NiSIS Competition 2006-
Soft Sensor for the adaptive Catalyst Monitoring of a Multi–Tube Reactor

Jens Strackeljan
Universität Magdeburg
39106 Magdeburg, Universitätsplatz 2, Germany
Phone: +49-391-6718437, Fax: +49-391-6711379
email: jens.strackeljan@mb.uni-magdeburg.de

1. INTRODUCTION

HISTORY

Scientific competitions are well-known instruments which have motivated scientists to develop new ideas for several centuries. Today too, large sums of money are sometimes offered as prizes for solutions to unsolved problems in various fields of natural science. The Clay Mathematics Institute has singled out seven so-called Millennium Problems, for which prizes totalling seven million dollars have been offered. Without offering monetary prizes we have made very good experiences with a series of competitions in the framework of former EU networks (NoE's EUNITE, ERUDIT) and consequently the TTE committee will organize new competitions in NISIS. The continuing receipt of enquiries and requests for data and solutions from the 1998 to 2000 competitions even today is highly encouraging. If scientists employ the competition data as a benchmark, this is a reliable indication that we have made a correct selection of problems.

MOTIVATION

What are the reasons for the motivation to participate in such a competition? Without any doubt a successful score in a competition can positively accentuate one's personal curriculum vitae. This feature is especially important for participants from universities. On the other hand, it may also prove to be beneficial for one's own career in business and industry. Above all, however, it thus became evident that the company's own R&D personnel perform at a high level, and that they need not be afraid of comparisons with specialists elsewhere. This aspect is of particular interest, since very few opportunities for benchmarking otherwise exist for such positions in a company. A competition offers a possibility of appraising the status of one's own algorithms, that is, how good one's own approach is in comparison with those of other specialists. Every serious scientist must have an interest in comparisons of this kind, since continuing development is feasible only on the basis of such position determinations. Whenever the task involves a real problem encountered in industry, a valuable contact with the associated industrial partner can result from participation in the competition. So much for the motivation of participants. However, what can a company expect if it participates in a competition and furnishes data for the purpose? The answer can be summarised in a very simple way. Whoever is willing to invest time in the preparation and evaluation can profit immensely from such a competition but without active cooperation of this kind, the result will not be satisfactory, and especially the assessment of the significance for one's own company will be difficult.

ADDED VALUE FOR NISIS

Besides the participants and industrial companies, the third partner must also be mentioned: NiSIS and the TTE Committee itself.

1. Internal technology transfer
First of all, a platform is created by processing an identical problem; such a platform can facilitate the exchange of methods and algorithms among the participants. In particular, the instrument of the competition team has considerably enhanced this exchange. The competitions could result in increased cooperation, which can be designated as genuine networking.

2. External technology transfer
If the competition results in a real application at a company, the objective of technology transfer could be achieved within the network.

3. Improvement in the external presence of the network
The former competitions have been continuously observed even by those members of the network who did not participate actively. Advertising for the competition by e-mails and the Internet have positively affected the publicity of the network.

If the competition task comes from industry and will lead to a real application at a company, the objective of technology transfer would be achieved within the network. Advertising for the competition by e-mails and the Internet have positively affected the publicity of NISIS. The problem task 'Soft Sensor for the adaptive Catalyst Monitoring of a Multi-Tube Reactor' and the data material was provided by the Company DEGUSSA, Germany, and represents an real world problem. In 2005 we have developed a new data exchange concept to ensure the adaptive component in a competition. Without this the competition task will be far away from all NiSIS concepts. The awards for the for the best model and the best 'Nature inspired concept' should motivate researchers and people from industry to participate. The problem is related to the task force ´NISSPI-Nature-inspired self healing soft sensors for process industry´. The solutions should inspire the task force activities and help to foster NiSIS cross-disciplinary networking.

2. PROBLEM TASK

2.1 DESCRIPTION OF THE PROBLEM

The reactor to be modelled consists of some 1000 tubes filled with catalyst, used to oxidize a gaseous feed (ethane is taken as example). It is cooled with a coolant supposed to be at constant temperature. The description of the reaction speed is taken from literature and depends strongly non-linearly from temperature.
Its exothermal reaction is counteracted by the cooling and leads to a temperature maximum somewhere along the length of the tube. As the catalyst decays, this becomes less pronounced and moves further downstream. The catalyst activity usually decays within some time to zero, a year is taken as example here. The process to be modelled takes input from other, larger processes, so that the feed will vary over the days. The operating personal reacts to this by choosing appropriate operating conditions. The catalyst decay is however much slower than these effects.
All measurable influences are considered as input variables for a mathematical multi-input-single-output-model describing relevant process variables (model outputs) representative for chemical industry:

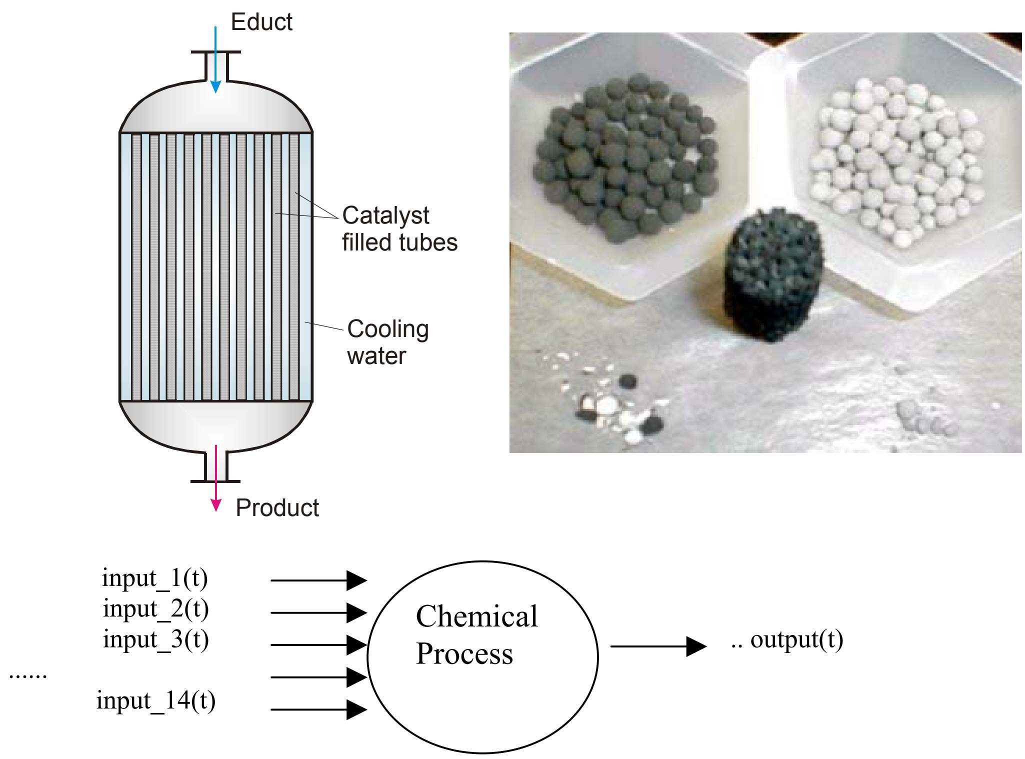

Figure 1: Principal description of the competition task

The process is equipped with measurements to log all the variations of the feed and the operating conditions. In addition, there are measurements showing some concentrations, flows and a lot of temperatures along the length of a characteristic tube to identify the processes state. All inputs and the output vary dynamically, and there might occur large time-delays.

2.2 OBJEKTIVE OF THE COMPETITION

The objective of the competition is to create a adaptive mathematical model describing the relationship between 14 input variables and one output variables, all of them varying with time. Such a model probably has to be adapted to process state changes resulting from non-measurable influences. After adaptation to the current working point of the process the model should be able to predict the output variable over a certain time horizon, supposed the future behaviour of the input variables is known. There was no restriction concerning methods and algorithms but an adaptive component is an obligatory requirement.

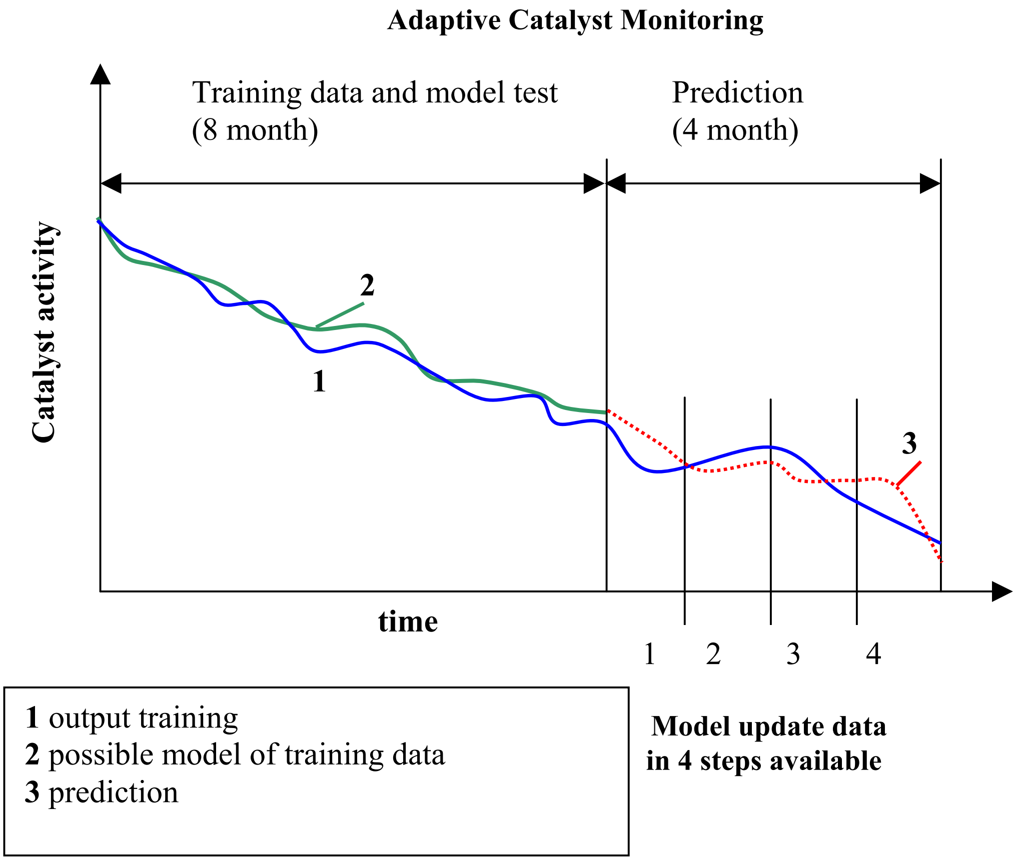

Figure 2: Adaptive modelling 4 month prediction time period

2.3 DESCRIPTION OF THE DATA FILE

For model identification a data file in EXCEL format containing process data will be provided. No filtering, smoothing or feature selection of data has been applied, since data pre-processing is regarded as part of the problem. In the data file the participant will find 16 columns, where the content of the columns is as follows:

column1: time in hours (1/24) since last catalyst change
column2-15: input data at this time
column16: output data at this time.

Input:

- Measured flow of air, kg/hr
- Measured flow of combustible gas, kg/hr
- Measured concentration of combustible component in combustible gas feed in mass fraction
- Total feed temperature, Cooling temperature
- Temperature at length 1/20, 2/20, 4/20, 7/20, 11/20, 16/20, 20/20 of reactor length in Celsius
- Product concentration of oxygen in mass fraction
- Product concentration of combustible component in mass fraction

Output:

- Catalyst activity

In the first phase of the competition the participants have access to two data files. The first data file covers 242 * 24 hours as a training data samples. This file corresponds to process data over a period of 8 month. For this time 5808 rows (one per hour) of data with input and all output values are given (training data set). For the last 4 month only input data will be provided in four steps, one file for each month (dates see below). The second file contains the input data for the first month (only column 1-15). The forecast for this first period have to be submitted until April 25th at the latest. Due to high complexity of the analysis the catalyst activity would be provided in a time interval of days only. Therefore the participants should not predict the output at each time interval but in two days steps. Each of the 15 rows where a predicted value is expected is marked with a string '???' in column 16.
The next three output files should be submitted two weeks after the publication of the input data via e-mail in an EXCEL – format mirroring the input data and showing the model output in the 16th column (second phase of the competition).

These four months represent the forecasting horizon where the model quality will be evaluated (comparison with output data in this period which are not given to the competitor). The criterion for model quality in this competition will be the closeness of model output to process output in a forecasting period of 4 month, where training data cover the preceding 8 month.Solutions will be ranked by accuracy *ERR* of the approximation of all output data during the forecasting horizon, according to the following formula:

$$ERR = \sum_{j=1}^{4} \frac{100}{N} \sum_{n=1}^{N} \frac{|O_{Rni} - O_{Fni}|}{|O_{Rni}|}$$

where

$N = 15$ (number of catalyst lab measurements in each month)
O_{Rni} = real output timestep n of output
O_{Fni} = forecasted output timestep n of output

2.4 RESULTS OF THE NISIS COMPETITION 2006

Finally we have received 11 solutions. The participation was time consuming because we expected solutions for four different time steps over a period of two month. Due to this organisational procedure the 2006 competition was not comparable to previous ones – not for the participants and not for the TTE committee as the organizer. We get 3 very promising approaches and with respect to the calculated error and the available report one solution seems to be outstanding (Fig. 3). Table 1 shows the result of all solution with the total error and the error of each indivudaul time step
The winner D. Ruta used a cross-trained ensemble of Neural Networks for his time series prediction. The best nature inspired model was suggested by M. Macas and his Elman network approach trained by a Particle Swarm Optimization (PSO).

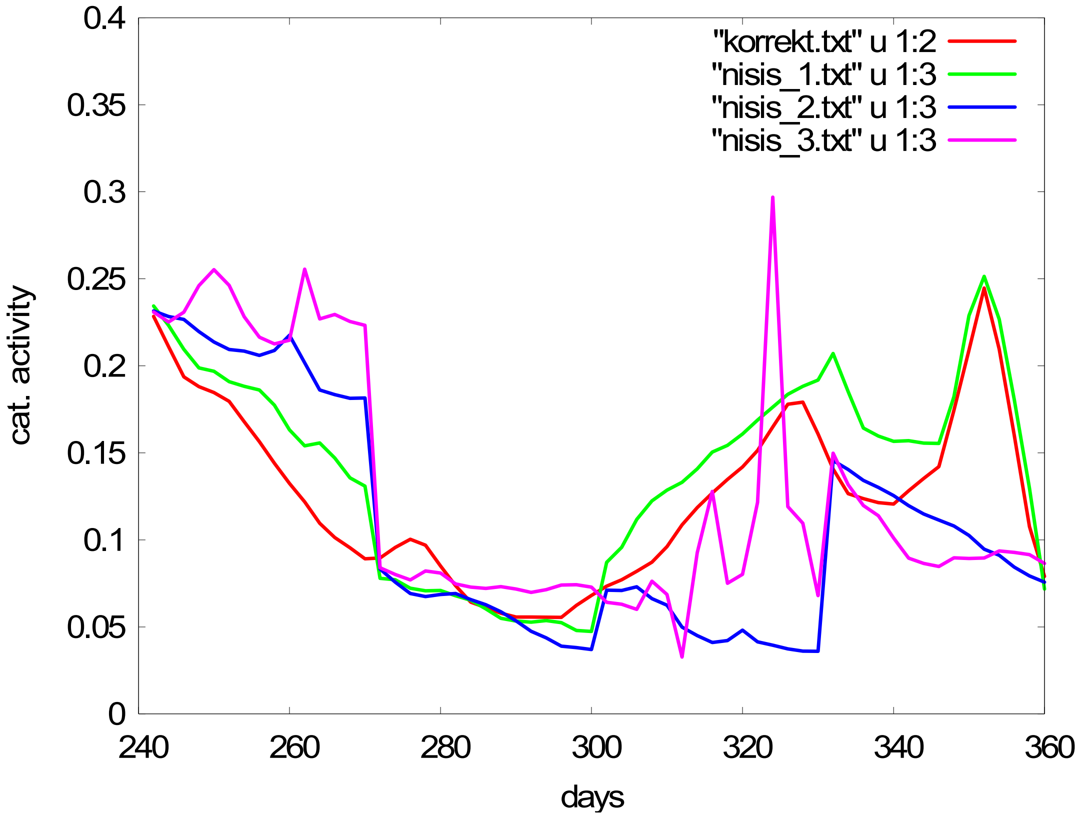

Figure 3: Three best solutions in comparison to the correct data set (red line). Best solution (green line) D. Ruta

	Sum	Test 1	Test 2	Test 3	Test 4
1	**73.06**	**21.01**	**12.87**	**19.14**	**20.13**
2	138.26	43.41	18.38	52.31	24.14
3	143.79	63.15	17.83	33.89	28.91
4	167.86	62.56	14.48	53.75	37.05
5	168.12	81.46	29.80	33.78	23.06
6	193.75	59.01	69,91	37,21	27.54
7	194.20	24.13	87.15	54.16	28.74
8	209.43	75.97	25,48	80,20	27.77
9	217.05	77.78	84.80	32.24	22.21
10	246.64	47.87	126.71	32.59	39.45
11	281.89	52.56	135.91	35.46	57.95

Table1 : Error of all 11 solutions. It is obvious that the winner solution is outstanding because the winner delivers in all for time steps the best model.

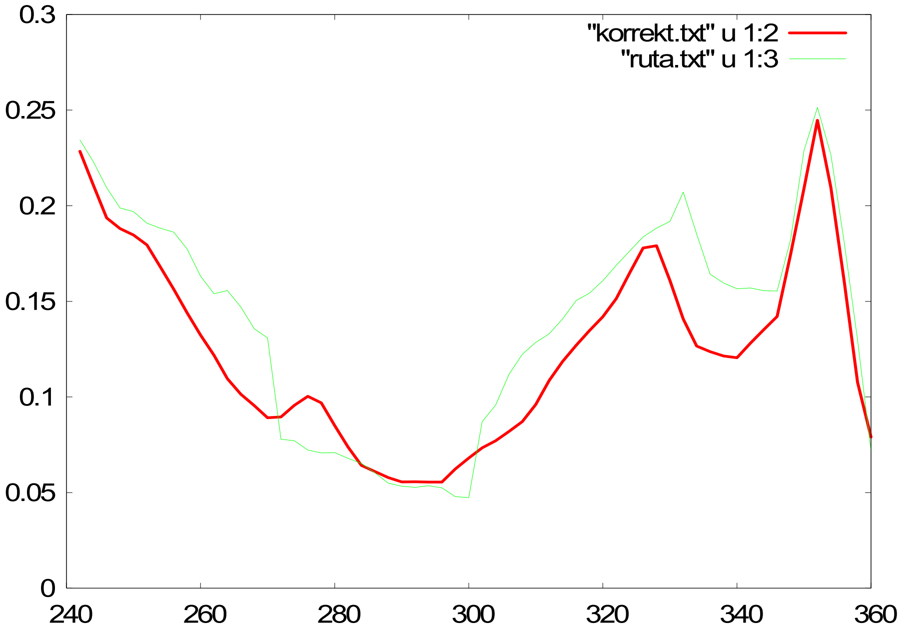

Figure 4: The winner solution given by D. Ruta shows an excellent prediction capability

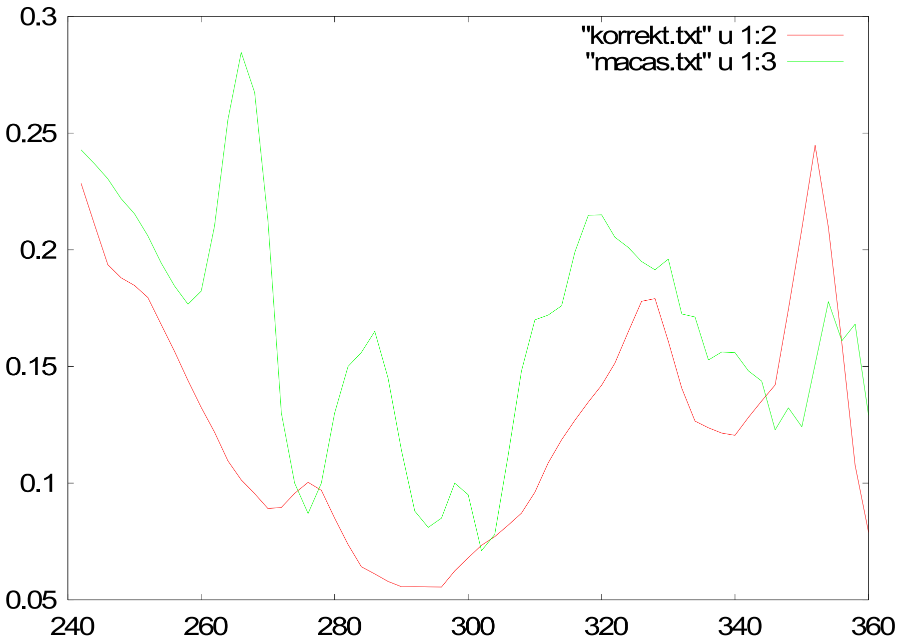

Figure 5: Prediction of the best nature inspired model by M. Macas

A more detailed description of both winner solutions is part of this booklet and is available online www.nisis.de