

DOI: 10.18721/JCSTCS.14302
УДК 004

DYNAMIC ENERGY CONSUMPTION RATIONING BASED ON MACHINE LEARNING ALGORITHMS FOR OIL REFINING TASKS

N.S. Kudriashov

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

Energy consumption rationing is necessary for high-quality production planning, and allows optimizing their use. This paper provides an analysis of various approaches to building a model of energy consumption, describes their limitations and new approaches to dynamic rationing. As the object of modeling the ELOU-AVT-6 (CDU/VDU-6) unit has been taken. Such units are intended for desalination and primary fractionation of oil. Functional requirements for the algorithms have been formed, based on real production needs. As the solution, models based on machine learning algorithms have been analyzed. These algorithms include CatBoost Regressor, Gradient tree boosting, Random Forest, ElasticNet and artificial neural networks. The analysis of the modeling results and comparison of the accuracy of the models is carried out. The paper also demonstrates a scenario of using a dynamic rationing model to analyze the causes of deviations of the actual consumption values from the planned ones.

Keywords: energy consumption rationing, machine learning, digital twin, oil refining, factor analysis.

Citation: Kudriashov N.S. Dynamic energy consumption rationing based on machine learning algorithms for oil refining tasks. Computing, Telecommunications and Control, 2021, Vol. 14, No. 3, Pp. 20–32. DOI: 10.18721/JCST-CS.14302

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ДИНАМИЧЕСКОЕ НОРМИРОВАНИЕ ПОТРЕБЛЕНИЯ ЭНЕРГОРЕСУРСОВ ДЛЯ ЗАДАЧ НЕФТЕПЕРЕРАБОТКИ НА ОСНОВЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Н.С. Кудряшов

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Нормирование потребления энергоресурсов необходимо для качественного планирования производства и позволяет рационализировать их использование. В статье приведен анализ различных подходов к построению модели потребления энергоресурсов, определены их недостатки и представлен новый подход к динамическому нормированию. В качестве объекта моделирования рассмотрен процесс суммарного потребления топлива для установки ЭЛОУ-АВТ-6, предназначенной для обессоливания и первичного фракционирования нефти. Сформированы функциональные требования к разрабатываемым алгоритмам, исходя из актуальных задач, диктуемых производством. В качестве решения рассмотрены модели на основе алгоритмов машинного обучения, такие как Catboost регрессор, Градиентный бустинг деревьев, Случайный лес, ElasticNet и искусственные нейронные сети. Проведен анализ результатов моделирования и сравнения точности моделей. Продемонстрирован сценарий использования модели динамического нормирования для анализа причин отклонения фактических значений потребления от плановых.

Ключевые слова: нормирование потребления энергоресурсов, машинное обучение, цифровой двойник, нефтепереработка, факторный анализ.

Ссылка при цитировании: Kudriashov N.S. Dynamic energy consumption rationing based on machine learning algorithms for oil refining tasks // Computing, Telecommunications and Control. 2021. Vol. 14. No. 3. Pp. 20–32. DOI: 10.18721/JCSTCS.14302

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

High speed of technological growth and global trends towards the production digitalization and Industry 4.0 concepts dictate an increasing volume of requirements for the industry. As an illustrative example, the oil refining industry poses a task of active integration of new technologies to optimize technological processes, increase the quality of the final product and reduce its production costs. Consequently, in the last 10 years engineers have begun developing so-called digital twins of technological processes. A digital twin is a complex program entity based on accurate process models, statistical data, regulatory values and machine learning algorithms [1, 2].

Rationing of energy consumption is not an exception of such digitization processes. On the one hand, a digital twin of production allows us to implement various scenarios, such as: forecasting, retrospective analysis, simulation of various operating conditions of equipment, etc. On the other hand, the consumption rationing model should allow us to calculate the necessary and sufficient resource consumption for a particular scenario. The ability to analyze the actual deviations in energy consumption is a key feature, which can lead us to more rational cost management.

One of the key issues of building a model for the regulation of energy consumption, is a large number of influencing external and internal factors. Such factors impact often cannot be detected analytically. These factors include the parameters of the technological process, the parameters of raw input materials, meteorological conditions, time interval etc. All existing methods and algorithms for the regulation of energy consumption are often based on some sort of regulatory values and do not allow them to be effectively applied in such digitized scenarios, described previously [3].

The main goal of this work is to develop an approach and determine the methods of developing the models for dynamic rationing of energy consumption using the example of oil refining problems. For this case, the existing approaches to energy consumption rationing were analyzed and new methods, based on machine learning algorithms, were developed and approved. This new approach provides the ability to dynamically recalculate the rates of energy consumption within the process or environment changes.

Energy consumption rations

Energy consumption rations are the calculated values that characterize the maximum allowable expenditure of certain resources. At the same time, during the calculation, the operating conditions of the technological equipment and the environment at a particular moment of time are taken into account. The norms determine the calculation basis for planning the consumption of fuel and energy resources, and also allow you to control their expenditure and identify any potential saving reserves [4].

For industrial consumers, ration is an indicator of the planned consumption of some resources for the production of a unit of final product. There are two key groups of resources used as a subject of rationing: the main resources and the operation of production facilities.

In this paper, the rations of resources consumption, associated only with the provision of the main production process (fuel and energy resources) were considered. Therefore, we focus our attention on them. These values are evaluated in a form of a generalized indicator, expressed in Tones of Equivalent Fuel (TEF) per unit of production.

Let's consider the key existing methods for calculating consumption rations [5]:

- Experienced method;
- Computational and analytical method;
- Computational and statistical method.

Experienced method is, as the name suggests, an experiment, the results of which are used to form individual rations. A significant disadvantage of this approach is that the object of modeling must perform its work exclusively within the regimes provided at the stage of its testing [6]. Also, the process of such development norms entails large labor costs. It consists of methodology development, testing, analysis of results, etc. All this makes it difficult to replicate the solution. Another disadvantage is that applying this method makes it impossible to quickly revise the existing consumption rations due to changes in the process parameters.

Another common method, that allows us to ensure high accuracy of rationing, is the computational and analytical method. This method is based on a thorough study of technical regulations and design and engineering documentation [7]. During such process of rations calculation, a sequential division of the modeling object into separate aggregates is performed. After this, the engineer analyses such aggregates interaction. A significant disadvantage of this method is that the quality of the developed values is directly proportional to the quality and accuracy of the description of the object in the technical documentation. Thus, this approach may not take into account the real state of the object, which will entail a decrease in the quality of energy consumption rationing [8].

In turn, the computational and statistical method allows you to deal with the problem of inconsistency between the state of the object and its technical description. It provides the determination of rations, based on the reports data of the actual consumption of fuel and energy resources during the past periods. In other words, when applying this method, the values are interpolated by forming a function that characterizes the relationship between operating conditions and the amount of energy consumption [9].

For this work, during the building of the machine learning models, we employed the latter method first, and in particular the so-called analytical model. Mathematical formulation can be written as follows.

The process of energy consumption can be described as:

$$Y = f\left(X_1^{(in)}(T_1), \dots, X_N^{(in)}(T_N), X_1^{(out)}(T_{(N+1)}), \dots, X_K^{(out)}(T_{(N+K)}), \tau, \dot{X}\right) + \varepsilon(\tau),$$

where $X^{(in)}$ – interior industrial parameters values; $X^{(out)}$ – outer parameters values; T – different times-tamps, depending on the parameter; τ – current datetime value; \dot{X} – parameters, which are not kept in manufacture; ε – measurement error.

According to the amount of data we can use – our rationing model should have such mathematical description:

$$N(W) = f\left(\left(X_1^{(in)}(T_1), \dots, X_N^{(in)}(T_N), X_1^{(out)}(T_{(N+1)}), \dots, X_K^{(out)}(T_{(N+K)})\right), \tau\right).$$

As we see, the main goal of such model development is decreasing of the number of \dot{X} and to minimize the difference between such model values. In this case the mathematical description of the model development task will be as follow:

$$E = \arg \min_W \left(J(Y, N(W)) \right),$$

where J is an error function, more detailed description will be provided when it is referred to in the next paragraphs.

Another task for the development is the approach of choosing the $X^{(in)}$ and $X^{(out)}$ according to the analytical model.

One of the key advantages of this approach is its universality and applicability for various objects. Also, this approach is more flexible compared to its alternatives, described earlier.

This approach is based on the analytical construction of the energy consumption model. Therefore, its disadvantages include a high probability of a mistake in description of the nature of the function, the volume of which is directly proportional to the complexity of the modeled object.

The idea of building an analytical model is consonant with the task of machine learning. In the classical approach, at the stage of exploratory data analysis, we can involve specialists in energy consumption to form a list of factors, based on their knowledge about the modeling object [10]. However, this approach entails high labor costs. Thus, we have set the goal of reducing the use of knowledge about the subject area with no loss in quality: it will serve as a guarantee of successful replication of the developed approach.

Modeling object – CDU/VDU-6 unit

As it has been already mentioned, in this paper we took the oil refinery process as the object for energy consumption modeling. Modern oil separation involves piping crude oil through a sequence of hot furnaces. The resulting liquids and vapors are discharged into distillation units. Such products of the plant operation include diesel fuel, tar, fuel oil, kerosene, gasoline, etc.

In particular, we analyzed the data from the unit built according to the standard design called CDU/VDU-6. The feed-stock for the unit is crude oil coming from oil pumping stations.

To illustrate the refinery process, including CDU and VDU, Fig. 1 presents a simplified oil refinery schematic [11].

As it is shown in the figure, CDU/VDU unit separates crude oil into different products by boiling point differences and prepares feed for secondary processing units. Two main units for the CDU/VDU are CDU, an electrical desalting and oil dehydration unit, and the VDU, an atmospheric and vacuum distillation of oil unit.

During the process, the following products are obtained at this unit: fuel and liquefied gas, straight-run gasoline fractions, a fraction of straight-run diesel fuel, vacuum gas oil and tar. In addition, in the process of heat recovery, water vapor generates, which, looking ahead, is one of the targets predicted in this work.

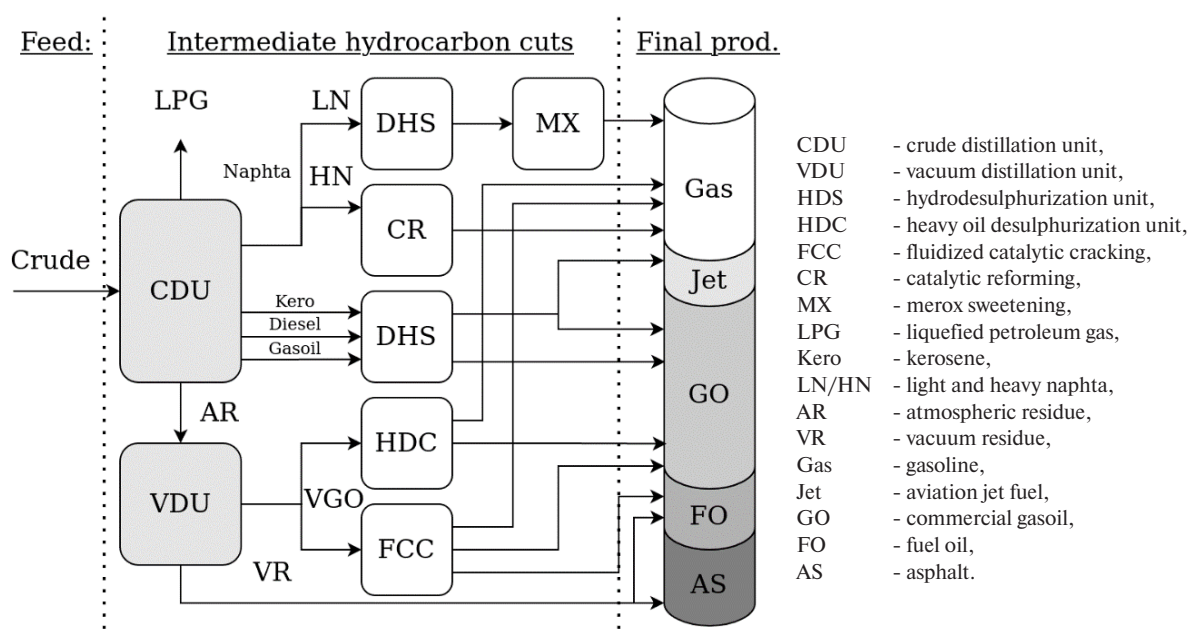


Fig. 1. Simplified schematic of an oil refinery

CDU/VDU-6 includes a gasoline stabilization unit and an intermediate tank farm for receiving, storing and dispensing raw materials of 6,000 cubic meters.

At the moment, for the modeled unit, an experimental method of rationing is used. The selection of norms is carried out from the corresponding reference book of norms with respect to the current mode of operation, month, and is multiplied by the actual load for raw materials. Such values are revised annually.

Such approach to rationing leads to ineffective plant management in terms of energy consumption. This is due to the low accuracy of the calculated consumption rates, since they do not take into account any real technological parameters, equipment degradation, environmental parameters, etc. Moreover, this approach does not allow us to analyze the factors and reasons of deviations in consumption from the calculated consumption rates. Such process is called factor analysis and it requires a more interpretative and transparent process of energy consumption.

Rationing model and digital twin

One of the main goals of using the dynamic rationing model is to ensure correct, well-grounded management of the technological process. This, in turn, can be achieved by the timely identification of deviations of the actual consumption from the planned one. Such approach would allow us to form operative corrective actions to improve the control quality.

To achieve this goal, it is necessary to establish a close interaction of the dynamic rationing model and a digital twin of the technological process [12]. Fig. 2 provides a more detailed description of such behavior.

In fact, there are two different scenarios of using the dynamic rationing models. The first one of them is the calculation of the planned y'' energy consumption rations, which are used as tasks for the plant operators. On the other hand, Fig. 2 shows a comparison between estimated consumption rations y' and actual consumption y'' . After reaching the point in time for which the consumption rate was planned, we can calculate its actual model value. This value should be close to the actual consumption of fuel and energy resources, otherwise, this algorithm will not be effective.

The algorithm of the deviation causes analysis (factor analysis) should demonstrate the degree of influence of each parameter on the total difference between planned and actual consumption. This approach allows us to carry out the retrospective analysis. With the knowledge we get from such analysis, we can adjust the operating plan in order to minimize expenses and costs.

To implement the algorithm for analysis of the causes of deviations, the dynamic rationing model must be interpretable. We must be able to numerically assess the degree to which each parameter affects the result. Such coefficients should be normalized in relation to the difference in consumption. Resulting delta values will characterize the fraction of delta justified by this or that parameter [13].

Thus, the main functional requirements for the development of the dynamic consumption rationing model are:

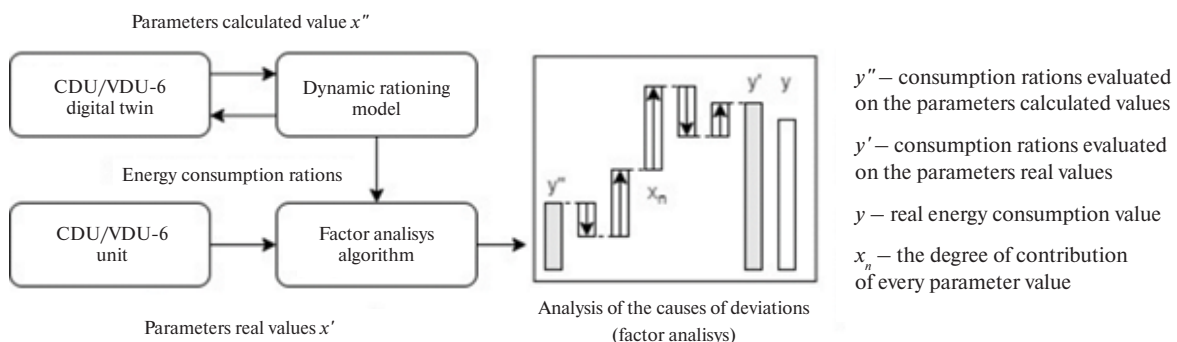


Fig. 2. Rationing model for the factor analysis issues

- interpretability of models or the possibility of using algorithms for analyzing the causes of deviations;
- the standardization error should not exceed 2 %, which is due to the absolute error of measuring devices for fuel and energy resources consumption;
- adaptability of the solution: the ability to revise consumption rates when critical changes in the process are detected;
- the model should take into account the degree of inertia of the process, since some parameters can contribute to consumption with different delta over time.

Initial modeling data

To start the dynamic rationing model development, it is necessary to assess the quality of the existing data and prepare them for transfer to the algorithm. For machine learning algorithms bad quality can be crucial, so we need to preprocess the data. Collecting, storage and primary generalization of data for the CDU/VDU-6 unit is carried out by PI Systems utilities. At the first step, ETL procedure was set up to download the data from the server.

For the analysis, a data interval of one and a half year was selected. Such choice is justified by the necessity of having a volume of data covering the entire annual cycle of the installation, as well as some reserve of data for validation and testing. Downloaded data contains technological information (unit temperatures, pressures, etc.), production data (plant load, quality characteristics), as well as environmental parameters (temperature, wind directions, etc.). The total fuel consumption was taken as an energy resource for rationing modeling.

The process of the control and monitoring of energy consumption is implemented in the context of hourly averaged values. Thus, the dynamic rationing model should also form the hourly average indicators of energy consumption. However, the data from the source were downloaded with a sampling rate of 1 minute, cleaned and averaged within an hour. This is due to the fact, that the data can contain a large number of anomalies. Such anomalies can be related to some functional failures of measuring instruments, breakdowns and production interruptions. In this case, it is necessary to filter the data somehow. For these purposes, the EllipticEnvelope algorithms [14], a high-pass filter and a moving average filter, were applied to achieve the best anomaly detection efficiency:

$$\hat{Y} = \theta(Y) * \frac{1}{60} \sum_{i=60} \vartheta(Y_{-i}),$$

where \hat{Y} – the resulting value of hourly consumption; θ – EllipticEnvelope operation giving [0, 1] values; ϑ – high-pass filter operation.

Fig. 3 shows an example of the preprocessed fuel consumption data.

Also, during the preparation, the data were preprocessed in order to reduce their volume with minimal loss in quality. Unfortunately, the use of dimensionality reduction methods (Principal component analysis, tSNE, etc.) is not further allowed for data interpretability, otherwise it would significantly complicate such an algorithm.

In such situation, to reduce the amount of data, a correlation analysis was carried out. The Spearman and Pearson coefficients were analyzed and the values were discarded according to the following criteria:

- parameters, the correlation coefficient of which with the target is more than 0.9, in order to avoid target leaks;
- to combat multicollinearity, all values with cross-correlations greater than 0.9 were also discarded.

Finally, the initial data volume contained 1083 parameters with 780 thousand values each. However, after their preprocessing, they were reduced to 562 parameters of 3 thousand averaged hourly values. For this publication, the data have been anonymized and replaced with tag_n for parameters and y for consumption not to violate the contribution rules. All the data were scaled from 0 to 1 for the same reason.

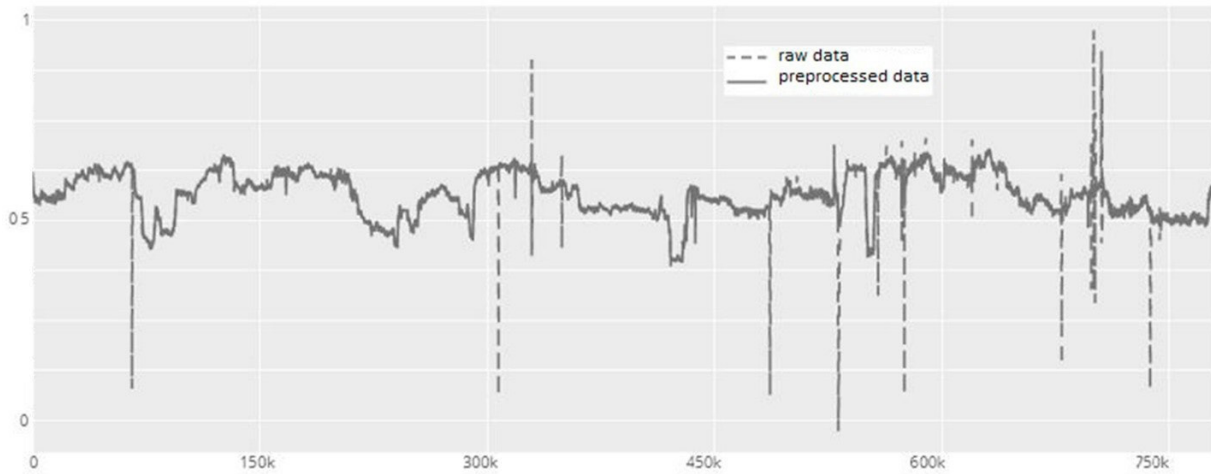


Fig. 3. The results of the consumption data preprocessing

Dynamic energy consumption rationing model

At the very beginning of the model development, the prepared dataset was divided into training, validation and test samples. The volume of the training dataset is 12 months of operation of the unit and the inner processes. Such dataset is used to train the model. The validation set serves to determine the generalizing ability of the model and is used to select its parameters. The validation process is based on the TimeSeries-Split cross-validation method with a sliding window of 1 month [15]. A test dataset with the length equal to 1 month is used to evaluate the simulation results.

Also, one of the important features of the modeled process is that some parameters can affect power consumption with some time delta. To handle such time deltas, we inserted a time lag parameters $\text{tag_n}^{(tk)}$ into our model. For this work, we have considered k value equal to 8, which implies that the maximum time lag can be 8 hours. Such delay is caused by the characteristics of the process, but have to be adjusted in future work.

In this research, the following machine learning approaches to the construction of a dynamic normalization model were considered [16]:

- linear models;
- models based on tree boosting algorithms;
- one-dimensional convolutional neural networks.

The Linear Regression algorithm is the most obvious and simplest solution to such problems. It is usually applied to create the baseline solution of the machine learning tasks. This algorithm is mainly suitable for linearly separable data. Although, it still often gives a satisfactory result for solving many real-world problems. While the generalizing ability of the model is usually weak, a significant advantage of the algorithm is its interpretability, which is crucial for the described task.

Simple Linear Regression was applied, but it showed low generalization ability. In this case, the ElasticNet algorithm was tested, which is a linear model with L1 and L2 regularizations. Fig. 4 shows a generalized diagram of the linear ElasticNet model and the result of forecasting consumption on a test sample.

The resulting vector of weights was pruned in case of finding the correct T (timestamp) for each parameter. After that, the model had been retrained. This operation can be described as follows.

From initial equation of Linear Regression:

$$N = W_{11}X_1(t_0) + \dots + W_{1n}X_1(t_{-n}) + \dots + W_{k1}X_k(t_0) + \dots + W_{kn}X_1(t_{-n}) + b,$$

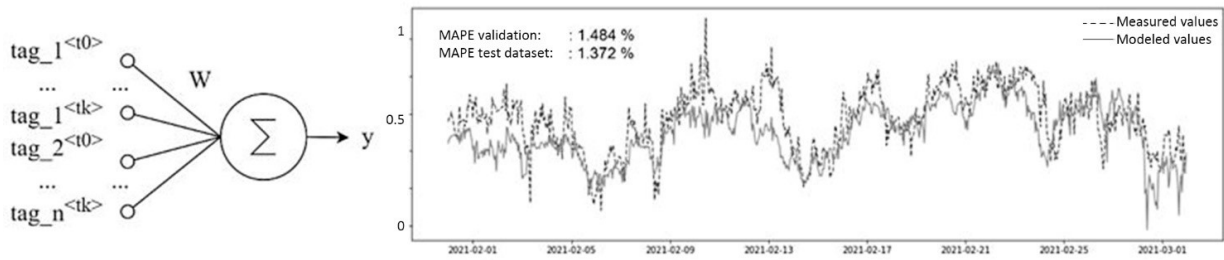


Fig. 4. Linear model of dynamic consumption rationing

where k is the number of parameters, n is the depth of time analysis, and b is a bias value. Model was transformed into:

$$N(W) = \sum_n W_i X_i(T_i) + b,$$

where T refers to different timestamps depending on the chosen parameter X . The same approach has been applied in all the models.

Another tested group of machine learning algorithms for consumption prediction is a family of tree-based boosting algorithms. Such algorithms are based on the concept of constructing a group of weak regressors to solve a more complex problem. In this work, a regressor model from the Catboost library was chosen. This algorithm has good generalization ability. Also, this algorithm has built-in functionality for analyzing the feature importance, based on the algorithm for analyzing and calculating trees (function `get_feature_importance`) [17].

Another realization of the boosting algorithms tested in this work is the classic Gradient boosting from scikit-learn library, which was applied to compare the results of modeling. However, the Random Forest algorithms were tested to compare the boosting to bagging approaches, but their further study was stopped due to the low accuracy. To solve the set task of analyzing lags in time, they are also added as parameters to the model. The resulting algorithm and mathematical description of the Catboost model development is described in Fig. 5.

An example of an element of the resulting tree and an assessment of the accuracy of the model on a test sample is shown in Fig. 6.

Another approach studied in this work is a family of artificial neural network (ANN) algorithms. Neural networks are accepted as more complex algorithms, due to huge number of parameters, that can be tuned for such models. In this way, ANNs have a high generalizing ability and are more flexible to use. Applying this group of algorithms provides us with a more elegant and appropriate approach of taking time lags into account due to the peculiarities of configuring the network architecture. In comparison with the previous approaches, the significant disadvantage of ANN is the considerable complexity and variability of its development process. At the same time, those models have a very complex structure, which complicates the process of interpreting the model. It entails the need of an analysis based on Shapley vectors, which can give us the appropriate information for the analysis of deviations cause [18].

In this work two architectures were tested as the solution of the research task. The first model is a simple sequential fully-connected neural network, for which all the time deltas were flattened into one input layer. The second tested model is the developed 1-dimensional convolutional neural network (CNN). The convolution is used to convolve timeseries data with several filters to define the time delta of the input-to-output dependency. It allows us to fold time intervals and transfer them to deeper layers of the neural network. Thus, the input to the model is not a vector of 562×8 values, but a matrix. Fig. 7 describes the generalized structure of the developed neural network.

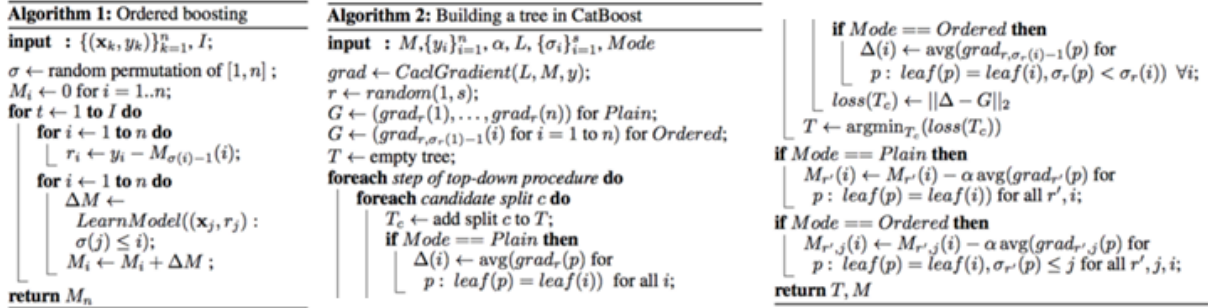


Fig. 5. CatBoost ordered boosting and tree building

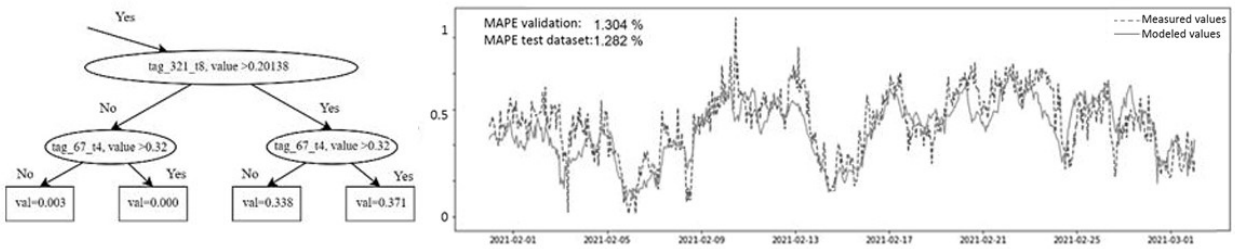


Fig. 6. CatBoost model of dynamic consumption rationing

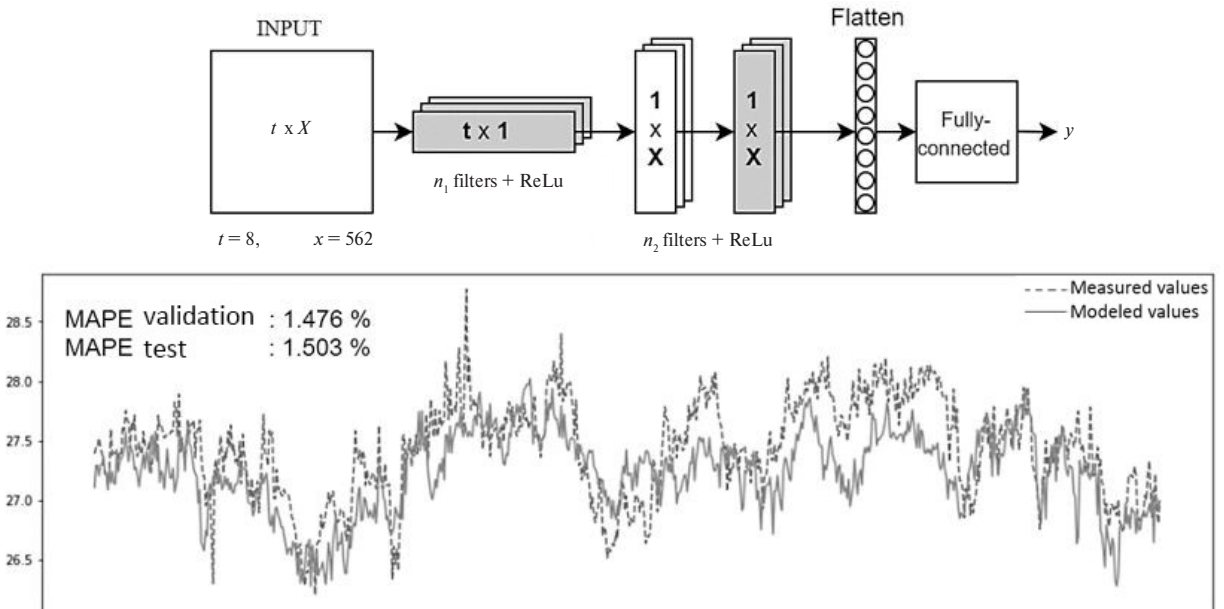


Fig. 7. CNN-based model of dynamic consumption rationing

Only three of the developed models produced satisfactory results on the described datasets. For this reason, only results of this three models consumption prediction was plotted in this publication. The analysis of the modeling results for all models is described in the next part of this work.

Modeling results analysis

During this research, a number of models of different nature have been developed. To test and validate those models, the mean absolute percentage error (MAPE) has been chosen.

The MAPE value calculation is similar to the absolute measurement error calculation [19]. Thus, we following formula was chosen to ensure the 2 % accuracy of developed algorithm:

$$M = \frac{1}{n} \cdot \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right|,$$

where n – number of measurements (data points) in the dataset; A_i – measured value; F_i – predicted value.

In this case, the task for the model development and evaluation can be rewritten as follows:

$$E = \arg \min_w \left(\frac{1}{n} * \sum_{i=1}^n \left| \frac{Y_i - N(W)_i}{Y_i} \right| \right).$$

Table 1 shows the results of evaluating the tested models. The models that performed the best were highlighted in the table.

Table 1

Model performance evaluation

Model type	Model name	MAPE validation, %	MAPE test, %
Linear	ElasticNet	1.484	1.372
	Linear Regression	1.013	2.422 (–)
Ensemble learning (Trees)	CatBoost	1.304	1.282
	Gradient Boosting	2.021 (–)	1.989
	Random Forest	3.199 (–)	2.551 (–)
Artificial Neural Networks	MLPerceptron	2.621 (–)	2.305 (–)
	Convolutional NN	1.476	1.503

Linear models performed well for this task. Classical linear regression shows the best generalizing ability on the validation set, however, the error increases dramatically on the test dataset. To solve this problem, L1 and L2 regularization was applied to the model, in the form of an ElasticNet model. This model performed well on both samples. A significant advantage of this model is also the simplicity of its interpretation in the form of regression coefficients.

On another hand, most of the ensemble learning models based on decision trees (Random Forest and Gradient Boosting) showed a slightly worse result for this task. Thus, Gradient Boosting and Random Forest showed a MAPE smaller than 2 %, which does not allow us to use them for efficient dynamic rationing.

As an exception, the model based on CatBoost Regressor showed the best result for all of the models. Also, this model can be easily interpreted by the tools of the CatBoost library itself. Further selection of hyperparameters and the use of other Gradient Boosting libraries can serve as possible ways of developing the model.

Models based on deep neural networks, generally, showed the worst prediction result among all models. The model developed on the basis of convolutional networks showed an acceptable result below 2 %. A feature of such models is a significantly larger number of parameters that can be configured. In this regard,

a further refinement of the architecture can serve as a possible way to improve the accuracy of the model. Also, the algorithm of deviation analysis is the most complex for this kind of models.

Thus, all three described models can be used to solve the problem. The priority at this stage is given to the model based on the Catboost regressor, since it showed a good generalization ability and low error rate. Moreover, this model allows us to calculate the significance of each parameter and use them to analyze the cause of deviations.

Conclusions

In this work, an approach to the dynamic rationing of energy consumption was developed. This approach is based on machine learning methods and algorithms of their interpretation.

Three main groups of algorithms have been tested, allowing us to implement dynamic normalization according to the described requirements: interpretability, adaptability, error lower than 2 %, taking into account the inertia of the process. Within these three groups, three models were built to achieve the modeling goals:

- Linear model based on EllipticEnvelope with regression coefficients as indicators of the significance of a parameter.
- CatBoost Regressor based model with parameter values (function `get_feature_importance`).
- Neural networks based model (one-dimensional convolutional neural network) with Shapley values.

Among those described models, the CatBoost Regressor based model produced the best result. This model performs the rationing of fuel and energy resources consumption with an accuracy of about 1.3 % MAPE. The built-in algorithm for calculating the significance of parameters allows us to characterize the reasons for the appearance of deviations in the data.

Despite the fact that the objectives of this study have been achieved, there is still a need of improvement of the accuracy for dynamic rationing models. Such improvement will cause the increasing of the economic benefits of the customer's manufacture.

There are two main ways to improve the obtained results. The first of them is further tuning of the parameters and hyperparameters of the developed models and expanding the training dataset. On the other hand, there are a number of algorithms that have not been tested in this work: Recurrent neural network based on GRU or LSTM; XGBoost library; LightGBM library.

REFERENCES

1. **Dozortsev V.M.** Tsifrovyye dvoyniki v promyshlennosti: genesis, sostav, terminologiya, tekhnologii, platformy, perspektivy. Chast 1. *Avtomatizatsiya v promyshlennosti*, 2020, No. 9, Pp. 3–11, (rus). DOI: 10.25728/avtprom.2020.09.01
2. **Kostenko D., Kudryashov N., Maystrishin M., Onufriev V., Potekhin V., Vasiliev A.** Digital twin applications: Diagnostics, optimisation and prediction. *Proceedings of the 29th DAAAM International Symposium*, 2018, Pp. 0574–0581, DOI: 10.2507/29th.daaam.proceedings.083
3. **Borky J.M.** Historical perspective: Energy monitoring and control systems. *The Military Engineer*, 2015, Vol. 107, No. 694, Pp. 91–92.
4. **Danilov O.L., Garyayev A.B., Yakovlev I.V.** *Energoberezheniye v teploenergetike i teplotekhnologiyakh*. Moscow: MEI Publ., 2017. (rus)
5. **Myo A.K., Portnov E.M., Kokin V.V.** Development of theoretical aspects of training systems construction for the basics of management and control over distributed power facilities. *2018 International Russian Automation Conference*, Sochi, IEEE, 2018. DOI: 10.1109/RUSAUTOCON.2018.8501758
6. **Gnatyuk V.I., Sheynin A.A.** ARS-Normirovaniye elektropotrebleniya infrastrukturykh ob'yektov. *Fedorovskiy Chteniya-2010. XL Vserossiyskaya Nauchno-Prakticheskaya Konferentsiya, Moscow, 16–19 nov. 2010*. Moscow: MEI Publ., 2010, Pp. 26–32. (rus)

7. **Gruntovich N.V., Kapanskiy A.A.** Raschetno-analiticheskiy metod normirovaniya raskhodov elektricheskoy energii v tekhnologicheskikh sistemakh vodosnabzheniya i vodootvedeniya. *Vestnik GGTU im. P.O. Sukhogo*, 2015, No. 2 (61). (rus)
8. **Kosharnaya Yu.V.** Razrabotka sistemy normirovaniya pokazateley elektropotrebleniya i otsenki ob'yemov energosberezheniya na primere metallurgicheskogo predpriyatiya. Chast 1. *Promyshlennaya Energetika*, 2015, No. 8, Pp. 13–17. (rus)
9. **Gofman I.V.** *Normirovaniye potrebleniya energii i energeticheskiye balansy promyshlennykh predpriyatiy*. Moscow: Energiya Publ., 1966. (rus)
10. **Tyralis H., Karakatsanis G., Tzouka K., Mamassis N.** Exploratory data analysis of the electrical energy demand in the time domain in Greece. *Energy*, 2017, No. 134, Pp. 902–918. DOI: 10.1016/j.energy.2017.06.074
11. **Galan A., De Prada C., Gutierrez G., Sarabia D.** *Implementation of RTO in a large hydrogen network considering uncertainty*. Springer, 2019. DOI: 10.1007/s11081-019-09444-3
12. **Kudriashov N., Markov S., Potekhin V.** Adaptive control system synthesis methods for complex manufacturing objects. *Proceedings of the 30th DAAAM International Symposium*. Publ. by DAAAM International, Vienna, Austria, 2019, Pp. 0493–0499. DOI: 10.2507/30th.daaam.proceedings.066
13. **Brown T.A.** *Confirmatory factor analysis for applied research*. 2nd ed. Guilford Press, 2015.
14. **Himeur Y., Ghanem Kh., Alsalemi A., Bensaali F., Amira A.** Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy*, 2021, Vol. 287(3). DOI: 10.1016/j.apenergy.2021.116601
15. **Ojala M., Garriga G.C.** Permutation tests for studying classifier performance. *Proceedings of the 2009 9th IEEE International Conference on Data Mining*, 2009, Pp. 908–913. DOI: 10.1109/ICDM.2009.108
16. **Prince Waqas Khan, Yung-Cheol Byun, Sang-Joon Lee, Dong-Ho Kang, Jin-Young Kang, Hae-Su Park.** Machine learning-based approach to predict energy consumption of renewable and nonrenewable power sources. *Energies*, 2020, No. 13(18). DOI: 10.3390/en13184870
17. **Prokhorenkova L., Gusev G., Vorobev, Dorogush A.V., Gulin A.** *CatBoost: Unbiased boosting with categorical features*, V5 2019.
18. **Molnar Ch.** *Interpretable machine learning. A guide for making black box models explainable*, 2021.
19. **de Myttenaere A., Golden B., Le Grand B., Rossi F.** Mean absolute percentage error for regression models. *Neurocomputing*, 2016 arXiv:1605.02541.

Received 23.05.2021.

СПИСОК ЛИТЕРАТУРЫ

1. **Дозорцев В.М.** Цифровые двойники в промышленности: генезис, состав, терминология, технологии, платформы, перспективы. Ч. 1 // Автоматизация в промышленности. 2020. № 9. С. 3–11. DOI: 10.25728/avtprom.2020.09.01
2. **Kostenko D., Kudryashov N., Mastrishin M., Onufriev V., Potekhin V., Vasiliev A.** Digital twin applications: Diagnostics, optimisation and prediction // Proc. of the 29th DAAAM Internat. Symp. 2018. Pp. 0574–0581. DOI: 10.2507/29th.daaam.proceedings.083
3. **Borky J.M.** Historical perspective: Energy monitoring and control systems // *The Military Engineer*. 2015. Vol. 107. No. 694. Pp. 91–92.
4. **Данилов О.Л., Горяев А.Б., Яковлев И.В.** Энергосбережение в теплоэнергетике и теплотехнологиях. М.: ИД МЭИ, 2017.
5. **Myo A.K., Portnov E.M., Kokin V.V.** Development of theoretical aspects of training systems construction for the basics of management and control over distributed power facilities // 2018 Internat. Russian Automation Conf. Sochi, IEEE, 2018. DOI: 10.1109/RUSAUTOCON.2018.8501758

6. **Гнатюк В.И., Шейнин А.А.** ARS-нормирование электропотребления инфраструктурных объектов // Федоровские чтения-2010. XL Всерос. науч.-практ. конф. с междунар. уч. с элементами научной школы для молодежи. М.: ИД МЭИ, 2010. С. 26–32.
7. **Грунтович Н.В., Капанский А.А.** Расчетно-аналитический метод нормирования расходов электрической энергии в технологических системах водоснабжения и водоотведения // Вестник ГГТУ им. П.О. Сухого. 2015. № 2 (61).
8. **Кошарная Ю.В.** Разработка системы нормирования показателей электропотребления и оценки объемов энергосбережения на примере металлургического предприятия. Ч. 1 // Промышленная энергетика. 2015. № 8. С. 13–17.
9. **Гофман И.В.** Нормирование потребления энергии и энергетические балансы промышленных предприятий. М.: Энергия, 1966.
10. **Tyralis N., Karakatsanis G., Tzouka K., Mamassis N.** Exploratory data analysis of the electrical energy demand in the time domain in Greece // Energy. 2017. No. 134. Pp. 902–918. DOI: 10.1016/j.energy.2017.06.074
11. **Galan A., De Prada C., Gutierrez G., Sarabia D.** Implementation of RTO in a large hydrogen network considering uncertainty. Springer, 2019. DOI: 10.1007/s11081-019-09444-3
12. **Kudriashov N., Markov S., Potekhin V.** Adaptive control system synthesis methods for complex manufacturing objects // Proc. of the 30th DAAAM Internat. Symp. Publ. by DAAAM International, Vienna, Austria, 2019, Pp. 0493–0499. DOI: 10.2507/30th.daaam.proceedings.066
13. **Brown T.A.** Confirmatory factor analysis for applied research. 2nd ed. Guilford Press, 2015.
14. **Himeur Y., Ghanem Kh., Alsalemi A., Bensaali F., Amira A.** Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives // Applied Energy. 2021. Vol. 287(3). DOI: 10.1016/j.apenergy.2021.116601
15. **Ojala M., Garriga G.C.** Permutation tests for studying classifier performance // Proc. of the 2009 9th IEEE Internat. Conf. on Data Mining. 2009. Pp. 908–913. DOI: 10.1109/ICDM.2009.108
16. **Prince Waqas Khan, Yung-Cheol Byun, Sang-Joon Lee, Dong-Ho Kang, Jin-Young Kang, Hae-Su Park.** Machine learning-based approach to predict energy consumption of renewable and nonrenewable power sources // Energies. 2020. No. 13(18). DOI: 10.3390/en13184870
17. **Prokhorenkova L., Gusev G., Vorobev, Dorogush A.V., Gulina A.** CatBoost: Unbiased boosting with categorical features. V5 2019.
18. **Molnar Ch.** Interpretable machine learning. A guide for making black box models explainable. 2021.
19. **de Myttenaere A., Golden B., Le Grand B., Rossi F.** Mean absolute percentage error for regression models // Neurocomputing. 2016 arXiv:1605.02541.

Статья поступила в редакцию 23.05.2021.

THE AUTHOR / СВЕДЕНИЯ ОБ АВТОРЕ

Kudriashov Nikita S.
Кудряшов Никита Сергеевич
E-mail: niki94@yandex.ru