Predictive Analytics on Data Streams

Petra Vrablecová*
Institute of Informatics, Information Systems and Software Engineering
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava
Ilkovičova 2, 842 16 Bratislava, Slovakia
petra.vrablecova@stuba.sk

Abstract

The growing amounts of data brought about the need for transition from offline to online data processing. In our thesis, we explored the utilization of incremental and online data processing in power engineering domain, where this need is imminent because of ongoing introduction of smart metering. We designed two stream processing power demand forecasting methods, which addressed the main requirements of stream mining, i.e., accuracy, timeliness and adaptability. The third requirement relates to changes that occur in stream data over time, i.e., concept drifts. We studied two ways of adaptation of prediction model to the drifts – informed and blind, which differ in incorporation of an explicit change detection mechanism. We proposed incremental forecasting method with informed adaptation and online method with blind adaptation. Both of our approaches equaled the standard batch approaches in accuracy with less computing resources.

Categories and Subject Descriptors

• **Mathematics of computing → Time series analysis**; *Regression analysis*; • **Information systems → Data streams**; **Data mining**; **Data stream mining**; *Data analytics*; • **Theory of computation → Online learning algorithms**; **Support vector machines**; Bio-inspired optimization; • **Computing Methodologies → Machine learning**; **Supervised learning by regression**; **Online learning settings**; • **Applied computing → Forecasting**.

Keywords

prediction, forecasting, data mining, stream mining, machine learning, incremental, online, adaptive prediction methods, concept drift, power demand forecasting

1. Introduction

The data volumes produced by various systems, devices and sensors are continually growing. Likewise, the need for methods, which help humans analyze and discover knowledge from the large volumes of data, is growing. The big data phenomenon is observed since the beginning of this century [12]. Contemporary technologies can already store large data quantities. Since the value of information hidden in data tapers over time, the data must be processed and analyzed online. Therefore, batch processing and analysis are in many cases no longer sufficient and the stream data processing is preferred.

Our thesis is aimed at prediction methods in dynamic environments with continually growing and changing data sets (data streams), which manifest the need for transition from batch to stream processing.

We focused on one application domain – power engineering, because of smart meters' introduction. Smart meters replace the annual physical power consumption meter readings by continual sending of interval measurements (e.g., every 15, 30 or 60 minutes). In the directive no. 2009/72/EC issued by the European Parliament and the Council, all European countries undertook to equip at least 80% of electricity consumers with a smart meter by 2020. Slovakia implemented this obligation in the Act no. 251/2012 on Energy Sector. According to terms in ordinance of Ministry of Economy of the Slovak Republic no. 358/2013, distribution network operators must install the smart meters.By 2020, all the consumers who are connected to the regional or local distribution system with low voltage and their consumption is higher or equal to 4MWh per year should be equipped with a smart meter. That is about 600,000 out of 2.38 million consumers connected to the low voltage [19].

Power suppliers will have at hand large volumes of streaming data for analysis. Analysis of smart metering data is beneficial for all stakeholders of the electricity market. Customers can monitor and lower their own consumption, suppliers can plan the power supply more precisely, invoice the exact amounts of consumed electricity, create tariffs tailored for customers, identify outages and illegal consumption easily, etc. [19].

Since it is hard to store electricity, it is essential to keep balanced voltage of the power grid. Power suppliers are financially penalized for deviations in the grid (i.e., power under/oversupply) the regulatory mechanisms, such as pumped hydroelectric energy storage, must cope with.

*Recommended by thesis supervisor: Assoc. Prof. Viera Rozinajová
To be defended at Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava on [TBA].

Therefore, they need to know ahead how much power will their customers (i.e. their balance group) consume. One way to get rid off power under/oversupply is trading at the short-term electricity market. The power can be sold no less than one hour ahead of its consumption. Online analysis of smart metering data can provide very accurate short-term power demand forecasts. For this purpose, automated resources and methods that consider the power consumption properties, are required. The improvement of current power demand forecasts in Slovakia by 1% can lower regulatory fees by millions of euros.

1.1 Open Problems

Several research areas overlaps in the selected research challenge. Therefore, we look at the open problems from several perspectives – from the more abstract to the specific ones:

- big data [2]:
  - adjustment of existing and design of new big data analysis methods,
  - easy interpretability of results;
- stream mining [11]:
  - automated and adaptive preprocessing,
  - evaluation of adaptive methods,
  - usability in practice;
- power demand forecasting [8]:
  - usability in practice and easy interpretability of results,
  - mapping and comparison of various existing power demand forecasting methods.

Open problems are based on requirements of individual research areas. Big data requirements are for example scalability, timeliness or human cooperation. In stream data mining, the requirements are mostly related to accuracy, time and memory complexity and adaptability to concept drifts. The advancements of forecasting in power engineering depends on rigorous evaluation, understanding of business needs and learning from many disciplines, such as data mining, statistics, meteorology, etc.

1.2 Thesis Goals

The ultimate goal of this thesis was to design a forecasting method that makes short-term forecasts, is adaptive to concept drifts, processes stream data incrementally and forecasts online; and to design a proper evaluation to verify the method's accuracy, ability to adapt to concept drifts and time and memory complexity. We wanted to answer whether incremental and online methods can achieve the accuracy of batch processing, whether incremental method can be still accurate in the presence of concept drifts, how does such method behave during concept drifts, and to what extent can concept drift detection improve the accuracy of forecasting methods.

2. Related Work

In the first part of our thesis, we analyzed the theoretical background of research areas that relate to the goal of our thesis – big data, knowledge discovery from data (KDD) and forecasting. We described the big data life cycle, current big data management tools, and we specified in detail the open problems in this area. We captured the KDD process and the basic data mining tasks, such as data summarization, anomaly detection, associations and patterns mining, prediction and clustering. We focused on prediction and divided the prediction methods by their theoretical foundations (regression, time series analysis, artificial intelligence). In the forecasting overview, we mentioned the division of the forecasting methods to subjective and objective ones, and the measures used for their evaluation.

In the second part, we focused on specific works in stream mining area, power engineering domain and short-term power demand forecasting. Data stream has similar properties as big data – (possibly infinite) volume, velocity and variety (data of various types or data changing over time). Therefore, the data should be analyzed as soon as they are collected. The main requirements for stream processing method are low time and memory complexity, single pass over data, and the ability to adapt to concept drifts. We based the analysis of current stream mining methods on the survey by Gama et al. [7], which defined the three steps of stream mining process (predict, diagnose, update) and divided the stream mining methods by approaches used in four parts of stream processing system diagram: memory, learning, loss estimation and change detection (see figure 1). It is the third step of the process (update) that addresses the need to adapt the prediction model in time. The second step (diagnose) is implemented in the change detection part, which alarms about concept drifts, which require model update. Based on the presence of explicit change detection mechanism, the adaptation can be divided into informed and blind. The blind adaptation updates the prediction model continuously, regularly without respect to the diagnostics.

Stream mining open problems specifies some of the big data open problems, they are related to privacy, variety and mostly to the human cooperation. The restrictions of stream processing (limited time, memory, concept drifts, . . .) suggest that stream methods operate in a more complex way. Despite of this, the methods are supposed to be more understandable to humans and comfortable to utilize in practice. Therefore, the thorough evaluation and analysis of their properties, especially their ability to adapt to concept drifts, are essential.

In power engineering domain, the analysis was aimed at current changes in legislation related to smart metering introduction, the way the power is produced in Slovakia, how is the power demand estimated nowadays, and which factors influence the power demand. From the short-term standpoint, the demand is affected mostly by time of the day, day of the week, holidays and weather, especially temperature. It is sufficient to use historical power

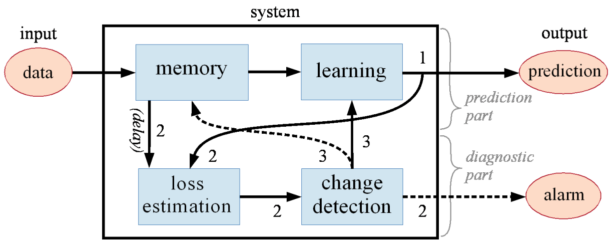


Figure 1: Stream data mining process [7] (1 – predict, 2 – diagnose, 3 – update; solid lines – mandatory steps, dashed lines – optional steps).

load records only for very short-term forecasts (up to 24 hours), because it is assumed that the demand depends only on the recent past and is not changing rapidly. The current trends and technologies in smart metering are subject of a number of both local and global conferences, e.g., *Smart metering*[1], *Energetika*[2], *Energofórum*[3], *Elektroenergetika*[4]. The best power demand forecasting methods are regularly compared in competitons, such as M-Competition[5] and GEFCom[6].

Prediction methods divide into two big groups by their theoretical foundations – statistical and artificial intelligence. The pros and cons of both groups are summarized in table 1. They are often combined into hybrid methods. With the repopularization of neural networks at the beginning of this century, they became the most utilized artificial intelligence method, even in the power demand forecasting. From the adaptability point of view, the blind adaptation with a sliding window is mostly employed, i.e., a new prediction model is regularly trained on data from the sliding window. The training of a new model can be quite time-consuming and not always necessary. Informed adaptation based on concept drift detection can potentially improve prediction accuracy and spare computation and memory resources. Informed adaptation can besides historical power load records consider other data sources, e.g., weather, holidays, consumer behavior, and separate them from the prediction model. This can be advantageous if the other data sources are not available/updated all the time unlike smart metering data. When we analyzed existing power demand forecasting methods, we did not encounter an approach with an informed adaptation. That is why we focused on utilization of such approach in the design of our stream processing prediction method.

3. Datasets

We used smart metering data from two countries: Slovakia and Ireland.

Slovak data were obtained in the project "International Centre of Excellence for Research of Intelligent and Secure Information-Communication Technologies and Systems"[7]. The data included smart measurements from all over the Slovakia from July 1, 2013 to February 16, 2015 (596 days). The measuring frequency was 15 minutes, most of the customers were small and medium enterprises. We divided the data into 19 parts by the region (according to the first two digits of postal codes). We filtered out the consumers without missing measurements and summated the consumption of each region (groups of 100 to 1,300 customers). In the end, we got 19 power load time series. We examined the series and selected four types of concept drift patterns (see figure 2). The data are described in more detail in publication [4].

Irish data came from Smart metering project of Irish regulatory office CER, that happened during 2007 to 2013 and its goal was to perform trials to assess the performance of smart meters, their impact on consumers' energy consumption and the economic case for a wider national rollout. The data contained 30-minute measurements of approximately 5,000 households and small and medium enterprises[8]. We used aggregated power load time series of 3,639 without missing measurements. We chose two test sets: one month (September 20 to October 20, 2009) and six months (july to december 2010) long. These test sets were used in similar papers [13, 18]. The data were normalized to $\langle 0, 1 \rangle$.

Table 1: Pros and Cons of Power Demand Forecasting Methods

	pros	*cons*
statistical	• easy interpretability • few parameters to estimate • better univariate models	• ability to model only linear relationships • low accuracy in longer term • statistical and domain knowledge required
artificial intelligence	• minimum statistical or domain knowledge required • ability to model also non-linear relationships between power demand and exogenous variables (e.g., weather) • better multivariate models	• difficult interpretability • over-parametrization • heavy computation without optimization

4. Incremental Power Demand Forecasting using Error-Driven In formed Adaptation

The adaptive prediction method for stream processing is based on the stream mining process, which consists of three steps: predict, diagnose, update (see figure 3). We chose double seasonal Holt-Winters exponential smoothing (DSWH) [16] as a prediction method, because of its simplicity, robustness and suitability for incremental processing as it is defined recursively. It can also model double seasonality (daily and weekly) of power load time series. DSHW also won in comparison with various prediction methods on European power load data [17].

Informed adaptation is based on monitoring of prediction error in time. The increasing prediction error suggests that prediction model becomes unsuitable for modeling of the incoming observations and it is necessary to update it. To detect changes in prediction error, we used the

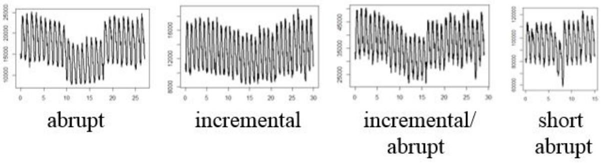


Figure 2: Concept drift types in Slovak power demand time series.

[1] https://konferencie.efocus.sk/konferencia/8.-rocnik-konferencie-smart-metering-smart-grid-nova-energetika-sme-na-nu-p
[2] http://www.power-engineering.sk
[3] http://www.energoforum.sk
[4] http://seen.fei.tuke.sk
[5] https://www.mcompetitions.unic.ac.cy
[6] http://www.drhongtao.com/gefcom
[7] http://ice-rise.sk/
[8] hhttp://www.ucd.ie/issda/data/commissionforenergyregulationcer/

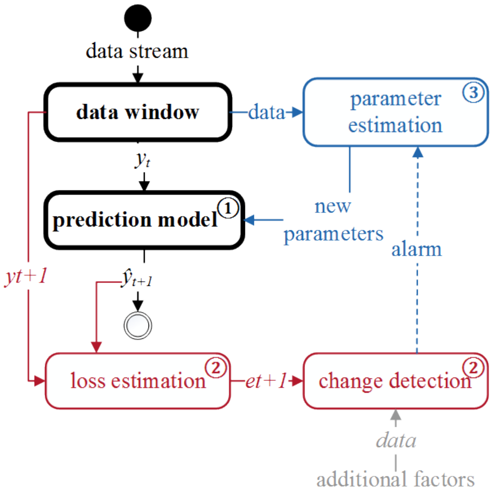


Figure 3: Abstract diagram of adaptive forecasting method consisting of 3 steps: 1 – predict (black), 2 – diagnose (red) and 3 – update (blue). y_t and $\hat{y}_t$ are real and forecasted power load values, e_t is prediction error at time t.

condition defined by equation 1. The change is detected when daily percentage error pe_t (for the last 24 hours, i.e., the last 96 15-minute measurements) exceeds 5%. e_t is prediction error and y_t is power load value in time t. The condition considers the specifics of power engineering domain, specifically the maximum acceptable 5% daily deviation of prediction from the actual power load and the fact that both positive and negative deviations are penalized.

$$pe_t = \frac{|e_{t-95}| + \cdots + |e_{t-1}| + |e_t|}{y_{t-95} + \cdots + y_{t-1} + y_t} \tag{1}$$

When a change is detected, the model is updated. To decrease the computing complexity, we decided not to update the whole model, instead we update only a part of its smoothing coefficients. We kept the coefficients related to daily and weekly seasonality, which were assumed not to change. The new coefficients were estimated from the most current data from the 2-week sliding window.

The precondition of our method is an initial prediction model trained on a longer chunk of data. This model is then continually monitored and updated.

4.1 Evaluation

We evaluated the method by four experiments, which compared the performance of the method on data streams with and without concept drifts, incremental and batch processing, the performance of the method on various types of concept drift patterns with other commonly used methods and the performance of the method on various types of time series (outside power engineering domain). Slovak data were used.

The prediction accuracy of our method was not significantly different when we tested it on data with and without concept drifts. The method worked good in both cases.

The second experiment compared batch and incremental approach with the same prediction method. In the batch approach, the training set (all historical data) was at the end of each day supplemented with measurements from the current day and a new model was trained to predict the next 24 hours. We found out that incremental processing required approximately half the updates the batch approach needed and at the same time its accuracy was not significantly worse. It achieved comparably good accuracy with much smaller amount of training data (2-week sliding window) and updates.

In the third experiment, mean absolute percentage error (MAPE) was measured on four types of concept drift patterns. The results were compared with 8 incremental methods that used blind adaptation (daily model training on data from a sliding window). The comparison is displayed on figure 4. The accuracy of our method was significantly better than accuracy of some blindly adapting methods and it was similar to the accuracy of an ensemble model [4]. Our method worked best on incremental or incremental/abrupt concept drift patterns. The informed adaptation struggled the most with the abrupt concept drift pattern.

The purpose of the last experiment was to find out for which type of time series is our method the most suitable. We used data from M3 competition [14], which contain 1,428 time series with monthly frequency from six domains: micro, industry, macro, finance, demography a other. Our method did not excelled in every domain, however, it worked well for macro time series. It is the most frequent type of time series, e.g., aggregated production, demand, prices data for regions, states, etc. Power demand also belongs to this type.

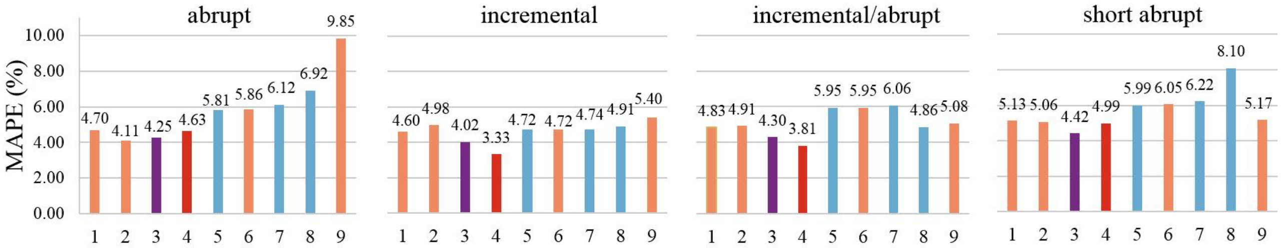
4.2 Discussion

The disadvantage of our method is that after the update of the prediction model, its diagnostics is again possible after 24 hours and great loss can occur during that time if the model update did not improve the accuracy. The prediction model should adapt more quickly and continually. Despite of this disadvantage, we consider the employment of informed adaptation in power demand forecasting as a contribution, since we did not encounter such approach in this domain yet and we showed its potential to achieve similar results as the batch approaches and to spare computing resources. Another challenge could be the investigation of detected concept drifts and their connection to events that might cause them. Because of the mentioned disadvantage, we focused on online methods in the next part of the thesis.

5. Smart Grid Load Forecasting Using Online Support Vector Regression

Online stream processing, unlike incremental stream processing, processes data one-by-one without the necessity of their storage in a database or a sliding window. From the adaptation standpoint, the model changes continually, after processing of each instance, without change detection mechanism, i.e., blindly.

As a prediction method, we chose support vector regression (SVR). Its short-term forecasts are very accurate (also in power demand forecasting) and it often outperforms neural networks [1, 9, 10, 15, 18]. SVR tries to find



methods: 1– random walk, 2 – seasonal and trend decomposition by local regression+ARIMA, 3 – ensemble model [4], 4 – informed DSHW, 5 – random forest, 6 – multiple linear regression, 7 – SVR, 8 – multilayer perceptron, 9 – DSHW

Figure 4: Comparison of incremental adaptive DSHW (no. 4) with 8 incremental blindly adaptive methods on 4 types of concept drifts. Blue columns are AI-based methods, the violet one is an ensemble model, the rest are statistical methods.

such function with ε margin, so most of the instances are covered (modeled) by it. Instances that lay on the edge of the margin are called support vectors. The distance of instances outside the margin (outer vectors) from the margin is penalized by cost C. Based on these principles we can define conditions for each group of vectors in the SVR model: inner, support and outer vectors. If a function with a margin can not be found, a kernel trick (function) is used. The instances are transformed into a space with more dimensions where such function exists. One flaw of SVR is its inability to forecast more than one value ahead. To forecast a longer horizon, multiple simultaneous SVR models must be used. To estimate suitable SVR parameters (ε, C and parameters of kernel function), biologically inspired optimization algorithms are utilized.

Online version of SVR has been already published in 2000 [5], but even nowadays it is not well-known, mostly because there is no standard library implementation. The method tries to assign each new instances (vector) to one of three vector groups so it satisfies the group's conditions. If it is not possible, the vectors in model migrate between groups until there is found a space for the new vector. There are three types of allowed migrations between groups (see figure 5). The maximum number of vectors in model is limited by a *threshold*. The algorithm of online SVR training is stated below (see algorithm 1).

5.1 Evaluation

We evaluated the method by six experiments. We compared its accuracy with standard SVR, examined its accuracy when various kernel functions were used, evaluated accuracy of very short-term and short-term forecast, compared its accuracy with selected traditional prediction methods, explored the optimization of its parameters by biologically inspired optimization algorithms, and evaluated its computing and memory complexity. We used Irish data in these experiments.

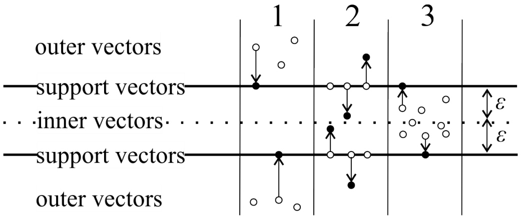


Figure 5: Geometric representation of possible vector migrations. 1 – from outer to support, 2 – from support to inner or outer, 3 – from inner to outer vectors.

Algorithm 1: Online SVR training.
Vector removal (2-5) and vector addition (6-12).

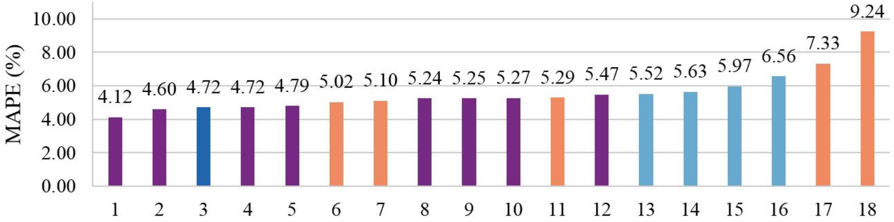
```
1  if model contains threshold number of vectors then
2    until weight of the oldest vector is not minimum do
3      find vector with minimum weight in model
4      update (migrate) the vector
5    remove the oldest vector

6  add new vector to SVR model
7  if it is an inner vector then
8    end of action
9  else
10   until the vector is not support or outer vector do
11     find minimum needed updates (migrations) of
         vectors in model
12     update model vectors
```

We found out that online SVR has the same accuracy as standard (batch) SVR. The best results for power demand forecasting are achieved when radial basis function is employed as the kernel function.

In the third experiment, we forecasted 30 minutes and one hour ahead for one month and for six months. We evaluated mean absolute percentage error (MAPE) and found out that the error lowers with increasing number of vectors in model, i.e., the longer the model forecasts, the more accurate it is. We achieved similar results (MAPE approx. 2.5%) as a similar existing method [18].

The fourth experiment was aimed at one-day ahead forecast for six months. We used 48 simultaneous online SVR models (one for each half-hour). The forecast error was compared with other 10 methods that utilized a sliding window. We also tried separate workday/weekend models for 7 methods. The results are shown on figure 6. Online SVR method achieved the third best accuracy. The best methods were based on ensemble learning (random forest, bagging, extremely randomized trees) and contained hundreds or thousands of decision trees. Quite good accuracy was achieved by statistical methods based on time series analysis that considered double seasonality.

We compared online SVR to a similar work that used the same one-month test set [13]. OS-ELM method clustered similar consumers at first, then forecasted the consump-



methods: 1 – random forest* , 2 – bagging, 3 – online SVR, 4 – extremely randomized trees*, 5 – bagging*, 6 – DSHW, 7 – seasonal and trend decomposition by local regression (STL)+ARIMA*, 8 – extreme gradient boosting*, 9 – extreme gradient boosting, 10 – extremely randomized trees, 11 – STL+exp. smoothing*, 12 – random forest, 13 – SVR*, 14 – SVR, 15 – deep learning, 16 – multilayer perceptron, 17 – STL+ARIMA, 18 – STL+ exp. smoothing

**separate workday/weekend model*

Figure 6: Comparison of online SVR (no. 3) with 10 forecasting methods with sliding window, one-day ahead forecast on 6-month test set. Blue columns are AI-based methods, the violet ones are ensemble learning models, the rest are statistical methods.

tion of each cluster and finally summated the forecasts. It utilized weather data in prediction as well. Average and maximum MAPE of the method was 2.47% and 4.21%. The results of our method were 3.00% and 4.01%. On average, online SVR was worse, but its error did not increased over time dramatically and behaved quite stable.

To estimate the parameters of SVR (ε, C and parameters of radial basis function) we employed particle swarm optimization (PSO). In the fifth experiment, we examined whether the accuracy can be further improved by utilization of other optimization method. We chose cuckoo optimization algorithm (COA). We let the models with optimized parameters forecast for five times on one-month test set. We discovered that parameters optimized by COA significantly lowered the forecast error (by 0.02%). The time of optimization by both of the algorithms was almost identical.

The last experiment evaluated the time and memory complexity of online SVR and compared it to standard (batch) SVR from LIBSVM library [6]. We found out that the time complexity of both methods is $O(n^2)$ to $O(n^3)$. Measured training times on one-month long test set are shown on figure 7. The duration of a new vector addition to the model depends on the number of vectors that have to migrate in order to find place for the new vector. If the new vector belongs to the inner vectors, the training time is very low (see the light blue curve in the bottom of upper graph in figure 7). If the maximum number of vectors in model is limited by a *threshold*, training time oscillates around mean value and the ascending trend stops. The training time also depends on the length of the vectors – the longer the vectors, the longer the training time. On the other hand, the forecast error of SVR model with longer vectors and higher *threshold* was lower. Memory complexity of online SVR depends on the number of vectors in model (i.e., on *threshold*). To create a new vector it is also necessary to keep in memory a sliding window of size of a vector. Batch SVR needs to keep in memory *threshold* number of observations, since it is always trained on all historical data from scratch.

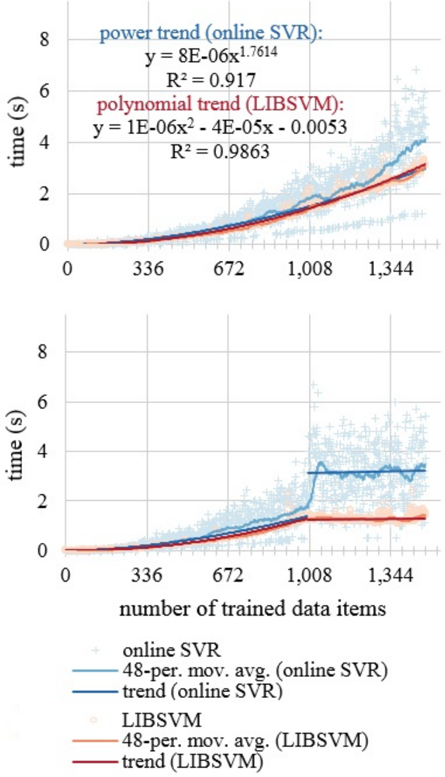


Figure 7: Comparison of online and batch (LIBSVM) SVR training time. Measured times, 48-period moving average and trendline. *threshold* = 2000 **(up) and** *theshold* = 1000 **(down).**

5.2 Discussion

The disadvantage of online SVR is that if we want to forecast a longer horizon at once, a separate SVR model must be used for each forecasted period, because the method has only one output. Iterative strategy or reformulation of the SVR model to multiple outputs, e.g., [3], could improve its accuracy. Other possible improvements leading to lower forecast error are inclusion of other data souces (e.g., weather), a kernel function designed specially for our data, usage of other optimization algoritms for SVR parameters estimation or clustering of consumers into similar groups before forecasting.

6. Conclusion

In this thesis, we analyzed areas related to predictive analysis on data streams, such as big data, knowledge discovery from data, forecasting or stream mining. The thesis is set in the power engineering domain and we focused on power demand forecasting. Our task was to create a short-term forecasting method of power demand of a group of consumers. The main motivation was the fact that inaccurate forecasts cause big deviations in power grid, which is damaged by them, and power suppliers are financially penalized for causing deviations by a regulatory office. The improvement of forecasts in Slovakia by 1% could lower the regulatory fees by millions of euros. The introduction of smart meters caused that much more data is generated than it was available until now, and therefore provided data for more accurate forecasts. That is why there emerged a need for transition from standard batch data processing to stream processing.

Data stream has similar properties as big data – *volume*, *velocity*, *variety.* According to these properties, we identified the main requirements for stream prediction method, i.e., *accuracy*, *timeliness* and *adaptability.* We put emphasis on *usability* of the method in practice, which is one of the very often mentioned open problem of big data and stream processing. In power engineering, these properties project in endless streams of measurements from a large number of smart meters. Regularly incoming power load values are influenced by consumers' behavior, weather, social events, etc. Requirements for power demand forecasting method are daily prediction error under 5%, known forecasts at least one hour ahead for electricity market trading purposes, and consideration of concept drifts in power load. Our task was one-day ahead power demand forecasting for a group of customers of one power supplier, i.e. a balancing group. Measurements from individual consumers were summated into one time series.

We looked at stream forecasting methods mainly from two standpoints – *data processing* and *adaptation technique.* We focused on incremental and online processing, which, unlike the standard batch processing, does not have to keep all historical data in memory. The most utilized approach is *incremental learning with partial memory*, usually a sliding window. After each slide of the window, the prediction model is trained on data from the window. Therefore, the model develops in time and accomodates some degree of adaptation. Since the adaptation happens independently from the changes in data (concept drifts), it is also called *blind adaptation.* On the other hand, *informed adaptation* is based on explicit change detection, otherwise it does not happen at all. This way of adaptation can be more suitable in cases, when concept drifts are not very frequent or distinct and regular blind adaptation would not bring anything new to the prediction model.

In this thesis we studied two forecasting methods – *incremental with informed adaptation* and *online with blind adaptation.*

We found out that informed adaptation needed significantly less computing resources than batch processing with blind adaptation and the accuracy did not significantly lower. This method is suitable for flow time series that measure activity over time, e.g., macroeconomic time series of aggregated variables, such as unemployment, industry production, export, import, etc. The time series of power demand of a larger group of consumers also belongs to this category. The disadvantage of informed adaptation was the required longer pause between two consequent concept drift detections. That is why we decided to explore online approach, which processes data in one-by-one fashion without keeping them in memory.

We showed that online forecasting method with blind adaptation is suitable for power demand forecasting and achieves the accuracy of ensemble models, which have in general much higher computing complexity. We evaluated its properties, pros and cons by a series of tests and our findings are further applicable in other application domains. At the same time, we compared a wider spectrum of power demand forecasting methods and provided a self-contained view on performance of forecasting methods in this area. We encountered in literature only a few similar reports, which compared more prediction methods, usually on private data sets.

Both of our approaches accomplished the daily forecast error under 5% and the accuracy of batch forecasting methods. Their design is aimed at constant computing resources, which is an important aspect in big data processing, low number of method parameters, easy understandability and interpretability. We addressed the open problems we mentioned at the beginning: the need for transition from batch to stream processing, usability in practice and evalution of stream mining methods.

In future, we want to focus on possible improvements of our approaches we mentioned in discussion sections of this paper, or on brand new forecasting methods with promising results, such as approaches based on analysis of patterns, which occur repeatedly in time series. These approaches can model well the various types of consumers or various concept drift patterns in power load. We would like to focus also on other smart grid components, e.g., photovoltaic panels, batteries, electric cars, which bring new open problems related to their optimal settings, localization, microgrid establishment, etc. The road to the future vision of smart grid must yet run through many legal, practical and research challenges.

Acknowledgements. This contribution was created with the support of the Research and Development Operational Programme for the project "International Centre of Excellence for Research of Intelligent and Secure Information-Communication Technologies and Systems", ITMS 26240120039, co-funded by the ERDF; and the Slovak Research and Development Agency under the contract No. APVV-16-0213.

References

[1] R. Achanta. Long term electric load forecasting using neural networks and support vector machines. *Int. J. Computer Science and Technology (IJCST)*, 3(1):266–269, 2012.

[2] D. Agrawal, P. Bernstein, E. Bertino, S. Davidson, U. Dayal, M. Franklin, J. Gehrke, L. Haas, A. Halevy, J. Han, H. Jagadish, A. Labrinidis, S. Madden, Y. Papakonstantinou, J. Patel, R. Ramakrishnan, K. Ross, C. Shahabi, D. Suciu, S. Vaithyanathan, and J. Widom. Challenges and opportunities with big data. Technical report, Computing Community Consortium, Washington, D.C., USA, 2012.

[3] Y. Bao, T. Xiong, and Z. Hu. Multi-step-ahead time series prediction using multiple-output support vector regression. *Neurocomputing*, 129:482–493, 2014.

[4] A. Bou Ezzeddine, M. Lóderer, P. Laurinec, P. Vrablecová, V. Rozinajová, M. Lucká, P. Lacko, and G. Grmanová. Using biologically inspired computing to effectively improve prediction models. *Int. J. Hybrid Intelligent Systems*, 13(2):99–112, 2016.

[5] G. Cauwenberghs and T. Poggio. Incremental and decremental support vector machine learning. In T. K. Leen, T. G. Dietterich, and V. Tresp, eds., *NIPS'00 Proc. 13th Int. Conf. Neural Information Processing Systems*, pp. 388–394, Cambridge, MA, USA, 2000. MIT Press.

[6] C.-C. Chang and C.-J. Lin. LIBSVM. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–27, 2011.

[7] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia. A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4):1–37, 2014.

[8] T. Hong. Energy forecasting : Past, present and future. *Foresight: The Int. J. Applied Forecasting*, 2014(32):43–48, 2014.

[9] Z. Hu, Y. Bao, and T. Xiong. Electricity load forecasting using support vector regression with memetic algorithms. *The Scientific World J.*, 2013:10, 2013.

[10] S. Humeau, T. K. Wijaya, M. Vasirani, and K. Aberer. Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households. In *2013 Sustainable Internet and ICT for Sustainability (SustainIT)*, pp. 1–6, Laxenburg, Austria, 2013. IFIP.

[11] G. Krempl, M. Spiliopoulou, J. Stefanowski, I. Žliobaitė, D. Brzeziński, E. Hüllermeier, M. Last, V. Lemaire, T. Noack, A. Shaker, and S. Sievi. Open challenges for data stream mining research. *ACM SIGKDD Explorations Newsletter*, 16(1):1–10, 2014.

[12] D. Laney. 3D data management: Controlling data volume, velocity and variety. Technical report, META Group, Stamford, CT, USA, 2001.

[13] Y. Li, P. Guo, and X. Li. Short-term load forecasting based on the analysis of user electricity behavior. *Algorithms*, 9(4):80, 2016.

[14] S. Makridakis and M. Hibon. The M3-competition: results, conclusions and implications. *Int. J. Forecasting*, 16(4):451–476, 2000.

[15] J. C. Sousa, H. M. Jorge, and L. P. Neves. Short-term load forecasting based on support vector regression and load profiling. *Int. J. Energy Research*, 38(3):350–362, 2014.

[16] J. W. Taylor. Short-term electricity demand forecasting using double seasonal exponential smoothing. *J. Operational Research Society*, 54:799–805, 2003.

[17] J. W. Taylor and P. E. McSharry. Short-term load forecasting methods: An evaluation based on European data. *IEEE Trans. Power Systems*, 22(4):2213–2219, 2007.

[18] T. K. Wijaya, M. Vasirani, S. Humeau, and K. Aberer. Cluster-based aggregate forecasting for residential electricity demand using smart meter data. In H. Ho, B. C. Ooi, M. J. Zaki, X. Hu, L. Haas, V. Kumar, S. Rachuri, S. Yu, M. H.-I. Hsiao, F. L. Jian Li, S. Pyne, and K. Ogan, eds., *2015 IEEE Int. Conf. Big Data (Big Data)*, pp. 879–887, Piscataway, NJ, USA, 2015. IEEE.

[19] Združenie Dodávateľov Elektriny. Posúdenie výhodnosti zavedenia inteligentných meračov elektriny v podmienkach SR [Examination of profitability of smart meters' roll-out in Slovak Republic's conditions]. Technical report, Accenture, Bratislava, Slovakia, 2012.

Selected Papers by the Author

P. Vrablecová, A. Bou Ezzeddine, V. Rozinajová, S. Šárik, and A. K. Sangaiah. Smart grid load forecasting using online support vector regression. *Computers & Electrical Engineering*, 65:102–117, January 2018.

P. Vrablecová, V. Rozinajová, and A. Bou Ezzeddine. Incremental adaptive time series prediction for power demand forecasting. In Y. Tan, H. Takagi, and Y. Shi, eds., *Data Mining and Big Data*, volume 10387 of *LNCS*, pp. 83–92, Cham, Switzerland, 2017. Springer.

P. Vrablecová, V. Rozinajová, and A. Bou Ezzeddine. Incremental time series prediction using error-driven informed adaptation. In C. Domeniconi, F. Gullo, F. Bonchi, J. Domingo-Ferrer, R. Baeza-Yates, Z.-H. Zhou, and X. Wu, eds., *2016 IEEE 16th Int. Conf. Data Mining Workshops (ICDMW)*, pp. 414–421, Piscataway, NJ, USA, December 2016. IEEE.

V. Rozinajová, A. Bou Ezzeddine, M. Lóderer, J. Loebl, R. Magyar, and P. Vrablecová. Computational intelligence in smart grid environment. In A. K. Sangaiah, Z. Zhang, and M. Sheng, eds., *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*, chapter 2, pp. 23–59. Elsevier, Cambridge, MA, USA, 1st ed., 2018.

G. Grmanová, P. Laurinec, V. Rozinajová, A. Bou Ezzeddine, M. Lucká. P. Lacko, P. Vrablecová, and P. Návrat. Incremental ensemble learning for electricity load forecasting. *Acta Polytechnica Hungarica*, 13(2):97–117, February 2016.

G. Grmanová, V. Rozinajová, A. Bou Ezzeddine, M. Lucká. P. Lacko, M. Lóderer, P. Vrablecová, P. Laurinec. Application of biologically inspired methods to improve adaptive ensemble learning. In N. Pillay, A. P. Engelbrecht, A. Abraham, M. C. du Plessis, V. Snášel, and A. K. Muda, eds., *Advances in Nature and Biologically Inspired Computing*, volume 419 of *Advances in Intelligent Systems and Computing*, pp. 235–246, Cham, Switzerland, 2016. Springer.

P. Laurinec, M. Lóderer, P. Vrablecová, M. Lucká, V. Rozinajová, and A. Bou Ezzeddine. Adapptive time series forecasting of energy consumption using optimized cluster analysis. In C. Domeniconi, F. Gullo, F. Bonchi, J. Domingo-Ferrer, R. Baeza-Yates, Z.-H. Zhou, and X. Wu, eds., *2016 IEEE 16th Int. Conf. Data Mining Workshops (ICDMW)*, pp. 398–405, Piscataway, NJ, USA, December 2016. IEEE.

A. Bou Ezzeddine, M. Lóderer, P. Laurinec, P. Vrablecová, V. Rozinajová, M. Lucká, P. Lacko, and G. Grmanová. Using biologically inspired computing to effectively improve prediction models. *Int. J. Hybrid Intelligent Systems*, 13(2):99–112, April 2016.