqrt{w_{u_4}}}$	$l_{u_3} = \frac{1}{\sqrt{w_{u_3}}}$	$l_{u_2} = \frac{1}{\sqrt{w_{u_2}}}$	$l_{u_1} = \frac{1}{\sqrt{w_{u_1}}}$	σ_f	σ_n
4	1628	$3.1 \cdot 10^{-3}$	-3.31	$5.1 \cdot 10^{-3}$	-3.22	9.2	6.3	4.8	977	406	7.2	7.0	8.9	2.5	$5.0 \cdot 10^{-4}$
3	1621	$2.3 \cdot 10^{-3}$	-3.38	$3.8 \cdot 10^{-3}$	-3.32	×	20.0	3.90	49.3	×	17.4	17.3	15.6	4.2	$5.1 \cdot 10^{-4}$
2	1612	$1.2 \cdot 10^{-3}$	-3.55	$1.9 \cdot 10^{-3}$	**-3.50**	×	×	5.8	29.8	×	×	13.3	14.7	5.3	$5.3 \cdot 10^{-4}$
2 ♠	1603	**$7.8 \cdot 10^{-4}$**	**-3.57**	**$1.2 \cdot 10^{-3}$**	-3.49	×	×	5.2	×	×	×	13.1	14.4	5.6	$5.4 \cdot 10^{-4}$
2 ♣	785	$3.5 \cdot 10^{-3}$	−2.36	$3.1 \cdot 10^{-4}$	−2.37	×	×	4.07	×	×	×	10.7	8.7	3.3	$2.1 \cdot 10^{-3}$

Notes:
The index $_1$ denotes the first set of data for identification
The index $_2$ denotes the second set of data for identification
‡ the identified standard deviation of noise σ_n
♠ reduced number of regressors by masking the regressor $y(k-1)$
♣ identified on the output signal with the increased standard deviation of noise, $\sigma_\nu = 0.002$

Kocijan (2016) Modeling and control of dynamic systems with Gaussian process models, Springer.

Bioreactor (4)

2^{nd} order model	Ident. data		
Covariance function	ℓ_1	$SMSE_2$	$MSLL_2$
Squared exponential + ARD	**1603**	$7.8 \cdot 10^{-4}$	**-3.57**
Matérn + ARD $d = \frac{3}{2}$	1591	$\mathbf{7.2 \cdot 10^{-4}}$	-3.54
Rational quadratic + ARD	**1603**	$4.1 \cdot 10^{-3}$	-3.31
Linear + ARD	1281	$1.7 \cdot 10^{-2}$	-1.46
Matérn $d = \frac{3}{2}$	1587	$1.2 \cdot 10^{-3}$	-2.03
Matérn $d = \frac{5}{2}$	1597	$1.7 \cdot 10^{-3}$	-1.58
Neural network	1596	$5.2 \cdot 10^{-3}$	-1.00
Squared exponential	1600	$1.6 \cdot 10^{-3}$	-1.35

Kocijan (2016) Modeling and control of dynamic systems with Gaussian process models, Springer.

5. Hyperparameter training

Bioreactor (5)

6. Model validation:
   - plausibility ("looks", "behaves" logically),
   - falseness (validation of input-output behaviour),
   - purposiveness (satisfying the purpose of modelling).

Bioreactor (6)

Model evaluation:

- Falseness:
  - qualitative (visual assessment of response)
  - quantitative - price lists:
    - mean squared error (MSE),
    - Mean Relative Square Error (MRSE),
    - The logarithm of the prediction error density (LD),
    - The negative logarithm of the optimization data likelihood (LL),
    - SMSE , MSLL , NRMSE etc.

variance

$$\mathrm{SE} = \frac{1}{N}\sum_{i=1}^{N} e_i^2,$$

$$\mathrm{MRSE} = \sqrt{\frac{\sum_{i=1}^{N} e_i^2}{\sum_{i=1}^{N} y_i^2}},$$

$$\mathrm{SMSE} = \frac{1}{N}\frac{\sum_{i=1}^{N}(y_i - E(\hat{y}_i))^2}{\sigma_y^2}$$

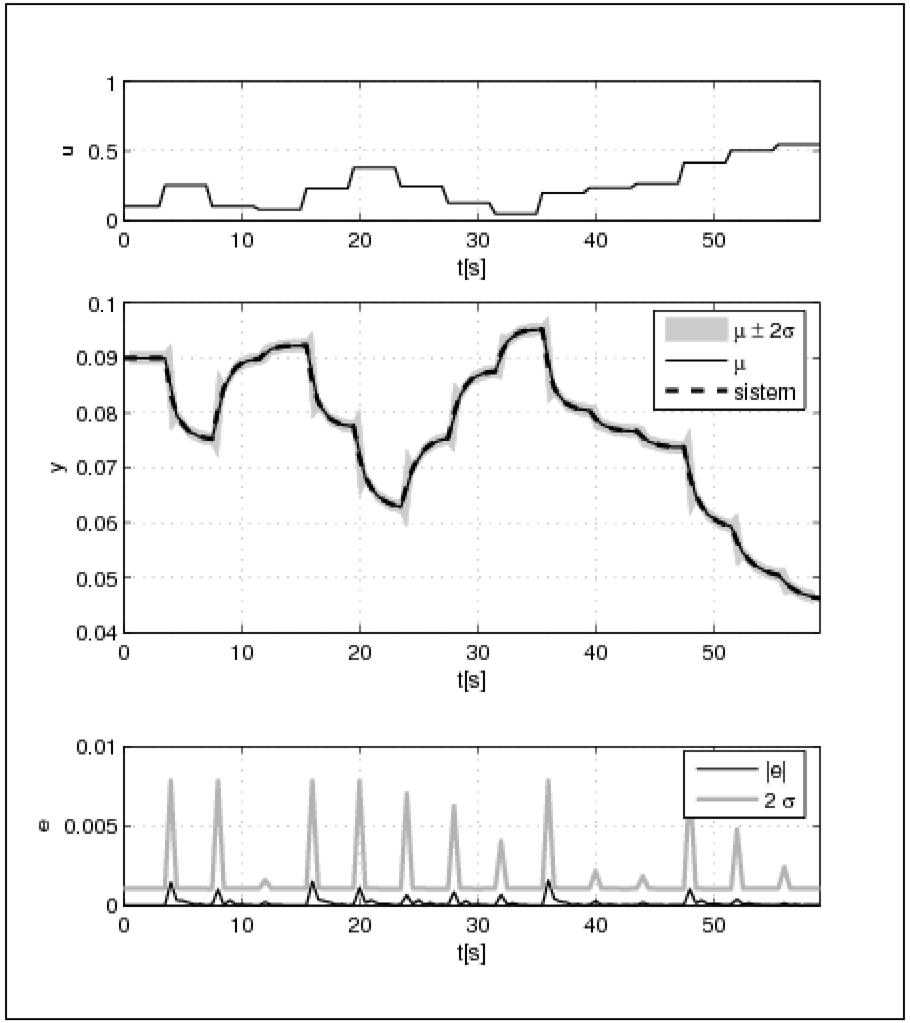
$$\mathrm{LD} = \frac{1}{2}log(2\pi) + \frac{1}{2N}\sum_{i=1}^{N}\left(log(\sigma_i^2) + \frac{e_i^2}{\sigma_i^2}\right),$$

$$\mathrm{LL} = \frac{1}{2}\log|\mathbf{K}| + \frac{1}{2}\mathbf{y}^T\mathbf{K}^{-1}\mathbf{y} + \frac{N}{2}\log(2\pi)$$

$$\mathrm{MSLL} = \frac{1}{2N}\sum_{i=1}^{N}\left[\ln(\sigma_i^2) + \frac{(y_i - E(\hat{y}_i))^2}{\sigma_i^2}\right] - \frac{1}{2N}\sum_{i=1}^{N}\left[\ln(\sigma_y^2) + \frac{(y_i - E(\mathbf{y}))^2}{\sigma_y^2}\right]$$

Bioreactor (7)

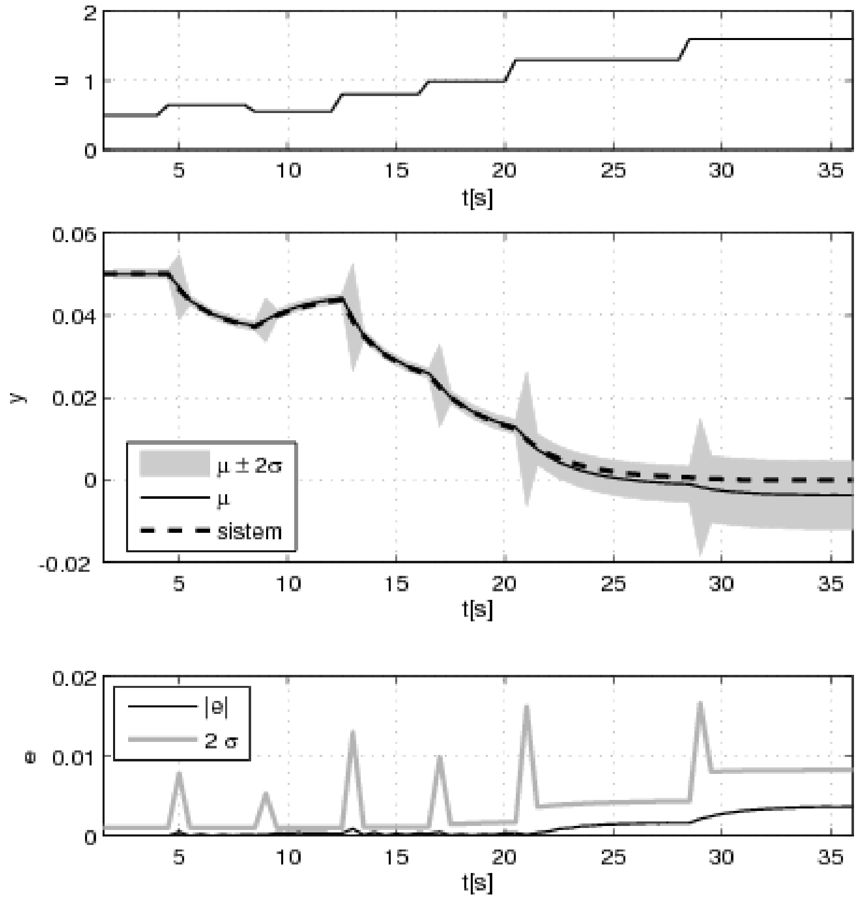
Simulation in the domain of identification data, but not with an identification signal.



Bioreactor (8)

Simulation outside the domain of identification data:

$u(k) > 0.7$.



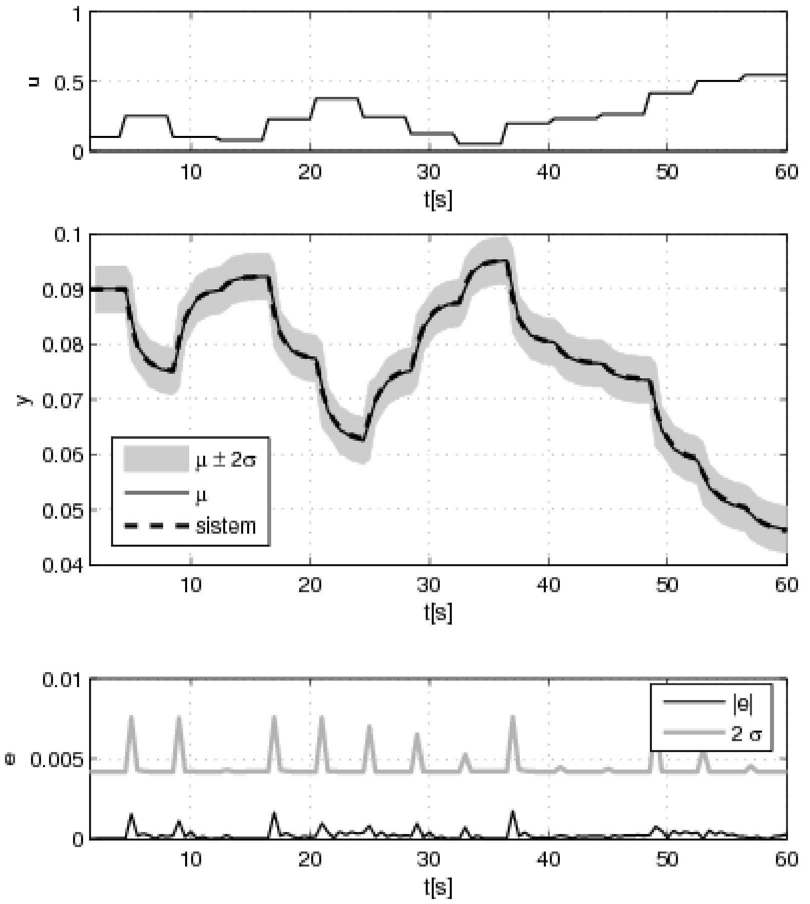
Bioreactor (9)

More noise:

$\sigma = 2\% \, (y_{max} - y_{min})$.

Increased variance:

- increased noise,
- insufficient data.

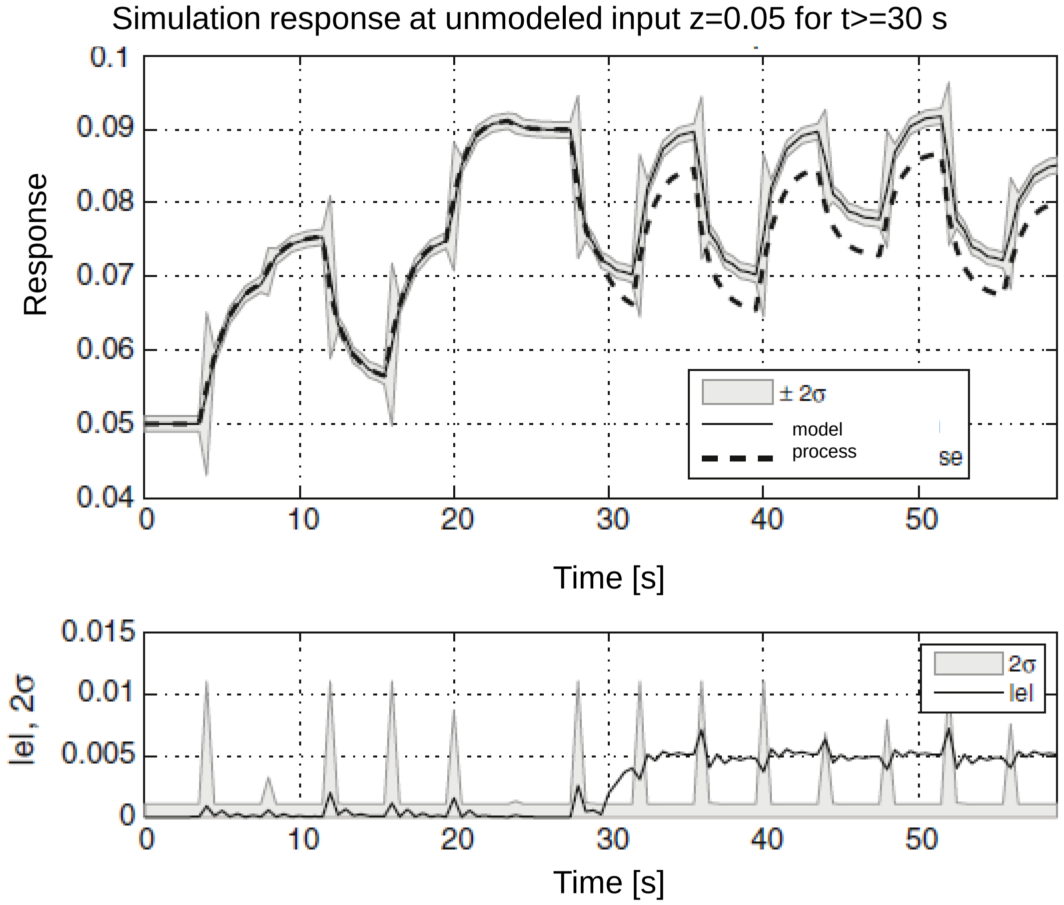


Bioreactor (10)

Unmodeled input

z =0.05 for *t* >=30s.

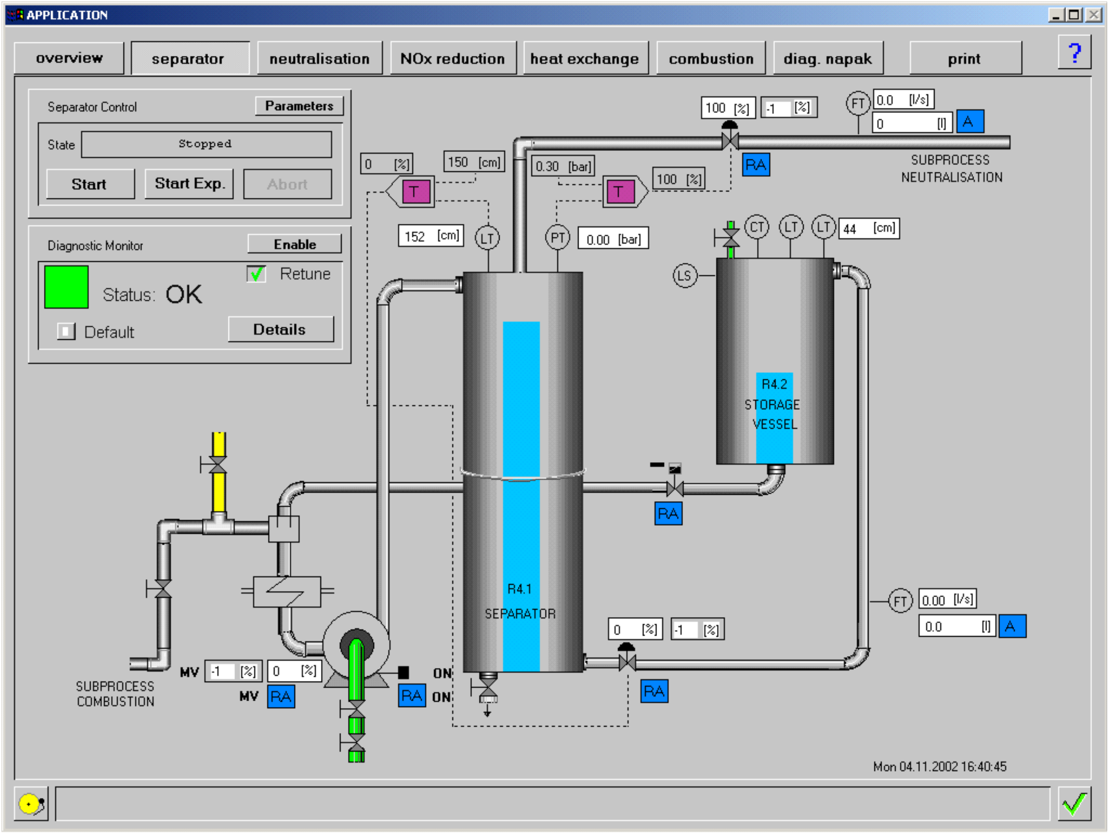
Confidence in the forecast does not change in the field of identification data.

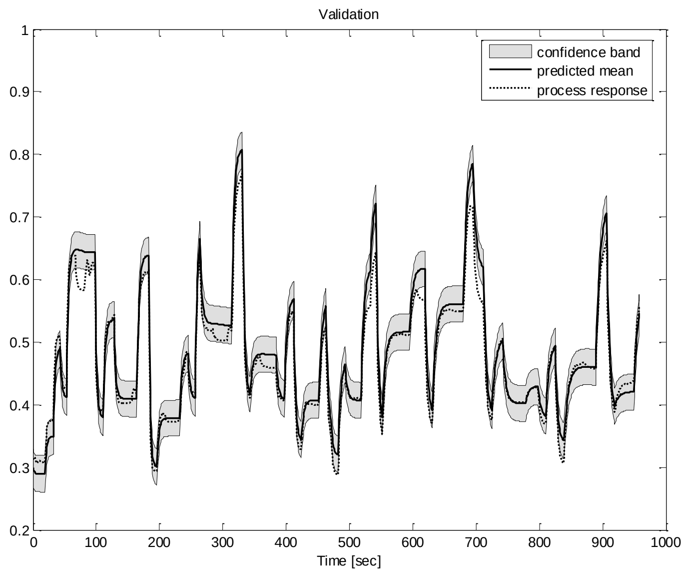


Selected examples of GP modeling applications

- Gas-liquid separation – process engineering.
- Gear life-span prognosis - mechanical production.
- Dispersion of nuclear pollution - environmental science.
- Online modelling with GP models - various fields.
- Modelling and forecasting of ozone values - environmental science.
- Predictive control of the gas-liquid separation – process engineering.
- Online adaptive control - automatic.
- Adaptive predictive control of bioreactor - biotechnology.

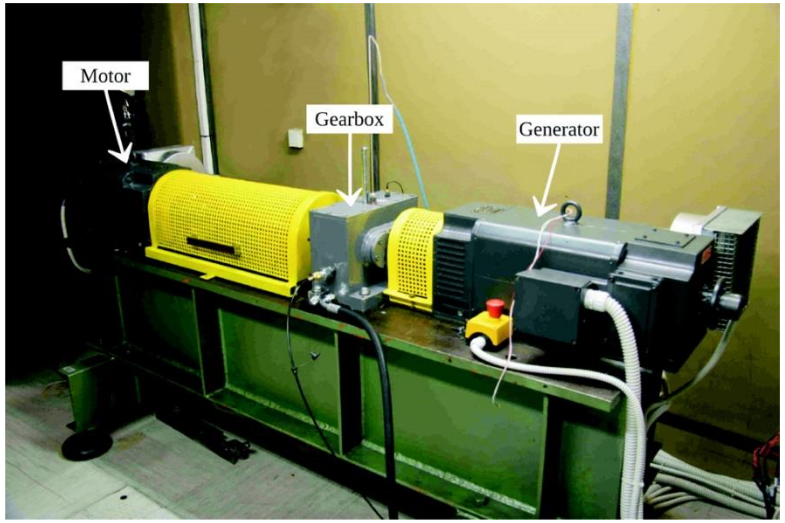
Modelling of the gas-liquid separator





KOCIJAN, Juš , LIKAR, Bojan. Gas - liquid separator modelling and simulation with Gaussian process models . Simulation modelling practice and theory , 2008, vol. 16, well. 18, p. 910-922.

Gear life-span prognostics



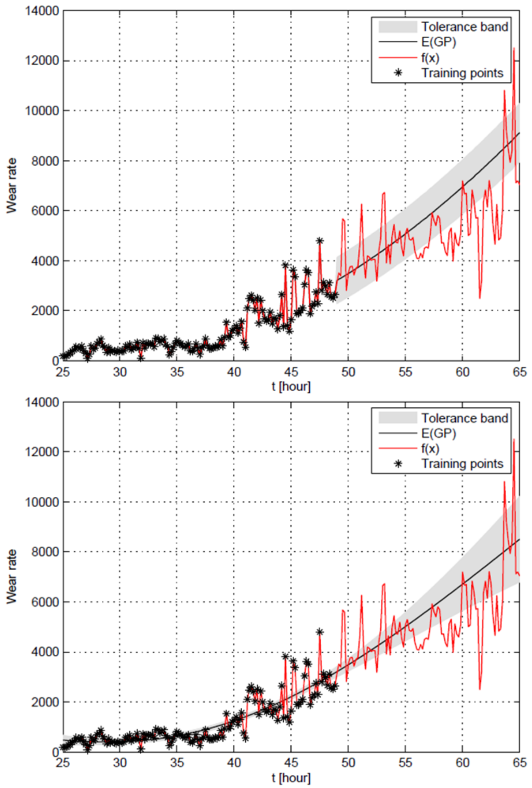


Fig. 4. The modelling of feature evolution for the gear with the model containing the sum of Matérn and polynomial covariance function (a) and neural network covariance function (b) for 144 training data points

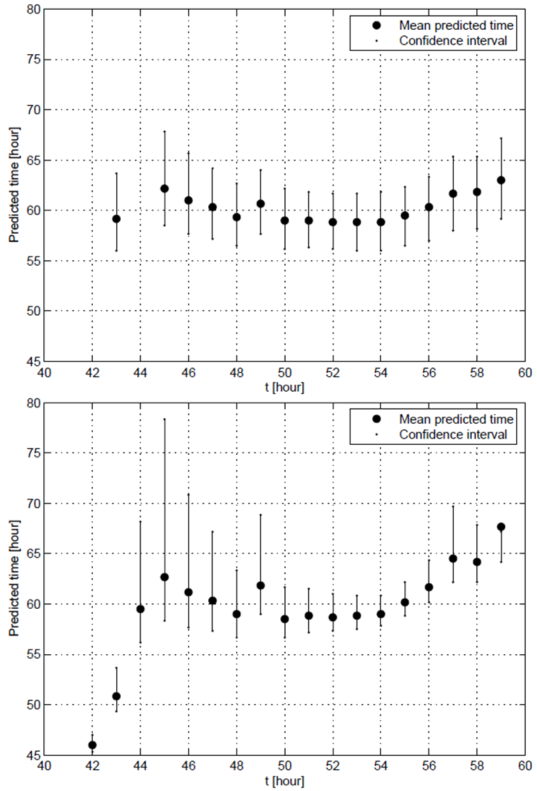
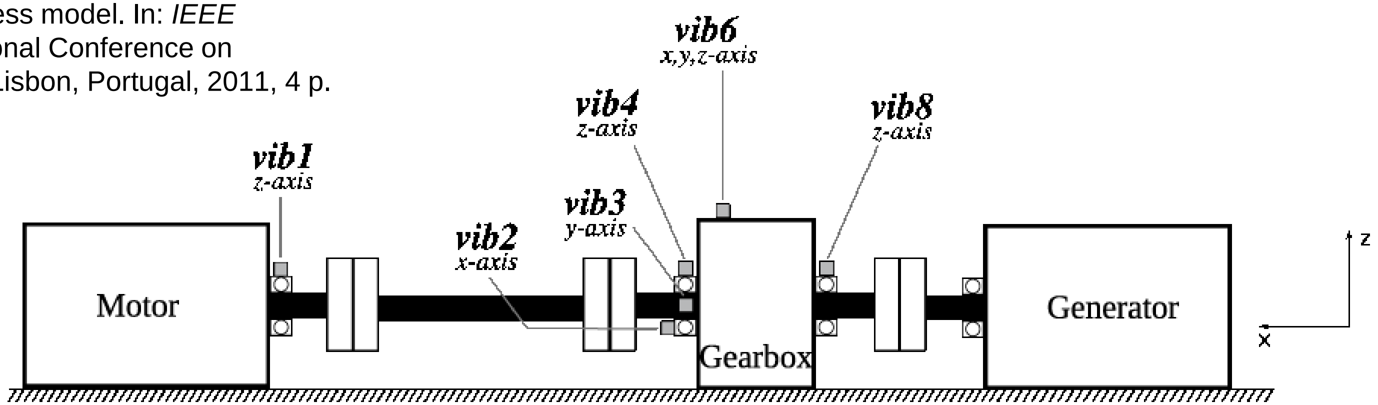
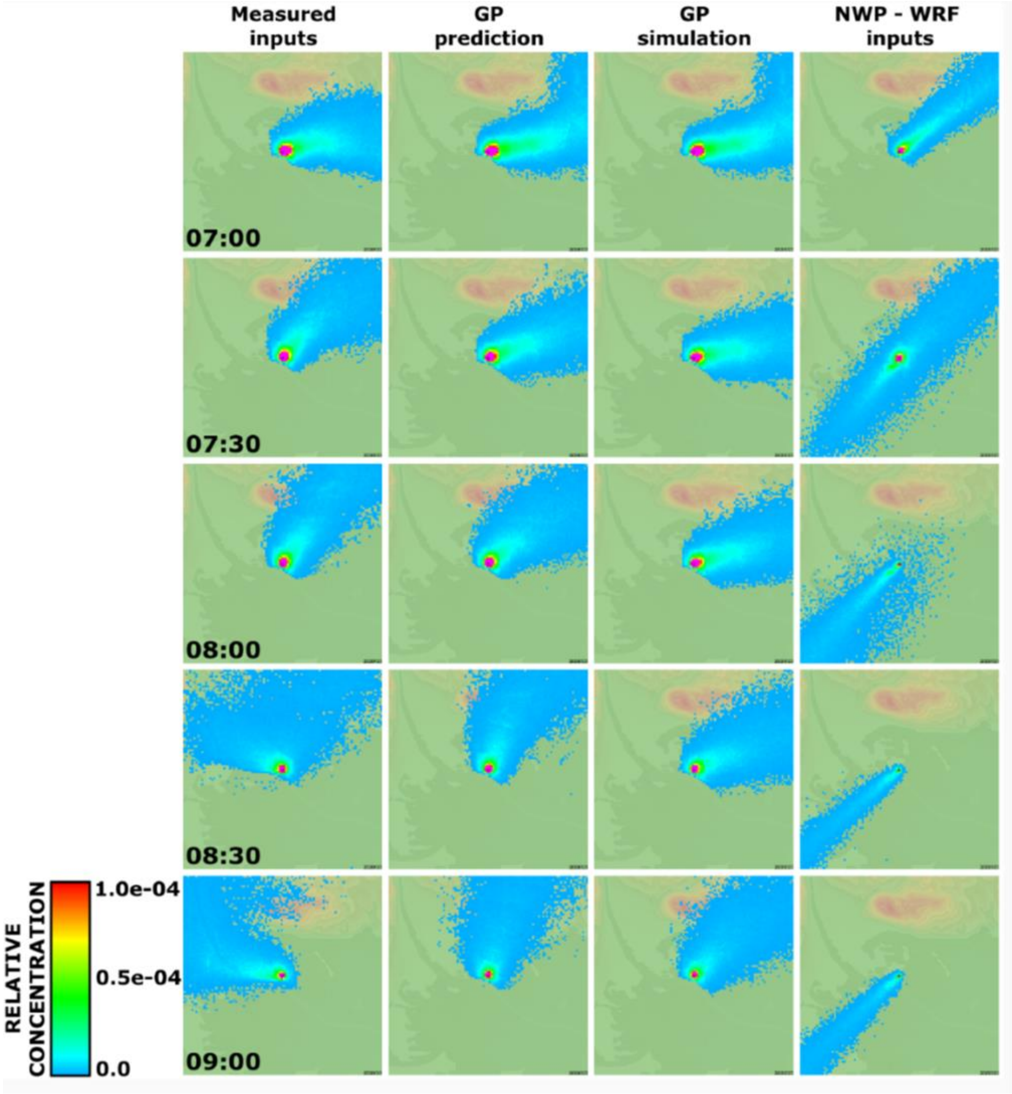


Fig. 5. Prediction of critical point using the eighth harmonic component feature with model containing Matérn and polynomial covariance function (a) and with neural network covariance function

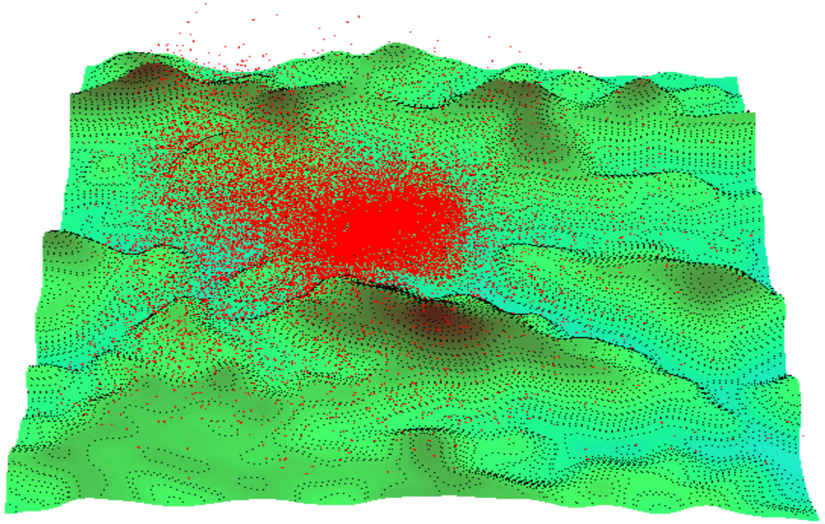
Kocijan , Juš , Tanko , Vesna . Prognosis of gear health using Gaussian process model. In: *IEEE EUROCON 2011* , International Conference on Computer as a Tool. 2011, Lisbon, Portugal, 2011, 4 p.



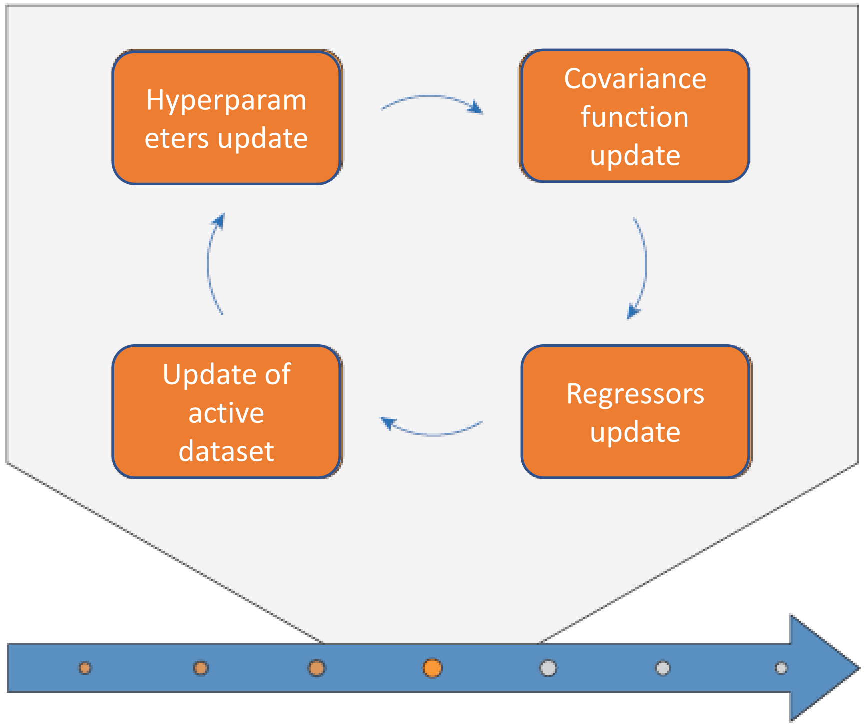
Modeling weather variables to analyze the spread of nuclear contamination

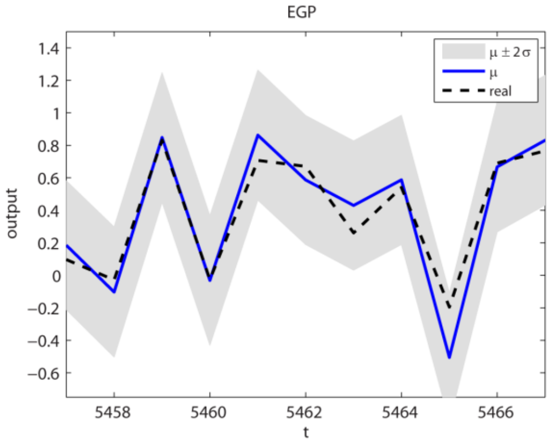


KRIVEC, Tadej, KOCIJAN, Juš, PERNE, Matija, GRAŠIČ, Boštjan, BOŽNAR, Marija, MLAKAR, Primož. Data-driven method for the improving forecasts of local weather dynamics. Engineering applications of artificial intelligence, Elsevier, 2021, vol. 105, 104423-1-104423-14.

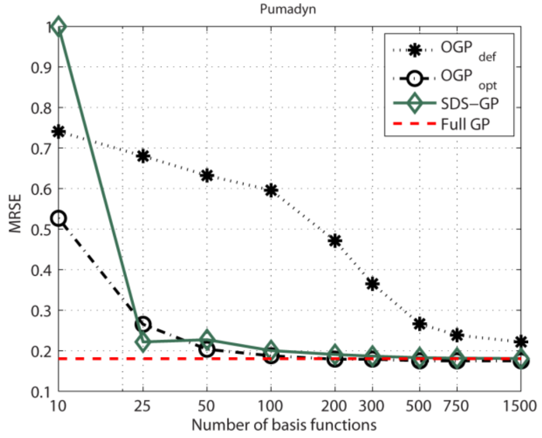


Online modelling - a scheme of the concept of online learning of an evolving GP-model

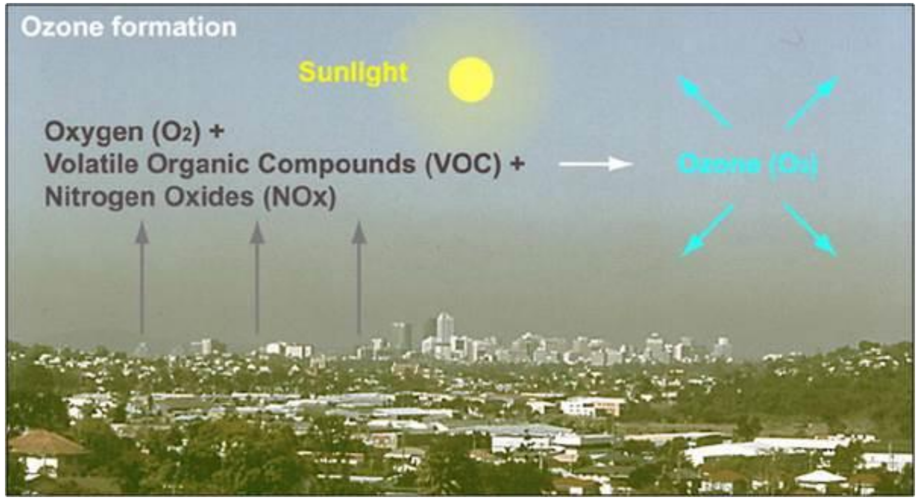




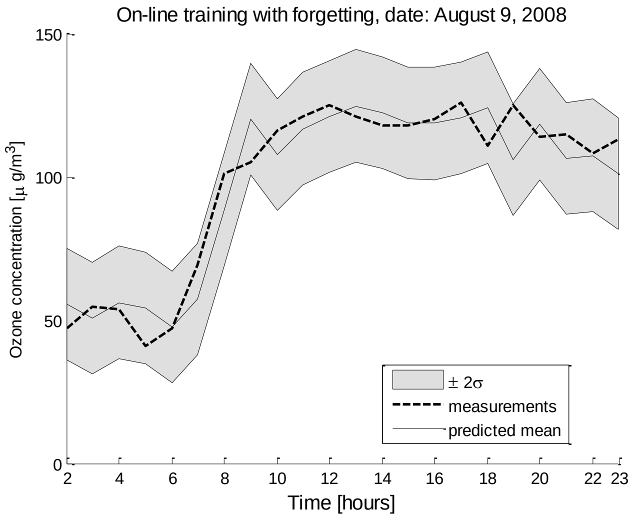
Petelin , Dejan , Kocijan , Juš . Streaming-data selection for Gaussian-process modeling . In: Borgelt , Christian. *Towards advanced data analysis by combining soft computing and statistics* . Heidelberg: Springer, 2013, p. 177-190.



Modelling and forecasting of ozone values - a study of a mobile pollution measurement station



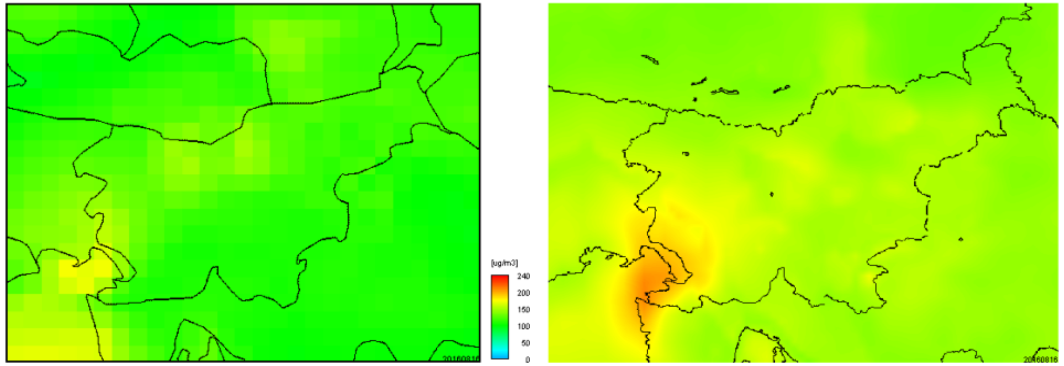
Courtesy of Saint Louis University



Ground level ozone forecast - hourly maximum for 16.08.2016

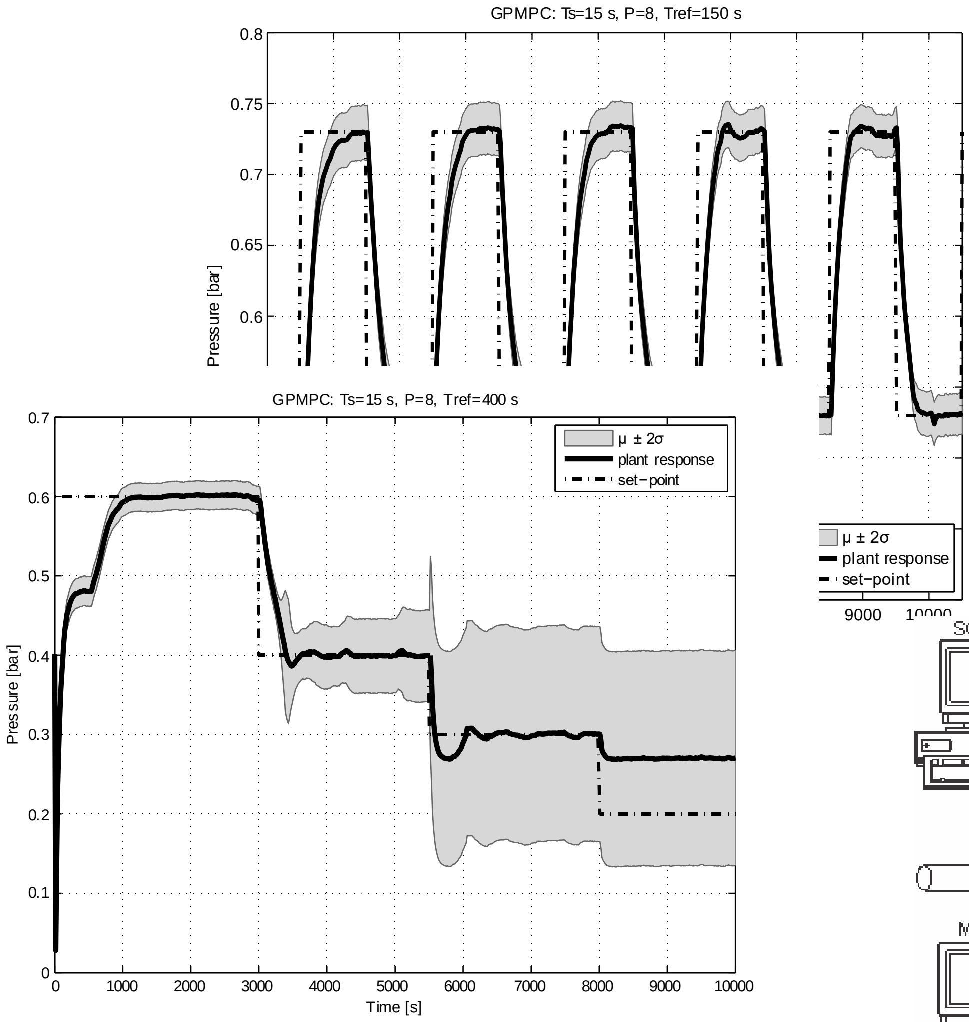
QualeAria photochemical model

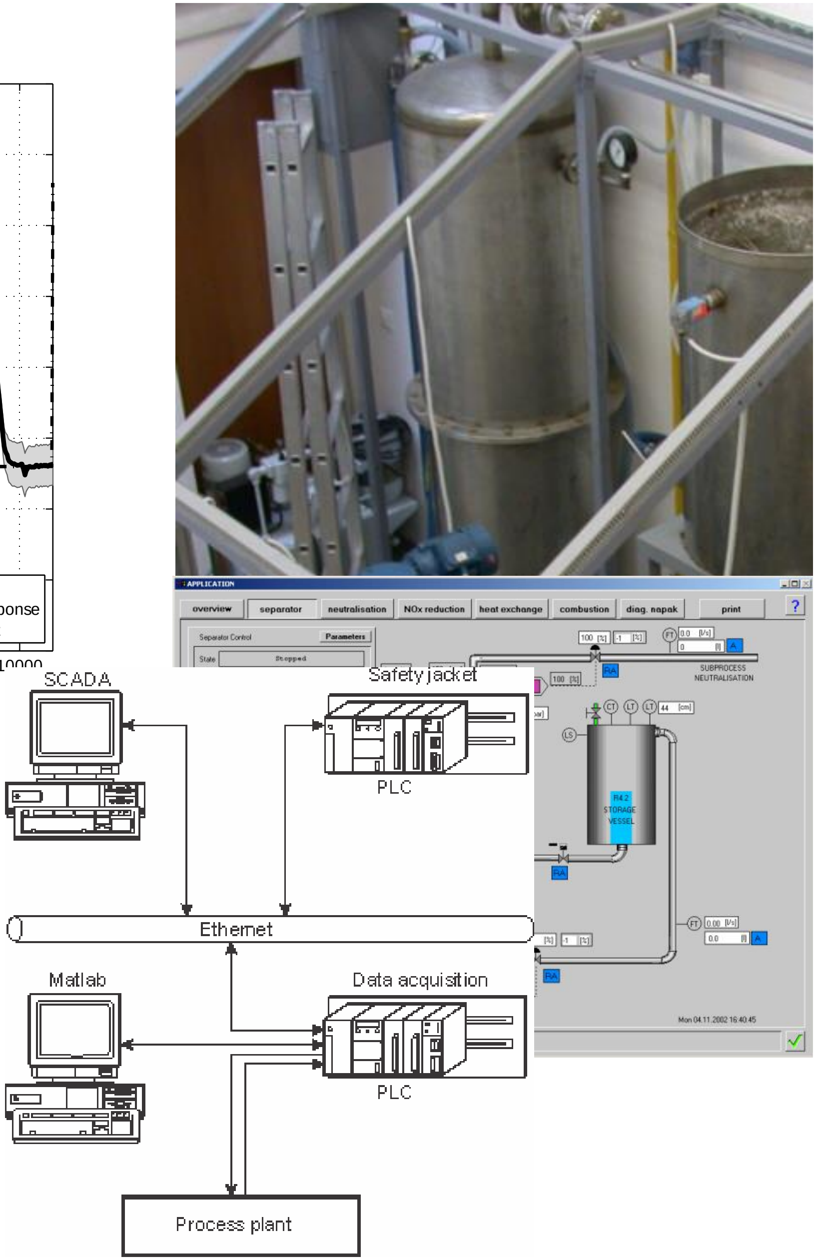
Nonlinear data assimilation



Kocijan , Juš , Gradisar , Dejan , Božnar , Marija , Grašič , Boštjan , Mlakar , Primož . On-line algorithm for ground-level ozone prediction with a mobile station. *Atmospheric environment* , 2016, vol. 131, p. 326-333

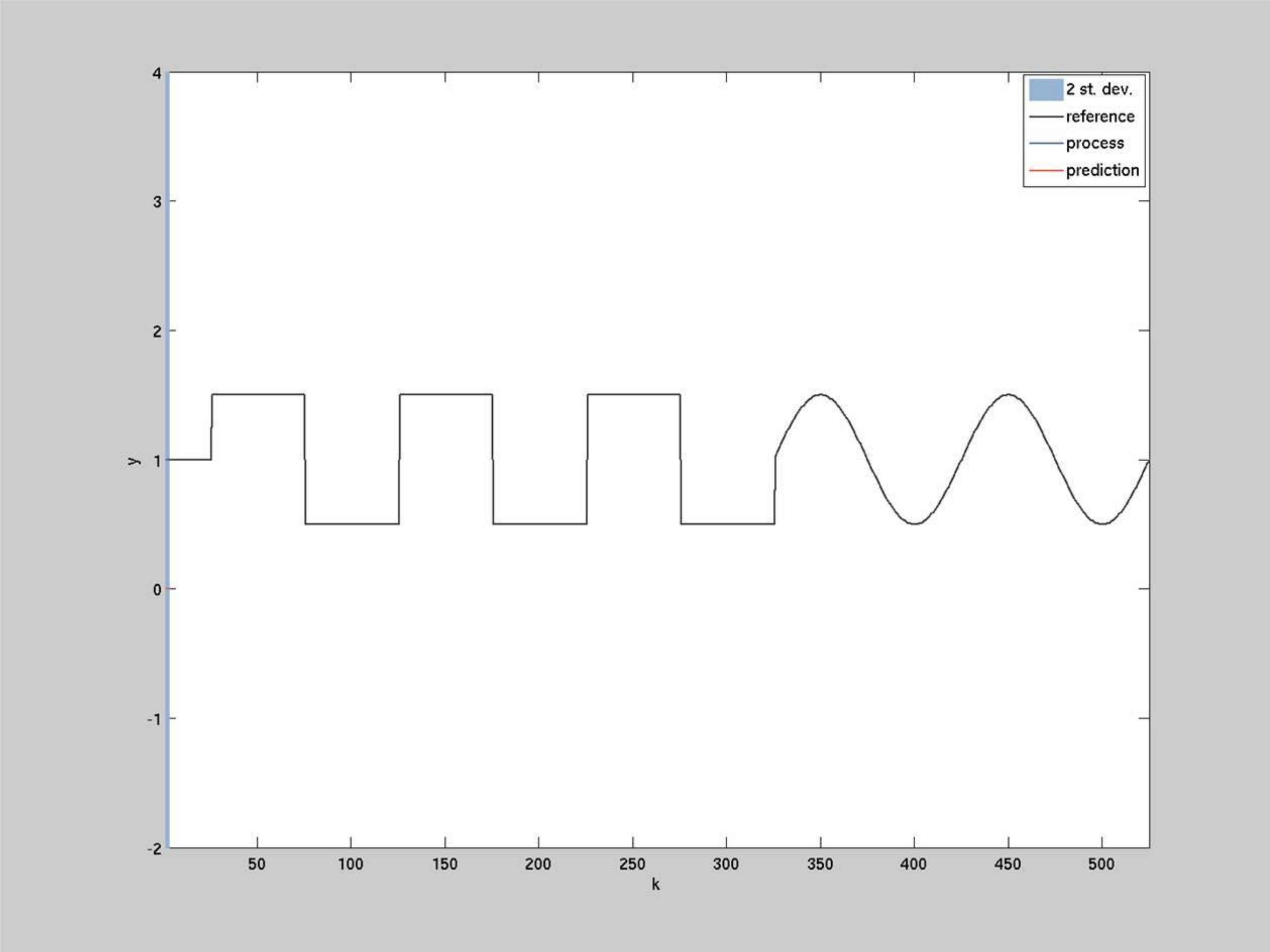
Predictive control of the gas preparation unit



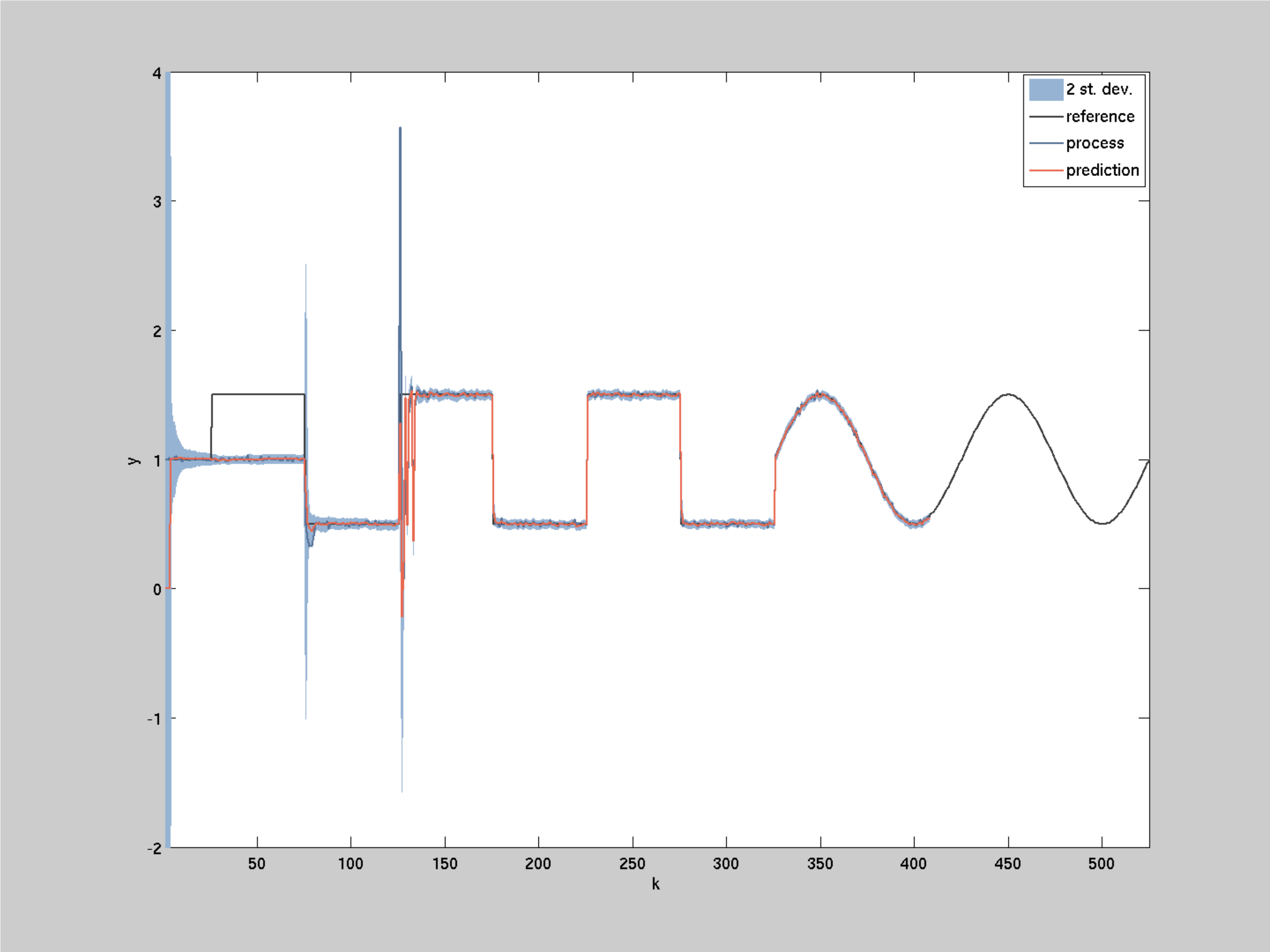


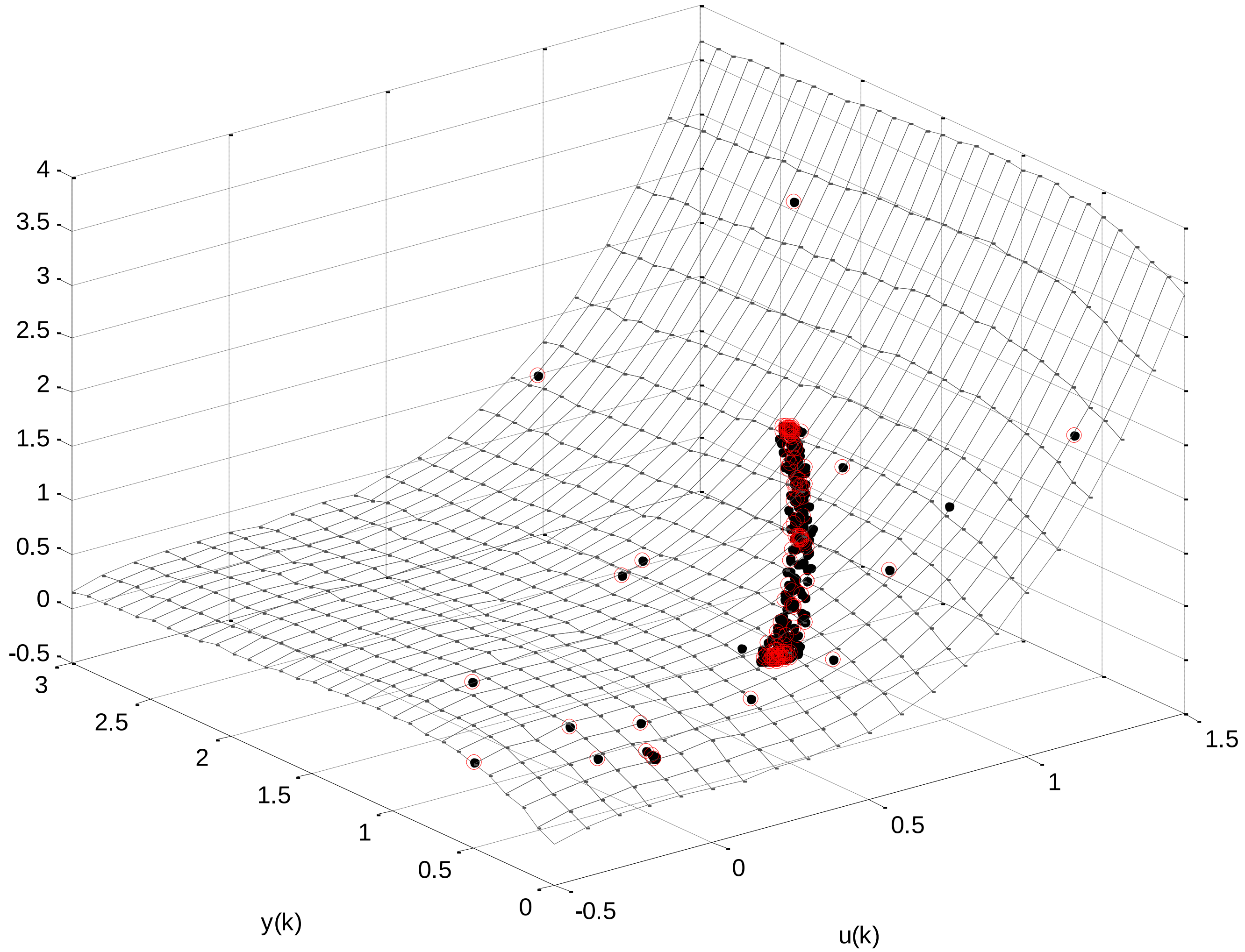
Kocijan (2016) Modelling and control of dynamic systems
with Gaussian process models , Springer.

Control demonstration using evolving GP-model



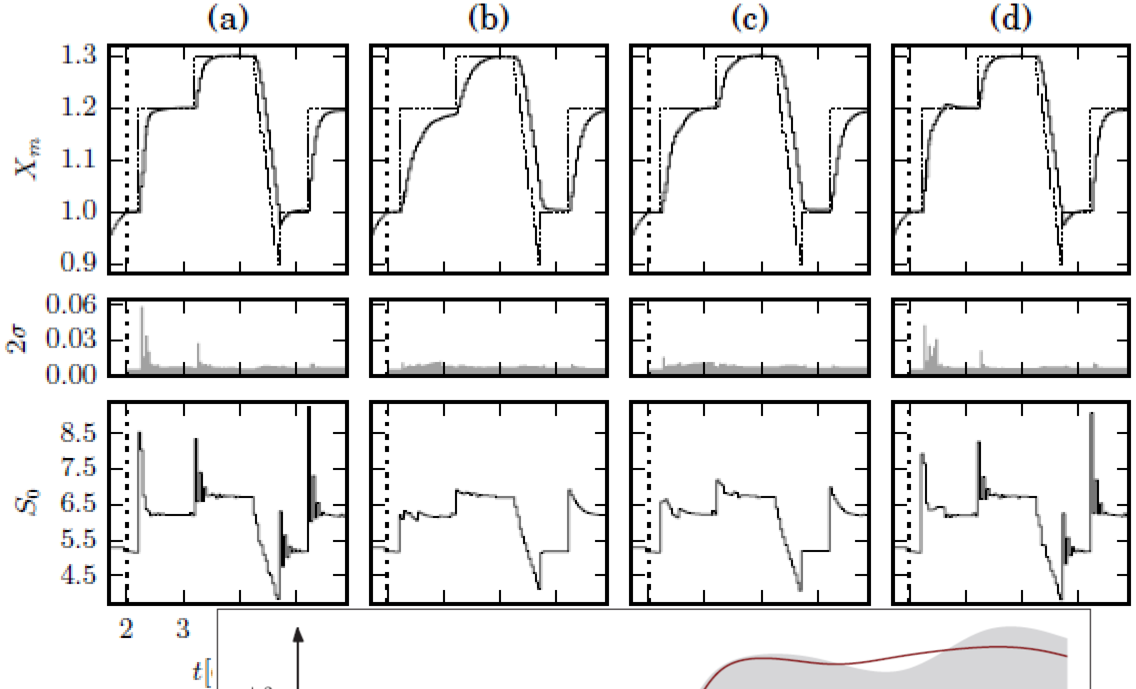
Control demonstration using evolving GP-model

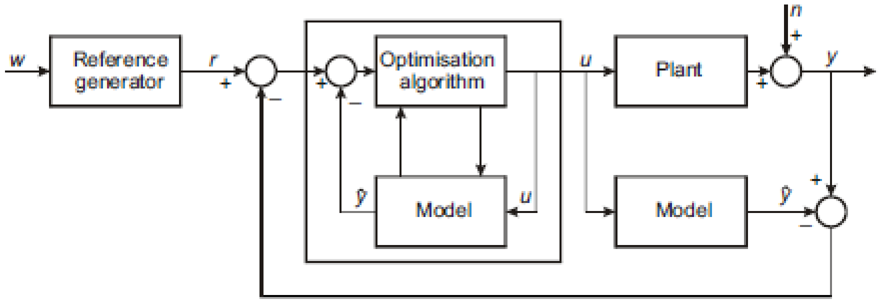


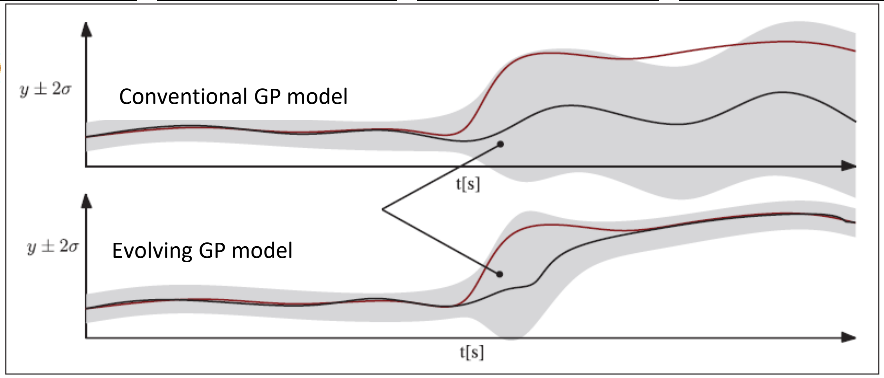


Identification data (black dots) and data with the most information (red circles) displayed on the surface of the selected nonlinear system

Adaptive predictive bioreactor control



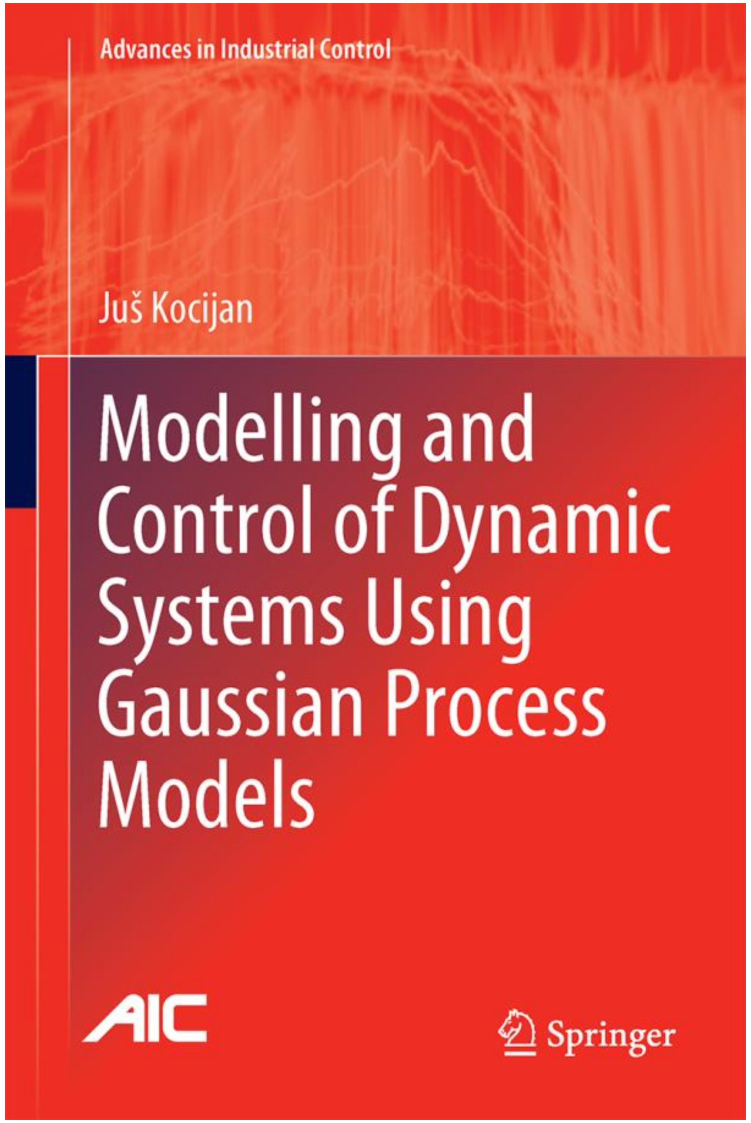




STEPANČIČ, Martin, GRANCHAROVA, Alexandra, KOCIJAN, Juš . Adaptive MPC based on probabilistic black-box input-output model. *Comptes rendus de l' Academie bulgare des Sciences* , ISSN 1310-1331, 2015, vol. 68, well. 6, p. 767-774.

Summary

- GP modelling is a flexible, probabilistic, approach to experimental modelling
- Applicable for uncertain, missing and corrupted data
- A promising method for engineering practice, especially for control design.



Springer International Publishing

eBook ISBN
978-3-319-21021-6

DOI
10.1007/978-3-319-21021-6

Hardcover ISBN
978-3-319-21020-9

Toolbox
https://github.com/Dynamic-Systems-and-GP/GPdyn.git
or
http://extras.springer.com

We are hiring!

PhD POSITION

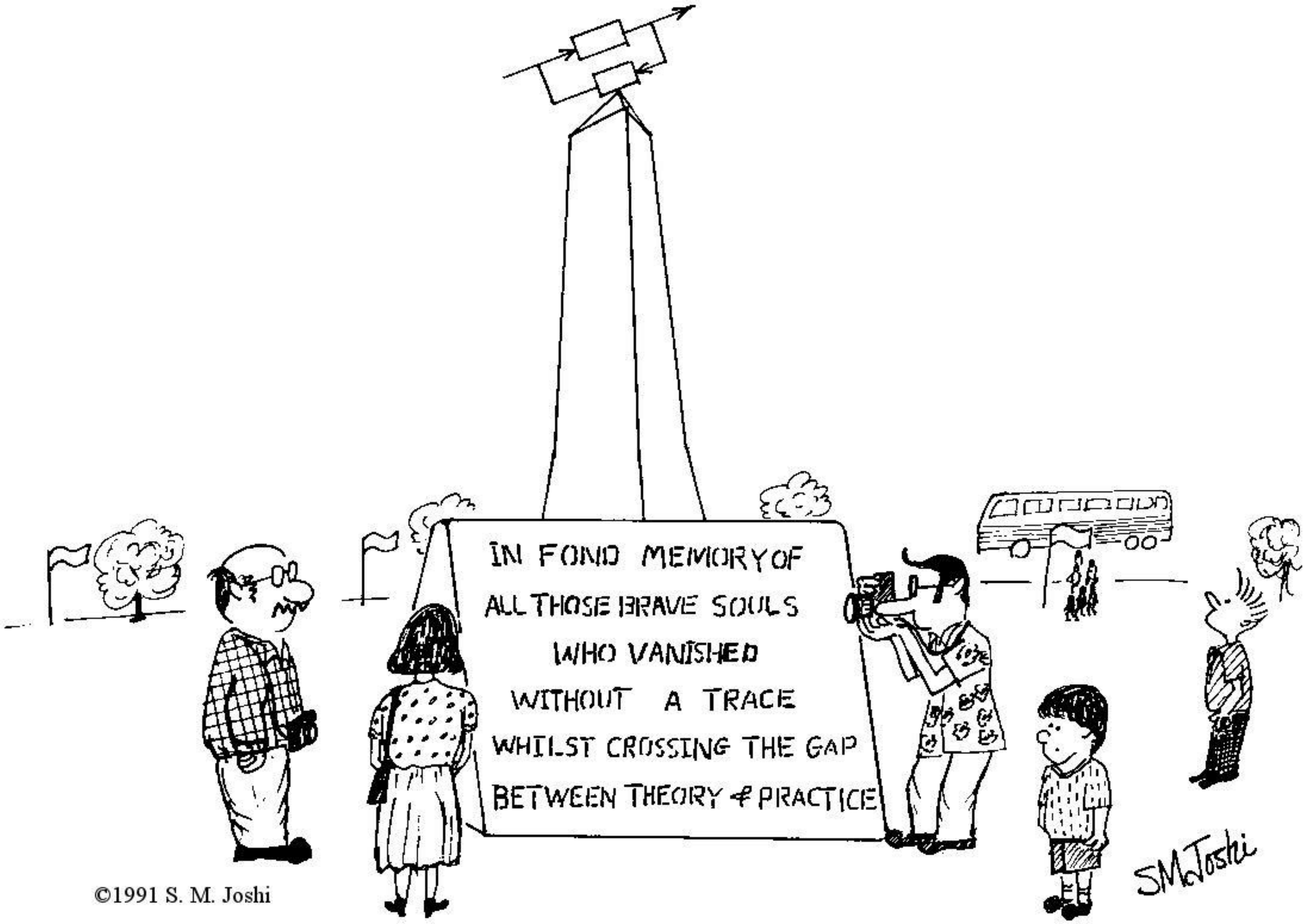
ABOUT US

The Jožef Stefan Institute is the leading Slovenian scientific research institute, covering a broad spectrum of basic and applied research. The staff of more than 950 specializes in natural sciences, life sciences and engineering. The subjects concern production and control technologies, communication and computer technologies, knowledge technologies, biotechnologies, environmental technologies, nanotechnologies, nuclear engineering, material sciences, etc. The mission of the Jožef Stefan Institute is the accumulation - and dissemination - of knowledge at the frontiers of natural science and technology to benefit society at large through the pursuit of education, learning, research, and development of high technology at the highest international levels of excellence.

RESEARCH DIRECTIONS

The objectives of the proposed research topic for PhD project are to develop methods for dynamic system modelling based on machine learning methods like kernel methods or to improve the existing data-driven modelling procedures to be better suited for the targeted application. The models obtained are meant for different applications, each representing a potential PhD project, which are as follows.

- Modelling of spatiotemporal systems using Gaussian processes and image or video forecasting, e.g., for modelling pollution dispersion
- Designing of adaptive and robust control systems that handle signal and safety constraints in complex dynamic systems.


IN FOND MEMORY OF
ALL THOSE BRAVE SOULS
WHO VANISHED
WITHOUT A TRACE
WHILST CROSSING THE GAP
BETWEEN THEORY & PRACTICE
SMJoshi