GECCO 2013 *Industrial Challenge: Optimizing Room Climates by Climate Data Forecasting*

Martina Friese, Oliver Flasch, Olaf Mersmann and Thomas Bartz-Beielstein [1]

Marchl 5, 2013

Goal of the GECCO 2013 Industrial Challenge is to develop accurate and computationally efficient forecasting methods for room climate data. Based on real-world room climate time series data, recorded by modern smart-home equipment, and ambient temperature data these methods should provide accurate predictions for room climate profiles. As the forecast should be seen as a basis for a intelligent heating control system, methods must be accurate and computationally efficient. This document provides a set of rules and regulations for the GECCO 2013 Industrial Challenge, a detailed problem description, as well as contact and submission information.

[1] Cologne University of Applied Science, 51643 Gummersbach, Germany (email: {martina.friese, oliver.flasch, olaf.mersmann, thomas.bartz-beielstein}@fh-koeln.de)

1 *Introduction*

When the words health, wellness and cost reduction go hand in hand, the topic is probably "smart home". The human feeling of warmth is determined by the surrounding temperature as much as by the humidity. At a humidity of about 60 percent, most people already feel comfortable warmth at 20°C. In contrast to this, a temperature of 23°C is needed for the same feeling of warmth at 30 percent humidity. A room climate with low humidity generated by strong heating encourages coughs and sneezes whereas a humidity of 60 percent or higher in combination with sparse ventilation eases the formation of mold. Furthermore, it is a matter of fact that each degree Celsius less heating shaves 6 percent off heating costs.

Smart home equipment records room climate data every time a value changes and checks the weather forecast at regular intervals to gain additional knowledge. The data collected can be correlated and analyzed to discover the particular heating properties of the room, the influence of the weather on the room climate and user specific patterns and preferences. This knowledge can be used to safe heating energy by only consuming energy that is really needed. Goal of the GECCO 2013 Industrial Challenge is to develop accurate forecasting methods for room climate profiles. Accurate forecasts build a sound basis for intelligent heating control. This reduces energy consumption and still provides a comforting and healthy room climate.

Goal of the GECCO 2013 Industrial Challenge is to develop accurate and efficient forecasting methods for room climate profiles. Accurate and fast room climate forecasting enables intelligent control of the room climate, by automatic regulation of the heating and giving proper advice about how and when to air the room to obtain

the most comfortable and healthy room climate possible. By giving this support in airing and regulating the heating only as needed it also helps to reduce heating cost and thus indirectly even contributes to a smaller carbon footprint. This time series forecasting problem will be referred to as the *room climate forecasting problem* in the remainder of this document.

The room climate forecasting problem belongs to the class of time series forecasting problems. [2] Although many different methods can be used for time series forecasting, Computational Intelligence (CI) methods, such as Evolutionary Computation and Artificial Neural Networks, offer an attractive option. CI methods are robust to changes in the underlying system, i.e.heating patterns or the ambient weather, which can preserve a comfortable room climate despite sudden changes in the weather situation. CI methods have been successfully applied to time series prediction problems in the past, which makes CI-based systems an interesting alternative to the classical time series analysis methods more widely applied in energy consumption forecasting, and motivated this competition. [3]

[2] P.J. Brockwell and R.A. Davis. *Introduction to Time Series and Forecasting*. Springer, New York NY, 2002. ISBN 0387953515

[3] J.D. Hamilton. *Time Series Analysis*. Princeton University Press, 1 edition, January 1994. ISBN 0691042896

Highlights of the GECCO 2013 Industrial Challenge include:

- *Interesting Problem Domain:* Forecasting room climate data based on smart home and ambient temperature data offers a challenging test case for modern time series prediction methods.
- *Real-world Data:* Real room climate time series, irregularly recorded when values change, are provided for training, testing, and assessing forecasting methods.
- *Realistic Quality Measurement:* Forecasting methods are assessed by their prediction accuracy for a continuum of prediction horizons, ranging from minutes to days into the future.
- *Fair Submission Assessment:* Prediction accuracy is determined on test data available to the organizers, which will be made public after the competition ends.
- *Direct Link to Industry:* GreenPocket GmbH will evaluate the winning submissions for real-world application and will be in direct contact with the winning participants, who will keep all rights to their forecasting systems.

The remainder of this document gives the information needed to take part in this competition. It is organized in three parts: Section 2 introduces the problem of making accurate forecasts of room climate profiles, based on smart-home and ambient tepmerature data, as well as the training time series data sets provided. Section 3 presents the set of rules and regulations. Finally, Section 4 gives information on how to participate in the industrial challenge.

2 *Problem Description*

The objective of this competition is to provide accurate time series predictions of room climate data based on real-world training data. The data provided for this competition consist of three independent time series. Two of those series, `Sensor_Temperature` and `Sensor_Humidity`, are recorded by a smart home device. The third time series, `Weather`, contains the outside temperature.

Based on these training data, participants of the challenge must provide a predicted room climate time series, with equidistant observations in 10 minute intervals, that extends the given room climate training time series into the future.

2.1 *Training- and Test-Datasets*

All three time series are series of timestamp and decimal number pairs, with timestamps being represented in an ISO 8601 derived text format. Temperature data is given in °C, humidity information in %.

While the ambient temperature time series comes with observations from equidistant time intervals, the time intervals between the observations from the room climate time series are rather iregular, since the values are only recorded when a change has been recognized. A first impression of the training time series data is given by figure 1.

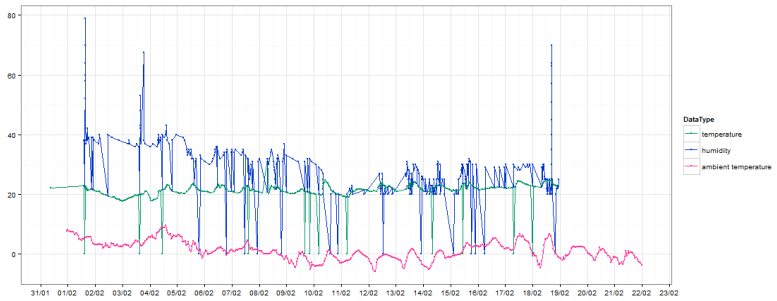

Figure 1: The Plot shows the complete training data time series. Room climate data is given for a timeframe of 3 weeks, while ambient temperature is known also for the time interval to-be-predicted

To participate in the competition, prediction time series for the room climate and room temperature over the time frame from `2013-02-19 00:00:00` to `2013-02-22 00:00:00` have to be supplied. These predicted time series should come with equidistant observations in 10 minute intervals, wich makes a total of 433 predicted observations.

As the true room climate time series during the prediction interval is known to the organizers, the score of a submission can be calculated based on prediction error, as described in the following subsection.

2.2 *Prediction Quality Rating*

In this competition, the quality of a predicted room climate time series is defined as the *root mean square error* (RMSE) between the predicted time series $\hat{t}$ and the respective linar interpolated true room climate time series t. For both, room temperature and humidity, the true time series are known to the organizers and will be published after the competition ends. Equation 1 defines the RMSE.

$$\mathrm{RMSE}(\hat{t}, t) := \sqrt{\frac{\sum_{i=1}^{n} (\hat{t}_i - t_i)^2}{n}} \tag{1}$$

Submissions are ranked by RMSE. See Section 4 for instructions on how to download reference material, source code, and documentation.

3 *Rules and Regulations*

The challenge is organized in a single round. In order to participate in the competition, prediction time series [4] for the room temperature and the humidity over the prediction interval given (see sec. 2.1) have to be supplied. These predicition time series are based on the training data given and feature the same data format. Submissions are accompanied by a two page report describing the algorithm that has been applied to generate the predicition time series. On request, algorithms have to be supplied as executable code, we expect the runtime of these to be reasonable. The organizers will rank the submissions by the total RMSE that is calculated as defined in Section 2.2. The winner of the GECCO 2013 Industrial Challenge will be the participant with lowest total RMSE.

[4] The concrete CSV format of the prediction files must match the format of the training data. See Section 4.1 for details.

4 *Submission*

Submissions to the GECCO 2013 Industrial Challenge should consist of:

- The predicition time series as described in Section 2.1 based on the training data sets given in Section 2.1,
- and a short report on the methods used (two pages maximum).

Please use the form provided on the website of the challenge at `http://www.spotseven.de/gecco-challenge/gecco-challenge-2013/` to send your submission. You can contact the organizers via email (`gecco@f10.fh-koeln.de`) if you have any questions.

4.1 *Software and Data*

Example predictions and the training data sets for the competition are available for download at `http://www.spotseven.de/gecco-challenge/gecco-challenge-2013/`.

4.2 *Important Dates*

- *Data Availability:* April 1, 2013
- *Industrial Challenge Submission Deadline:* May 31, 2013
- *GECCO 2012 Conference:* July 6-10, 2013

4.3 *Organizing Committee*

- Martina Friese, Cologne University of Applied Sciences
- Oliver Flasch, Cologne University of Applied Sciences
- Olaf Mersmann, Cologne University of Applied Sciences
- Thomas Bartz-Beielstein, Cologne University of Applied Sciences
- Jens Neuhalfen, GreenPocket GmbH

4.4 *Acknowledgements*

This work was sponsored by the by the Federal Ministry of Education and Research (BMBF; Germany), Förderlinie Ingenieurnachwuchs, projects FIWA (AiF FKZ 17N2309) and MCIOP (AiF FKZ 17N0311).

List of References

P.J. Brockwell and R.A. Davis. *Introduction to Time Series and Forecasting*. Springer, New York NY, 2002. ISBN 0387953515.

J.D. Hamilton. *Time Series Analysis*. Princeton University Press, 1 edition, January 1994. ISBN 0691042896.