Lars Dannecker, Robert Lorenz, Philipp Rösch, Wolfgang Lehner, Gregor Hackenbroich

Efficient Forecasting for Hierarchical Time Series

SLUB
Wir führen Wissen.

TECHNISCHE
UNIVERSITÄT
DRESDEN

QUCOSA
Quality Content of Saxony

Efficient Forecasting for Hierarchical Time Series

Lars Dannecker
SAP AG
lars.dannecker@sap.com

Robert Lorenz
SAP AG
robert.lorenz01@sap.com

Philipp Rösch
SAP AG
philipp.roesch@sap.com

Wolfgang Lehner
TU Dresden
Database Technology Group
wolfgang.lehner@tu-dresden.de

Gregor Hackenbroich
SAP AG
gregor.hackenbroich@sap.com

ABSTRACT

Forecasting is used as the basis for business planning in many application areas such as energy, sales and traffic management. Time series data used in these areas is often hierarchically organized and thus, aggregated along the hierarchy levels based on their dimensional features. Calculating forecasts in these environments is very time consuming, due to ensuring forecasting consistency between hierarchy levels. To increase the forecasting efficiency for hierarchically organized time series, we introduce a novel forecasting approach that takes advantage of the hierarchical organization. There, we reuse the forecast models maintained on the lowest level of the hierarchy to almost instantly create already estimated forecast models on higher hierarchical levels. In addition, we define a hierarchical communication framework, increasing the communication flexibility and efficiency. Our experiments show significant runtime improvements for creating a forecast model at higher hierarchical levels, while still providing a very high accuracy.

Categories and Subject Descriptors

G.3 [**Mathematics of Computing**]: Probability and Statistics—*Time series analysis*; G.1.6 [**Numerical Analysis**]: Optimization—*Nonlinear programming*

Keywords

Forecasting; Hierarchies; Time Series; Optimization

1. INTRODUCTION

Time series forecasting is an important statistical analysis technique used as a basis for manual and automatic planning in many application domains [12]. Forecasts are calculated using mathematical models that capture a parameterized relationship between past and future values to express behavior and characteristics of a historic time series. The parameters of these forecast models are estimated on a training data set to fit the specifics of the time series by minimizing the forecast error (i.e., difference between predicted and actual values). This estimation is typically conducted using local (e.g., LBFGS [16]) or global (e.g., Simulated Annealing [15])

optimization algorithms and is typically very time consuming, due to a parameter search space that increases exponentially with the number of model parameters. Additionally, many application domains exhibit a hierarchical data organization, with time series and forecast models on multiple levels. Here, the time series are aggregated along the hierarchy based on dimensional attributes such as location [14, 7, 8]. Forecasting in these environments is especially complex since it is necessary to involve data and entities across hierarchical levels and to ensure forecasting consistency among them.

In the energy domain accurate forecasts are of special importance, because they are a fundamental pre-requisite for balancing energy consumption and production. Previously, with a predominant usage of conventional energy sources, the energy supply could be almost perfectly matched to the energy demand. However, with an increasing share of renewable energy sources new requirements are posed on the balancing and forecasting process [24, 5, 3]. Intermittent renewable energy sources such as wind and solar power are subject to frequent and strong fluctuations and thus, their production is hard to predict [24, 3]. In addition, the amount of energy storage is rather limited, which means that the energy from intermittent renewable energy sources must be directly used when it is available [3]. As a consequence, a continuous and more fine-grained balancing in real-time is required. Research projects such as MIRABEL [19] and MeRegio [18] develop advanced technologies such as demand response systems and flexible energy requests [1, 2] to further address the challenges of real-time energy balancing. A fundamental pre-requisite for these approaches is the availability of current and accurate forecasts at any point in time.

We tackle the issue of constantly available accurate forecasts by introducing a novel hierarchical forecasting system that exploits the hierarchical organization of the energy market. The approach builds upon and enhances some early ideas [17]. We propose to decentralize the forecasting calculation by deploying forecast models directly to the smart meters of customers. These individual customer forecast models reflect the energy consumption or production of its respective entity. Companies on higher hierarchical levels can utilize the individual customer models to form a global forecast model representing the measurements of all connected lower level entities. This merging process is several magnitudes less expensive than the (re-)estimation of the forecast model parameters. Thus, with the help of our hierarchical forecasting system balancing companies are able to rapidly calculate very accurate forecasts and with that we enable them to balance energy in real-time. Overall, this paper makes the following contributions that also reflect the organization of the paper. First, we discuss the organization of the European energy market exploited in our hierarchical forecasting system in Section 2. Second, we introduce an efficient forecasting approach for hierarchical time series in Section 3. Third, we pro-
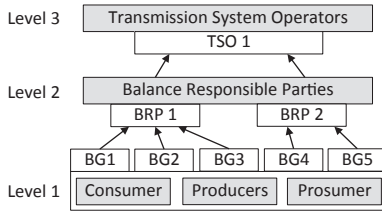
Figure 1: Hierarchical EDMS from the MIRABEL project [2]

pose a communication protocol enabling hierarchical forecasting in Section 4. Fourth, we present our experimental evaluation that substantiates the benefits of our approach in Section 5. We also discuss related work in Section 6 and conclude the paper in Section 7.

2. THE EUROPEAN ENERGY MARKET

In this paper we focus on the European energy market, which is currently changing very drastically with the goal of creating a competitive, sustainable and integrated single European energy market by 2014 [10, 21]. Since the energy market is hierarchically organized, energy data management systems (EDMS) are typically organized accordingly. Figure 1 illustrates the hierarchical EDMS of the research project MIRABEL[19], which reflects a simplified version of the target role model for the European energy market [2, 6]. The lowest level nodes are made up by industrial energy producers as well as private and industrial consumers. Furthermore, customers consuming and producing energy (e.g., solar panels) at the same time are called prosumer. At this level, smart meters continuously monitor the consumed or produced energy. Producers, consumers and prosumers are pooled into balance groups, with the goal of balancing the energy consumption and production within these groups. Companies taking the role of a balance responsible party (BRP) are managing the balance groups. The third level comprises the transmission system operators (TSOs), which account for a stable operation of the grid and the transportation of electricity.

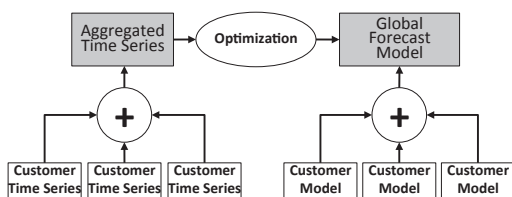## 3. FORECASTING IN HIERARCHICAL ENVIRONMENTS

Figure 2: Estimation vs. Aggregation Approach

To allow efficient forecasting in hierarchical environments, we propose to exploit the hierarchical organization of time series and to decentralize the forecasting process. More specifically, the responsibility to build and maintain a forecast model is pushed to the base level of the hierarchy. In the energy domain this is motivated by the fact that smart meters are expected to be widely deployed and thus, we can exploit their available computing capacity. Hence, next to their task of recording current measurements, smart meters gain the capability to predict the future consumption or production of a customer. The employed forecast models can be seen as a compact representation of the customer data and in an ideal case the forecast model describes values that are identical to the

actual time series. While, in the real-world forecast models always exhibit a certain error, this error is typically reduced with an increasing aggregation level, since in the entirety single fluctuations are neutralized to a large extend. Accordingly, the idea is to transmit the forecast models of the single customers and aggregate them to directly determine a global forecast model describing the energy consumption and production of the entire group. Forecast model aggregation in this sense means, that each of the global forecast model components (i.e., the combination of a forecast model coefficient x and its parameter p) is calculated using a weighted linear combination of the respective single customer components. Figure 2 compares our forecast model aggregation on the right-hand side, with the conventional model construction using time series aggregation on the left-hand side. Despite using the energy domain as an example, our approach can also be adapted to other application domains exhibiting hierarchical time series such as sales or retail.

The proposed forecast model aggregation is far less time consuming compared to the conventional forecast model estimation. As a result, our hierarchical forecasting approach provides a very efficient way for higher-level entities to create accurate global forecast models in several orders of magnitudes less time (experiments in Section 5). In addition, our hierarchical forecasting approach enables a more dynamical communication between the entities on both layers. Instead of transmitting values in a fixed interval as typically done in the conventional approach, a lower level entity initiates an information exchange depending on the accuracy of its forecast model; as long as the base model is up-to-date there is no need to exchange information between the hierarchy levels.

3.1 Classification of Forecast Model Coefficients and Parameters

Forecast model coefficients can be divided into endogenous and exogenous coefficients depending on their characteristics and their handling in forecast models. Endogenous coefficients describe the main time series that is aggregated through the hierarchical levels: $x = \sum_{k=1}^{K} x_k$. Typical endogenous coefficients in the energy domain are energy consumption, seasonal information and past forecast errors. In contrast, exogenous coefficients describe an external factor influencing the main time series. Exogenous coefficients occur as additional time series or dummy variables (existence/non-existence using 0 or 1), where the external influence is determined by the combination of time series value and parameter value when using an additional time series and by the parameter value only when using a dummy variable. Typical examples for exogenous coefficients are the current day of the week (dummy variable, e.g., Monday=0 or 1) or the temperature (time series, e.g., temperature at customer k is $15\,^\circ\mathrm{C}$). External influences are typically aggregated by using the average of all entities: $x = \frac{1}{k} \sum_{k=1}^{K} x_k$.

3.2 Aggregation in Detail

Creating aggregated global time series $x(t_i)$ on a higher level means to sum the time series of all lower level entities $x_k(t_i)$, with k referring to the k-th customer and t to the current point in time. Accordingly, following our assumption that a forecast model approximately describes its underlying time series, we create the next level global forecast model $M(t_i)$ by aggregating the forecast models of all lower level entities $M_k(t_i)$ as illustrated in Equation 1.

$$x(t_i) = \sum_{k=1}^{n} x_k(t_i) \approx \sum_{k=1}^{n} M_k(t_i) = M(t_i). \tag{1}$$

In the following, we omit the notation (t_i) for readability reasons.

A forecast model M can be divided into separate components $M(C_1, ..., C_n)$, where we define a model component C as a forecast model coefficient x combined with its parameter p so that $C = p \cdot x$. A linear model $a \cdot x + b$ for example consist of two model components $a \cdot x$ and b (where the coefficient is 1). To create the global forecast model M, we separately aggregate the components C_k of all corresponding lower level forecast models ($M_k, k \in K$), including their coefficients x_k and parameters p_k. Accordingly, the calculation of the global forecast model component C, which likewise consists of a global parameter p and a global aggregated coefficient $x = \sum_{k=1}^{K} x_k$, is defined as shown in Equation 2.

$$C = p \cdot x = \sum_{k=1}^{K} p_k x_k. \tag{2}$$

Since the aggregated time series of the coefficient x is available to the upper level entity, the only missing factor for calculating a global forecast model component is the global parameter p. We adapt Equation 2, to define a general aggregation rule for determining the parameter of a global forecast model component and arrive at Equation 3. Dividing the single entity time series value x_k by the global aggregated time series coefficient x, as done in Equation 3, is equivalent to determining the current share of a single entity value on the global value. Thus, the global forecast model parameter p is derived by creating a linear combination of the single entity parameters multiplied by their share on the global time series.

$$p = \sum_{k=1}^{K} \frac{p_k x_k}{x}. \tag{3}$$

Equation 3 has the disadvantage of considering the coefficient values of the individual entities x_k and the values of the respective global coefficient x. Thus, we would have to request and process the current measurements of all lower level entities each time the global forecast model is adapted, which clearly contradicts to our target of a more efficient communication. To solve this issue, we approximate the current share (x_k/x) of an individual entity by its average historic share (δ). Since the average share of an entity is typically relatively stable over time, it is not necessary to recalculate it for every adaptation of the global forecast model. We define the historic share of an individual entity as:

$$\delta_k = \frac{\bar{x}_k}{\sum_{k=1}^{K} \bar{x}_k} \text{ where, } \bar{x}_k = \frac{1}{N} \sum_{i=1}^{N} x_k(t_i). \tag{4}$$

Thus, the average historic share of an entity is the ratio between the arithmetic mean over a time series sample ($1...N$) of the single entity to the summed arithmetic mean of all entities. The sample size is configuration-specific, where we found a sample size of one day as most beneficial. The necessary historic values for creating the average historic share δ_k are transmitted in conjunction to the forecast models of individual entities. Accordingly, the initial share is determined during the system initialization, where the higher-level entities receive the forecast models and average historic values from all lower level entities. During runtime, we adapt the share of an individual entity, whenever it transmits an adapted forecast model. In the energy domain the determination of the shares is even simpler, because for billing purposes all lower level entities transmit their actual measurements once per accounting period anyway. Since this transmission is done asynchronous to the forecasting process, it still increases the communication efficiency compared to transmitting values for every model adaptation. Using δ_k in Equation 3, we calculate the model parameters on the next level as a weighted linear combination as shown in Equation 5.

$$p = \sum_{k=1}^{K} p_k \delta_k. \tag{5}$$

The global forecast model is now created by applying the aggregation rule to all components of the model $M(C_1, ..., C_n)$. Thus, for aggregating an individual global forecast model component we plug in equation 5 into $C = p \cdot x$ and finally arrive at Equation 6.

$$C = \sum_{k=1}^{K} p_k \delta_k x. \tag{6}$$

For exogenous coefficients we adapt Equation 3 to the fact that for this coefficient type customer values are aggregated instead of summed. Thus, $x = \bar{x} = \frac{1}{K} \sum_{k=1}^{K} x_k$ and the aggregation rule is:

$$p = \sum_{k=1}^{K} \frac{p_k x_k}{\bar{x}} = \sum_{k=1}^{K} p_k \sigma_k. \tag{7}$$

This means that we scale the parameters of exogenous coefficients by the ratio of the single entity value to the global average value $\sigma = x_k/\bar{x}$. Like for endogenous parameters we approximate the current ratio by the historic ratio based on entity values communicated during the last forecast model transmission. For exogenous coefficients exhibiting the same value for all entities (in most cases dummy variables), we can further simplify the aggregation rule. In this case the global parameters p can be calculated by simply summing the parameters of the individual lower level entities.

With both aggregation rules, our hierarchical forecasting works for models involving either endogenous or exogenous coefficients. However, when combining both coefficient groups we found a lack in accuracy. The reason is a varying influence of each coefficient group that is usually automatically addressed during the conventional parameter estimation. Since we do not estimate the parameters in our hierarchical forecasting approach, we address this issue by introducing an influence weight for each coefficient group. Accordingly, we adapt Equations 5 and 7 for endogenous coefficients $p(end)$ and exogenous coefficients $p(exo)$ respectively leading to:

$$p(end) = \omega_1 \sum_{k=1}^{K} p_k \delta_k \text{ and } p(exo) = \omega_2 \sum_{k=1}^{K} p_k \sigma_k \tag{8}$$

The weights ω_1 and ω_2 have to be estimated using optimization algorithms, but the estimation of two parameters is computationally very cheap, especially when compared to the large number of parameters typically estimated during conventional forecasting.

With our hierarchical forecasting approach it is possible to create higher-level forecast models without a time-intensive parameter estimation. Our approach does not only work between consecutive hierarchical levels, but across all hierarchy levels. Thus, in the energy domain balance responsible parties (BRP) as well as transmission system operators (TSO) are likewise enabled to rapidly provide accurate forecasts as needed for real-time balancing. Also customers might benefit from the availability of forecast models in their smart meters, since smart home and smart building systems might use the forecasting capabilities to provide improved and enhanced functionalities. Even the usage of concepts such as the MIRABEL flex-offers or other demand-response systems [11] can be better implemented with the availability of local forecasts.

4. HIERARCHICAL COMMUNICATION

In this Section we specify the communication protocol for transmitting information between hierarchy levels. Previously the lower

level entities provided measurements in fixed intervals (e.g., every 5 minutes). With our protocol we establish a more flexible communication, where information are only transmitted when they are needed. The protocol is illustrated in Figure 3. First, an initialization is conducted when using our hierarchical forecasting for the first time or when a re-initialization is necessary (e.g., large organizational changes). (1) During the initialization the upper level entity requests the individual forecast models M_k of all corresponding lower level entities. (2) The lower level entities estimate their forecast models and afterwards transmit the model parameters to the next level entity. (3) The higher-level entity creates its global forecast model M by aggregating all transmitted forecast models.
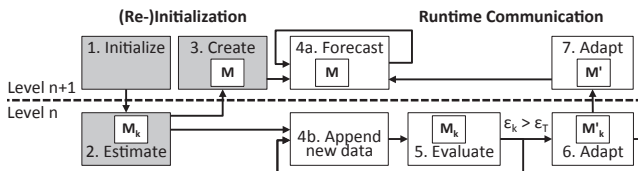
Figure 3: Communication protocol for hierarchical forecasting

(4a) During runtime the higher-level entities calculate forecasts based on the global forecast model M, (4b) while the lower level entities append new measurements to their time series. (5) The new values are used to evaluate the accuracy of the individual forecast models M_k. (6) Whenever the forecast error ϵ_k exceeds a specified error threshold ϵ_T, the respective entity adapts its forecast model M'_k and transmits the adapted model to the next level entity. (7) The next level entity uses the received forecast model to adapt the global forecast model M. Our communication protocol also works between multiple hierarchy levels in a cascading way. The lowest level entities provide information to their second level entity, which adapts its global forecast model accordingly. Afterwards the entity on the third level of the hierarchy is informed by the respective second level entity and so on.

For our approach to work we assume that higher-level entities still have access to the most recent aggregated time series (e.g., balance group consumption, company revenue, global sales figures). This is the case in many application domains, since the time series at higher aggregation levels are recorded separately. In the energy domain for example, BRPs have metering devices at the level of their balance group. Recordings at this level are necessary to guarantee the balancing of energy demand and supply for the entire group. In addition, with an increasing employment of smart grids, TSOs have much finer monitoring capabilities and could also provide the necessary information to entities on other levels.

To sum up, our hierarchical communication protocol provides an efficient way for exchanging information within a hierarchy. In most cases it is sufficient to just transmit data when forecast models are adapted at lower level entities. While there might still be the need to transmit the actual measurements for accounting purposes, this transmission can be handled asynchronously and independent of the forecasting calculation. Data can for example be transmitted and processed in a bulk mode when free resources are available.

5. EXPERIMENTAL EVALUATION

In our experimental evaluation we show that our hierarchical forecasting approach can rapidly calculate forecasting results, while still providing a high accuracy. In this evaluation we use two forecast models the Double Seasonal HWT (DSHWT) and the multi-equation EGRV forecast model [22]. The single-equation DSHWT

model as introduced by Taylor et al. is a domain-specific adaptation of the general Holt-Winters exponential smoothing model [23]. It exhibits four parameters and includes the daily and the weekly season. The EGRV model uses a separate sub-model per hour and different models for working days and weekends. This results in a total number of 48 sub-models. For our evaluation we did not consider temperature data. Altogether, our implementation of the EGRV model exhibits 28 parameters in total. The parameter estimation was conducted using the local Nelder Mead Downhill Simplex algorithm [20] and the global simulated annealing [15] algorithm. For the global simulated annealing algorithm we empirically determined a time budget, where we found a steady state of the forecast after around 90 seconds for the EGRV model and 17 seconds for the DSHWT model. Accordingly, we set the time budget to 100 seconds and 20 seconds respectively.

The basis for our evaluation is the real-world data set from the MeRegio project [18]. The data set contains the hourly energy demand of 86 private customers from the 1st Novermber 2009 to the 22nd March 2010. For the lowest level in the hierarchy we directly used the 86 customers from the MeRegio data set. To form the second level entity, we aggregated the time series of all 86 customers. While this data set is rather small, it still shows the applicability and advantages of our approach on a real-world data set. Larger datasets will lead to an even larger benefit.

The results are presented with respect to forecast error and efficiency. We measured the forecast error using the SMAPE error metric [13]. The efficiency is represented by the execution time as well as the number of iterations the respective optimization algorithms needed. The results are the average of multiple runs from different starting points and 30 repetitions. For the experiments we used an Intel Core 2 Duo P8400 (2.26 GHz), 8 GB RAM, Mac OSX 10.7. The prototype was implemented in C++ using gcc 4.2.1.

5.1 Model Aggregation using DSHWT

In this experiment we compare our hierarchical forecasting to conventional forecasting using the DSHWT model. For the hierarchical forecasting, we first estimated the parameters of all 86 customer forecast models and aggregated them to form the second level model. The DSHWT model contains endogenous parameters only, which means that solely Equation 5 applies. For the conventional forecasting we estimated the second level model using the local Nelder Mead and the global Simulated Annealing algorithm.

The results are illustrated in Figure 4, where Figure 4(a) compares the runtime for the Nelder Mead algorithm and Figures 4(b) and 4(c) compare the final accuracy. We can see that our hierarchical forecasting is multiple orders of magnitudes faster than the conventional parameter estimation. The model aggregation takes only 0.16 ms compared to 3,400 ms when using the Nelder Mead algorithm. For a larger number of values the runtime difference is expected to even increase, since the conventional parameter estimation iterates over the dataset multiple times. In contrast, the model aggregation does not iterate over the time series and thus, remains stable with an increasing number of values. However, an increasing number of entities slightly increases the time for solving a weighted linear combination, but this remains computationally inexpensive, compared to a parameter estimation using optimization algorithms.

With respect to the accuracy, our hierarchical forecasting provides almost as accurate results as the conventional forecasting. The forecast error increases only slightly by 0.04% compared to the regular estimation using the Nelder Mead algorithm. For Simulated Annealing, the accuracy also differs only marginally. Our hierarchical forecasting exhibits an error of 3.73% and the conventional estimation provides an error of 3.62%. We also evaluated the
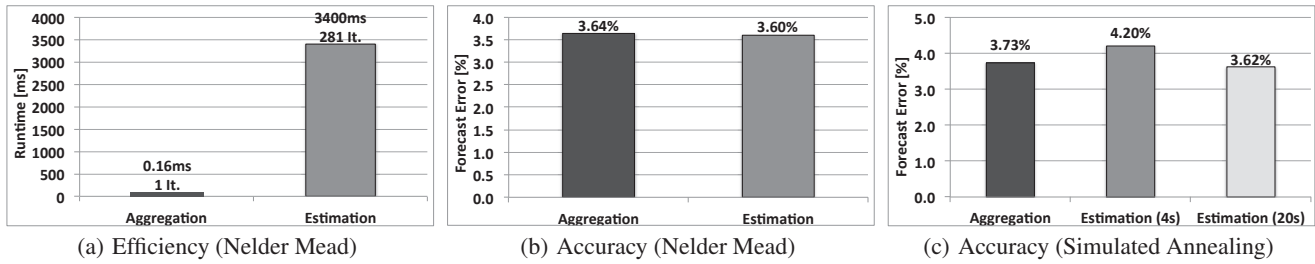
(a) Efficiency (Nelder Mead)   (b) Accuracy (Nelder Mead)   (c) Accuracy (Simulated Annealing)

Figure 4: Results for the hierarchical forecasting with the DSHWT model

(a) Efficiency (Nelder Mead)   (b) Accuracy (Nelder Mead)   (c) Accuracy (Simulated Annealing)
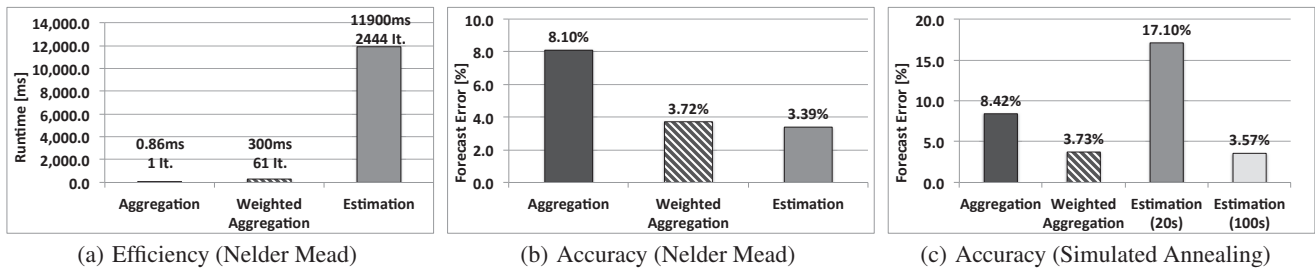
Figure 5: Results for the hierarchical forecasting with the EGRV model

accuracy, when significantly reducing the time budget for the Simulated Annealing algorithm to only 4 s (hierarchical forecasting took 0.16 ms). In this case, the forecast error increases to 4.20%. Thus, it is not possible to conform the runtime of Simulated Annealing to our hierarchical forecasting without sacrificing accuracy. As a result, for the single-equation DSHWT model we can significantly reduce the runtime, while still providing a high accuracy.

5.2 Results EGRV Parameter Adjustment

In this experiment we compare our hierarchical forecasting to the conventional parameter estimation using the EGRV model. The general settings for this experiment are the same as for the DSHWT model. However, since the EGRV model contains both exogenous and endogenous coefficients, we also compared the accuracy when using our weighted aggregation described in Equation 8.

The results are illustrated in Figure 5, where Figure 5(a) compares the runtime when using the Nelder Mead algorithm and Figures 5(b) and 5(c) present the accuracy results. The hierarchical forecasting (Aggregation) exhibits the best runtime with only requiring 0.86 ms, while the conventional estimation using Nelder Mead required 11,900 ms. However, the sole aggregation caused a substantial increase of the forecast error to 8.10% compared to 3.39% when estimating the parameters. The results are similar when using Simulated Annealing (hierarchical forecasting: 8.42%, estimation: 3.57%). The weighted aggregation significantly reduces the forecast error to 3.72% and 3.73% respectively, which brings it very close to the conventional estimation again. Hence, the experiment clearly shows the need for the weighted aggregation when dealing with forecast models comprising both coefficient types. At the same time the weighted aggregation requires to estimate two weights for the two parameter groups, which increases the runtime to 300 ms. However, the weighted aggregation is still several magnitudes cheaper in both the runtime (300 ms vs 11,900 ms) and the number of iterations (61 iterations vs. 2,144 iterations). We also again evaluated the accuracy when reducing the time budget for the Simulated Annealing algorithm from 100 s (326,983 iterations) to 20 s (63,475 iterations). This reduction sig-

nificantly increased the forecast error to 17.10%, which means a drastic drop in accuracy.

As a result, for the EGRV model our hierarchical forecasting again provides a significant runtime improvement, allowing an almost instant calculation of accurate forecasts. However, to still allow for a high accuracy, the separate weighting of both parameter groups is very important. Overall, the accuracy difference is slightly higher than for the DSHWT forecast model, but still well in range given the significant runtime benefits.

6. RELATED WORK

Forecasting of hierarchical time series is an emerging issue that gains more and more attention in research and industry. Multiple studies [4, 9, 25] analyzed this topic and especially examined the most beneficial aggregation. In general, they distinguish between bottom-up and top-down approaches. Using the bottom-up approach forecasts are calculated at the lowest level and aggregated through the hierarchy. In contrast, using the top-down approach forecasts are calculated at the top level and disaggregated to lower levels. The studies do not arrive at a common solution since it is not ultimately decidable which approach is more beneficial.

Hyndman et al. [14] propose a hierarchical forecasting that aims at providing a better accuracy than either a complete top-down or bottom-up approach. They calculate independent forecasts on all hierarchical level and combine them in accordance to the hierarchical structure using a regression model. In contrast to our approach, Hyndman et al. aim at increasing the forecasting accuracy instead of optimizing the efficiency. In particular, their approach is more expensive, since they estimate models on all hierarchical levels.

Recently, Fischer et al [7] published a hierarchical forecasting approach that uses only a sample of base models. Forecasts on a specific hierarchy level may be based on a subset of optimized models on other levels. Despite limiting the number of base models they achieved a reliable accuracy comparable to aggregating all base models. In contrast to our approach, the concept of Fischer et al. is based on aggregating time series, instead of aggregating forecast models. Thus, all forecast models involved are estimated us-

ing numerical optimization algorithms. In addition, we only transmit information through the hierarchy, when forecast models were changed, while Fischer et al. transmit data for any forecasting calculation. In general, they aim at increasing the forecasting accuracy rather than the forecasting efficiency.

7. CONCLUSION AND FUTURE WORK

In this paper, we presented a novel hierarchical forecasting approach that substantially increases the forecast calculation efficiency for higher-level entities. The core idea of our approach is to decentralize the forecast process and to distribute the responsibility for building and maintaining forecast models to the lowest level of the hierarchy. Our approach is motivated by the energy domain, where smart meters are more and more deployed to private and industrial customers. Smart meters provide some data processing capabilities that we exploit for maintaining individual customer forecast models. These individual forecast models are utilized on higher hierarchy levels to rapidly create global forecast models without conducting a time-consuming parameter estimation. Thus, with the help of our approach we significantly reduce the time for calculating forecasts. In addition, we defined a communication protocol, increasing the flexibility of the forecasting related communication within the hierarchy. In our experiments we demonstrated that our hierarchical forecasting approach significantly reduces the time for calculating forecasting results, while still providing a very high accuracy. Despite describing our approach exemplarily by means of the energy domain, it can be easily adapted to other application domains such as the retail domain.

In the future we want to enhance our approach by allowing the use of heterogeneous forecast models on the different hierarchical levels. There, we want to capture the gradient of forecast model adaptations and exploit this information to ease the estimation of different forecast models on higher hierarchical levels. Additionally, we plan to examine if we can determine the global forecast model when only using a subset of the base models.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] H. Berthold, M. Boehm, L. Dannecker, F.-J. Rumph, T. B. Pedersen, C. Nychtis, H. Frey, Z. Marinsek, B. Filipic, and S. Tselepis. Exploiting renewables by request-based balancing of energy demand and supply. In *11th IAEE European Conference*, 2010.

[2] M. Boehm, L. Dannecker, A. Doms, E. Dovgan, B. Filipic, U. Fischer, W. Lehner, T. B. Pedersen, Y. Pitarch, L. Siksnys, and T. Tusar. Data management in the mirabel smart grid system. In *1st EnDM (15th EDBT)*, 2012.

[3] M. Cailliau, J. A. Ogando, H. Egeland, R. Ferreira, H. Feuk, F. Figel, S. G. Jensen, J. Karas, C. Kawann, C. M. Villar, J. McManus, S. Naletto, H. Nosei, A. Poullikkas, J. Sundell, R. Vrolijk, B. Walter, and M. Zoglauer. Integrating intermittent renewable sources into the eu electricity system by 2020: Challenges and solutions. Industrial analysis and report, Eurelectric: Electricity for Europe, 2011.

[4] D. M. Dunn, W. H. Williams, and T. L. DeChaine. Aggregate versus subaggregate models in local area forecasting. *Journal of the American Statistical Association*, 71:68–71, 1976.

[5] ENTSO-E. *Scenario Outlook and System Adequacy Forecast 2011 - 2025*, 2011.

[6] ENTSO-E - European Network of Transmission System Operators. *The Harmonised Electricity Market Role Model*, 2011.

[7] U. Fischer, F. Rosenthal, and W. Lehner. Sample-based forecasting exploiting hierarchical time series. *Proceedings of the 16th IDEAS*, pages 120–129, 2012.

[8] U. Fischer, C. Schildt, C. Hartmann, and W. Lehner. Forecasting the data cube: A model configuration advisor for multi-dimensional data sets. In *29th ICDE*, 2013.

[9] G. Fliedner. Hierarchical forecasting issues and use guidelines. *Industrial Management & Data Systems*, 101:5–12, 2001.

[10] General Secretariat of the Council. European council 4 february 2011: Conclusions. Governmental publication, European Commission, 2011.

[11] B. Gnauk, L. Dannecker, and M. Hahmann. Leveraging gamification in demand dispatch systems. *Proceedings of the 1st EnDM workshop (EDBT 2012)*, pages 103–110, 2012.

[12] J. G. D. Gooijer and R. J. Hyndman. 25 years of time series forecasting. *International Journal of Forecasting*, 22(3):443–473, 2006.

[13] R. J. Hyndman. Another look at forecast-accuracy metrics for intermittent demand. *The International Journal of Applied Forecasting*, (4):43–46, 2006.

[14] R. J. Hyndman, R. A. Ahmed, G. Athanasopoulos, and H. L. Shang. Optimal combination forecasts for hierarchical time series. *Computational Statistics and Data Analysis*, 55(9):2579–2589, 2008.

[15] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[16] D. Liu and J. Nocedal. On the limited memory bfgs method for large scale optimization methods. *Mathematical Programming 45*, 45(3):503–528, 1989.

[17] R. Lorenz, L. Dannecker, P. Roesch, W. Lehner, and G. Hackenbroich. Forecasting in hierarchical environments. In *25th SSDBM*, to appear 2013.

[18] MeRegio Project. http://www.meregio.de/en/, 2012.

[19] MIRABEL Project. http://www.mirabel-project.eu, 2013.

[20] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[21] I. Pineda and P. Wilczek. Creating the interal energy market: a report by the european wind energy association. Industrial analysis and report, EWEA - European Wind Energy Association, 2012.

[22] R. Ramanathan, R. Engle, C. W. Granger, F. Vahid-Araghi, and C. Brace. Short-run forecasts of electricity loads and peaks. *Int. Journal of Forecasting*, 13(2):161–174, 1997.

[23] J. W. Taylor. Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204(1):139–152, 2009.

[24] J. Wilkes, J. Moccia, and M. Dragan. Wind in power: 2011 european statistics. Industrial analysis and report, EWEA - The European Wind Energy Association, 2012.

[25] A. Zellner and J. Tobias. A note on aggregation, disaggregation and forecasting performance. *Journal of Forecasting*, 19:457–469, 2000.