Immunoassay for Estrogens in the Environment based on Fluorescence and Neural Networks

F. Dieterle, S. Reder, G. Gauglitz

Institute of Physical Chemistry, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen
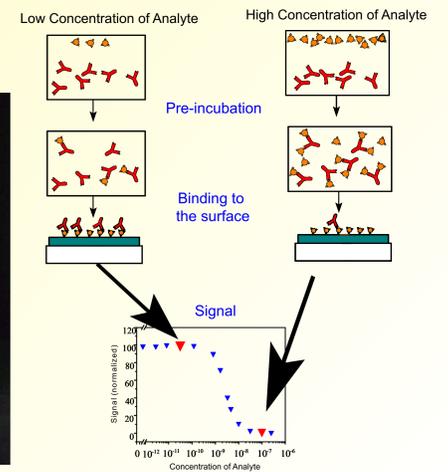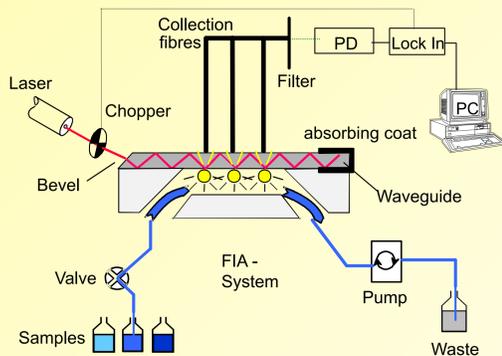
http://barolo.ipc.uni-tuebingen.de

Introduction

- Our environment is contaminated by an increasing number of substances that show estrogenic activity.

- A water monitoring system is needed. This instrument should be easy to use, small, and inexpensive.

- Our set-up is based on a heterogeneous immunoassay. Two estrogenic substances are detected simultaneously due to cross-reactivity.

- The signal strength is corrected (bleaching effects) and the number of samples is virtually increased to avoid overfitting. Finally the evaluation is performed by artificial neural nets

Set-up

- The device consists of an optical detection unit, a flow cell and an integrated fluid handling, based on flow injection (FIA)

- The detection is based on Total Internal Reflection Fluorescence (TIRF) as transducer principle

- Analyte specific antibodies are labelled with a fluorescence dye (Cy 5.5) and react with the analyte in the sample. The unoccupied antibodies can bind to the modified surface of a glass transducer.

Pretreatment of Data

- A loss of signal strength due to bleeching effects and irreversible bonds has to be corrected.

- Nonlinear fit of blank samples for each spot:

$$f(x) = a_1 e^{-\frac{x}{a_2}} + a_3 + a_4 x$$

x = number of sample

$f(x)$ = signal strength

- The signals of all samples are divided by this fit.
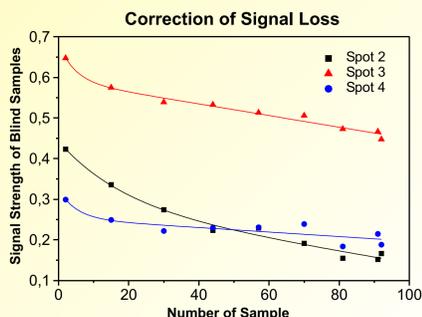
Avoiding Overfitting

1. Two methods for increasing the amount of data

A) Adding noise to replica:

$$y_{i,k,i} = y_{k,i} + 0,62666 \ln\left(\frac{1}{random(0;1)} - 1\right) r$$

B) Linear combination of 2 replica:

$$y_i = y_1 + \frac{y_2 - y_1}{2} \cdot 2 \cdot \frac{i}{n} \quad \text{for } i = 0...n$$

The number of samples was increased from 80 to 300. Method B gave slightly better results

2. Pruning of links and units
Magnitude Based Pruning reduced number of links

3. Early stopping of training
Training was stopped when the mean square error of cross-validation started to increase.
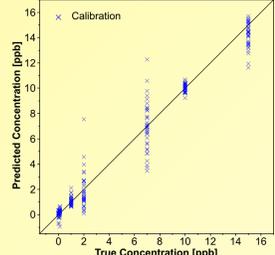
Data Evaluation

- Input variables were centered and standardized
- Output variables were scaled from -0.9 to 0.9
- One neural net per analyte
- Activation functions of units: tanh
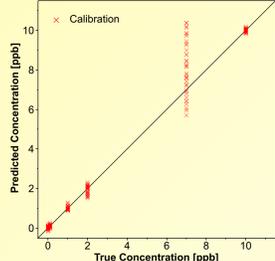- Training algorithm: Rprop

Input Layer   Hidden Layers   Output Layer
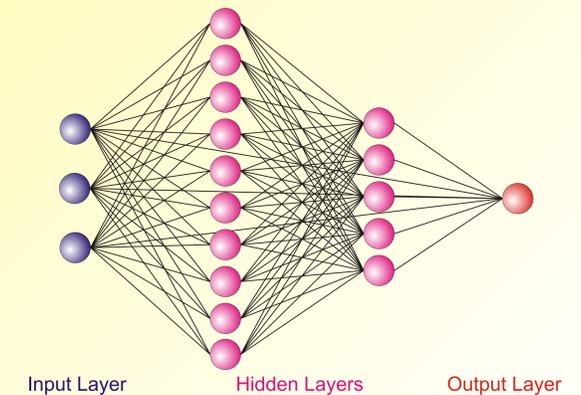
Results

Calibration by 10 neural nets

10 different mixtures as independent test samples
Each mixture was evaluated by 10 neural nets
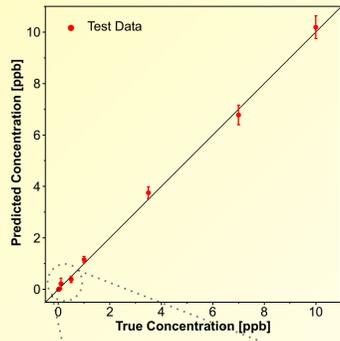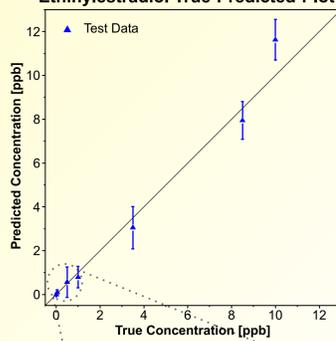
Summary and Outlook

- The Prediction of 10 independent test mixtures (not used for calibration) was very exact for estradiol and quite good for ethinylestradiol.

- The variance of predictions during calibration is caused by outliers in combination with the algorithms for increasing the amount of data. The variance of prediction of the test substances is better, as these have been measured several times.

- Enhancements of the stability of the fluorescence dye and a more effective regeneration should improve the prediction (less loss of signal strength).

- Simultaneous assays for Bisphenol A, Estradiol and Ethinylestradiol will be established.