Gaussian Process Regression

Chimiometrie 2018 , Paris

Ali Gahkani
Aunir, UK

AUNIR
MAKING LIGHT WORK

Gaussian Process Regression

Introduction

- Gaussian Process (GP) is a *non-parametric probabilistic* approach.
- GP extends the concept of Gaussian distribution of numbers to *functions*!

- A GP is defined by a mean function (often zero) and a *covariance function* or *kernel:*

$$k(x, x') = \boxed{\sigma_f^2} \exp\left[\frac{-(x-x')^2}{\boxed{2l^2}}\right]$$

Optimizable parameters

- $y = f(x)$ and y_* (prediction for new x_*) make a joint Gaussian distribution:

$$\begin{array}{r} \textit{cal} \\ \textit{val} \end{array} \begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & K_*^{\mathrm{T}} \\ K_* & K_{**} \end{bmatrix}\right)$$

Covariance matrices

- *Conditional* probability $p(y_*|y)$:

$$y_*|\mathbf{y} \sim \mathcal{N}(\boxed{K_* K^{-1}\mathbf{y}}, \boxed{K_{**} - K_* K^{-1} K_*^{\mathrm{T}}})$$

Best estimate: $\bar{y}_*$ *Uncertainty: s^2*

AUNIR
MAKING LIGHT WORK

Gaussian Process Regression

Introduction

- Gaussian Process (GP) is a *non-parametric probabilistic* approach.
- GP extends the concept of Gaussian distribution of numbers to *functions*!
- A GP is defined by a mean function (often zero) and a *covariance function* or *kernel:*

$$k(x,x') = \boxed{\sigma_f^2} \exp\left[\frac{-(x-x')^2}{\boxed{2l^2}}\right]$$

Optimizable parameters

- $y = f(x)$ and y_* (prediction for new x_*) make a joint Gaussian distribution:

$$\begin{array}{l} \text{cal} \\ \text{val} \end{array} \begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & K_*^{\mathrm{T}} \\ K_* & K_{**} \end{bmatrix}\right)$$

Covariance matrices

- *Conditional* probability $p(y_*|y)$:

$$y_*|\mathbf{y} \sim \mathcal{N}(\underbrace{K_*K^{-1}\mathbf{y}}_{\text{Best estimate: } \bar{y}_*}, \underbrace{K_{**} - K_*K^{-1}K_*^{\mathrm{T}}}_{\text{Uncertainty: } \sigma^2})$$

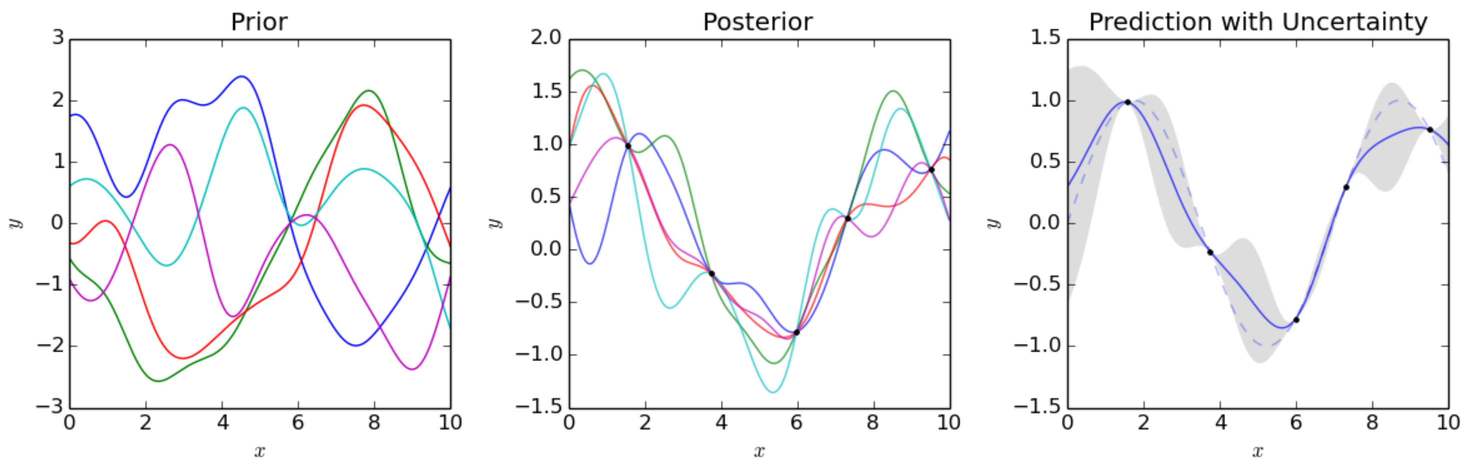
AUNIR

MAKING LIGHT WORK

Gaussian Process Regression

Introduction

GP begins with a prior distribution and updates this prior as new data are observed, producing the posterior distribution over functions.



Great demo: http://www.tmpl.fi/gp/

Gaussian Process Regression

Benefits

- Non-linear
- Kernel parameters automatically learned from data
- Different kernels can be combined together
- Automatic feature selection (ARD)
- Fully probabilistic predictions
- Fairly resistant to over-fitting

GPR on the Challenge data

Settings

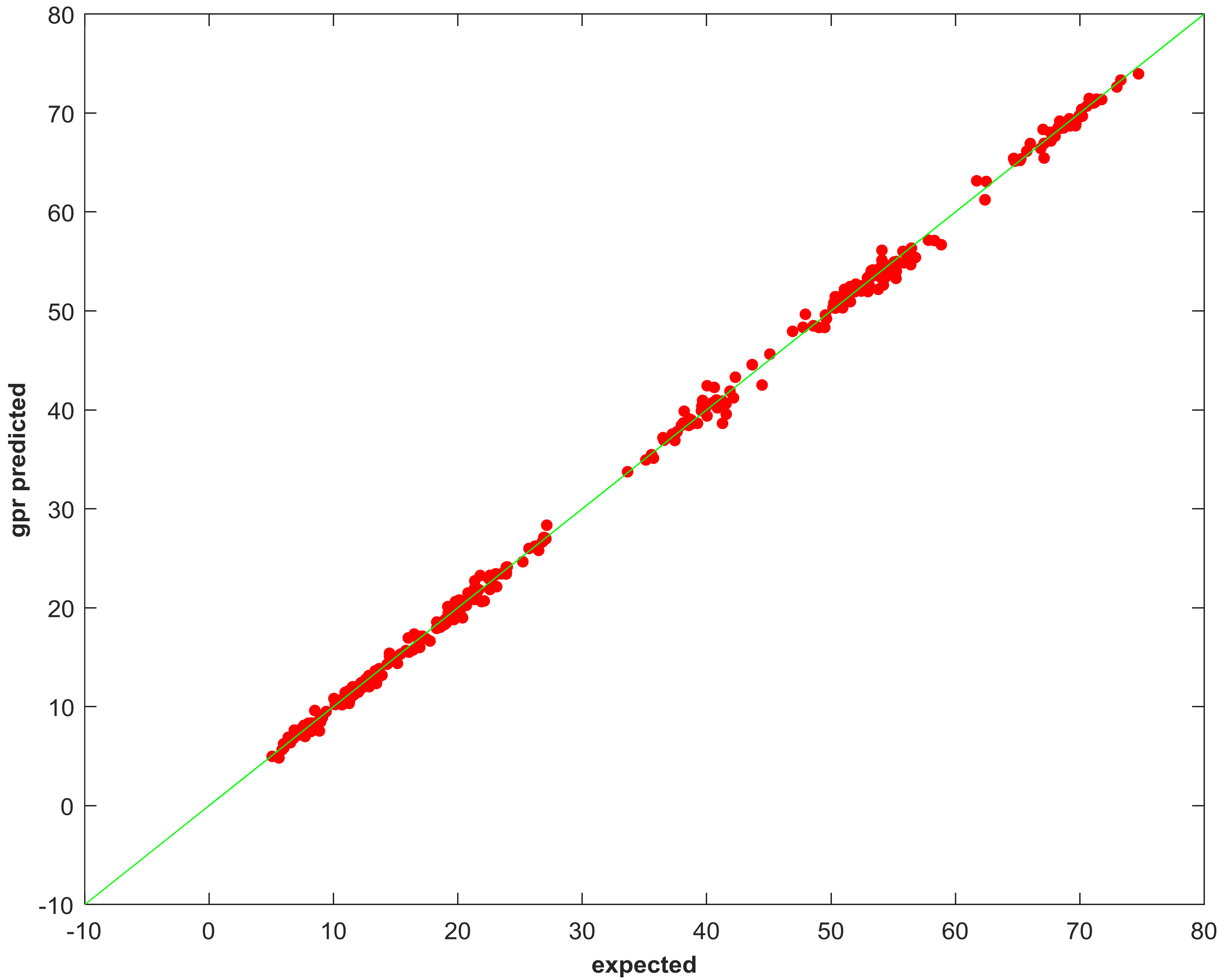
- 207 redundant samples removed from calibration set
- Pre-processing: ***SG(1,15,2) + SNV***
- Kernel: ***Matérn v=5/2***
- Kernel Parameters optimized by maximizing marginal likelihood
- Inference method: ***Gaussian likelihood***

GPML *Matlab Code version 4.1*
(by Carl Edward Rasmussen and Hannes Nickisch)
http://www.gaussianprocess.org/gpml/

AUNIR
MAKING LIGHT WORK

GPR on the Challenge data

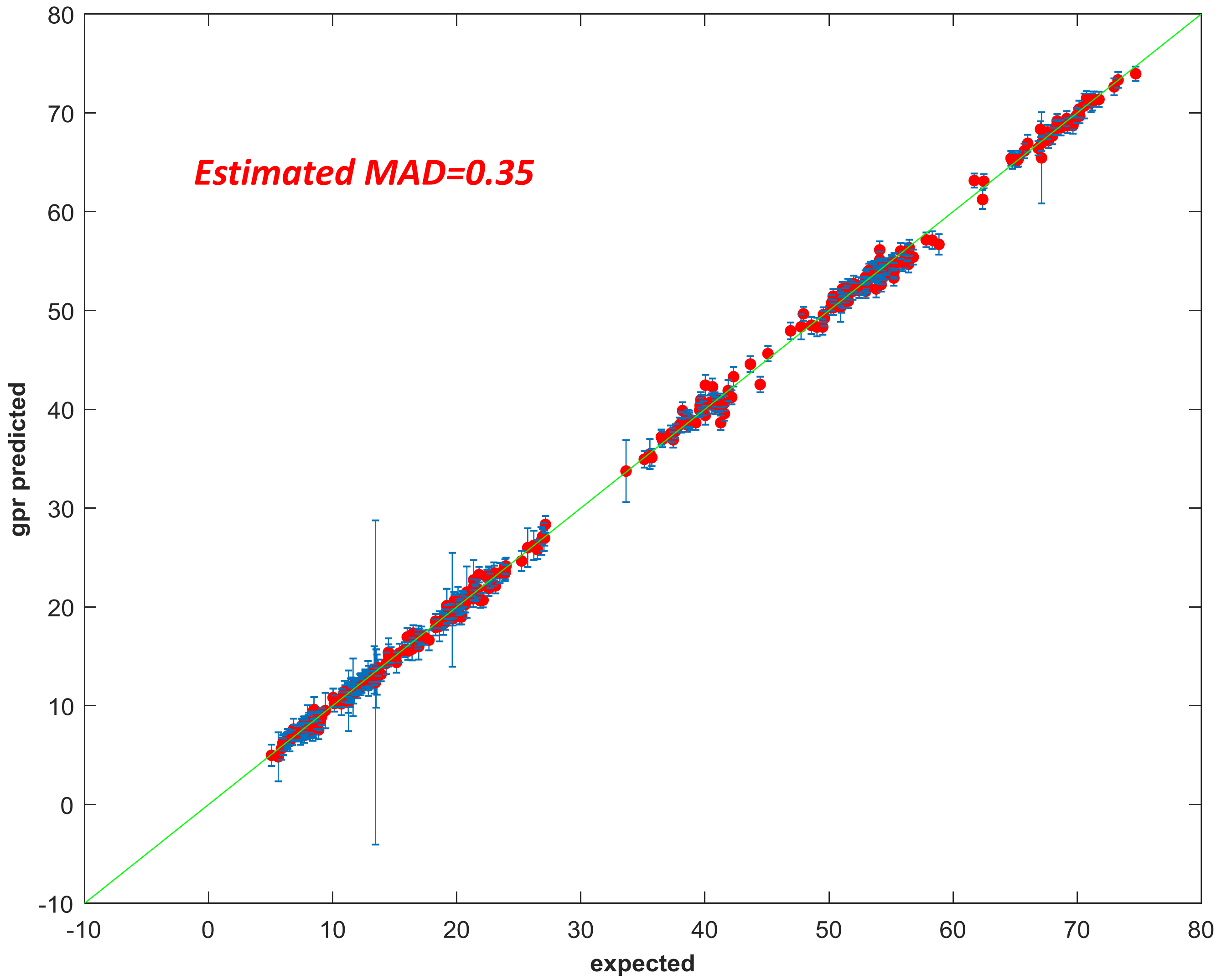
Internal validation



AUNIR
MAKING LIGHT WORK

GPR on the Challenge data

Internal validation

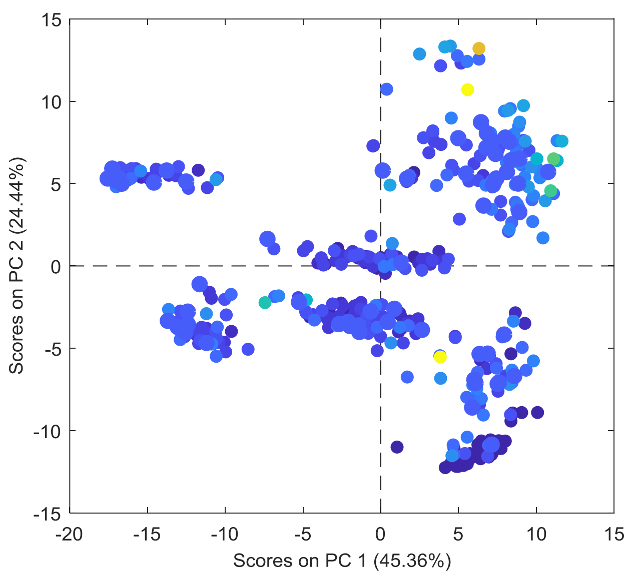


AUNIR
MAKING LIGHT WORK

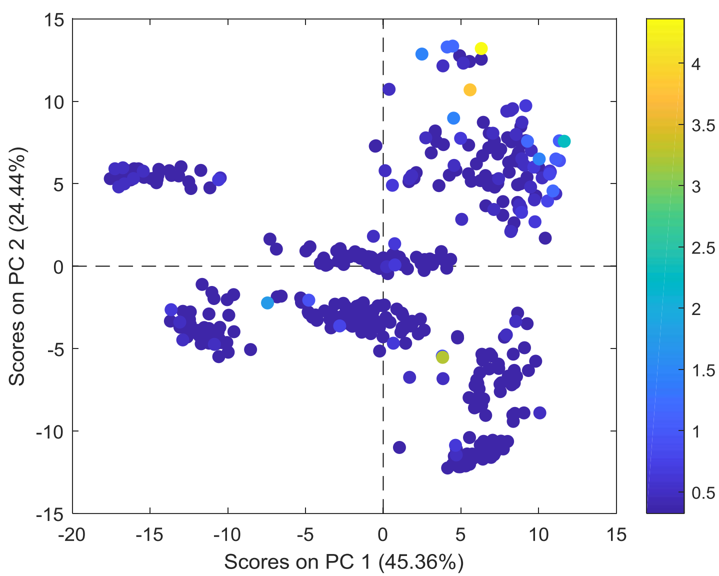
GPR on the Challenge data

Validation set

Nearest neighbour distance



GPR estimated standard deviation



AUNIR
MAKING LIGHT WORK


Merci