Doctoral Thesis

Developing the Basic Method for Quality Assurance and Accelerated Commissioning of Energy Monitoring Systems

submitted in satisfaction of the requirements for the degree of
Doctor of Science in Civil Engineering
of the TU Wien, Faculty of Civil Engineering

Dissertation

Entwicklung der grundlegenden Methode zur Qualitätssicherung und beschleunigten Inbetriebnahme von Energiemonitoringsystemen

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Doktors der technischen Wissenschaften
eingereicht an der Technischen Universität Wien, Fakultät für Bauingenieurwesen
von

Alexander David, BSc MSc
Matrikelnummer 01652653

Supervisor: Univ. Prof. Dipl.-Ing. Dr.techn. Thomas Bednar
Institute of Material Technology, Building Physics, and Building Ecology
TU Wien

Reviewer: Univ. Prof. Dipl.-Ing. Dr.techn. Wolfgang Kastner
Institute of Computer Engineering
TU Wien

Reviewer: Univ. Prof. Dipl.-Ing. Dr.techn. Michael Monsberger
Institute of Building Physics, Services, and Construction
Graz University of Technology

Wien, Oktober 2022

Danksagung

Ich möchte mich herzlich bei all jenen Menschen bedanken, die diese Arbeit überhaupt erst ermöglicht haben. Allem voran möchte ich mich bei meinem Betreuer Prof. Thomas Bednar bedanken. Lieber Thomas, danke, dass du in den letzten Jahren mein Mentor warst und mir stets das Gefühl gabst, wertgeschätzt zu sein.

Weiters möchte ich mich bei Prof. Wolfgang Kastner und Prof. Michael Monsberger für die Begutachtung meiner Arbeit bedanken.

Dank gebührt auch meinen Kolleginnen und Kollegen des Forschungsbereichs. Durch das angenehme Arbeitsklima am Forschungsbereich bin ich noch an keinem einzigen Tag widerwillig zur Arbeit gegangen. Viele von euch sind nicht nur Kolleginnen und Kollegen sondern auch Freunde.

Eine ehemalige Kollegin, der besonderer Dank gebührt, ist Naomi, die mir wertvolles Feedback dazu gab, wie ich das Englisch dieser Arbeit verbessern kann.

Weiters gilt mein Dank meiner Lebensgefährtin Astrid, die nicht nur meine Launen und unsere reduzierte gemeinsame Freizeit ertrug, sondern mir auch ihr wertvolles Feedback zur Arbeit gab. Liebe Astrid, danke für deine Geduld, deine Unterstützung und für die vielen wunderbaren gemeinsamen Jahre.

Abschließend möchte ich mich bei meiner Familie bedanken, die stets für mich da war, mich durch alle Höhen und Tiefen des Lebens begleitet hat und somit wesentlich zu meinem Fortschritt und Erfolg im Leben beigetragen hat. Ich kann mir keine besseren Eltern und keine bessere Schwester vorstellen als euch.

Abstract

Buildings are among the largest energy consumers and thereby significantly contribute to global greenhouse gas emissions. Reducing the energy consumption of buildings is one key element for achieving the European Union's climate goals.

Modern designs of newly erected or refurbished buildings proved to be highly energy-efficient while still offering a comfortable indoor climate. While some of these highly energy-efficient buildings meet their design values for energy consumption, others showed a significant deviation. An energy monitoring system with deeper submetering can help identify the reasons for a deviation, e.g., errors in the building's systems or a building usage that deviates from the planning. Moreover, energy monitoring is a useful tool for continuous commissioning, which aids in ensuring an energy-efficient building operation.

Before an energy monitoring system can be used to detect errors in the building operation, the monitoring system itself needs to be commissioned and validated – especially if the system is with deeper submetering and thus more complex. Ensuring that the energy monitoring system provides reliable data and is free of errors is a time-consuming but necessary task. By using different computer-aided methods and algorithms, this task could be aided and sped up. Considering the recent advances in computer science, a set of methods that can automatically detect almost all errors in the energy monitoring system solely by analyzing the data provided by the system appears possible. A key element that is needed as a prerequisite for such a set of methods is a method that can determine whether an energy meter is subordinate to another meter or not. The objective of this thesis is the development of such a method.

The developed novel method determines the main meter-submeter relationship between two energy meters solely by analyzing the time series of the meter's energy measurements. It is shown how the time series data must be transformed and which features must be derived from it to use them as input for random forest classifier machine learning algorithms. The application limits of the method are evaluated through development and validation with monitoring data of varying time resolutions and varying time slot lengths. As the method primarily focuses on detecting similar changes in load profiles, its performance depends on the measured energy consumers and the volatility of their load profiles. For the case of monitoring data from electricity meters with a time resolution of 5 min, it can be expected that the method can infer approximately 40% of the electricity meter hierarchy. For the cases 10 min and 15 min time resolution, the rate drops to approximately 34% and 26%.

Kurzfassung

Gebäude sind einer der größten Energieverbraucher und tragen daher stark zu den globalen Treibhausgasemissionen bei. Die Reduzierung des Energieverbrauchs von Gebäuden ist ein Schlüsselelement zur Erreichung der Klimaziele der Europäischen Union.

Moderne Designs von neu errichteten oder sanierten Gebäuden haben sich als höchst energieeffizient erwiesen und bieten dennoch ein angenehmes Raumklima. Während einige dieser höchst energieeffizienten Gebäude ihre Auslegungswerte für den Energieverbrauch einhalten, zeigen andere eine deutliche Abweichung. Ein Energiemonitoringsystem mit Submetering kann helfen, die Gründe für Abweichungen zu identifizieren, z.B. Fehler in der Gebäudetechnik oder eine von der Planung abweichende Gebäudenutzung. Darüber hinaus ist das Energiemonitoring ein nützliches Tool für Continuous Commissioning, welches dabei hilft, einen energieeffizienten Gebäudebetrieb sicherzustellen.

Bevor ein Energiemonitoringsystem zur Erkennung von Fehlern im Gebäudebetrieb eingesetzt werden kann, muss das Monitoringsystem selbst in Betrieb genommen und validiert werden – insbesondere, wenn das System mit Submetering und damit komplexer ist. Sicherzustellen, dass das Energiemonitoringsystem vertrauenswürdige Daten liefert und fehlerfrei ist, ist eine zeitaufwändige aber notwendige Aufgabe. Durch den Einsatz verschiedener computergestützter Methoden und Algorithmen könnte diese Aufgabe unterstützt und beschleunigt werden. Angesichts der jüngsten Fortschritte in der Informatik erscheint ein Set von Methoden möglich, das fast alle Fehler im Energiemonitoringsystem automatisch erkennen kann, indem es allein die vom System bereitgestellten Daten analysiert. Ein Schlüsselelement, das als Voraussetzung für ein solches Set von Methoden benötigt wird, ist ein Verfahren, das bestimmen kann, ob ein Energiezähler einem anderen Zähler untergeordnet ist oder nicht. Das Ziel dieser Arbeit ist die Entwicklung einer solchen Methode.

Die entwickelte neuartige Methode bestimmt die Hauptzähler-Subzähler-Beziehung zwischen zwei Energiezählern allein durch die Analyse der Zeitreihen der Energiemessungen der Zähler. Es wird gezeigt, wie die Zeitreihendaten transformiert und welche Merkmale daraus abgeleitet werden müssen, um sie als Input für Random Forest Classifier Machine Learning Algorithmen zu verwenden. Die Einsatzgrenzen der Methode werden durch Entwicklung und Validierung mit Monitoringdaten unterschiedlicher zeitlichen Auflösungen und unterschiedlicher Zeitspannen evaluiert. Da sich das Verfahren primär auf die Erkennung gleichartiger Änderungen in Lastprofilen konzentriert, hängt seine Leistungsfähigkeit von den gemessenen Energieverbrauchern und der Volatilität ihrer Lastprofile ab. Für den Fall von Monitoringdaten von Stromzählern mit einer zeitlichen Auflösung von 5 min ist zu erwarten, dass die Methode ca. 40% der Stromzählerhierarchie identifizieren kann. Für die Fälle mit 10 min und 15 min zeitlicher Auflösung sinkt die Rate auf etwa 34% und 26%.

Table of Contents

1 Introduction

As buildings account for approximately 40% of the EU's final energy consumption [1], the building sector is crucial for achieving the EU's energy and climate goals. With the Energy Performance of Buildings Directive (EPBD) 2018/844/EU [2] in its most recent version, the EU member states aim to improve the energy performance of buildings. One of the key elements of the EPBD is to ensure that all new buildings are nearly zero energy buildings (nZEB) by 31st December 2020.

The numerous nearly zero energy buildings built before that deadline proved that the energy efficiency demanded by the EPBD is achievable [3]. Some highly-efficient buildings even harnessed enough energy from local resources to achieve net zero energy. Moreover, it is possible to design high-rise office buildings as net zero energy buildings, as exemplified by the Elithis Tower in France [4] and the (Plus-)Plus-Energy Office High-Rise Building in Austria [5] – see Fig. 1.

Fig. 1: Examples of net zero energy buildings. Left: Elithis Tower in Dijon, France [4]. Right: (Plus-)Plus-Energy Office High-Rise Building in Vienna, Austria [5].

Even though achieving a nearly or net zero energy balance is feasible in theory, there are often issues when implementing the concepts. Several studies showed that the real energy consumption of highly efficient buildings often exceeds their theoretical consumption [6]. The difference between actual and theoretical energy consumption is referred to as "Energy Performance Gap" (EPG). If the actual consumption exceeds the theoretical consumption, the EPG is positive. In cases where the theoretical consumption exceeds the actual consumption, the EPG is negative.

The EPG was also observed in the Elithis Tower and the (Plus-)Plus-Energy Office High-Rise Building. Fig. 2 illustrates both buildings' non-renewable primary energy demand and the resulting EPGs due to higher actual demands observed during monitoring. It can be observed that in both cases, the energy consumption measured by the energy monitoring significantly exceeded the design value. In both cases, there was additional energy consumption in the category “usage” – the consumption caused by plug-loads, which heavily depends on the building’s users. In the case of the (Plus-)Plus-Energy Office High-Rise Building, additional consumption could also be observed in the category “building” – the consumption needed to operate the building and make the building useable.

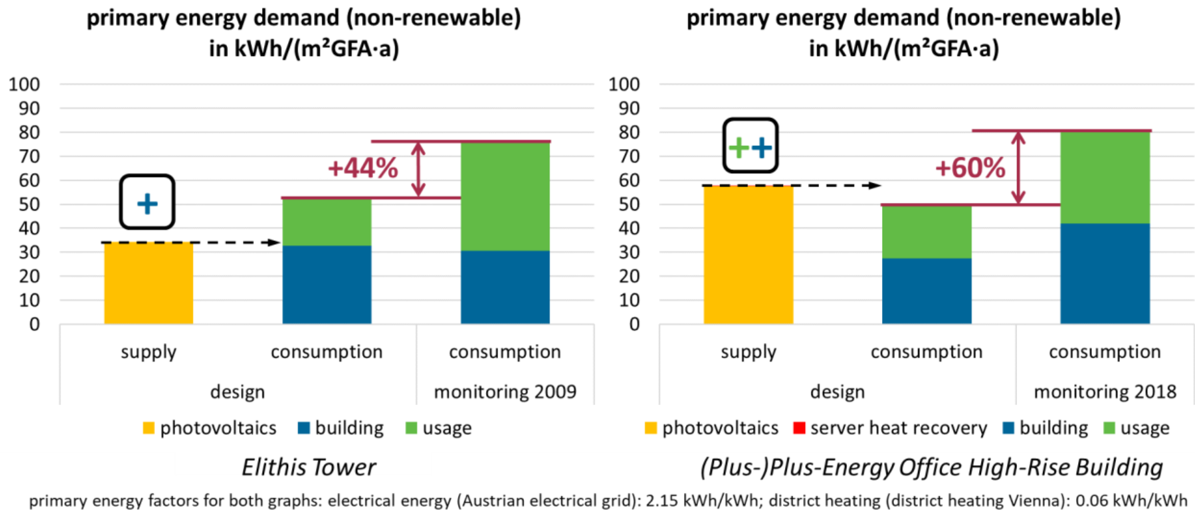


Fig. 2: Performance gaps of the Elithis Tower and the (Plus-)Plus-Energy Office High-Rise Building. The primary energy factors of the (Plus-)Plus-Energy Office High-Rise Building [5] were applied to the original data of the Elithis Tower [4].

There are various reasons for the EPG, ranging from inaccuracy of inputs and assumptions for building modeling to malfunctioning equipment [7]. Especially for the cases with a positive EPG, an investigation should be conducted to identify the causes behind observed differences.

The energy monitoring system plays a crucial role in such an investigation. Energy monitoring systems with deep submetering on system- and equipment levels allow for a detailed analysis of energy flows inside the building. Depending on the structure and depth of the monitoring system, the energy consumption can be traced back to certain building parts and systems or even single components [8].

An example of an energy monitoring system with deep submetering is the system installed in the (Plus-)Plus-Energy Office High-Rise Building. It is an extensive energy monitoring system with three levels of submetering. Thus, by analyzing the monitoring data, it was possible to pinpoint the consumers who cause additional energy consumption and draw conclusions regarding the underlying reasons – see Fig. 3.

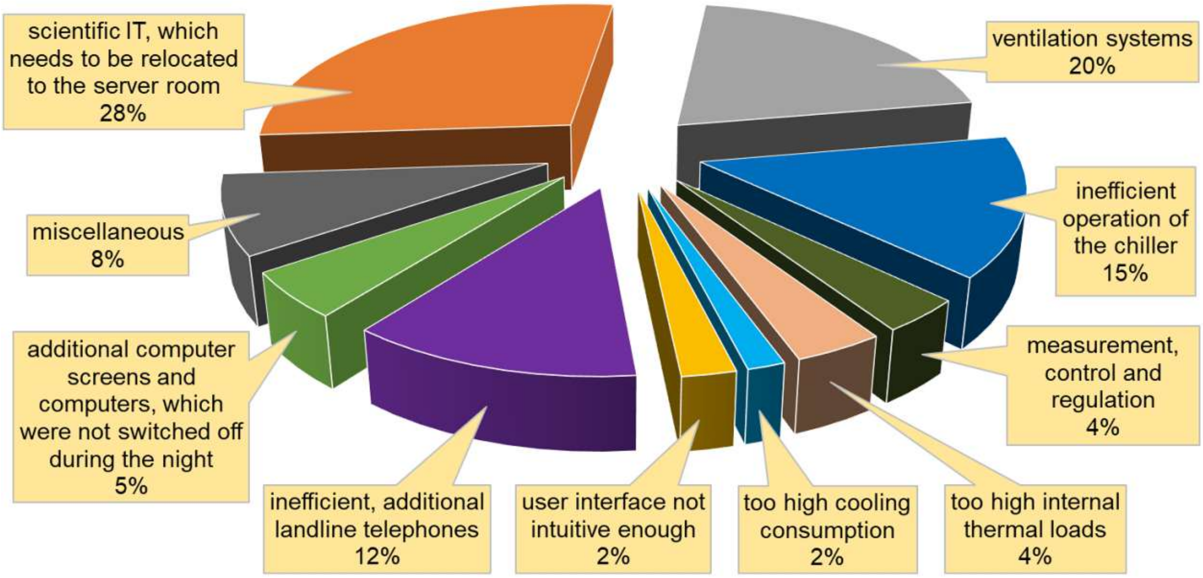


Fig. 3: Reasons for the (Plus-)Plus-Energy Office High-Rise Building's performance gap in the year 2018, presented as percentage of the additional non-renewable primary energy demand.

Fig. 3 illustrates the reasons for the EPG at the end of the building's monitoring and optimization project – the end of the year 2018. In the years before that, the EPG was even higher. Identifying the causes of the high energy consumption during the monitoring and optimization project was the first step in reducing the consumption [9]. Knowing the causes for the increased energy consumption that is still there after the optimization is an important knowledge gain for future near or net zero energy buildings. Without the building's extensive energy monitoring system, it would not have been possible to determine the reasons for the excessive energy consumption.

An energy monitoring system is not only useful to optimize the building and thus reduce its consumption during or after the commissioning phase, but it is the key element to ensure the correct operation of the building's equipment and controls and is the foundation for continuous commissioning [10,11]. By ensuring that buildings operate as efficiently as possible and stay that way, energy monitoring plays an essential role in reaching the EU's energy and climate goals.

1.1 Motivation

Before an energy monitoring system can effectively be put to use, the system itself must be validated, and errors in the system must be fixed. The deeper the submetering, the more components are part of the monitoring system, and therefore the higher the probability of errors. There is comprehensive research regarding fault detection and diagnosis in building *energy* systems [12–14] but not in the case of energy *monitoring* systems themselves.

The few methods that use energy monitoring data to detect faults in the energy monitoring system focus mainly on detecting faulty smart meters energy theft [15–17]. None of the methods addresses errors in the structure and topology of the energy monitoring system, e.g., meters placed at a wrong location or meters measuring other consumers than they should. The commissioning of the energy monitoring of the (Plus-)Plus-Energy Office High-Rise Building in Vienna showed that there is indeed a need for methods that can help to deal with such errors – several structural and topological errors were detected.

Fig. 4 shows the timeline of the building's refurbishment process. The energy monitoring system should have been fully operational at the beginning of 2015, during the commissioning phase of the building itself. Instead, it took almost a year before the errors in the monitoring system were identified and corrected by the responsible contractors.

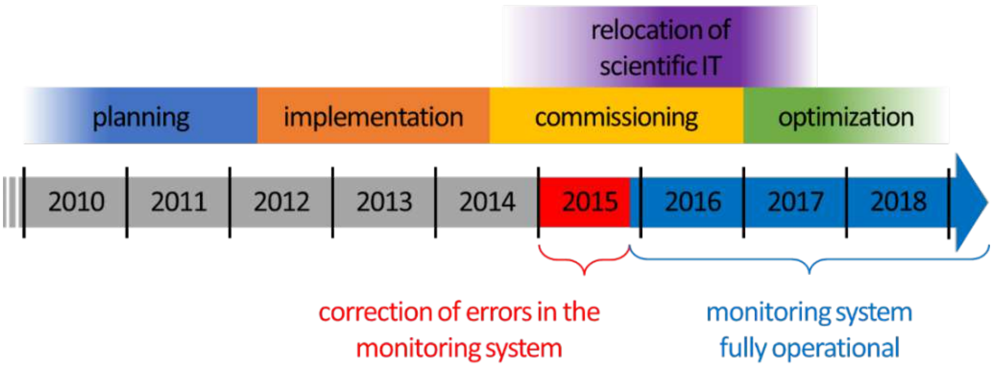


Fig. 4: Timeline of the (Plus-)Plus-Energy Office High-Rise Building's refurbishment process.

There are several reasons why it took so long to identify and correct the errors:

- The process of reporting errors to contractors needed to run through multiple administrative levels, and it was not possible for the research team to report the errors directly.
- Four different contractors were commissioned to install the energy meters. It was not precisely defined if one contractor was responsible for the parametrization of all energy meters or whether each contractor was responsible for their own energy meters.
- There was no overview of the hierarchy of the energy monitoring system. The research team had to derive the hierarchy from the circuit diagrams of the building's switch cabinets, the heating and cooling schematics, and through site visits.
- Some errors could only be discovered after other errors were fixed.

No heat meter errors were found apart from the issue that the straight length of pipe in the inlet and outlet sections was less than advised by the heat meter manufacturer's guidance. As this is not an issue that can be fixed easily and as metering heat is not the main purpose of the piping, this issue is not considered a real error in the context of this work.

Thus, practically all encountered errors were related to electricity meters. These errors can be summarized as follows:

- Wrong parametrization of the electricity meter’s current transformer values
- Current transformers were installed incorrectly
- Electricity meter measured other consumers than expected
- The wrong type of electricity meter was installed
- Electricity meter was faulty
- Electricity meter was not installed at all

The identification and correction of those errors were relatively time-consuming. Instead of using the energy monitoring system to analyze the building performance, the research team had to verify that the energy monitoring system works properly first. Particularly, the preparation of an overview of the hierarchy of the energy monitoring system was labor-intensive, but it was a major prerequisite for identifying the stated errors.

To accelerate this time-consuming process of validating the energy monitoring system, a set of methods that aids in the tasks of generating an overview of the real energy meter hierarchy and identifying errors would be beneficial. The analysis of the energy meter data suggested that it might contain enough information to aid with both tasks. Considering the actual state of data science, the *vision* depicted in 5 steps in Fig. 5 seems possible:

Step 1: After an energy monitoring system has been initialized by the contractor, the system accumulated some weeks of monitoring data and stored it in a data storage. The energy meter data is then automatically identified and extracted from the data.

Step 2: By analyzing the energy meter data, it is determined whether an energy meter is the submeter of another energy meter. Using this information, parts of the energy meter hierarchy are automatically identified without providing additional information about the expected hierarchy. Some energy meters, e.g., those which measure consumers that measure almost constant energy consumption, cannot be placed in the energy meter hierarchy automatically.

Step 3: All those energy meters which could not be placed in the energy meter hierarchy automatically are then positioned in the hierarchy manually. This step relies on additional information, which might be gathered by the designation of the data points, circuit diagrams, or site visits.

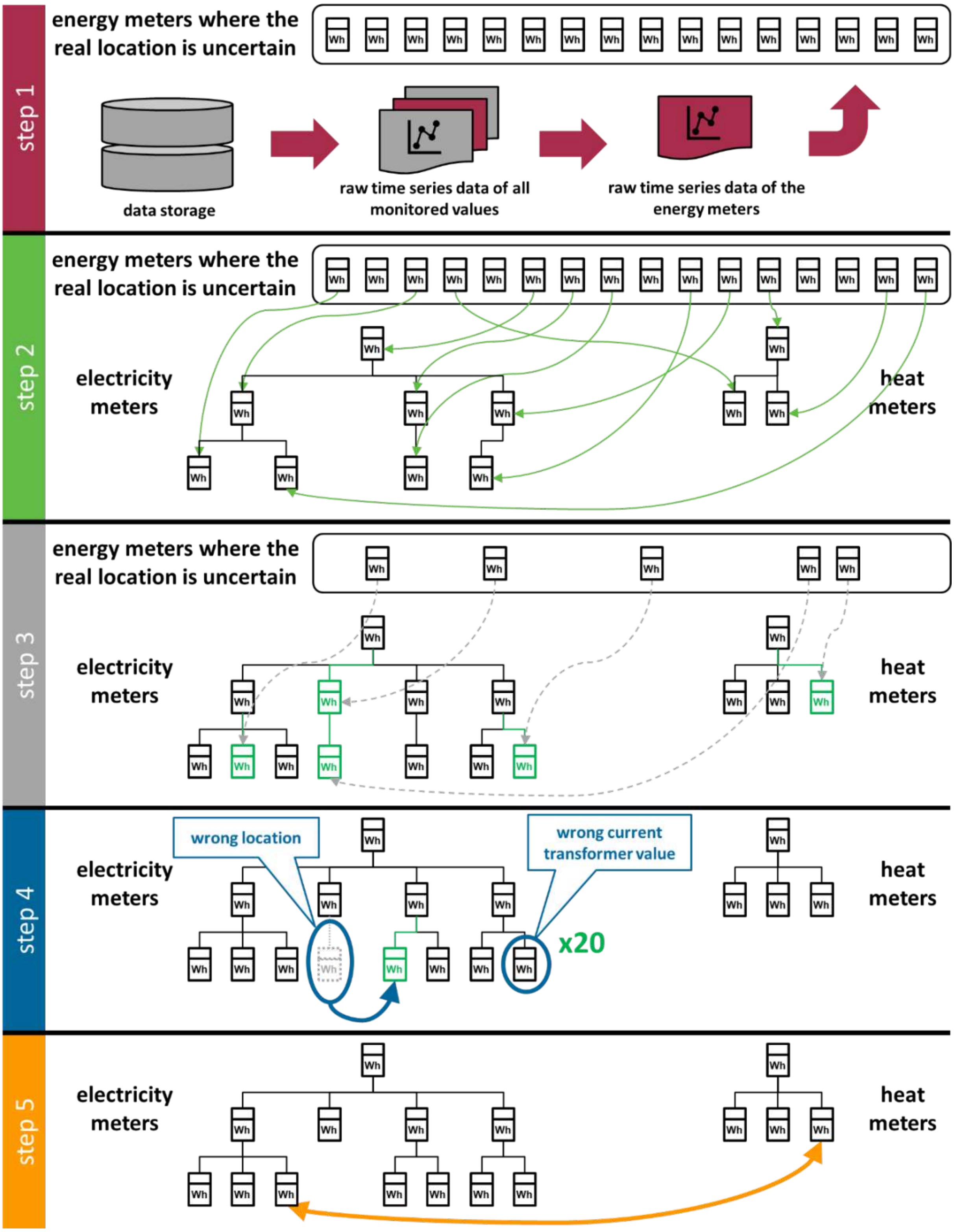


Fig. 5: Vision of how the process of validating energy monitoring systems can be accelerated – step 1-5.

Step 4: After the energy meters were placed in the hierarchy automatically or manually, further algorithms check the plausibility and try to identify errors. If possible, the algorithms indicate how an error can be fixed, e.g., the correct current transformer value.

Step 5: The information contained in the energy meter data, combined with the knowledge about the energy meter hierarchy, might be enough to determine causal relations between energy meters, e.g., if the relation between an electricity meter which measures the pump that transports the water which is measured by a heat meter is determined, it would be highlighted as a possible causal relationship.

The key element to enable the process described in steps 1-5 in Fig. 5 is a method to determine whether an energy meter is subordinate to another energy meter. Even though the main application of this method is in step 2, it might also contribute to step 4, e.g., by using the method in a loop where the main meter-submeter relation of two meters is evaluated repeatedly while the data of one meter is manipulated according to typical current transformer values. Suppose the method only detects a main meter-submeter relation when the data of one meter is manipulated as it came from a meter with a particular current transformer value. In that case, this value should be the correct one that the meter must be parameterized with.

1.2 Objective

The objective of this doctoral thesis is the development of a method that can determine whether an energy meter is subordinate to another meter. The method shall be applicable to both types of energy meters – electricity meters and heat meters. As described in the *vision* in chapter 1.1, this method is intended to be the key element for a process to accelerate the validation of energy monitoring systems – the final step of the commissioning of an energy monitoring system.

The ambition of the developed method is that the determination of whether an energy meter is subordinate to another meter is made based on the meters' raw time series data only. There shall be no need to provide additional information. All necessary steps to automatically process the raw time series data are part of the method.

As energy monitoring data can have different time resolutions, the method shall be capable of handling 5 min, 10 min, 15 min, 30 min, and 60 min time resolutions.

Following the idea of the *vision* formulated in chapter 1.1, it is intended that the method is used shortly after the energy monitoring system has been initialized by the contractor. That means it should work with monitoring data from a few weeks. Using a time span of five weeks as the baseline, the impact of shorter and longer time spans shall be investigated. Thus, the method is to be tested with data from time slots with the following lengths: 1 week, 2 weeks, 5 weeks, 3 months, 6 months, and 1 year.

The prediction performance of the method and its application limits shall be evaluated for all the mentioned time resolutions and time slot lengths. Finally, the method is to be validated by providing it with data from another data source.

2 State of the art

Fault detection and diagnostics of energy monitoring systems is not a broadly discussed or researched topic. There are various papers discussing the application of energy monitoring systems to detect errors in *other systems*, but none of them addresses the issues that arose during the commissioning of the *energy monitoring system*. Still, the plausibility of errors and faults in the energy monitoring system itself is documented in guidelines [18,19] and reports [20] regarding the installation of energy meters.

Most of the available research that can be attributed to the topic "fault detection and diagnostics of energy monitoring systems" is part of at least one of these three research fields:

- Smart meters
- Fault detection and diagnostics for building systems
- Fault detection and diagnostics for power systems

Fig. 6 illustrates the overlap of these research fields and their sub-topics. It is based on the results of the literature research that was conducted for this thesis. The objective of this thesis, the detection of main meter-submeter relationships, is a subject at the intersection of several disciplines. The topic can be found under the umbrella of "detection of structural and topological errors", which in turn is used for both "fault detection and diagnostics for power systems" and "fault detection and diagnostics for building systems".

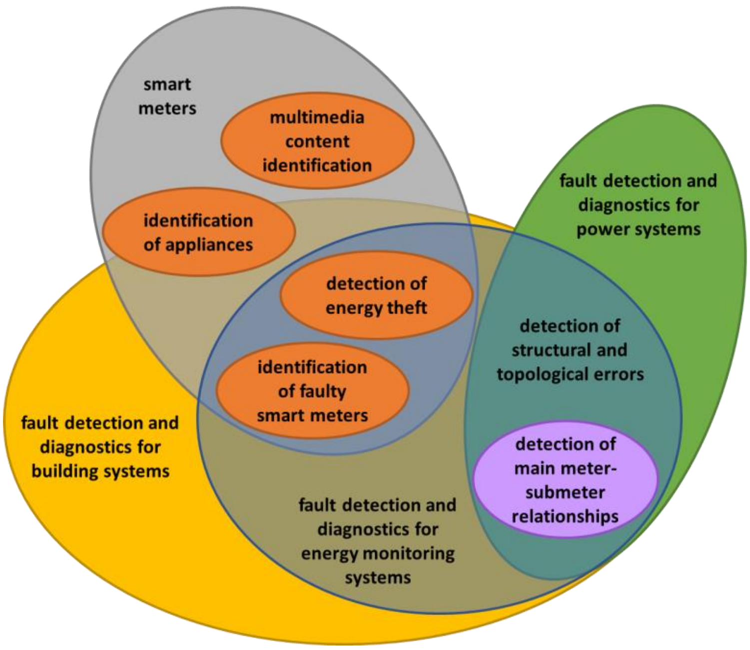


Fig. 6: Overview showing the topics related to the detection of main meter-submeter relationships.

2.1 Smart meters

The EU has specified the legal framework for the installation of smart meters in the directive 2009/72/EC [21]. The directive states that, to facilitate energy efficiency, as many consumers as possible shall be equipped with smart meters. This directive is already implemented in the national legislation of the member states, and the roll-out of smart meters is ongoing – e.g., in Austria, approximately 30% of consumers were equipped with smart meters in 2020 [22].

With the increasing number of smart meters, there is also an increasing amount of research related to them. Smart meters promise advantages for the operation of power systems. Still, there are also concerns about issues regarding privacy and security, e.g., that the smart meters might reveal information about the presence of people at their residences [23].

One of the examples of research, which highlights the danger that smart meters pose to consumer privacy, is a study on multimedia content identification through smart meter power usage profiles. In this study, it was shown that it is possible to identify what channel the TV in a household was displaying under the precondition that the smart meter's sample rate is set to 0.5 s^{-1} and there is at least a five minute-chunk of consecutive viewing without major interference by other appliances. The method used in this study relies mainly on calculating the correlation between the expected energy consumption calculated for the TV content and the real energy consumption measured by the smart meter [24].

Another study investigated the correct identification of appliances by analyzing their load profiles. The load profiles used in this study had a varying time resolution between one and five seconds. Out of these profiles, metadata was calculated, then used as features for a classification process. Different types of classifiers were investigated, and their cross-validation accuracy for 25 folds was calculated. The random committee and random forest classifiers achieved the highest cross-validation accuracy, which is slightly above 95% in both cases. In this study, the classifiers were only trained with and used on clean load profiles of single appliances. It was not investigated how single appliances can be identified in a load profile of a group of multiple appliances, which is the load profile that a smart meter would provide in a real setting [25].

Yip et al. went further and analyzed the load profiles of a group of multiple smart meters individually and as a group to detect energy theft and defective smart meters in smart grids. The basic principle behind the method presented in this study is that the load profile measured at the supply node of the utility provider must equal the sum of the load profiles of the smart meters of the supplied consumers. For the method to work, all the supplied consumers must be equipped with a smart meter. Besides the transportation losses between the supply node and the consumers and usual measurement errors, unmetered energy consumption is not permitted. The main method used in this study is a multiple linear regression, where the load profiles of all smart meters are the predictors, and the difference profile of their sum profile and the load profile of the supply node are the target. The regression coefficients of this multiple linear regression provided an indication if and which consumers seem to steal energy as well as the presence of defective smart meters. This method has been proven to work with load profiles with a time interval of 30 min [15].

A similar setup was used in another study, which focused solely on identifying defective smart meters. The method presented in this study also processes the load profile of a group of multiple smart meters and the load profiles of each of the meters to find the defective ones. Again, besides transportation losses and measurement errors, unmetered energy consumption is not permitted. The method was tested with load profiles with 15 min and 60 min time intervals. Instead of linear regression, a combination of a long short-term memory artificial recurrent neural network (LSTM) and a modified convolutional neural network (CNN) was used. The LSTM predicted the load profile of a group of multiple smart meters. If the prediction had a too high deviation from the real load profile of the group, a diagnostic process was initiated, where the CNN then classified the defective smart meters [17].

2.2 Fault detection and diagnostics for building systems

Fault detection and diagnostics (FDD) for building systems is a broadly researched field. A review by Kim et al. showed that close to 200 studies investigating automated fault detection and diagnostics of building systems were published between 1990 and 2018. 83 of those studies discuss FDD methods for ventilation systems, 33 methods for chillers and cooling towers, and 32 methods for air conditioners and heat pumps. Other building systems, as well as FDD for the overall building, were discussed in 49 studies. The majority, namely 62%, of the methods used in the studies are process history-based, i.e., they derive behavioral models from measurement data obtained from the process over time. 26% of the methods are qualitative model-based, i.e., the models consist of qualitative relationships derived from knowledge of the underlying physics. The remaining 12% of methods are quantitative model-based, i.e., the models are sets of quantitative mathematical relationships based on the underlying physics of the process [13].

Zhao et al.'s review focused on artificial intelligence-based FDD methods for building systems. They found that 135 papers discussing these methods were published between 1998 and 2018. The studies can be classified into two major sub-categories: studies that discuss knowledge-driven methods and studies that discuss data-driven methods. About 21% of the methods are knowledge-driven, and a majority of 79% are data-driven. With the rise of computational power and the increased research in the field of machine learning, the number of studies investigating data-driven methods increased over the years. Approximately half of the studies covering data-driven methods were published between 2014 and 2018. Of the data-driven methods, about 44% were unsupervised learning methods, 30% classification methods, and 26% regression methods [14].

For the literature research that was conducted for this thesis, some of the studies cited in [13] and [14] were investigated further. Still, none of them discussed a method that seemed useful for identifying main meter-submeter relationships.

2.3 Fault detection and diagnostics for power systems

In the research field of fault detection and diagnostics for power systems, there are several studies [26–30] that address topology identification of power systems. The authors of those studies state that the main motivation for topology identification of power systems is that especially the topology of low-voltage distribution networks recorded by the network operators is often inaccurate. This is due to the high complexity of those systems, continuously undergoing changes, and the fact that the data is basically entered into the operator’s recording system manually. Some of the studies concern the identification of relationships between substation, feeder, and transformer nodes (short: supply nodes), others the identification of relationships between supply nodes and electricity meters, and some the identification of relationships between electricity meters.

In most of the investigated studies [26–29], the topology is identified by processing the time series of the voltages measured at supply nodes or electricity meters. Basically, the methods used in these studies all combine a method that compares the time series of voltages of the different supply nodes or electricity meters with other methods or algorithms.

In [26], the Pearson correlation coefficient is used to determine the strength of the linear association between the different voltage time series of electricity meters and their one shared transformer node. The pairs of voltage time series with the highest correlation coefficients indicate that they should come from neighboring electricity meters or the transformer node. By taking advantage of the fact that the general voltage level drops the further away a meter is from the energy supply, the hierarchy is determined. The method proposed in this study was tested with real monitoring data from one day with 5 min time resolution. Between 87.5% and 93.5% of the topology could be identified.

The Pearson correlation coefficient is also the main metric used in [27]. In this study, the goal is not to identify the metering topology below one transformer node but instead to determine which electricity meter is supplied by which transformer node. For this purpose, an algorithm combining the Pearson correlation coefficient and a Fuzzy C-Means clustering was proposed. The algorithm was tested with real monitoring data from 48,220 users that are supplied by 500 transformers. It was data from a whole day with a time resolution of 15 min. In that case, 91.3%

of the transformer node-electricity meter relationships were identified correctly by the algorithm.

Another study that uses the coefficient of determination R^2 as main metric is [28]. Similar to [27], the goal of this study is to find out which electricity meter is supplied by which feeder node. For each set of possible combinations of electricity meters and feeder node, a univariate linear regression is performed and the respective R^2 is calculated. Combinations with R^2 values above 0.95 are seen as correctly identified feeder node-electricity meter relationships. Energy meters that only achieve R^2 values below this threshold are automatically assigned to feeder nodes by using K-means clustering. This method was not tested on real data but instead on the results of a MATLAB/Simulink simulation of a power grid consisting of two feeder nodes and 25 electricity meters. With a sampling rate of 50 µs and a simulation period of 0.2 s could correctly identify 75% of the topology.

The last of the investigated studies, in which the voltage time series are processed to identify the topology, is [29]. In this study, the goal is to find out which electricity meter is supplied by which transformer node phase. The proposed method also utilizes R^2 as main metric, but instead of applying univariate linear regression on the voltage time series, it is applied on characterization indicators that were derived from each voltage time series. Contrary to the method used in [28], there is no threshold for R^2. Thus, an electricity meter is automatically assigned to the transformer node phase with the highest R^2 value. The method was tested on Simulink simulation data of a power grid consisting of two transformer nodes with three phases each. Three or four electricity meters are connected to each phase, resulting in an overall number of 20 electricity meters. For the given simulation data, which has a time resolution of 15 min, the proposed method could correctly identify 100% of the topology. The time span of the simulation data was not stated in the study.

Only one of the investigated studies uses time series of energy measurements instead of time series of voltage measurements for the topology identification. The study proposes two algorithms that utilize principal component analysis. Algorithm one aims to identify the phase connectivity between electricity meters and phases of transformer nodes. Algorithm two is an extended version of algorithm one, which aims to identify the whole topology of the power grid – from the electricity meters over transformer nodes from all kinds of voltage levels up to the power supply with the highest voltage level. The algorithms were tested with synthetically generated data, which was generated by randomly (uniformly) sampling from certain pre-defined ranges of possible energy measurements. Different distances between the nodes and meters were randomly sampled and modeled, and random errors were introduced to account for losses and limited measurement accuracy. As soon as there are at least twice as many energy values from different points in time as there are transformer nodes or electricity meters, both algorithms perfectly fulfill their purpose by identifying 100% of the phase connectivity or topology. The main prerequisites for these algorithms are that there is no energy theft and that there are no un-metered loads in the power network [30].

2.4 Machine learning

Most of the methods used in the studies that are described in subchapters 2.1 to 2.3 utilize a machine learning algorithm or statistical approach that is related to machine learning.

Machine learning can be seen as a sub-field of artificial intelligence. It aims at designing algorithms that allow a computer to learn in order to fulfill a particular task – e.g., the prediction of values or the classification of new samples to certain predefined groups. There are several types of machine learning techniques, but as the objective of this thesis is a typical classification problem, this chapter focuses only on the type that is usually used for classification: the “supervised learning”. In “supervised learning”, an algorithm generates a function that maps input to desired outputs. The available compilation of input and output data is split into a training dataset and a test dataset. The training dataset is used to set up a model for the purpose of fulfilling the desired task, and the test dataset is then used to evaluate the model’s performance at fulfilling the task. This process of training and testing is usually repeated several times while fine-tuning the model [31,32].

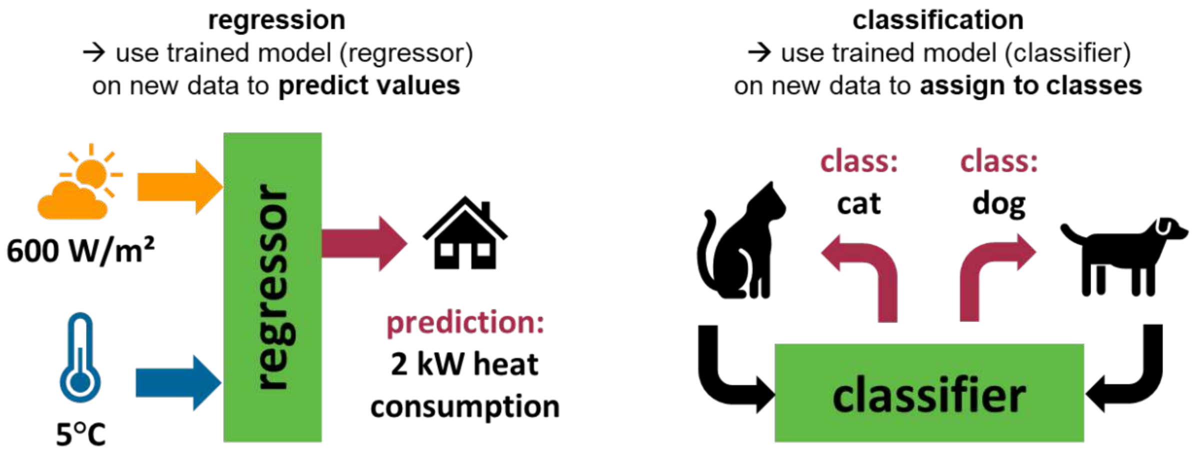


Fig. 7: The two main prediction problems, regression and classification, with examples.

Fig. 7 illustrates the two typical prediction problems that are addressed in “supervised learning”: the regression and the classification. In the regression case, a model is trained, which predicts continuous values, while in the classification case a model is trained, which predicts discrete classes. A model trained for the regression problem is called “regressor”. A model trained for the classification problem is called “classifier”. Depending on the number of classes that a classifier shall predict, the classification problem is either a “binary classification” or a “multiclass classification”. Binary classification indicates if an entry of data is either part of a class or not and thus classifies data into two classes. In the case of the multiclass classification, there are more than two classes [33].

As the objective of this thesis is to develop a method that can determine whether an energy meter is subordinate to another meter, the problem is a typical binary classification problem. This problem shall be solved by analyzing the time series data that is provided by the energy meters. For such time series classification problems, there are two types of methodologies: “feature-based” and “distance-based”. Since time series are usually of varying length, varying quality, and probably have gaps, the time series data must be processed before a classifier is trained.

In the case of the feature-based methodology, features – typically statistical indicators – are extracted from the original time series signal. After the feature extraction, the classifier is trained with the features instead of the time series. All new time series data that the classifier shall predict to classes also must undergo the feature-extraction process. Thus, the classifier never directly processes the original time series data. The basic idea behind that is that the classifier shall capture information hidden in the signal statistics that can be used to separate data into classes.

The distance-based methodology avoids the feature extraction process. Instead, suitable distances between time series signals are calculated and used as metric for separating data into classes. For this methodology to work, the time series signals either must be cut to the same time length, the time axis must be adjusted, or the signal must be transformed, e.g., to a parametric representation [34].

The distance-based methodology can only classify one signal at once because the distances are calculated between the time series signal that shall be classified and other time series signals that the classifier was trained with. However, the classification problem that shall be solved in this thesis involves two time series signals from two energy meters. As the features can be calculated from both signals, the feature-based methodology is the logical choice.

Basically, the feature-based methodology can be split into four steps: Preprocessing, Feature extraction, Dimensionality reduction, and Classification [34].

Embedding these four steps into the general workflow for classification of time series data, the whole classification process can be summarized as follows:

1. Define the problem: analyze the setting, specify the classes, and specify the cases that need to be classified
2. Algorithm selection: select a suitable feature-based classifier algorithm
3. Preprocessing: reduce measurement noise, e.g., outliers

4. Feature extraction: calculate features for all cases out of the time series data
5. Clustering: automatically or manually assign the entries to the classes
6. Split the dataset into training and test data: automatically or manually assign the entries to the training dataset or test dataset
7. Dimensionality reduction: remove the features that do not contribute to the classification
8. Classification: set up classifiers in different configurations
    a. (Repeatedly) train the classifiers with the training dataset
    b. (Repeatedly) test the classifiers with the test dataset
9. Model selection: evaluate the classification performance that was achieved with the test dataset and select the best classifier
10. Validation: validate the classifier performance with other data

There are several machine learning methods that could be used as feature-based classifier algorithm for the classification of time series data: Ranging from the application of the very fundamental statistical technique of regression, over the memory-based k-nearest-neighbor classifiers, to the computation-intensive neural networks [32]. In this thesis, only two of those methods were used: (i) linear regression of an indicator matrix (short: linear classifier) and (ii) random forest classifier.

The *linear classifier* is one of the most basic machine learning algorithms. It facilitates the well-established method of multivariate linear regression and adjusts it for the classification purpose. In most cases, other types of classifiers will outperform the linear classifier. Still, one of its main advantages is that its function is relatively easy to comprehend, as one simply has to analyze its coefficient matrix to understand the impact of the different parameters of the training data on the classifier model [32].

A *random forest classifier* is basically a set of decision tree classifiers that were trained with different datasets, which were created by randomly sampling data from the original training data. During the training of a decision tree, the parameters available in the training data are analyzed, and then they are used to partition the data with the aim that the different classes are separated as good as possible from each other. When a decision tree is fully grown, there are only entries of one of the classes in each partition. As a fully-grown decision tree always perfectly fits its training dataset, it is overfitted and might not fully grasp the underlying connections hidden in the data. To counteract that, several decision trees are grown from data that was bootstrapped (randomly sampled with replacement) from the original training dataset.

When tasked with the classification of new data, all decision trees are provided with the new data, and each of the trees assigns the data to one class. Finally, the class which was assigned by the majority of the trees is the classification output of the random forest classifier [32,35].

2.5 Research gap

The review of the literature showed that there are no examples of studies investigating the detection of structural and topological errors in building energy monitoring systems. Instead, several studies in the research field of fault detection and diagnostics for power systems address this topic. Since both building energy and power system monitoring process time series data related to energy measurement, the methods used for power systems should also be applicable to building energy monitoring. The main difference is that power system research focuses on the distribution of energy to buildings or single consumers, while building energy monitoring research focuses on the distribution of energy inside a building.

Tab. 1 gives an overview of the studies from the research field of fault detection and diagnostics for power systems that were discussed in section 2.3. They all investigate methods that shall detect relationships between supply nodes and/or electricity meters based on time series data from an electricity measurement. Most of the investigated methods have in common that they have a very high accuracy between 87.5% and 100%. Only one of them has an accuracy of 75%, which can still be interpreted as success.

Tab. 1: Study overview of methods that detect relationships between supply nodes and/or electricity meters.

<table>
<tr><th>source</th><th>processed time series</th><th>method allows for unmetered consumers</th><th>time resolution (sampling rate)</th><th>time series length (analyzed interval)</th><th>origin of the data that the method was tested with</th><th>correctly identified relationships between supply nodes and/or energy meters (accuracy)</th></tr>
<tr><td>[26] Zhang et al.</td><td rowspan="4">voltage measurements</td><td rowspan="4">yes</td><td>5 min</td><td>1 day</td><td rowspan="2">monitoring</td><td>87.5% - 93.5%</td></tr>
<tr><td>[27] Chao et al.</td><td>15 min</td><td>1 day</td><td>91.3%</td></tr>
<tr><td>[28] Qian et al.</td><td>50 µs</td><td>0.2 s</td><td rowspan="3">simulation</td><td>75.0%</td></tr>
<tr><td>[29] Liu et al.</td><td>15 min</td><td>not stated</td><td>100.0%</td></tr>
<tr><td>[30] Pappu et al.</td><td>energy measurements</td><td>no</td><td colspan="2">408 different points in time</td><td>100.0%</td></tr>
</table>

Four of the methods presented in Tab. 1 process time series data of voltage measurements [26–29], while only one processes time series data of energy measurements [30]. The processing of voltage measurements has the advantage that not only the profiles of the time series can be used to infer relationships between supply nodes and/or energy meters, but also the average voltage levels. The deeper the voltage level, the further the supply node/energy meter must be away from the electricity source.

There is no such advantage in the case of the method that processes only data of energy measurements. The detection of relationships between supply nodes and/or electricity meters must be conducted only based on the information hidden inside the profiles of time series. That this is more difficult than detecting relationships based on voltage measurements is indicated by the fact that the methods presented in [26–29] allow for some consumers to be unmetered, while the method of [30] does not.

As each consumer must be billed, the prerequisite that all consumers are metered should usually be met in power systems if there is no energy theft. In the case of building energy monitoring systems, there is a higher chance that consumers remain unmetered because sometimes only consumers of higher interest are metered, and a high number of meters increases the system's costs and complexity.

The reduction of costs and complexity is also the reason why in building energy monitoring systems, voltage, although measured by the electricity meters, is often not recorded. As the main goal of building energy monitoring is the observation of energy consumption, the recording of voltage is considered unnecessary. Moreover, the dynamics of the measured loads can be investigated by recording instantaneous power values.

Counter values and instantaneous power values are also the information that can be expected to be provided by heat meters. For heat meters, there is not even one study that investigates the hierarchical relationships between two heat meters.

A general method that can identify main meter-submeter relationships for both types of energy meters – electricity meters and heat meters – must work with the time series of counter values and eventually the time series of instantaneous power values. Thus, the methods that are based on the processing of voltage measurements [26–29] are not applicable. As the prerequisite that there are no unmetered consumers is often not met in building energy monitoring systems, the method presented in [30] is also not applicable. There is a need for a new method that processes time series of energy measurements and allows for some consumers to be unmetered – see Fig. 8.

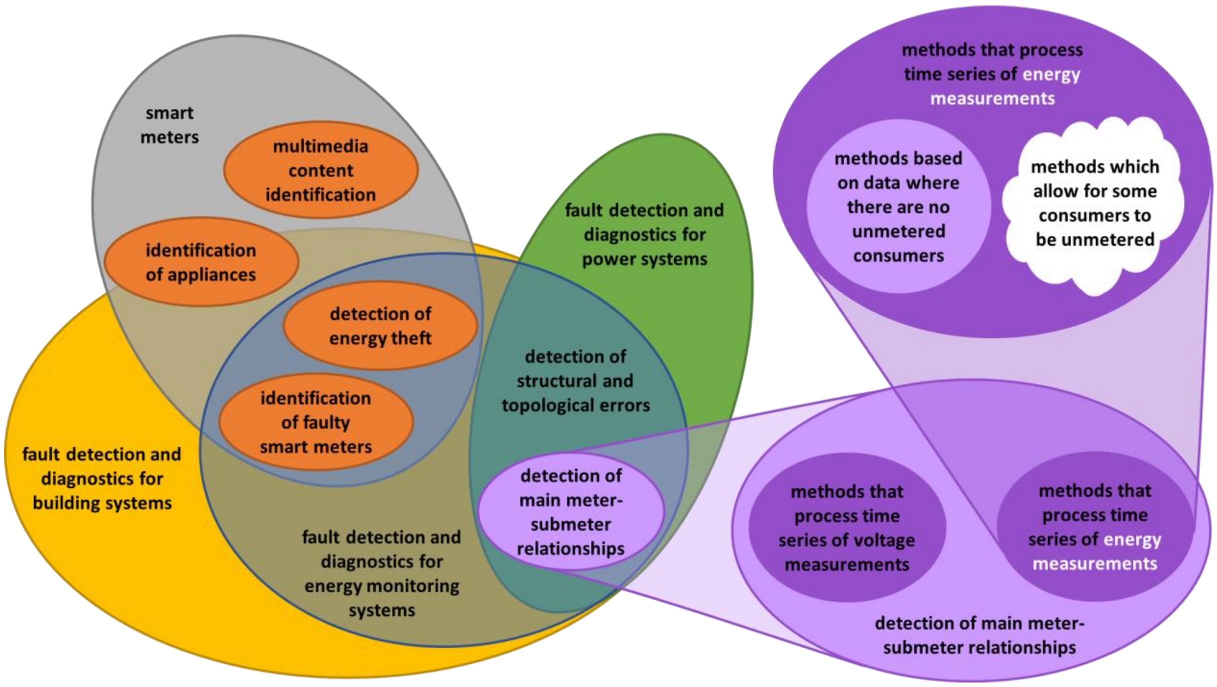


Fig. 8: Overview showing the topics related to the detection of main meter-submeter relationships with an illustration of the research gap.

Fig. 8 illustrates the research gap that needs to be filled by the new method by integrating it into the overview of the topics related to the detection of main meter-submeter relationships that was initially presented in Fig. 6. The detection of relationships between supply nodes and/or electricity meters is not separately shown for better readability. It is considered a part of “detection of main meter-submeter relationships”. The same applies to the detection of relationships between heat meters.

This doctoral thesis aims at filling the research gap. The task of automatically analyzing energy monitoring data to get an indication of whether an energy meter is subordinate to another meter is a possible application for machine learning algorithms. That machine learning is the right tool is indicated by the high accuracy rates achieved by the existing methods – see Tab. 1. As the task that needs to be handled by the new method is a typical classification problem, and as the data the method shall work with is time series data, a feature-based classification methodology is the most reasonable choice for the foundation of the new method.

3 Data sources

3.1 Case study

The (Plus-)Plus-Energy Office High-Rise Building is located near the center of Vienna. It is owned by the Austrian federal real estate company Bundesimmobiliengesellschaft (BIG) and used by TU Wien. The high-rise building is part of a building complex from the 1970s that was refurbished between 2012 and 2014 – see Fig. 9 for photographs of the building before and after the refurbishment.

Fig. 9: Case study building "(Plus-)Plus-Energy Office High-Rise Building" before and after the refurbishment.

The building complex consists of several parts with different usages, the main parts being: (i) offices, (ii) lecture halls, (iii) library, and (iv) event hall. The term “(Plus-)Plus-Energy Office High-Rise Building” only refers to the office part of the building complex [36]. The office part is designed as a net zero energy building. This means the building’s annual electricity and heat production on site is at least equal to the annual energy consumption of the building. The building was equipped with an extensive building energy monitoring system to substantiate whether the planned building performance could be achieved in reality and to aid with the commissioning of the building [37].

During the commissioning of the energy monitoring system itself, several errors were found in the system. As described in section 1.1, these errors were the main motivation for the development of the method presented in this thesis. Moreover, the energy monitoring system provided the necessary data for the development and validation of the method.

For better readability, the (Plus-)Plus-Energy Office High-Rise Building is referred to as “case study” in the following sections of this work.

Fig. 10 gives an overview of the case study’s energy monitoring system. It illustrates the different components of the system and how they are connected to each other.

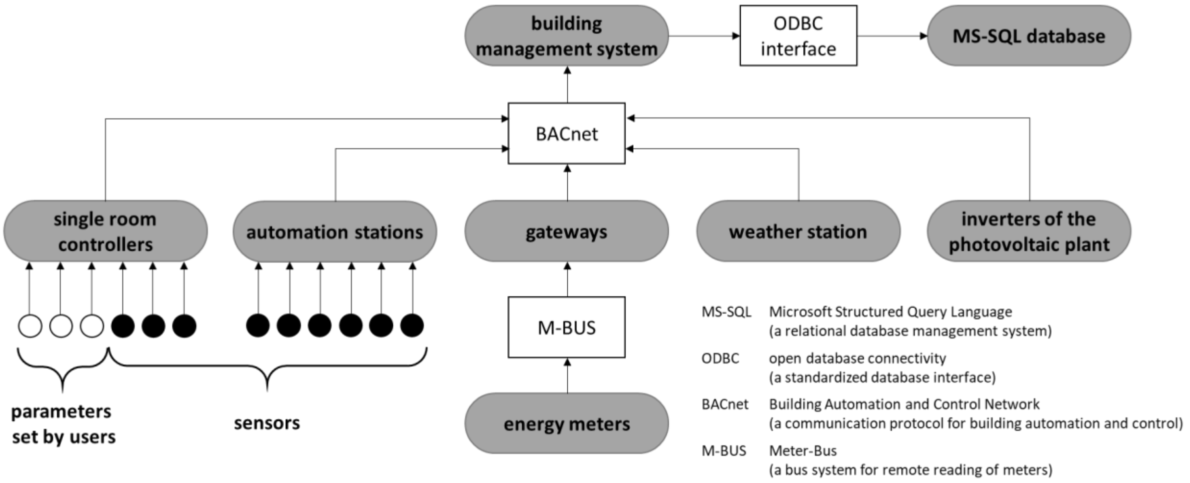


Fig. 10: Scheme of the case study's energy monitoring system.

As shown in Fig. 10, the energy monitoring system is interconnected with the building management system. Most of the collected data comes from energy meters but also selected data from the single room controllers, the automation stations, the weather station, and the inverters of the photovoltaic (PV) plant is collected. All components are connected to the building management system, which receives their data and then transfers it to the MS-SQL database.

Including the integrated energy meters of the photovoltaic plant inverters, the total number of energy meters amounts to 150. Tab. 2 shows this number broken down.

Tab. 2: Overview of the case study’s energy meters.

number	type of energy meter
28	heat meters with one counter
4	heat meters with two separate counters
96	electricity meters with one counter
3	electricity meters with two separate counters
19	electricity meters with one counter integrated into the inverters

Each of the 150 energy meters provides at least the counter value and the instantaneous power value of the measured energy. Some of the electricity meters also provide information about the instantaneous current values. Only four electricity meters are more advanced meters that offer detailed information about each of the measured phases. Those four are also the only meters that record information about the measured instantaneous voltage values. All measurement values that are used by the heat to calculate the counter value and the instantaneous power value internally are recorded, i.e., volume flow, volume count, flow temperature, and return temperature.

The total sum of the data points that originate from the electricity meters and heat meters is 537. Further 1,159 data points originate from other sources in the building, e.g., the single room controllers or the weather station. All data points are recorded in a 5 min time resolution.

Not considering gaps in the data, the 537 data points from the electricity and heat meters deliver 154,656 daily measurement values.

Fig. 11 illustrates the quantitative quality of the energy monitoring data by displaying the number of daily measurement values from the electricity and heat meters over time. Especially during 2015, 2016, but also 2017, there are several larger gaps in the data. Most of them were caused by software updates and bugfixes of the building management system. The building management system sometimes failed to reconnect to the MS-SQL database automatically. After 2017 there was only one larger gap in 2019. Some of the energy monitoring's data points were added in the first quarter of 2017, which is indicated by the slightly less saturated daily values before that. As the monitoring data of 2020 and 2021 is not processed yet, the heatmap ends at the beginning of 2020.

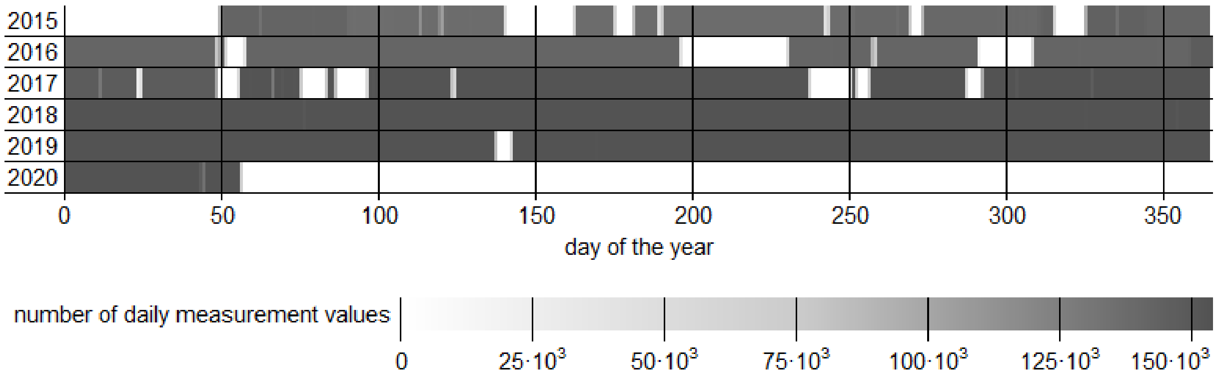


Fig. 11: Heatmap depicting the quantitative quality of the case study energy monitoring data.

Fig. 12 depicts the energy meter hierarchy of the case study's energy monitoring system. It shows the energy meters of each hierarchy level and their main meter-submeter relations. Meters that have two separate counters are displayed as two separate meters, e.g., the eight meters in the category "heat meters (heating & cooling)" are actually only four real meters.

The electricity meters on hierarchy level 1 that are connected to the PV supply are the integrated meters of the PV inverters. Thus, they are not consumers but suppliers whose combined electricity supply is metered by the "PV supply" meter on hierarchy level 0. All other electricity meters on hierarchy level 1 are supplied by three sources: the main supply, the emergency supply, and the PV supply. As the PV supply simply reduces the supply from the other two sources, its connections to the consumer electricity meters on level 1 are neither displayed nor considered a real main meter-submeter relation.

The connections between the server waste heat recovery and the cooling and heating systems are not displayed as well, and they are also not considered real main meter-submeter relations. This is because the server waste heat recovery simply reduces the cooling consumption of the server room while simultaneously reducing the heating consumption of the underfloor heating.

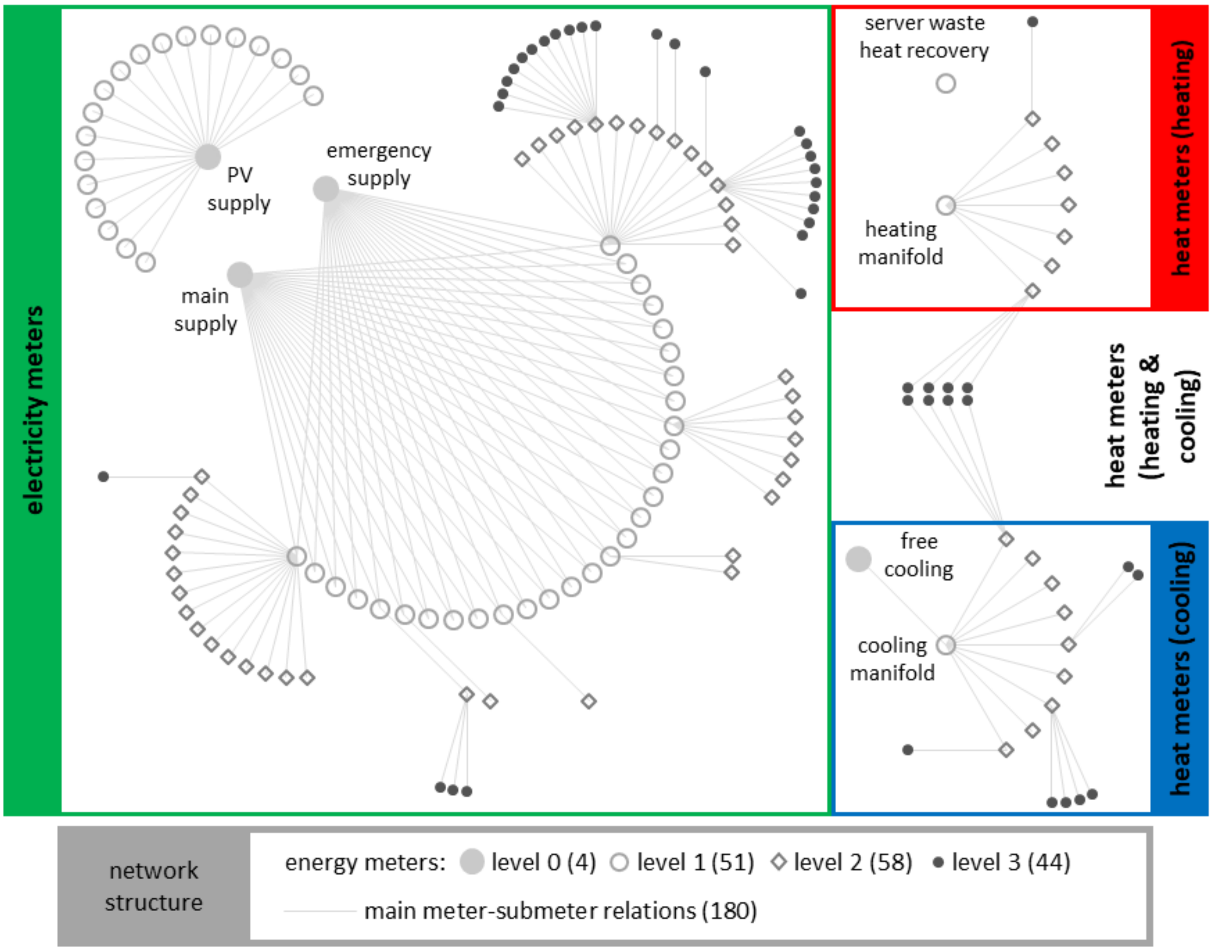


Fig. 12: Network graph depicting the case study's energy meter hierarchy.

A more detailed depiction of the energy meter hierarchy can be found in appendix A. There the energy meters are labeled with their short designation, indicating the measured energy consumers or energy supply.

Fig. 13 illustrates selected statistical data of the load profiles that are measured by the case study's energy meters: (i) the mean and (ii) the relation of the standard deviation to the mean, the coefficient of variation (COV). While the mean indicates the magnitudes of the energy flows measured by the meters, the COV shows the extent of the variability of these energy flows. The data is differentiated according to a meter's level in the hierarchy and according to the meter type – electricity meter or heat meter. 19 meters which did not measure consumption at all, i.e., have a mean load of 0 W, are not shown in the figure.

Fig. 13 shows that the higher the mean load, the smaller the COV. Energy meters on a higher hierarchy level measure higher mean load, as is to be expected because they are often the superior meters of meters with a lower level. The graph indicates that the mean loads of the meters on hierarchy levels 1 to 3 span over several magnitudes: from near zero (less than 10 W) up to more than 10 kW. For the case of the meters on level 3, the measured maximum mean load is approximately 15 kW; for the meters on level 2, approximately 50 kW; and for meters on level 1, approximately 130 kW.

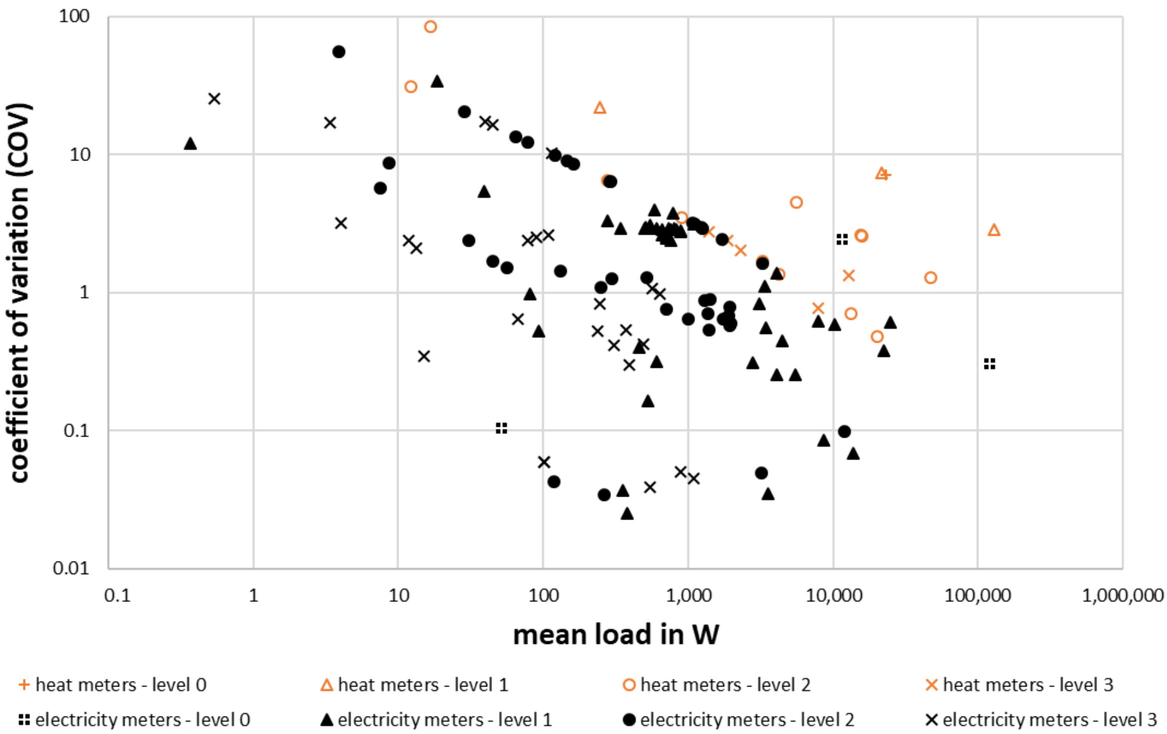


Fig. 13: Illustration of the statistical data (mean load and coefficient of variation) of the load profiles measured by the case study's energy meters. Basis for the graph: the monitoring data of 2018 in 5 min time resolution; load profiles calculated out of the first order difference-quotient of the counter values.

Compared to electricity meters, heat meters show a trend to load profiles with higher mean loads and higher COV values. This can be explained by the case study building having less submetering for heat meters than for electricity meters, e.g., only one meter measures the entire heat consumption of floors 5 to 10. In contrast, in the case of the electricity meters, each floor is metered separately. The reason for the higher COV values is the relatively low counter resolution of the installed heat meters. Given the case study data with a time resolution of 5 min, this often leads to load profiles that jump between few values, in the worst case, only between zero and one other value. This issue can also be observed in the load profiles of some electricity meters, which leads to elevated COV values.

Another reason for high COV values is that often meters feature low to no consumption but regularly measure sudden spikes in the consumption. For example, the load profiles of the case study's electrical under-counter water boilers in the accessible toilets showed such characteristics.

Meters with a COV below 0.1 measured a practically constant consumption with less than ten noticeable spikes with a duration below 5 min during the whole year.

3.2 Synthetically generated data

To ensure that the method that is developed in this thesis does work on data besides the monitoring data from the case study, another data source is needed for its validation. As the method shall work with real monitoring data, the second data source should be real monitoring data or at least replicate the attributes of real monitoring data.

The open data platform “Open Power System Data” offers the data package “Household Data” (OPSD household data), which encompasses real energy monitoring data from 68 electricity meters in a time resolution of 1 min. This data is one of the outputs of the EU research project “CoSSMic” (Collaborating Smart Solar-Powered Microgrids)” that was conducted between 2013 and 2016 [38].

Tab. 3: Overview of the 68 electricity meters included in the “OPSD household data” dataset.

category	description	code	industrial			public		residential					
			i1	i2	i3	p1	p2	r1	r2	r3	r4	r5	r6
supply	battery	s_b		X									
	public grid	s_g	X	X	X	X	X	X	X	X	X	X	X
	photovoltaic plant	s_pv1	X	X	X			X		X	X		X
		s_pv2	X		X								
export	public grid	e_g								X	X		X
consumption	battery charging	c_b		X									
	circulation pump	c_cip							X	X			X
	cooling aggregate	c_coa			X								
	compressor	c_com			X								
	cooling pumps	c_cop			X								
	dishwasher	c_d			X			X	X	X	X	X	X
	electrical vehicle	c_ev			X						X		
	freezer	c_f						X	X	X	X		X
	heat pump	c_hp						X			X		
	industrial or research machine	c_m1			X								
		c_m2			X								
		c_m3			X								
		c_m4			X								
		c_m5			X								
	several smaller loads in an office area	c_o			X								
	several smaller loads in a room	c_r1			X								
		c_r2			X								
		c_r3			X								
		c_r4			X								
	refrigerator	c_ref			X					X	X	X	
	ventilation	c_v			X								
	washing machine	c_wm						X	X	X	X	X	X

Tab. 3 gives an overview of the 68 electricity meters that are included in the OPSD household data dataset. Even though the dataset is labeled “household”, it also includes the monitoring data from industrial and public buildings. Regardless of the type of building, from each building, there is at least data from the main electricity meter that counts the electricity that is supplied by the public grid. From seven buildings, there is also data from the energy meter for each building's PV plant. The industrial building "i3" is monitored in detail – there is data from 20 electricity meters that measured different rooms, machines, and system components in the building. In the case of the residential buildings, most of the meters measured the electricity consumption of typical household appliances, e.g., dishwashers and washing machines.

For each of the electricity meters, there is only one time series, namely the time series of the electricity that was consumed within the 1 min time interval – basically the first difference quotient of an electricity counter value. In the case of the 20 electricity meters of the industrial building "i3", the total number of daily measurement values amounts to 28.800 – given that there are no data gaps during the day. As each of the public buildings only has one meter, the expected number of daily measurement values amounts to 1.440 for each building.

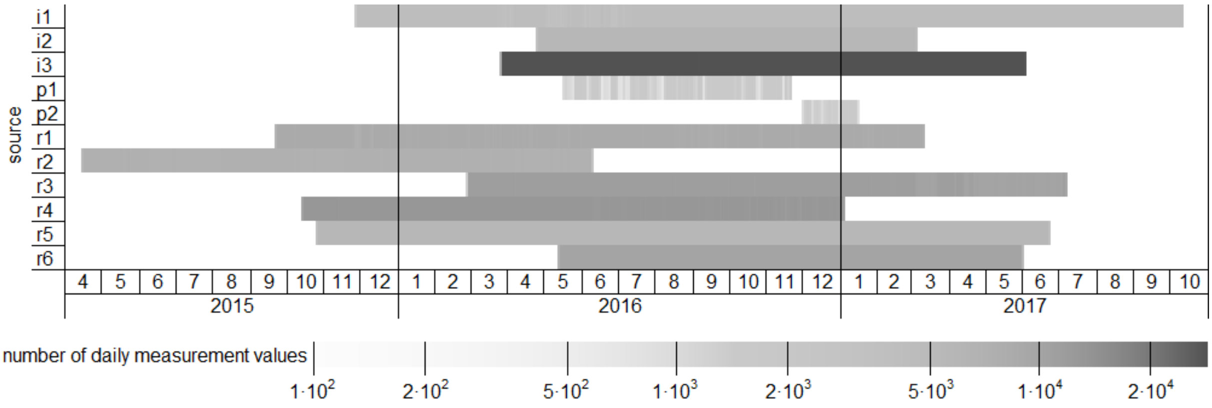


Fig. 14: Heatmap depicting the quantitative quality of the "OPSD household data" dataset.

Fig. 14 illustrates the quantity of available data points as the number of daily measurement values from the electricity meters over time. It also shows the time span during which the data was collected in each of the buildings. The almost constant shading of the industrial buildings i1-i3 and the residential buildings r1-r6 indicates that there were practically no gaps during the investigated time spans. On the other hand, the public buildings p1 and p2 had several gaps, at least during some of the days.

Since the OPSD household data lacks hierarchical levels and instead of electricity counter values and instantaneous power values, it only offers the first difference quotient of an electricity counter value, the data cannot be used directly for the validation of the method. The OPSD household data was therefore subjected to a transformation process to generate synthetical data that fits the following criteria:

- The time resolution of the data should be the same as the time resolution of the case study data, namely 5 min. Time resolutions with larger time intervals can be calculated out of the 5 min data later.

- Time series data of counter values and the instantaneous power values of the meter should be available for every meter.

- The underlying load profiles that are used to calculate the counter values and instantaneous power values of each meter shall have a realistic course. This means they must not be a series of randomly sampled values.
- There should be as least as many energy meters as there are in the case study.
- The energy meters shall be distributed over the same number of hierarchical levels as in the case study. This shall be done in a realistic way, i.e., energy meters that count a large consumption should not be placed on levels where there would usually only be small consumers.

The generation of synthetical data is conducted in two parts: (i) random synthetic load profiles are created, and (ii) a random energy meter hierarchy is generated. Fig. 15 illustrates the first part and Fig. 16 the second one.

As shown in Fig. 15, the first step is the preprocessing of time series data of the original OPSD household data. Within this step, outliers are removed, small gaps with one missing value are filled by linear interpolation, and larger gaps are filled with the last value before the gap. At the end of the preprocessing, the data of each building is stored in a separate data frame. The second step is the random selection of one of the building data frames, from which a time series (load profile) is then randomly selected in the third step. Within the fourth step, it is randomly decided whether the time series order of the load profile is to be reversed or not. The resulting load profiles are then used in the fifth step to create synthetic load profiles with a time span of one year.

Fig. 15 shows that there are two options to create the synthetic load profiles in the last step. For the first option, whole days of the load profile that is available after step 4 are selected at random and then strung together to form a combined new synthetic load profile. With this option, it is ensured that the daily cycle of the load profile is not lost, and therefore the synthetic load profile has a realistic course.

In the case of the second option, time slots with a random length of between 20 min and a whole day are selected at random and then strung together to form a combined new synthetic load profile. The idea behind this option is to add more randomness while keeping the local structure of the selected time slots intact.

Steps 2 to 5 are then repeated 300 times to generate 300 random synthetic load profiles with a time span of a whole year and a time resolution of 1 min. 150 of those profiles are created according to option 1 and the other 150 according to option 2.

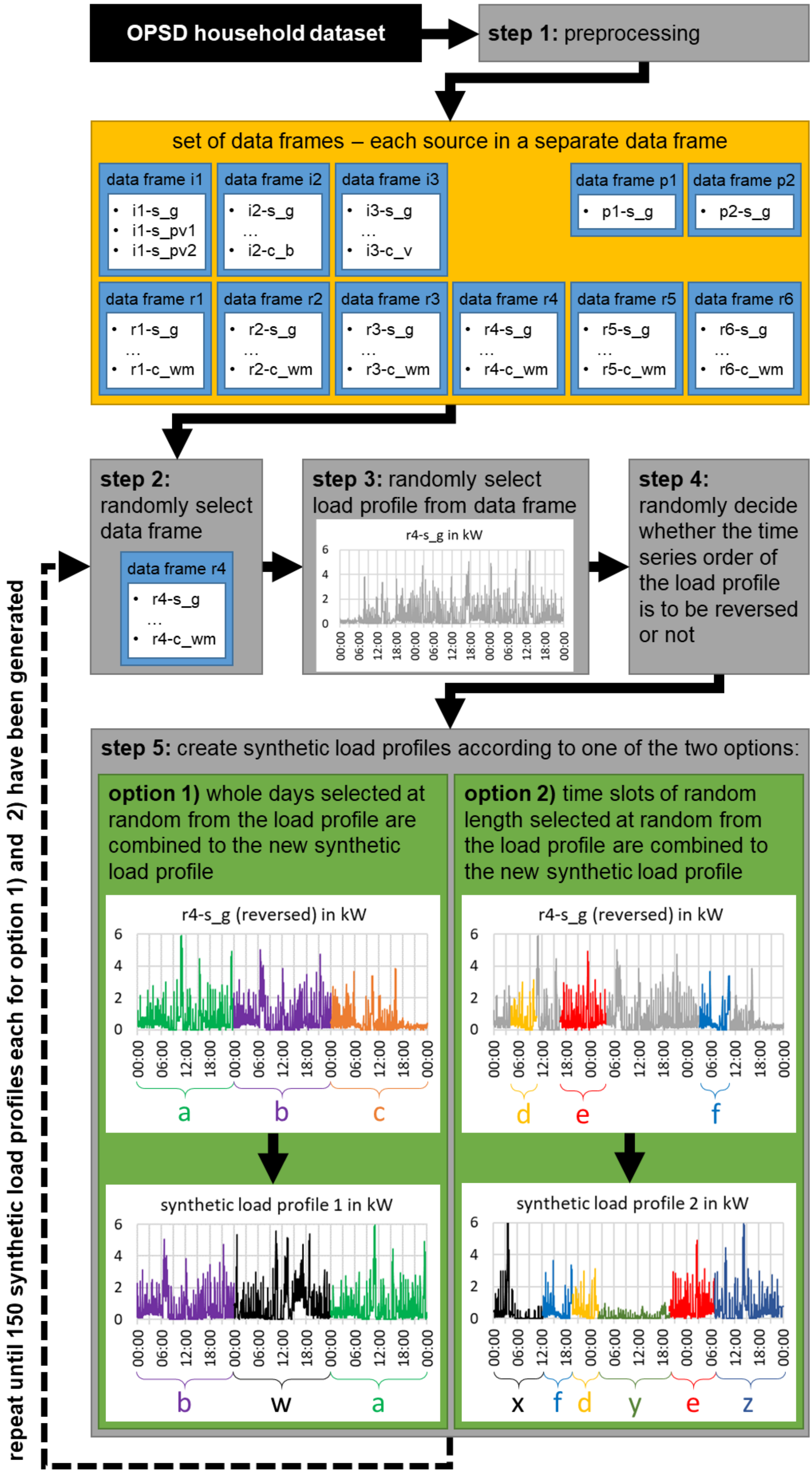


Fig. 15: Steps of the process to generate 300 random synthetic load profiles.

This pool of 300 random synthetic load profiles is then used as the basis for the process of the generation of a random energy meter hierarchy – see Fig. 16. First, each profile is assumed to be the load profile that is measured by an electricity meter – neglecting that there might be measurement errors and further inaccuracies. As the origin of the load profiles is real electricity meters, the data was already affected by measurement errors and inaccuracies. The fact that the synthetic load profiles have a time resolution of 1 min while the goal is to generate data with a time resolution of 5 min is exploited for the calculation of the 5 min time series data that meters would measure:

- For the calculation of the time series of one meter's instantaneous power values, the values of the 1 min synthetic load profile at the 5 min time interval are extracted. Thus, these values are indeed the instantaneous values of the load profile.
- For the calculation of the time series of one meter's counter values, the 1 min synthetic load profile is cumulated, and out of this cumulated profile, the values at the 5 min time interval are extracted. If the first-order difference quotient were to be calculated out of these counter values, the resulting time series would equal the average power values that occurred between the 5 min time intervals. Thus, this time series is slightly different than the time series of the instantaneous power values – generally, it is less volatile.

At the end of the first step, for each of the 300 random synthetic load profiles, there is a virtual energy meter (electricity meter) featuring time series data of a counter and instantaneous power values for a time span of one whole year with a time resolution of 5 min.

In the second step, the virtual energy meters are sorted into three groups according to their mean loads: (i) a mean load below 15 kW, (ii) a mean load between 15 kW and 50 kW, and (iii) a mean load above 50 kW. These limits were chosen according to the maximum mean loads in the different hierarchy levels of the case study meters, as they can be observed in Fig. 13. Those observations also influenced the assignment of the load profiles to the hierarchy levels of the random energy meter hierarchy, as depicted in Fig. 16 in the third step: load profiles with mean loads above 50 kW were only assigned to hierarchy level 1, load profiles between 15 kW and 50 kW to hierarchy level 1 and 2, and load profiles below 15 kW were assigned to all three hierarchy levels.

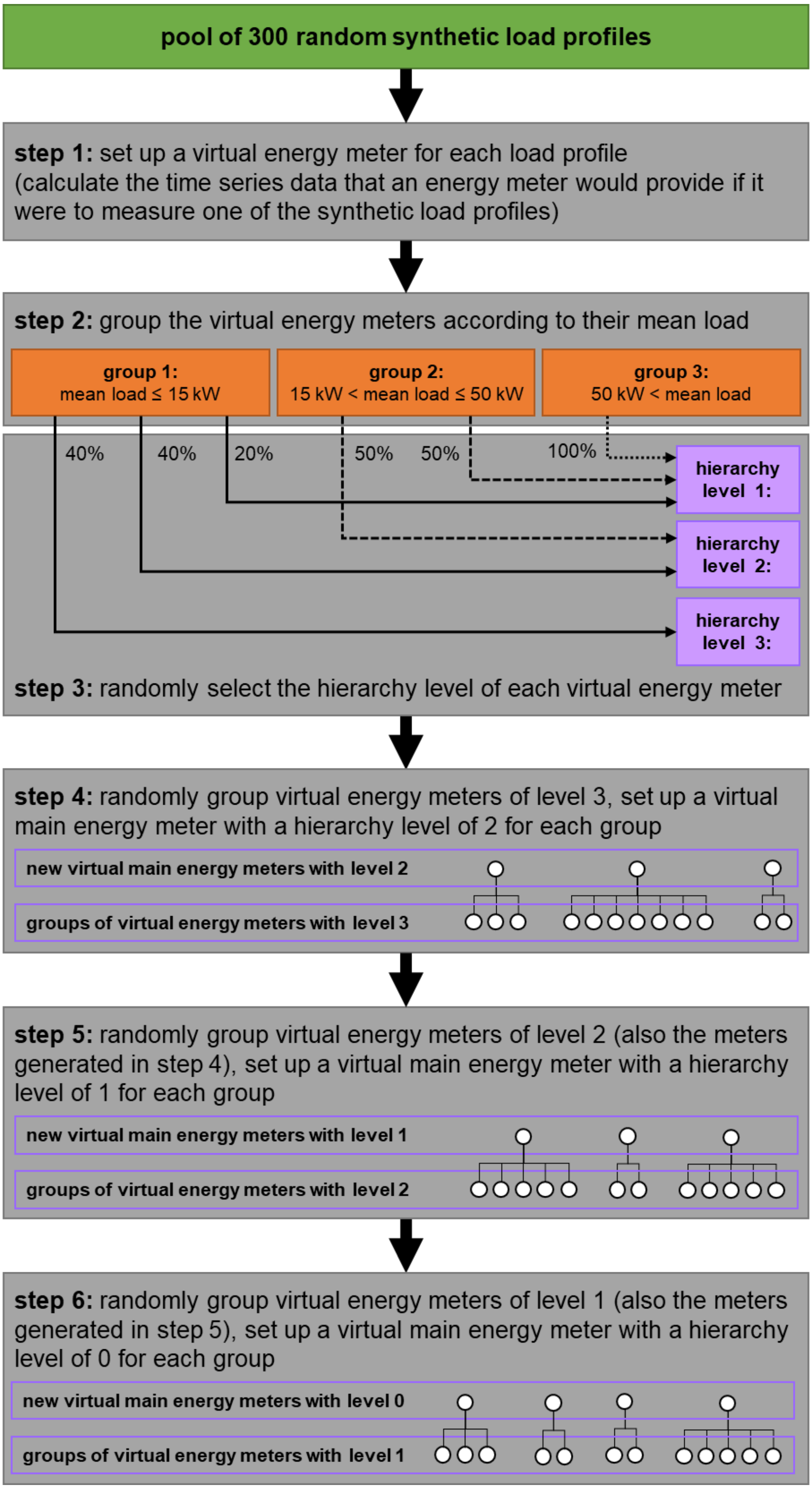


Fig. 16: Steps of the process to generate a random energy meter hierarchy.

The structure of the case study's energy meter hierarchy (see the network graph in Fig. 12) was used as a role model for the random grouping in steps 4 to 6, shown in Fig. 16. For each of the hierarchy levels of the case study meter hierarchy, the maximum number of meters that share the same main meter was determined. This value times two and then rounded to a multiple of five was then used as the maximum random group size for the random grouping in the random energy meter hierarchy:

- In step 4, the virtual energy meters assigned to hierarchy level 3 were randomly grouped in groups of 1 to 25 meters. By adding up the time series of the counter values and instantaneous power values of the meters in each group, the counter values and instantaneous power values of a virtual energy meter on hierarchy level 2 were created.
- In step 5, the virtual energy meters assigned to hierarchy level 2 were randomly grouped in groups of 1 to 30 meters. By adding up the time series of the counter values and instantaneous power values of the meters in each group, the counter values and instantaneous power values of a virtual energy meter on hierarchy level 1 were created.
- In step, 6 the virtual energy meters assigned to hierarchy level 1 were randomly grouped in groups of 1 to 50 meters. By adding up the time series of the counter values and instantaneous power values of the meters in each group, the counter values and instantaneous power values of a virtual energy meter on hierarchy level 0 were created.

At the end of the process of generating a random energy meter hierarchy, there are 322 virtual meters that are arranged in the random structure depicted in the network graph in Fig. 17.

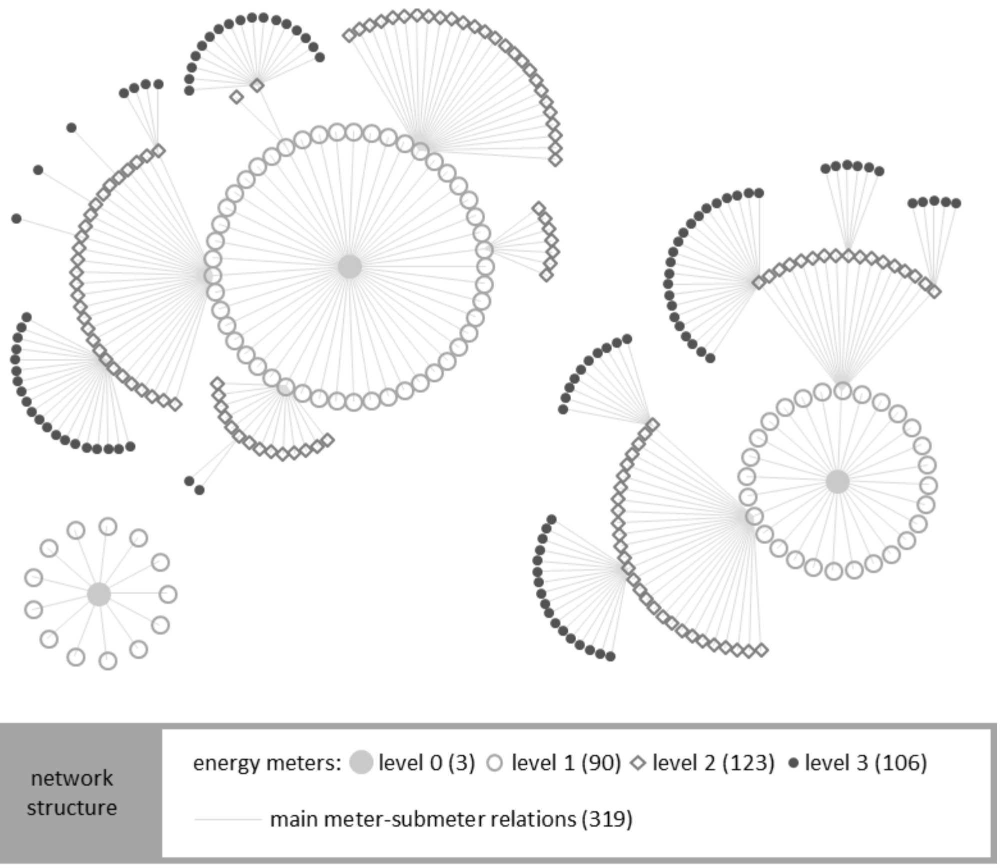


Fig. 17: Network graph depicting the randomly generated energy meter hierarchy.

Since the OPSD household data only consists of electricity meters, the virtual energy meters are all electricity meters as well. There are approximately twice as many virtual energy meters in this random hierarchy than there are real energy meters in the case study hierarchy (see Fig. 12). To keep the structure of the random hierarchy simple, it was set up in a way so that each virtual energy meter has only one main meter. Further, it was also refrained from integrating energy sources into the randomly generated energy meter hierarchy.

As the process of generating a random energy meter hierarchy uses the characteristics of the case study monitoring system as a role model, the characteristics of the synthetic load profiles should be like the characteristics of the case study load profiles.

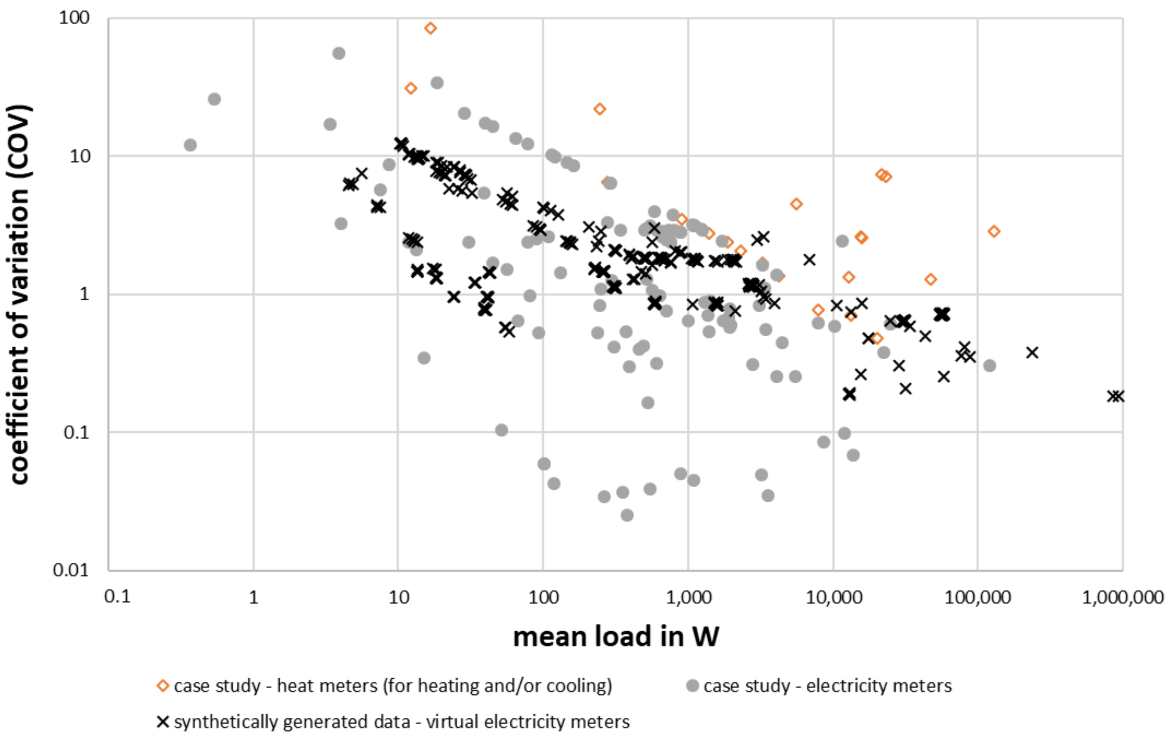


Fig. 18: Comparison of statistical data (mean load and coefficient of variation) of the load profiles measured by the case study's energy meters and the load profiles of the virtual energy meters in the randomly generated energy meter hierarchy.

Fig. 18 shows the comparison of the statistical data of the case study's energy meters and the statistical data of the virtual electricity meters, the synthetically generated data. It shows that the statistical data of the synthetically generated data fits well within the range that was observed in the case study. There is a tendency towards the area in the middle – there are no virtual electricity meters with a COV below 0.18 and none above 12.5. Since the load profiles of the case study's meters with a COV value below 0.1 showed an almost constant course, this area is not of interest for a method that uses data hidden in the course of load profiles to infer the meter hierarchy. As in the case study higher COV values were often caused due to too low counter resolutions, the area above the points of the synthetically generated data is also deemed as negligible for the method developed in this thesis.

4 Method

As stated in section 1.2, the objective of this thesis is the development of a method that uses solely the monitoring data of energy meters to evaluate whether a meter is subordinate to another meter. Considering the state of the art elaborated in section 2, utilizing a machine learning algorithm appeared to be the right choice to tackle the objective. Since determining whether a meter is subordinate to another meter is a typical classification problem and because the monitoring data is time series data, the development of the method basically follows the general workflow for time series data classification as laid out in section 2.4. Fig. 19 shows the steps taken to develop the method, the corresponding chapters where the steps are described, and how they are related to the general workflow for time series classification.

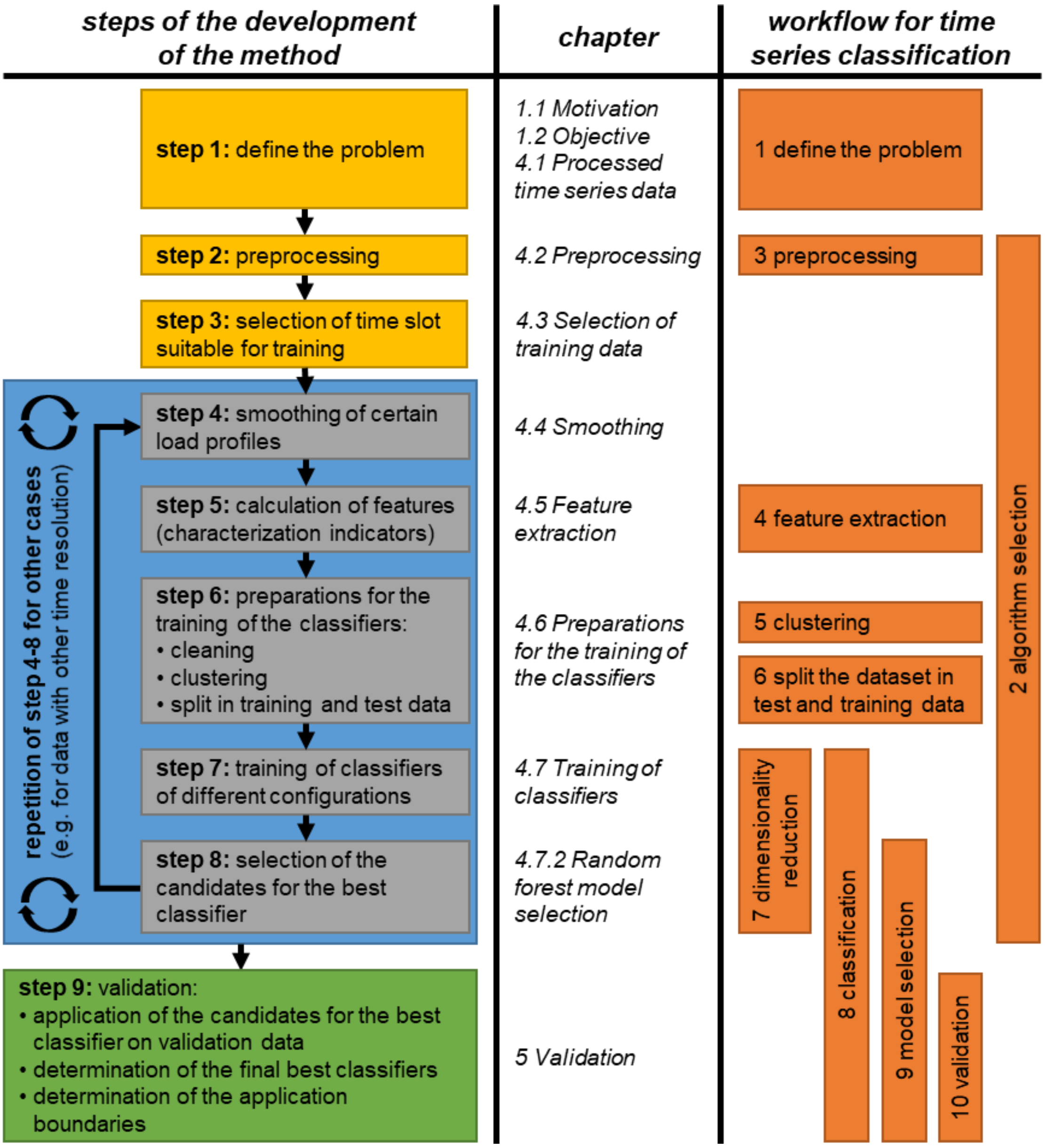


Fig. 19: Overview of the steps taken to develop the method, the corresponding chapters, and the relation to the general workflow for time series classification.

As illustrated in Fig. 19, the development of the method partially differs from the general workflow for time series classification. Different processes of the workflow extend over several development steps: the processes “algorithm selection”, “dimensionality reduction”, “classification”, and “model selection”.

The process “algorithm selection” extends over steps 2 to 8 because the method was initially developed to work with a linear classifier: a linear regression of an indicator matrix. Since such a linear classifier is relatively easy to comprehend and as its coefficient matrix indicates how different input parameters influence the classification, the linear classifier was deemed the right choice for the first development steps. After processing the data to be suitable to be used by the linear classifier, the data indicated that a tree-based method like a random forest should be better suited for the classification task at hand. Thus, the method was then developed to work with a random forest classifier.

The switch to the random forest classifier happened at an early stage of method development. As the random forest classifier promised better prediction performance than the linear classifier, the linear classifier was not investigated further. Nevertheless, the linear classifier influenced the choice of features that were extracted from the time series data. This is discussed further in section 4.5 “Feature extraction”.

Within the 9 steps of the development of the method, there are several challenges that needed to be overcome:

- Determination of the necessary preprocessing of the raw energy data. The data must be cleaned of outliers while ensuring that the information hidden in the time series is kept intact.
- Calculation of suitable features that can be used as input for the classifier. Features are metadata derived from the meter’s time series data. In the given case, the goal was to identify features that indicate the presence or absence of a main meter-submeter relation between two energy meters.
- Identification of strategies for:
    - clustering,
    - splitting the data into training and test data,
    - dimensionality reduction, and
    - the selection of the best classifiers.

This thesis does not address measurement errors or inaccuracies of the energy meters themselves. The method was developed from the point of view of a person, which did not actively participate in the design and installation of the energy monitoring system, but instead only received monitoring data from the system.

4.1 Processed time series data

The presented method aims at using the information contained in the monitoring data of energy meters to assess whether an energy meter is subordinate to another energy meter. This section provides an overview of the available time series data that is used as the basis for this assessment.

Regardless of the type of energy meter – heat meter or electricity meter – all of the case study's meters provide time series data of the meter's counter value and the instantaneous power value. Fig. 20 shows an example of the time series data of the counter value.

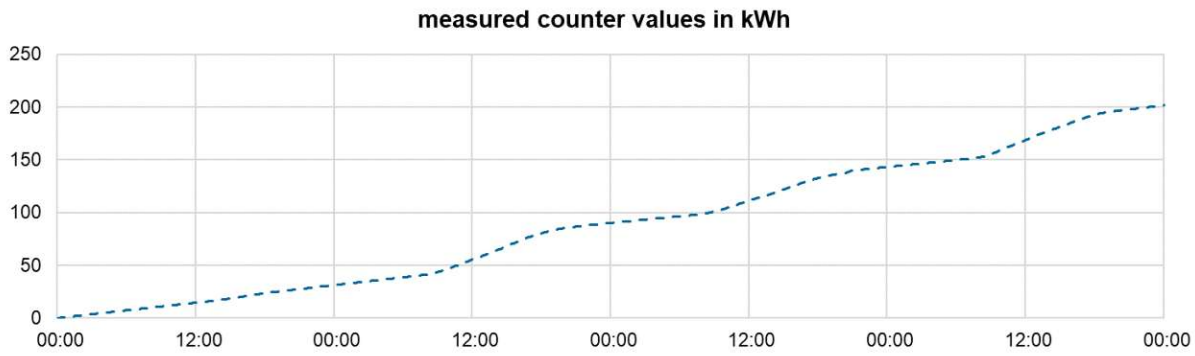


Fig. 20: Example of a typical time series of counter values of energy meters.

Provided that the counter of one meter either exclusively records energy consumption or energy supply, the curve of the counter value should be a monotonically non-decreasing function. Fig. 20 depicts such an expected curve. In the case that the meter is parameterized to count in the negative direction, the function should be monotonically non-increasing – i.e., the depicted example curve would be mirrored along the x-axis. Deviations from this monotonic behavior are usually measurement errors. Only if the counter counts energy flow in both directions – consumption and supply – the curve would be not monotonic.

Fig. 21 shows the time series data of the instantaneous power value of the same meter, whose counter value is illustrated in Fig. 20. The curve follows a pattern that indicates daily cycles. With the additional knowledge that the first day presented in the figure is a Sunday, it could be inferred that the daily cycles might differ from day to day – which is indeed true, as the measured consumer is the electricity consumption of one office floor.

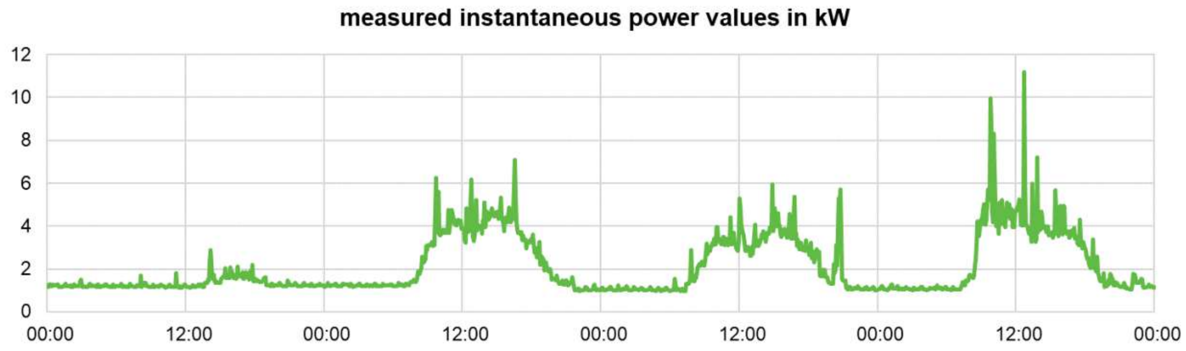


Fig. 21: Example of a typical time series of instantaneous power values of energy meters.

The curve shows a base consumption that is never undershot – even during the night. During the last three days (Monday, Tuesday, and Wednesday), the curve indicates a consumption pattern that would be expected to be found in a typical office: an increase in power consumption in the morning, during the day the power oscillates around a plateau, and in the afternoon the consumption drops to the base consumption. The spikes indicate larger consumers that were switched on for a short period of time, e.g., hotplates in the office kitchen.

As the instantaneous power values are the values that were recorded during the time interval of the monitoring system (in this case, it is a 5 min interval), the connecting lines between the measured values are just an approximation. In between the measurement time interval, the load profile could be very different. Therefore, the curve of the measured instantaneous power values should be considered an approximation of the true load profile.

Another approximation of the true load profile, which is at least true to the total energy that was consumed between two measurement points, is the first-order difference quotient of the time series of the counter values. Fig. 22 shows the first-order difference quotient of the curve of Fig. 20 atop the curve of Fig. 21. The first-order difference quotient curve basically follows the same course, but it appears smoother; the spikes are generally smaller.

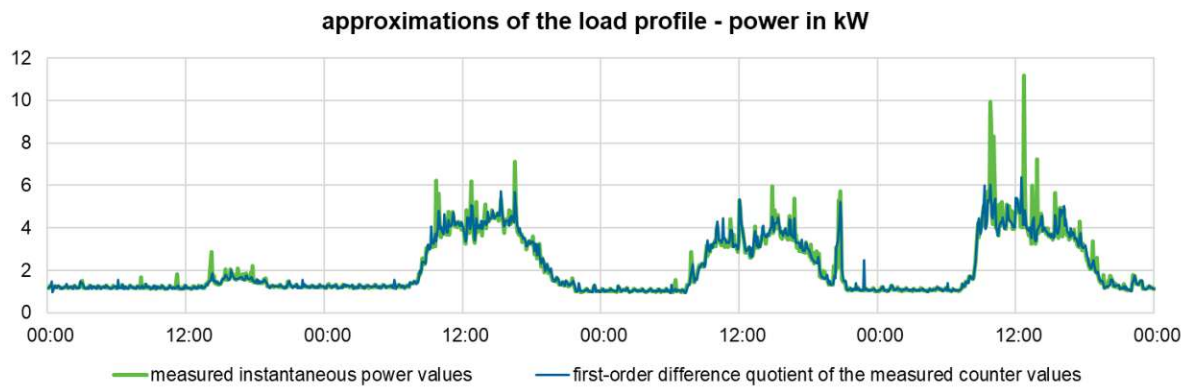


Fig. 22: Comparison of the time series of instantaneous power values and the timeseries of the first-order difference quotient of the counter values – both are approximations of the real load profile.

The two approximations of the load profile illustrated in Fig. 22 – one calculated out of the counter values and one being the measured instantaneous power values – are the two data sources that are utilized by the method.

Since the energy measured by one meter must also pass through a superior meter, changes of the load profile should also be observable in the load profile of the superior meter. The search for similarities in the changes of load profiles is the main principle of the developed method. Thus, it is evident that the method will struggle to determine the superior meter for the case that subordinate meters measure load profiles without notable changes. Moreover, measurement errors or the consumption of other consumers, which are supplied by the same superior meter, also pose a challenge for the search for similarities.

As the data provided by energy meters might vary, e.g., not include time series data of the instantaneous power values or not be in the time resolution of 5 min, other common cases are considered. This is done by resampling the available 5 min time series data to time resolutions with a larger interval. Tab. 4 gives an overview of the different cases and time resolutions that were evaluated.

Tab. 4: Overview of the evaluated cases.

utilized data sources	time resolution of the training data
counter values and **instantaneous power values**	5 min
	10 min
	15 min
	30 min
	60 min
only **counter values**	5 min
	10 min
	15 min
	30 min
	60 min

Almost all features are calculated for both the load profile derived from the counter values as well as the curve of the measured instantaneous power values. To highlight the underlying data source used for the feature calculation, the following codes are used as tokens and prefixes in this work:

- **c_d1** – The curve of the first-order difference quotient of the counter values of an energy meter: the load profile in W.
- **c_d2** – The curve of the second-order difference quotient of the counter values of an energy meter: the changes of the load profile in W/(5 min).
- **p_m** – The curve of the measured instantaneous power values: the approximation of the load profile in W.

- **p_d1** – The curve of the first-order difference quotient of the measured instantaneous power values: the changes of the approximation of the load profile in W/(5 min).

As later described in section 4.5.2, a filter was applied to c_d2 and p_d1, which is indicated by extending the prefix by the letter “t”:

- **c_d2t** – The curve of the second-order difference quotient of the counter values of an energy meter after applying a filter: the changes of the load profile in W/(5 min) after applying a filter.

- **p_d1t** – The curve of the first-order difference quotient of the measured instantaneous power values after applying a filter: the changes of the approximation of the load profile in W/(5 min) after applying a filter.

4.2 Preprocessing

Before the time series data can be used for time series classification, the data must be preprocessed. Irregularities and obvious measurement errors must be eliminated without compromising information deducted from load profile change information.

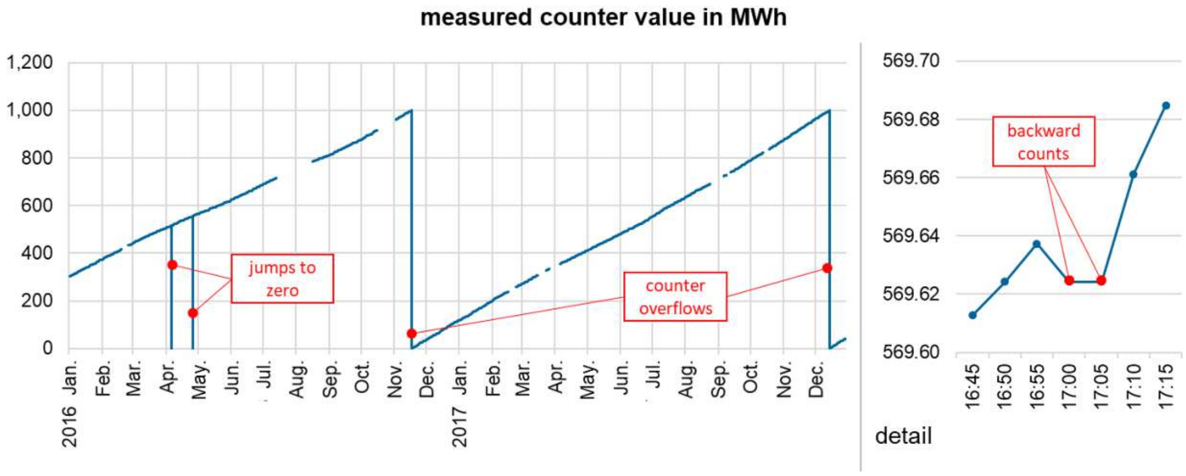


Fig. 23: Time series of the counter values of the case study’s main electricity energy meter illustrating the typical errors and occurrences.

Fig. 23 illustrates three of the issues that were encountered in the raw time series data of some energy meters:

- Sudden jumps to zero: The counter value suddenly dropped to zero for one or more consecutive points in time and then continued its course. This could be attributed to communication issues between the meters and the building management system.

- Counter overflows: The counter value dropped from a value just under 1.000 MWh to almost zero and started continuously increasing again from the new level. Such a counter overflow is a common occurrence that happens when the maximum value of the meter's counter is exceeded. In the case of the monitoring data of the case study, up until the beginning of 2020, counter overflows have only been observed in the electricity meters that measure the main supply and the meter that measures the entire PV supply.

- Backward counts: As practically all the meters only count consumption or supply, their counter values should show monotonically non-decreasing or non-increasing curves. Sometimes the curves deviate from this behavior – the counter counted in the other direction. It is suspected that these occurrences were due to measurement inaccuracies or communication issues between the meters and the building management system.

Another issue that can be observed in Fig. 23 is the gaps in the curves. These data gaps were already discussed in section 3. Since the method is developed to operate with information hidden in the changes of the load profiles, gaps are not an obstacle if there is enough time series data that can provide the needed information.

For the jumps to zero and the backward counts, the solution is to set the values of the affected data points to "NaN" and thus create artificial gaps.

The solution for the counter overflows is to shift all data points after an overflow by 1.000 MWh. When a counter had multiple overflows, the data points past those overflows are subject to multiple shifts.

Fig. 24 shows the same time series that is illustrated in Fig. 23 after correcting the discussed issues: Counter overflows were corrected by shifts, and the jumps to zero and backward counts were cut out of the data creating artificial gaps.

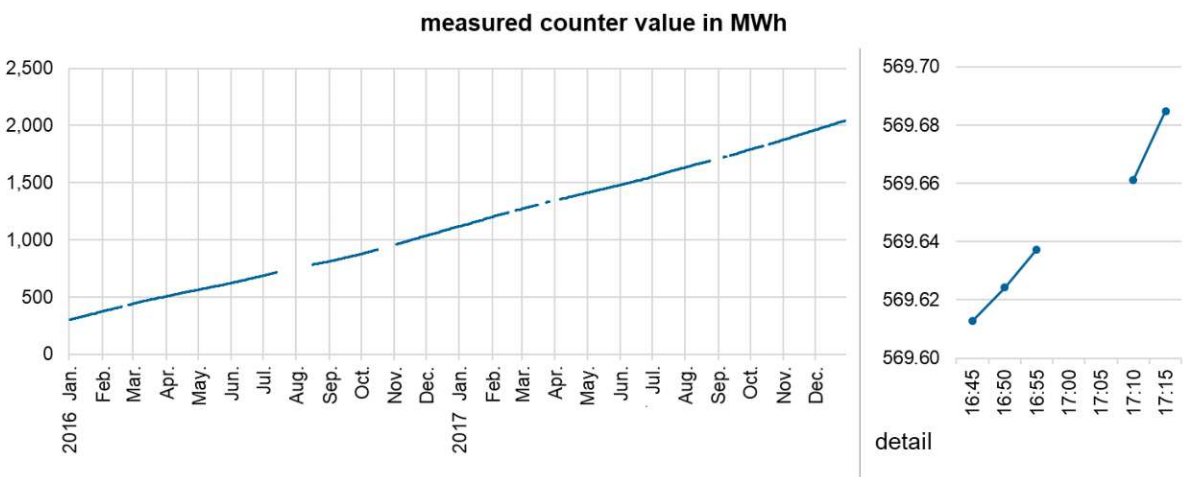


Fig. 24: Time series of the counter values of the case study's main electricity energy meter after the preprocessing.

The solutions presented in this section were applied to the entire available time series of the counter values of all energy meters. The analysis of the instantaneous power values showed that there are no obvious outliers. If there are some errors within the values, they seem to be within the range of the usual values. All the algorithms and data manipulations that are described in the following sections were only applied to data from specifically selected time slots.

4.3 Selection of training data

Updates, bugfixes, and the resulting communication issues during the first years of the case study's energy monitoring were the main reasons for the gaps in the energy monitoring data. Only after the final updates in the fall of 2017 the energy monitoring system operated continuously without larger errors. In the data of the year 2018, there are basically no gaps. Thus, this year was reserved as the source for the first part of the validation of the method – see section 5.2. Five consecutive weeks of 2017 without data gaps were selected as the source for the training data for the classifier development. Fig. 25 highlights both selected time slots.

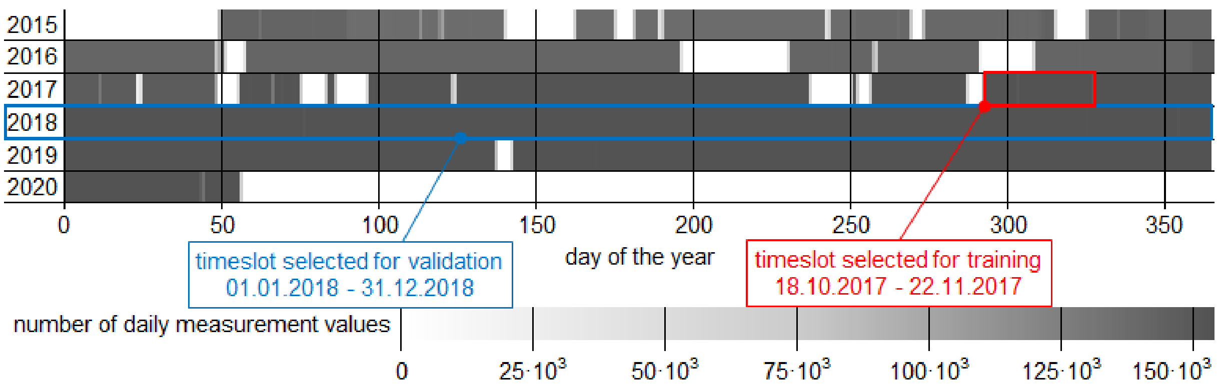


Fig. 25: Heatmap depicting the quantitative quality of the case study monitoring data and the time slot which was deemed suitable for training of classifiers.

4.4 Smoothing

Several misclassifications indicated that the low resolution of some meters – mainly heat meters – were the reason for the misclassifications. To cope with that, a gaussian filter was used on the load profiles, which were calculated as the first-order difference quotient of the counter values. After analyzing all load profiles, a simple ruleset was set up to decide whether a load profile needs smoothing or not.

For the first step, the load profiles were manually clustered into two groups: One group with low-resolution load profiles that need additional smoothing and a second group where smoothing is unnecessary.

In the next step, for each load profile, the changes of the load profile – the second-order difference quotient of the counter values – were calculated, and then the following metadata was derived:

- **c_d1_unique**: Number of unique values occurring in the load profile.
- **c_d2_all**: Number of all data points of the changes of the load profile.
- **c_d2_jumps**: Number of jumps of the changes of the load profile. In this case, a jump is the occurrence of a change of the load profile where the next following data point indicates a change of identical magnitude but in the other direction.
- **c_d2_zero**: Number of data points of the changes of the load profile that have a value of zero.
- **c_d2_relation_to_std**: The difference between the maximum value of the changes of the load profile and the minimum value of the changes of the load profile in relation to the standard deviation of the changes of the load profile.

Finally, by analyzing the metadata of the two groups, the ruleset, which is depicted as the decision tree in Fig. 26, was developed. This ruleset was applied to all load profiles derived from the counter values. The load profiles which were put into the group “smoothing necessary” were then smoothed by a Gaussian filter with a sigma of 3.

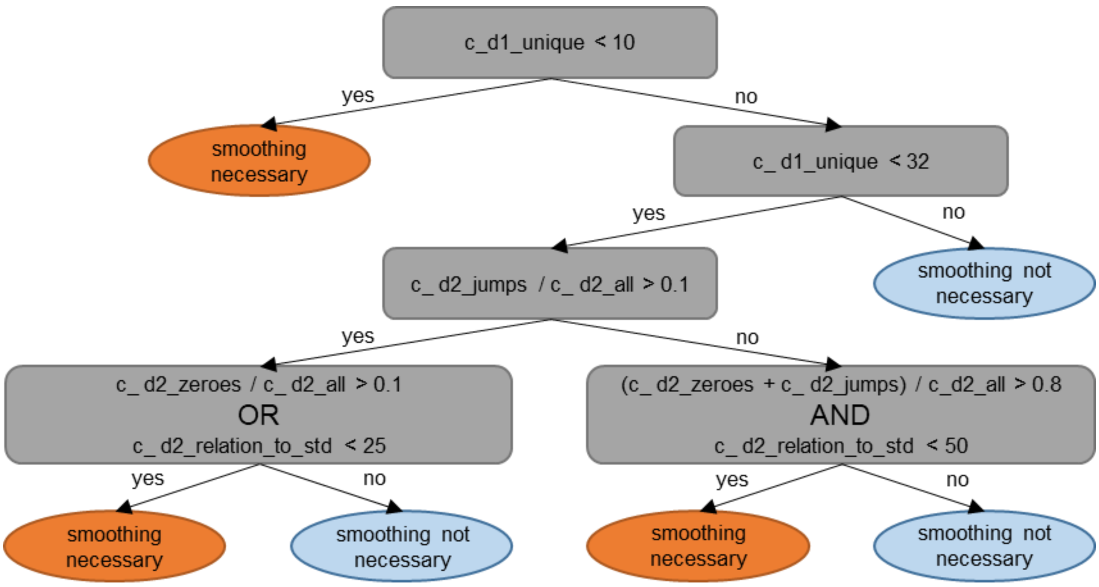


Fig. 26: Decision tree for determining whether smoothing is necessary or not.

In all following steps of the method, the smoothed load profiles were treated the same way as the unaltered load profiles from meters where no smoothing was necessary. The smoothed load profiles replaced the according original load profiles, and thus, their features (characterization indicators) were calculated out of the smoothed profiles.

4.5 Feature extraction

As the classifier cannot process the energy meter's time series data directly, suitable metadata must be derived from the time series data. The derived metadata can be seen as characterization indicators. One single indicator is commonly referred to as "feature" and several indicators as "features."

The energy monitoring data of the case study contains data of 150 energy meters. As seven energy meters feature two counters, the data from the counters are stored in 157 data points. The data from the instantaneous power values measured by the meters are stored in 150 data points.

Assuming that there is no further information except which measured instantaneous power values relate to which energy counter, the 157 counters with their corresponding instantaneous power values can be viewed as 157 individual energy meters where seven pairs of them have identical instantaneous power values.

There are 157 x 156 = 24,492 possibilities to combine these 157 energy meters. Each of these combinations represents a situation where one energy meter is assumed to be the superior main meter and the second one is assumed to be the subordinate submeter.

For all these possible 24,492 combinations, the features described in the following subchapters are calculated for both the load profiles derived from the energy counters as well as the load profiles approximated by the measured instantaneous values.

In this thesis, the general workflow for selecting the features is to first calculate several different features that might somehow reflect whether there is a main meter-submeter relation and subsequently reduce the number of features by step-wise backward selection. This diminishes the need for a detailed analysis per feature, as it can be assumed that features that do not carry relevant information for the classification are filtered out automatically.

4.5.1 Features derived from the load profiles

As first step for the identification and decision on features, the cases that shall be separated are compared. Since the classification problem is a binary one, there are two cases: (i) the case where there is a real main meter-submeter relation, and (ii) the case where there is no main meter-submeter relation. Examples for both cases are illustrated in Fig. 27 and Fig. 28.

Fig. 27 shows the load profiles of two energy meters where there is indeed a main meter-submeter relationship. The load profile of the main meter is always larger than the profile of the submeter. The difference between those profiles is energy that is consumed by other consumers – possibly metered by other submeters or unmetered.

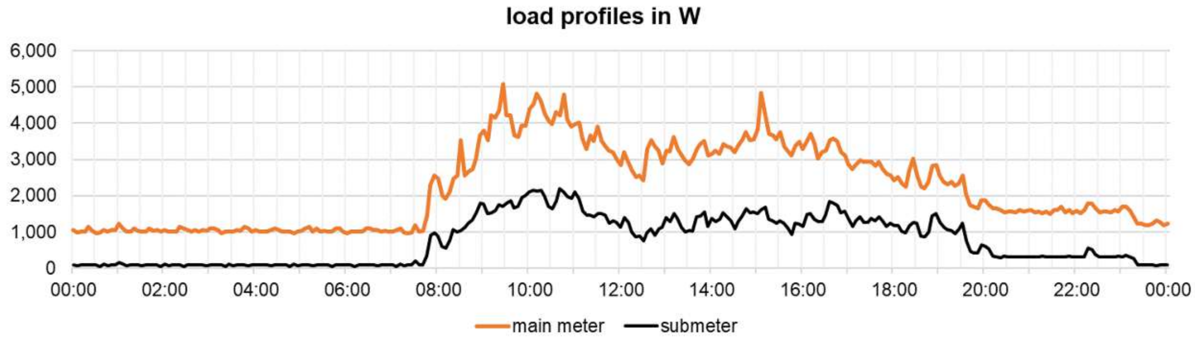


Fig. 27: Load profiles of two meters where there is a real main meter-submeter relation between both meters.

Changes in the load profile of the submeter can also be seen in the profile of the main meter, but the changes do not perfectly match. Measurement inaccuracies and errors aside, the other consumers are the reason for this mismatch. Spikes in the profile of the main meter, which are not visible in the profile of the submeter, are caused by the other consumers. Moreover, spikes in the profile of the submeter, which are not visible in the profile of the main meter, are hidden due to a drop in energy consumption of other consumers that occurred at the same moment.

Fig. 28 depicts the load profiles of two energy meters with no main meter-submeter relationship. There is no restriction regarding the level of the load profile – for any given point in time, any profile can be larger than the other profile. There are also no similarities in the changes of the profiles of both meters, or similarities are coincidental and possibly caused by other causal connections, e.g., similar time schedules of the consumers.

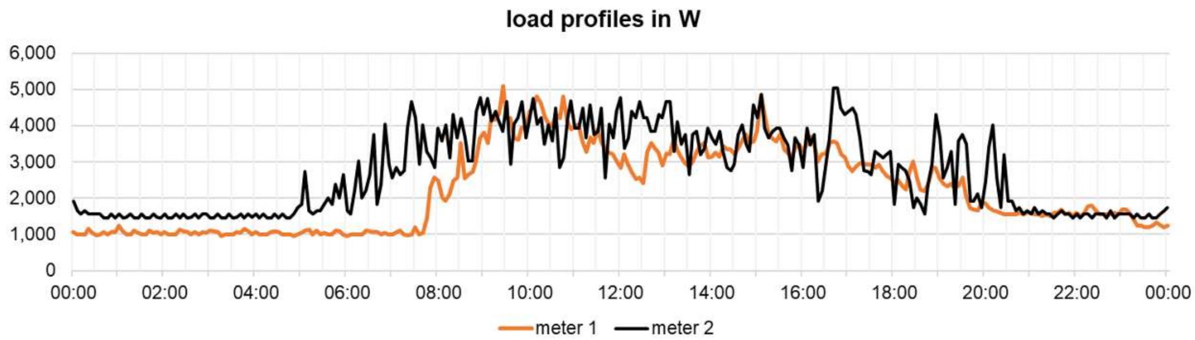


Fig. 28: Load profiles of two meters where there is no main meter-submeter relation between meter 1 and 2.

Following the observation that the load profile of the main meter should generally be larger than the load profile of the submeter, the first feature was defined as follows: The relation of the means of the load profiles, expressed by the symbol r_{pos} or the code **profile_relation**. With the load profile of the assumed main meter as f_{mm} and the profile of the assumed submeter as f_{sm}, the feature is calculated as shown in (1).

$$r_{pos} = \frac{\overline{|f_{mm}(t)|}}{\overline{|f_{sm}(t)|}} \quad (1)$$

When the assumed submeter is indeed a submeter of the assumed main meter, this indicator should yield values above 1. Values below 1 are usually a strong argument against the suspected main meter-submeter relation, except for when energy is supplied to the main meter via another source parallel to the submeter. As such cases are uncommon in single buildings, no special routines were developed to handle those cases.

The means' absolute values are used to account for the possibility that an energy meter might count in the negative direction. This alteration will affect the meaningfulness of r_{pos} only for the case that a submeter counts energy supply while the main meter predominantly counts consumption. But since such a submeter could theoretically be subordinate to every other meter, the information lost due to the alteration is deemed negligible.

The basis for the second feature is the curve of the differences of the load profiles $f_{mm} - f_{sm}$. Fig. 29 depicts an example of how such a difference profile could look like when the load profiles are not related and have a similar magnitude but a different course.

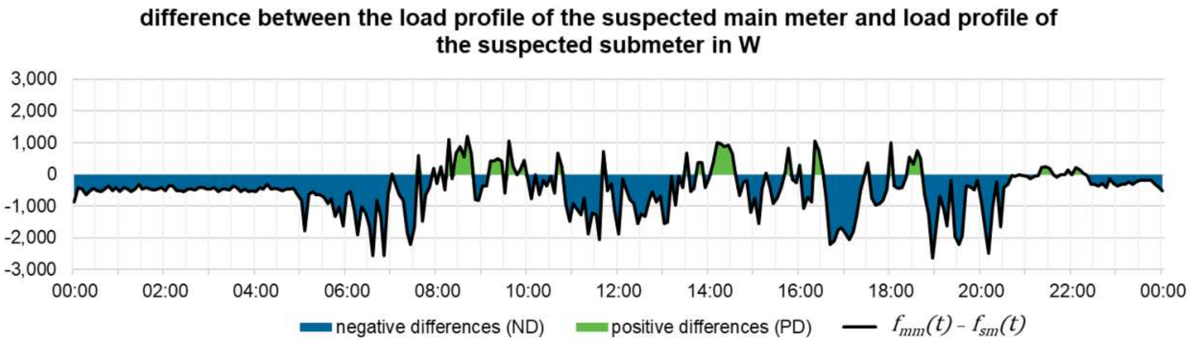


Fig. 29: The difference profile calculated by subtracting the load profile of the suspected submeter from the load profile of the suspected main meter.

The second feature, r_{neg} or **profile_relation_neg**, is calculated as the sum of the negative profile differences to the sum of all profile differences. The simplified formula for this calculation is shown in (2) and the corresponding formula featuring f_{mm} and f_{sm} in (3).

$$r_{neg} = \frac{ND}{PD + ND} \tag{2}$$

$$r_{neg} = \frac{1}{2} \frac{\sum_{t=0}^{n} |f_{mm}(t) - f_{sm}(t)| - f_{mm}(t) + f_{sm}(t)}{\sum_{t=0}^{n} |f_{mm}(t) - f_{sm}(t)|} \tag{3}$$

For correctly identified submeters, this feature should yield the value 0 or – to account for measurement inaccuracies – values slightly above 0. Values significantly above 0 are usually a strong argument against the suspected main meter-submeter relation.

In the case that there is a submeter that counts in the negative direction, the curves of f_{mm} and f_{sm} are interchanged to be also applicable for the case of two meters counting in the negative direction. This alteration impacts the case where a submeter counts energy supply while the main meter predominantly counts consumption so that the calculated value will always be exactly 1.

4.5.2 Features derived from the changes of the load profiles

Similar load change behavior is a strong indication of a main meter-submeter relationship. When calculating the time series of the changes of the load profiles that were depicted in Fig. 27, the result is the curves shown in Fig. 30.

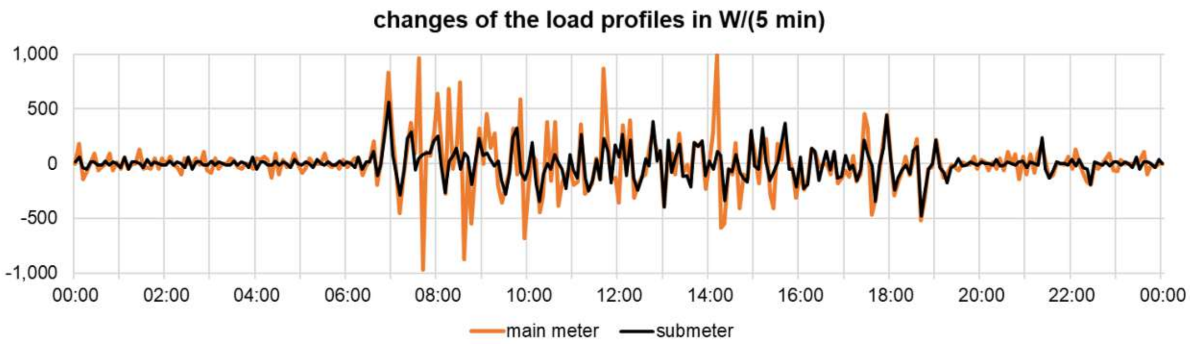


Fig. 30: The changes of the load profiles of two energy meters displayed over time.

Given that a main meter meters the only source of supply of the consumers that are metered by a submeter, every load change of the submeter's load should result in a similar load change of the main meter. As a main meter will most certainly register several submeters and consumers, the change of load in the main meter will be distorted. Moreover, distortions will also be caused by measurement uncertainties.

For cases with only minor distortions, it is possible to see a linear relationship between the load change of both meters, and the Pearson correlation coefficient should reflect this relationship. For the calculation of the Pearson correlation coefficient, the pre-defined function "pandas.DataFrame.corr" of the Python package "Pandas 1.0.1" [39] was used. This function was also utilized to calculate the Kendall tau rank correlation coefficient and the Spearman's rank correlation coefficient.

The mentioned three correlation coefficients are chosen as features. The symbols and code used to refer to them in this thesis are:

- Pearson correlation coefficient: symbol ρ, code **pearson**
- Kendall correlation coefficient: symbol τ, code **kendall**
- Spearman correlation coefficient: symbol r_s, code **spearman**

In cases where there is indeed a main meter-submeter with only minor distortions, all three features should tend towards a value of 1. For all other cases, these features are expected to tend towards a value of 0.

Further features are computed after applying an ordinary least squares regression on the scatterplot of the changes of the load profile of the assumed main meter against the changes of the load profile of the assumed submeter. If a meter is really a subordinate of another meter and its profile is distinctive enough, the ordinary least squares regression should yield an intercept of approximately 0 and a slope near 1. Fig. 31 depicts two examples calculated with the case study's monitoring data: On the left side, a case where there is no relationship between the two analyzed meters, and on the right side, a case where there is indeed the main meter-submeter relationship.

Fig. 31: Load profile change behavior of two energy meters displayed as scatterplots. Left: Exemplary with no submeter-main meter relation between meter 1 and 2. Right: Exemplary case with a real main meter-submeter relation between the two analyzed energy meters.

As indicated by the metrics of the example displayed in Fig. 31, in both cases, the R^2 value and the intercept value of approximately 0 do not seem to carry much information regarding the main meter-submeter relationship. Nevertheless, their values might still be relevant when combining them with other indicators.

From the least squares regression, the metrics R^2, slope, and intercept are then chosen as features. The symbols and code used to refer to those three features in this thesis are:

- R^2: symbol R^2, code **r_squared**
- slope: symbol β, code **slope**
- intercept: symbol ε, code **intercept**

To emphasize the impact of situations where the assumed submeter had a large load profile change, the data points that portray the large change and the corresponding value of the load profile change of the assumed main meter were filtered out and each stitched together to form sequences of peaks. To be registered by the filter, the change in the load profile of the submeter has to exceed a certain threshold. This threshold was defined as a quarter of the difference between the maximum value and the minimum value of the curve of the submeter's changes of the load profile. Fig. 32 shows the changes of the load profiles of an exemplary case of a real main meter-submeter relationship with the depiction of the corresponding submeter thresholds. Circles highlight the data points where the submeter values are outside the thresholds.

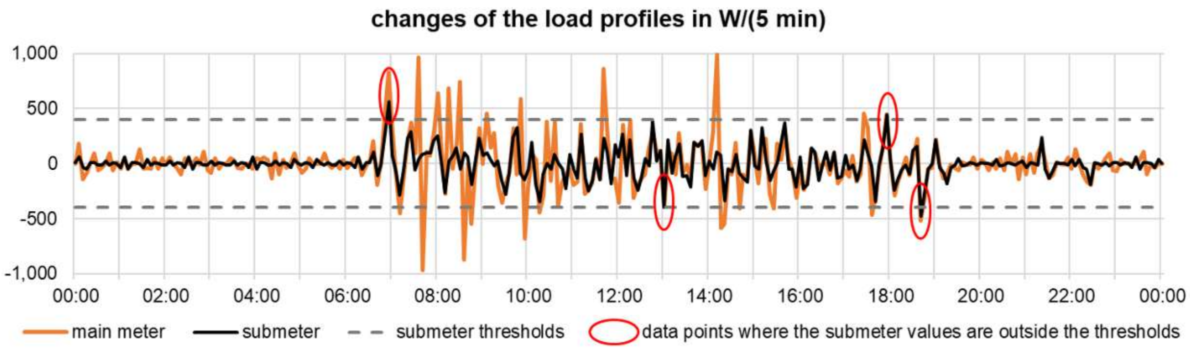


Fig. 32: The changes of the load profiles of two energy meters displayed over time including the depiction of the thresholds.

Fig. 33 shows the resulting sequences after stitching together the data points that were outside the thresholds. Note that the curve is based on the data points of a 5-week time slot and not just the 16-hour time slot displayed in Fig. 32.

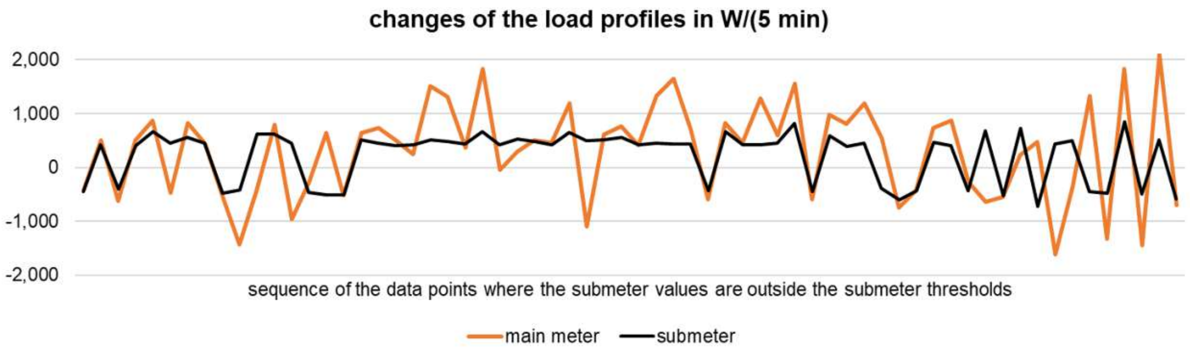


Fig. 33: The changes of the load profiles of two energy meters displayed over time after applying the threshold-filter.

This filter is the one that was already mentioned at the beginning of this chapter. When it is applied to the **c_d2** data source, the code for the filtered data is **c_d2t**, and when it is applied to the **p_d1** data source, the code for the filtered data is **c_d1t**.

In the same way, as with the unfiltered data, the curves of the load profiles of two energy meters are plotted against each other, and an ordinary linear regression is applied. Fig. 34 depicts the same examples as those shown in Fig. 31, but this time after applying the filter: On the left side, the case where there is no relationship between the two analyzed meters, and on the right side, the case where there is indeed the main meter-submeter relationship. The gap in the middle of both examples clearly shows the impact of the threshold filter.

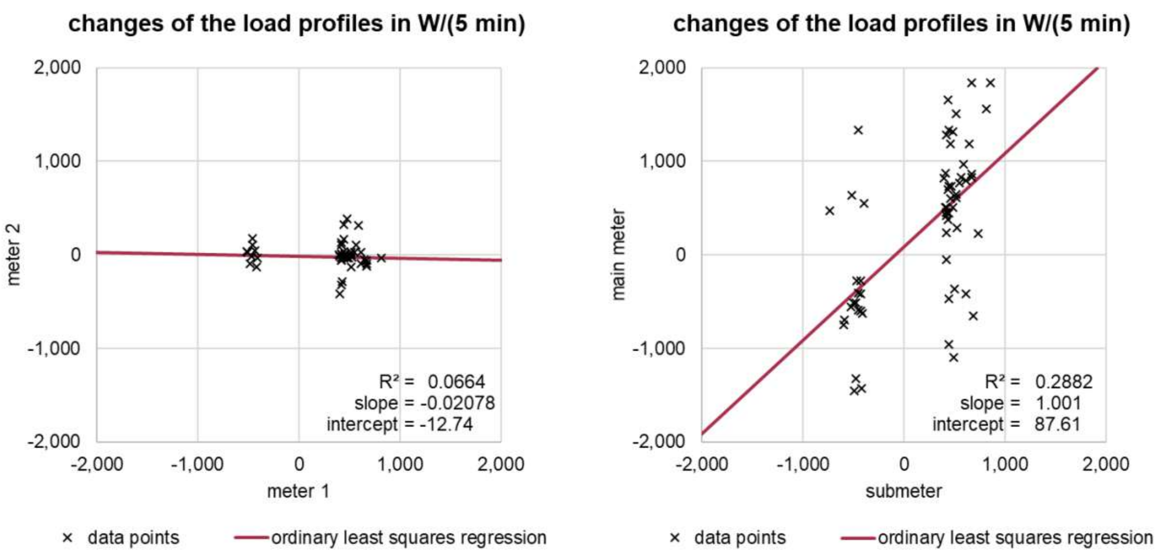


Fig. 34: Load profile change behavior of two energy meters after applying the threshold-filter displayed as scatterplots. Left: Exemplary case with no submeter-main meter relation between meter 1 and 2. Right: Exemplary case with a real main meter-submeter relation between the two energy meters.

Besides the metrics from the ordinary least squares regression, R^2, slope, and intercept, there are two additional features that are calculated out of the filtered data:

- The number of data points with submeter values exceeding the submeter thresholds: symbol n_{sm}, code **sum_sm**
- The number of data points with submeter values exceeding the submeter thresholds with main meter values unequal to zero: symbol n_{mm}, code **sum_mm**

The **sum_sm** value provides a measure of the volatility of load profile changes of the assumed submeter. In combination with **sum_mm**, it can give an indication of the main meter-submeter relationship: The difference between **sum_mm** and **sum_sm** yields the number of occurrences where the load profile of the assumed main meter showed no change while the assumed submeter showed a change that was larger than the threshold. Therefore, a high difference could be an argument against the main meter-submeter relation.

4.5.3 Overview over the primary features

In this section, all previously discussed features are presented together in Tab. 5. The “x” entries in the table mark the data sources used for the calculation of the respective feature. The code of the data source is used as a prefix together with the code of the feature, e.g., **c_d2_spearman** refers to the Spearman correlation coefficient calculated out of the second-order difference quotient of the counter values.

The 32 features presented in the table are the primary features. Further features that can be derived from those primary features are presented in section 4.5.4.

Tab. 5: Overview of the primary features and the data sources which were used to calculate the features

features				data source of the profiles which were used to calculate the features					
				counter values			instantaneous power values		
short description		symbol	code	first-order difference quotient	second-order difference quotient	second-order difference quotient with threshold	measure-ment	first-order difference quotient	first-order difference quotient with threshold
				c_d1	c_d2	c_d2t	p_m	p_d1	p_d1t
relation of the means of the profiles		f_{pos}	profile_relation	x			x		
relation of the absolute sum of the negative profile differences to the entire sum of the absolute profile differences		f_{neg}	profile_relation_neg	x			x		
Pearson correlation coefficient		ρ	pearson		x	x		x	x
Kendall correlation coefficient		τ	kendall		x	x		x	x
Spearman correlation coefficient		r_s	spearman		x	x		x	x
results of ordinary least squares regression	coefficient of determination	R^2	r_squared		x	x		x	x
	slope	β	slope		x	x		x	x
	intercept	ε	intercept		x	x		x	x
amount of data points where the submeter values are outside of the submeter thresholds	all data points	n_{sm}	sum_sm			x			x
	and where the main meter values were not zero	n_{mm}	sum_mm			x			x

4.5.4 Features derived from the primary features

Tab. 6 shows the features derived from the primary features. The table set-up is comparable to Tab. 5, with “x” marking the data sources used for the calculation of the respective feature. The total number of derived features amounts to 30. The “formula” column provides the information on how a derived feature was computed.

Many of those derived features implement a subtraction of a primary feature from the value 1. As the method was initially developed for a linear classifier, these features were intended to be used as building blocks for the calculation of interaction values. Because these derived features were already computed, they were kept for the training of the random forest classifier.

Moreover, the derived features are multiplied with the natural logarithm of the number of data points with submeter values exceeding the submeter thresholds: $\ln(n_{sm})$. Thereby a feature calculated out of many data points carries more weight than a feature calculated out of only a few data points.

Tab. 6: Overview of the features which were derived from the primary features and the data sources which were used to calculate the derived features.

primary feature that the derived feature is based upon		formula	code	data source of the profiles which were used to calculate the features			
				counter values		instantaneous power values	
				second-order difference quotient	second-order difference quotient with threshold	first-order difference quotient	first-order difference quotient with threshold
				c_d2	c_d2t	p_d1	p_d1t
results of ordinary least squares regression	coefficient of determination (R^2)	$1 - R^2$	r_squared_sub	x	x	x	x
		$(1 - R^2) \cdot \ln(n_{sm})$	r_squared_sub_ln		x		x
	slope	$\lvert 1 - \beta \rvert$	deviation_slope_1	x	x	x	x
		$\lvert 1 - \beta \rvert \cdot \ln(n_{sm})$	deviation_slope_1_ln		x		x
	intercept divided by the standard deviation	$\lvert \varepsilon / \sigma \rvert$	intercept_to_std	x	x	x	x
		$\lvert \varepsilon / \sigma \rvert \cdot \ln(n_{sm})$	intercept_to_std_ln		x		x
amount of data points where the submeter values are outside of the submeter thresholds	all data points / and where the main meter values were zero	$n_{sm} - n_{mm}$	diff_sm_mm		x		x
		$(n_{sm} - n_{mm}) \cdot \ln(n_{sm})$	diff_sm_mm_ln		x		x
Pearson correlation coefficient		$(1 - \rho) \cdot \ln(n_{sm})$	pearson_sub_ln		x		x
Kendall correlation coefficient		$(1 - \tau) \cdot \ln(n_{sm})$	kendall_sub_ln		x		x
Spearman correlation coefficient		$(1 - r_s) \cdot \ln(n_{sm})$	spearman_sub_ln		x		x
threshold value divided by the standard deviation		t / σ	threshold_to_std		x		x

One special feature that does not fit into the design of the overview tables Tab. 5 and Tab. 6 is the feature with the code **smoothed**. In this feature, the information, whether the load profile of either the assumed main meter or the assumed submeter was smoothed, is encoded as 0 (false) and 1 (true).

4.6 Preparations for the training of the classifiers

4.6.1 Cleaning

As the method shall be independent of the unit of the provided load profiles and the length of the processed time series, the primary features **intercept**, **sum_sm**, and **sum_mm** had to be excluded from the following steps. As these features were the basis for the calculation of certain derived features, their information is retained, e.g., parts of **intercept** are retained in its standardized versions: **intercept_to_std** and **intercept_to_std_ln**.

With the remaining 24 primary features, the 30 derived features, and the smoothing feature, 55 features are available for each of the 24,492 possible combinations of two energy meters. As the main principle behind most of the calculated features is to compare similarities in the changes of the load profiles, not all those combinations have useable information stored in their features; e.g., an energy meter, which has an almost constant load profile, could theoretically be a submeter of several meters as long as their load profiles never drop below the constant value. Or an energy meter, which did not count any consumption at all during the analyzed time slot, could be a submeter of any other meter. To deal with that, energy meter combinations were excluded from the dataset of possible combinations if they fit at least one of the following criteria:

- The assumed main meter or the assumed submeter counted no consumption at all.
- The profile of the measured instantaneous power value of the assumed main meter was equal to the profile of the measured instantaneous power value of the assumed submeter. These combinations were the meters which have two separate counters.
- Cases where there was no or only one jump of the assumed submeter (**sum_sm** ≤ 1). The reason for that is that several features are derived from the metrics of the least squares regression (**r_squared**, **slope**, **intercept**, and all their derivatives). For the metrics to be well-defined, the least squares regression requires at least two data points.
- Cases where there was no or only one jump of the assumed main meter while there were jumps of the assumed submeter (**sum_mm** ≤ 1). The reason for that is that such occurrences are considered a strong argument against a main meter-submeter relation.

After applying these criteria, the remaining possible combinations are 17,968. Of these 17,968 combinations, only 168 are cases where there is a real main meter-submeter relation, i.e., those are the combinations that can be seen in the case study's network graph in Fig. 12 and that were not filtered out.

Fig. 35 illustrates the described break-down of the 24,492 possible combinations. It visualizes that the classification problem at hand is an extremely unbalanced one. The vast majority of the 17,968 meter combinations are cases where there is no main meter-submeter relation. In contrast, only 1% of them are cases with real main-meter submeter relations.

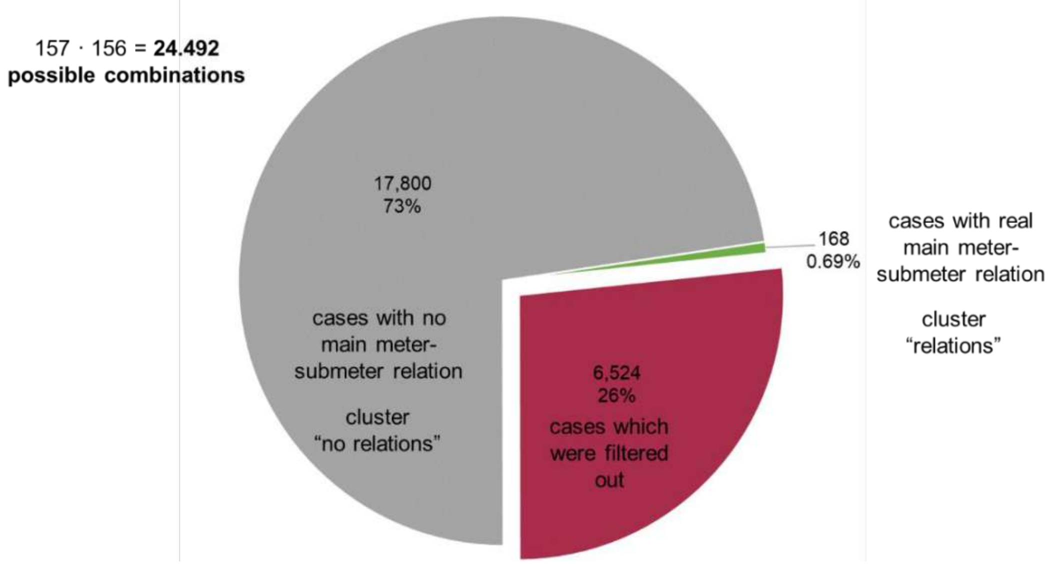


Fig. 35: Overview of the cases of the classification problem. Time resolution of the underlying data: 5 min

These two groups of cases are the two classes that the (binary) classifier must separate. In the following parts of this work, the cases with no main meter-submeter relation are referred to as cluster "**no relations**", and the cases with real main meter-submeter relation are referred to as cluster "**relations**".

4.6.2 Clustering

As the classification problem at hand is an extremely unbalanced one, a generally poor prediction performance of the classifiers is expected. To devise a strategy for dealing with this, the features of the cases of the clusters "**no relations**" and "**relations**" were analyzed in detail.

After the analysis, it became clear that the combinations with real main meter-submeter relationships are not one homogenous group where their features share similar values. Instead, these combinations could be split into three sub-clusters:

- cluster **"obvious relations"** – The features of these cases strongly tend towards the values that would be expected for real main meter-submeter relationships.

- cluster "**rare partial supply**" – Some energy meters have not only one superior main meter but two or more. This is the case when consumers have more than one energy supply – e.g., the emergency electricity supply from another main connection. In such cases, some of the features show values that would usually be a strong argument against a main meter-submeter relationship.

- cluster "**inconspicuous relations**" – The features of these cases don't show strong tendencies. They look like the features of cases where there is no real main meter-submeter relationship with no obvious distinctive pattern that could be used for classification.

With the knowledge of these three clusters, the random forest classifier was developed as a multiclass classification problem with four clusters: the three clusters mentioned before and the one large cluster with no main meter-submeter relations – cluster "**no relations**". Tab. 7 shows the four clusters and typical values for some selected features that can be expected in those clusters.

Tab. 7: Overview over the identified clusters and their typical values for selected features.

cluster	classification features							
	calculated out of counter values				calculated out of instantaneous power values			
	first-order difference quotient		second-order difference quotient	second-order difference quotient with threshold	measurement		first-order difference quotient	first-order difference quotient with threshold
	c_d1_profile_relation	c_d1_profile_relation_neg	c_d2_slope	c_d2t_slope	p_m_profile_relation	p_m_profile_relation_neg	p_d1_slope	p_d1t_slope
obvious relations	> 1	≈ 0	≈ 1	≈ 1	> 1	≈ 0	≈ 1	≈ 1
rare partial supply	< 1	≈ 1	≈ 1	≈ 1	< 1	≈ 1	≈ 1	≈ 1
inconspicuous relations	varying							
no relations	varying							

Fig. 36 illustrates how the meter combinations with real main meter-submeter relation were broken down into the three sub-clusters for the multiclass classification. Most of the cases with real main meter-submeter relation are part of the cluster "inconspicuous relations". That means that their features showed no distinct pattern that might help to distinguish these cases from the cases in the cluster “no relations”. Detecting subtle patterns in the data and using them to separate cases into classes is one of the main functions of classifiers.

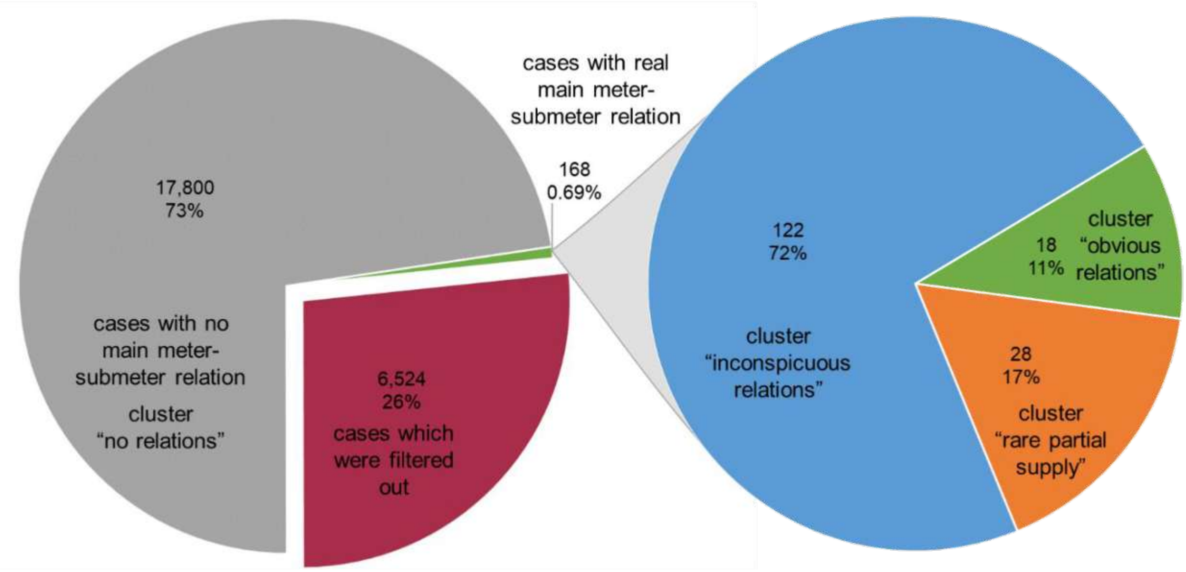


Fig. 36: Overview of the cases of the classification problem after splitting the cluster “relation” in the three separate clusters “obvious relations”, “rare partial supply” and “inconspicuous relations”. Time resolution of the underlying data: 5 min.

The binary classification problem was artificially transformed into a multiclass classification problem to reduce the misclassifications and use the different clusters as prediction quality indicator; e.g., one could be quite certain that there is a real main meter-submeter relationship if a combination is predicted to the cluster “obvious relations”. Otherwise, if a combination is predicted to the cluster “inconspicuous relations”, the probability might be higher that there is no real main meter-submeter relationship.

To prove the assumption that in this case multiclass classification outperforms binary classification, both types of classification are investigated in this work. The binary classifiers were trained with data that were grouped into the two clusters “**no relations**”, and “**relations**” and the multiclass classifiers were trained with the data that were grouped into the four clusters “**no relations**”, “**obvious relations**”, “**rare partial supply**”, and “**inconspicuous relations**”.

4.6.3 Split in training and test data

Since the clusters are very unbalanced and given the background knowledge about the specific structure and issues of the case study’s energy monitoring system, the splitting of the data in training and test data was not conducted completely at random. Instead, some combinations were assigned manually. The goals behind the manual assignments were:

- Ensure that there are enough entries of each cluster in the training dataset and the test dataset.

- Ensure that the entries of the cluster "**obvious relations**" that represent meters measuring different types of electrical consumers are evenly distributed between the training dataset and the test dataset.
- Ensure that entries of the cluster "**no relations**" which show features that would usually be expected for entries of the cluster "**obvious relations**" are not in the training dataset.

Tab. 8 gives an overview of the different members of each cluster were assigned to the training and the test dataset.

Tab. 8: Composition of the clusters and depiction of how they were assigned to the training and test datasets. Time resolution of the underlying data: 5 min.

cluster	sub-category within cluster	training dataset	test dataset
obvious relations	combinations of the 4th floor electricity main meter and a 4th floor electricity submeter	8	
	combinations of the 9th floor electricity main meter and a 9th floor electricity submeter		8
	rest	1 *	1 *
rare partial supply	all entries	14 *	14 *
inconspicuous relations	combinations of the 4th floor electricity main meter and a 4th floor electricity submeter	1	
	combinations of the 9th floor electricity main meter and a 9th floor electricity submeter		2
	combinations of the electricity meters which measure the energy recovery of the elevators and their superior electricity meter		3
	rest	58 *	58 *
no relations	combinations featuring at least one of the electricity meters of the 4th floor	2,274	
	combinations featuring one electricity meter of the 4th floor and one electricity meter of the 9th floor		193
	combinations featuring at least one of the electricity meters of the 9th floor		2,353
	anticipated blinds misclassifications		22
	anticipated PV inverter misclassifications		324
	rest	12,018 *	616 *
	sum	**14,374**	**3,594**
	proportion	**80.0%**	**20.0%**

* randomly sampled

As most of the combinations in the cluster "**obvious relations**" come from the electricity meters that are installed on the 4th and 9th floor – two very similar floors that are monitored in detail – it was decided that the combinations related to one floor shall exclusively be put into the training dataset, and the ones related to the other floor shall exclusively be put into the test dataset.

The electricity meters for energy recovery of the building's elevators measure negative load profiles. As some features are less meaningful when they are calculated out of negative load profiles than in the case of positive load profiles, combinations with the meters of the elevator energy recovery were excluded from the training dataset and instead manually put into the test dataset.

Moreover, the energy meters measuring the building's automatic blinds of the 4th and the 9th floor (the only blinds that are separately measured) were assigned to the test dataset. As the only difference between the load profiles of the blinds on each floor are the consumptions caused by manual interventions by users, there was a high probability of misclassifications.

Some of the building's PV inverters have very similar load profiles, which also involves a high probability of misclassification. Therefore, those anticipated misclassifications were also manually allocated to the test dataset.

The same test and training dataset that was used for the development of the multiclass classifiers was also used for the development of the binary classifiers. For this, the cluster membership of cases of the clusters "**obvious relations**", "**rare partial supply**", and "**inconspicuous relations**" was simply changed back to cluster "**relations**".

4.7 Training of classifiers

The steps elaborated in sections 4.4 to 4.6 were applied to the same data from the five-week time slot described in section 4.3, but with different time resolutions. Tab. 9 shows the overview of the 20 cases that were evaluated in this work.

Tab. 9: Extended overview of the evaluated cases.

type of classification	clusters	utilized data sources	time resolution of the training data
multiclass	"obvious relation" / "rare partial supply" / "inconspicuous relation" / "no relations"	features derived from **counter values** and **instantaneous power values** are used (CP)	5 min
			10 min
			15 min
			30 min
			60 min
		only features derived from **counter values** are used (C)	5 min
			10 min
			15 min
			30 min
			60 min
binary	"relations" / "no relations"	features derived from **counter values** and **instantaneous power values** are used (CP)	5 min
			10 min
			15 min
			30 min
			60 min
		only features derived from **counter values** are used (C)	5 min
			10 min
			15 min
			30 min
			60 min

It was not necessary to repeat the steps for each of the cases presented in Tab. 9. Instead, they were only repeated for each of the five different time resolutions. The variation of the different utilized data sources and different classification types could be done afterward by manipulating each time resolution's training and test dataset. The procedure was as follows:

1. For time resolutions different than 5 min: Resample the monitoring data (load profiles derived from the counter values and instantaneous power values) from the five-week time slot to the desired time resolution.
2. Apply the smoothing algorithm as described in section 4.4.
3. Calculate all 55 features as described in section 4.5, i.e., both the counter values and the instantaneous power values are used as source (CP).
4. Clean, cluster, and split the data in training and test data as described in section 4.6, i.e., prepare the data for multiclass classification.
5. To set up the variant where only the counter values are used as source (C), duplicate the training and test dataset and remove all features that were calculated out of the instantaneous power values. Thus, only 28 features remain in the duplicate.
6. To set up the variants of the binary classification, duplicate all training and test datasets and change the class memberships of cases of the clusters “obvious relations”, “rare partial supply”, and “inconspicuous relations” to cluster “relations” in those duplicates.

After this procedure was conducted for each of the five different time resolutions, there were 20 pairs of training and test datasets, which were ready for the training and test of the random forest classifiers.

Due to the relatively large number of 55 respectively 28 available features, with the possibility of some features containing little to no information useful for classification, a backward-selection process was chosen. Within this process, classification evaluation metrics were calculated to evaluate the performance of the classifiers. These metrics were then used to decide upon the best classifier for each case.

4.7.1 Classification evaluation metrics

Given the two clusters of the binary classification case, the confusion matrix for the classification problem at hand is a 2 x 2 matrix. It is depicted in Tab. 10.

Tab. 10: Confusion matrix between the actual cluster and the predicted cluster for the binary classification.

		predicted cluster	
		relations	no relations
actual cluster	relations	*tp*	*fn*
	no relations	*fp*	*tn*

The entries in each of the cells in Tab. 10 correspond to the number of entries that fit the scheme illustrated by the matrix. The letter pairs are the tokens that are used to refer to a certain position in the confusion matrix:

- *tp* (true positive) – Entries that are from the cluster "relations" and were correctly predicted to this cluster.
- *tn* (true negative) – Entries that are from the cluster "no relations" and were correctly predicted to this cluster.
- *fp* (false positive) – Entries that are from the cluster "no relations" and were incorrectly predicted to the cluster "relations".
- *fn* (false negative) – Entries that are from the cluster "relations" and were incorrectly predicted to the cluster "no relations".

Based on the entries of the confusion matrix, several metrics can be calculated to determine the quality of the prediction. A common metric is the so-called "accuracy" (*ACC*), which measures the ratio of correct predictions over the total number of instances evaluated [40]. In the binary case, it is usually calculated as shown in (4).

$$ACC = \frac{tp + tn}{tp + fp + tn + fn} \tag{4}$$

As the cluster "no relations" is extremely overrepresented, the ACC value is expected to be close to 1. Even if all entries of the test dataset would be manually assigned to the cluster "no relations", the ACC would be 0.991. To cope with that, the metric is altered by neglecting all "no relations" entries and thus forming a new metric, namely the "custom accuracy" (*CACC*), as presented in (5).

$$CACC = \frac{tp}{tp + fn} \tag{5}$$

Another common metric is the so-called "misclassification rate" (*MCR*) [40]. In the binary case, it is usually calculated as shown in (6).

$$MCR = \frac{fp + fn}{tp + fp + tn + fn} \tag{6}$$

Again, due to the overrepresentation of the “no relations” cluster, the *MCR* formula is altered to calculate a new metric. But in this case, all entries that were predicted to the cluster “no relations” are neglected. The new metric is called “custom misclassification rate” (*CMCR*), and its formula is (7).

$$CMCR = \frac{fp}{tp + fp} \tag{7}$$

Given the four clusters of the multiclass classification case, the confusion matrix for the classification problem at hand is a 4 x 4 matrix. It is depicted in Tab. 11. The single letters and the letter pairs are the tokens that are used to refer to a certain position in the confusion matrix. As in the binary case shown in Tab. 10, the entries of each cell correspond to the number of entries that fit the scheme illustrated by the matrix. The rows indicate the actual cluster membership and the columns the predicted cluster membership. Single letter tokens (*O*, *P*, *I*, and *N*) represent correct predictions to the actual cluster, like the true positive (*tp*) and true negative (*tn*) predictions in the binary case. Letter-pair tokens (*OP*, *PI*, *IN*, …) represent the incorrect predictions to another cluster beside the actual cluster, like the false positive (*fp*) and false negative (*fn*) predictions in the binary case.

For easier readability of the following formulas (8-13), the confusion matrix was extended by another column, which shows the tokens that can be used to refer to the sum of each row (*AO*, *AP*, *AI*, and *AN*).

Tab. 11: Confusion matrix between the actual cluster and the predicted cluster for the multiclass classification.

		predicted cluster				**Σ** (amount of all entries of one actual cluster)
		obvious relations	rare partial supply	inconspicuous relations	no relations	
actual cluster	obvious relations	*O*	*PO*	*IO*	*NO*	*AO*
	rare partial supply	*OP*	*P*	*IP*	*NP*	*AP*
	inconspicuous relations	*OI*	*PI*	*I*	*NI*	*AI*
	no relations	*ON*	*PN*	*IN*	*N*	*AN*

For multiclass classification, there is usually another formula for the accuracy, where the accuracy of each class is calculated, and their average is computed [40]. As the underlying issue is a binary case, the *ACC* formula of the binary case is used as the basis. In the case of the given multiclass classification, the *ACC* formula (4) translates to (8) or (9).

$$ACC = \frac{(O + P + I) + N}{(O + P + I) + (OP + OI + ON + PO + PI + PN + IO + IP + IN) + N + (NO + NP + NI)} \quad (8)$$

$$ACC = \frac{O + P + I + N}{AO + AP + AI + AN} \quad (9)$$

The *ACC* metric is altered to cope with the extreme overrepresentation of the cluster “no relations”. Comparable to the binary case, all “no relations” entries are neglected, and thus, the new metric “custom accuracy” (*CACC*) for the multiclass case is calculated as shown in (10).

$$CACC = \frac{O + P + I}{AO + AP + AI} \quad (10)$$

For multiclass classification, there is usually another formula for the misclassification rate, where the misclassification rate of each class is calculated, and their average is computed [40]. As the underlying issue is a binary case, the *MCR* formula of the binary case is used as the basis. In the case of the given multiclass classification, the *MCR* formula (6) translates to (11) or (12).

$$MCR = \frac{(ON + PN + IN) + (PO + IO + NO + OP + IP + NP + OI + PI + NI)}{(O + P + I) + (OP + OI + ON + PO + PI + PN + IO + IP + IN) + N + (NO + NP + NI)} \quad (11)$$

$$MCR = \frac{(ON + PN + IN) + (PO + IO + NO + OR + IP + NP + OI + PI + NI)}{AO + AP + AI + AN} \quad (12)$$

Comparable to the binary case, the *MCR* metric is altered to cope with the extreme overrepresentation of the cluster “no relations”. All entries that were predicted to the cluster “no relations” and the entries that were misclassified between the other three clusters are neglected. Additionally, the weights $w_1 = 2$ and $w_2 = 1$ are introduced. Thus, the new metric “custom misclassification rate” (*CMCR*) for the multiclass case is calculated as shown in (13).

$$CMCR = \frac{w_1 \frac{ON}{O + OP + OI + ON} + w_2 \frac{PN + IN}{P + PO + PI + PN + I + IO + IP + IN}}{w_1 + w_2} \quad (13)$$

The idea behind *CMCR* is to express how many entries from the “no relations” cluster were predicted to be part of the other clusters. As it is especially undesirable that such a member is classified to the cluster “obvious relations”, those cases are penalized with the weight w_1.

All combination cases which are members of the anticipated PV and blinds misclassifications are excluded from the calculation of *CMCR* – regardless of the classification type being binary or multiclass.

4.7.2 Random forest model selection

With the decision on evaluation metrics *CACC* and *CMCR* as indicators for the prediction performance of the classifiers, the process of training random forest classifiers in different configurations and deciding upon the best ones could be initiated. The starting point for all of the evaluated 20 cases was always a random forest classifier that uses all of the features that are available for the case:

- 55 for the cases where both the counter values and the instantaneous power values were used as data source (CP)
- 28 for the cases where only the counter values were used as data source (C)

Starting from this classifier utilizing all available features, a backward-selection process, where features were dropped one by one, was applied. The measure used to decide which feature shall be dropped from the features used in the random forest is the so-called “Gini importance”. This measure is named “feature importance” in the random forest implementation used for this work, “sklearn.ensemble.RandomForestClassifier” of the Python package “scikit-learn 0.22.1” [41]. It is automatically calculated for each feature used in the random forest during its training. The higher the feature importance value of a feature, the larger its relevance as splitting variable inside the random forest. So, this metric is indeed an expression of the importance of a feature.

There is some research regarding the fact that the Gini importance is biased towards features with many possible split points, e.g., continuous variables or variables with a high cardinality [42]. Since the 55 respectively 28 features all have many unique values – most of them show numbers between 10,000 and 20,000 – they can be considered continuous. Thus, the impact of the bias should be negligible.

The backward-selection process was combined with a variation of the maximum allowed depth of the trees in the random forest. Analyzing the trees of the random forest model featuring all 55/28 features without limitation of the maximum allowed tree depth revealed that none of those trees had a depth larger than 25 – thus, the maximum tree depth was varied between 1 and 30. Therefore, within the backward-selection process, there are 30 x 55 = 1,650 respectively 30 x 28 = 840 different configurations of random forests that were repeatedly trained and evaluated according to the following steps:

1. At the start, the selected features are all 55/28 features, and the selected maximum tree depth is 1.
2. A random forest model consisting of 1,000 trees featuring the selected features and the selected maximum tree depth is trained 50 times with the training dataset.

3. Each of these 50 models is used to predict the test dataset´, and for each of these predictions, the custom accuracy (*CACC*) and custom misclassification rate (*CMCR*) are evaluated, and afterward, their respective means are calculated.
4. The selected maximum tree depth is increased by one, and steps 2 to 4 are repeated if the maximum tree depth is below or equal to 30.
5. When the maximum tree depth is equal to 30, for each feature, the mean of the corresponding feature importances of the 50 models is computed.
6. If there is more than one feature left: The feature with the lowest mean feature importance is dropped from the group of selected features, the selected maximum tree depth is reset to 1, and the preceding steps are repeated starting from step 2.

As randomness is one of the basic characteristics of random forests, the models need to be trained and evaluated several times to get an indication of the prediction performance that can generally be expected from the models. It is common to choose 100 or 200 as number for the repetitions. Due to limited computational resources, the number was set to 50.

The process described in steps 1 to 6 was conducted for each of the 20 evaluated cases by applying it to the corresponding pair of training and test data. After this process, for each case, the resulting mean *CMCR* and *CACC* values were analyzed to select candidates for the best classifiers.

Fig. 37 and Fig. 38 illustrate the results of the backward-selection process of one of the 20 evaluated cases and visualize how the candidates were selected. The case that is presented as an example is the case of multiclass classifiers that utilize features derived from both the counter values and the instantaneous power values from data with a 5 min time resolution. The heatmaps in both figures share the same structure: The X-axis shows the number of features that are used by the random forests is indicated. The Y-axis shows the maximum allowed depth of the trees in the random forests.

As the backward-selection process started with all 55 features and a maximum tree depth of 1, the result for the first evaluated random forest configuration is the entry in the top right corner of the heatmaps. For the following configurations, the maximum tree depth was increased by one until the configuration with all 55 features and a tree depth of 30 was reached. Out of the feature importances of all 50 random forests of this configuration, the mean feature importance was calculated. The feature with the lowest mean feature importance was dropped for the next random forest configurations with 54 features and a maximum tree depth varying from 1 to 30. In the same manner, the features of all following random forest configurations were selected.

Thus, even though the X-axis only displays the used features as a number, the number refers to a specific set of features. In appendix B, there are tables that show the features ordered according to the number-reference for each of the 20 evaluated cases.

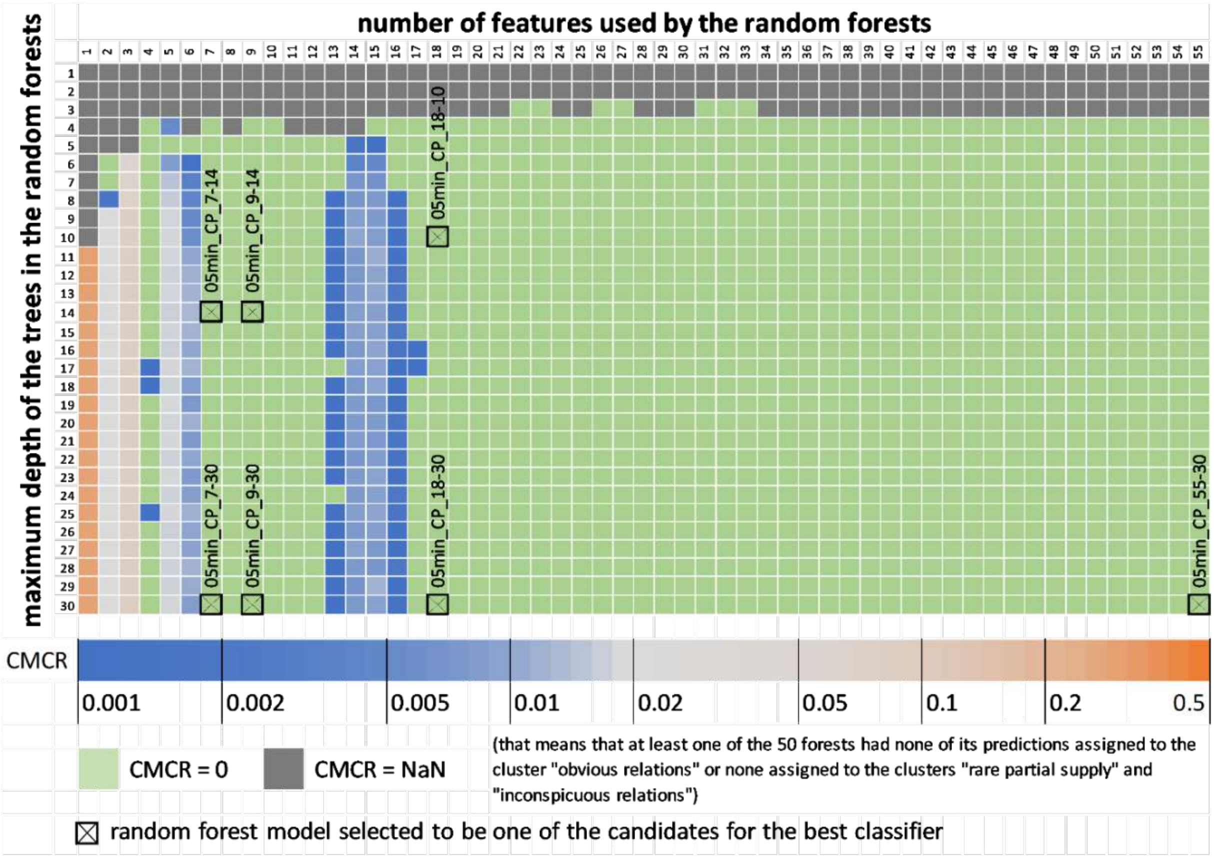


Fig. 37: Heatmap depicting mean values of the custom misclassification rates (CMCR) of 50 repetitions for different random forest configurations. Time resolution: 5 min. Used data sources: CP. Classifier: multiclass.

As the results of the mean *CMCR* have different magnitudes, the coloring scheme of the heatmap in Fig. 37 was adjusted accordingly:

- A mean *CMCR* value of exactly 0 is especially desirable as it means that with the given random forest model configuration, no members of cluster “no relations” were misclassified to the other clusters during each of the 50 repetitions. This case is highlighted in the heatmap in a separate color.
- Mean *CMCR* values above 0 are colored according to a continuous color scheme tailored for the exponential scale.
- No *CMCR* value could be computed for cases with no entries classified to the cluster “obvious relations” or to the clusters “rare partial supply” and “inconspicuous relations. The Python implementation of the *CMCR* formula yielded “NaN” in these cases. As these cases are especially undesirable, they were highlighted by a separate color.

As shown in Fig. 37, there are several possible random forest configurations with a mean *CMCR* value of the 50 repetitions of exactly 0. Most of those combinations utilize 18 to 55 of the 55 features and have a maximum tree depth of at least 4. Reducing the number of features below 18 shows that there is one area where the mean *CMCR* value is slightly above 0. Adjacent to this patch, there is again an area with mean *CMCR* values of exactly 0. Models with less than 7 features seem to be an unreasonable choice, as they show mean *CMCR* values of a much higher magnitude – except for models with 4 features.

While values as small as possible are the goal for the *CMCR* metric, the goal for the CACC metric is to achieve values as large as possible. Analyzing the heatmap of the mean *CACC* values, see Fig. 38, reveals that the highest values can be found near the area with the mean *CMCR* values slightly above 0 and in the area between 7 and 9 features. This knowledge led to the decision that the candidates for the best random forest configuration must come from these areas.

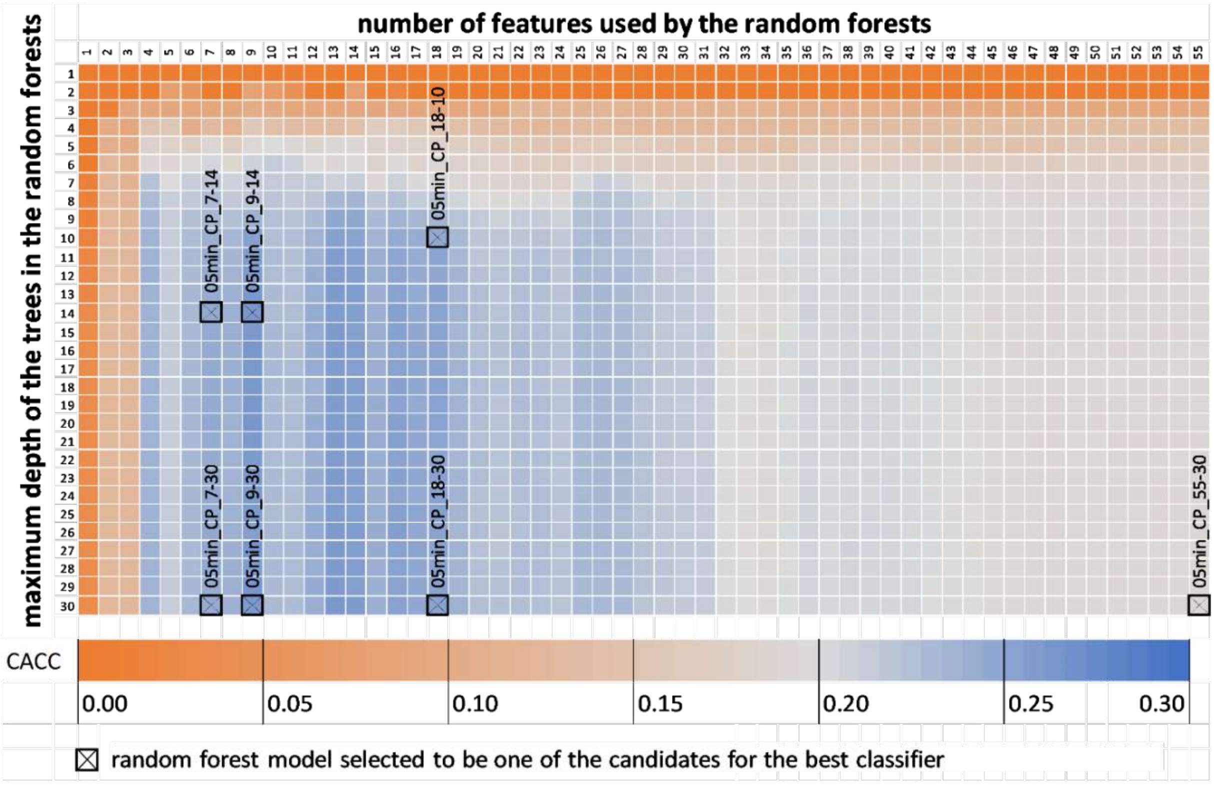


Fig. 38: Heatmap depicting mean values of the custom misclassification rates (CACC) of 50 repetitions for different random forest configurations. Time resolution: 5 min. Used data sources: CP. Classifier: multiclass.

For the unique identification of the selected candidates for the best classifier, the naming scheme illustrated in Fig. 39 is used. It summarizes all major information of a classifier model.

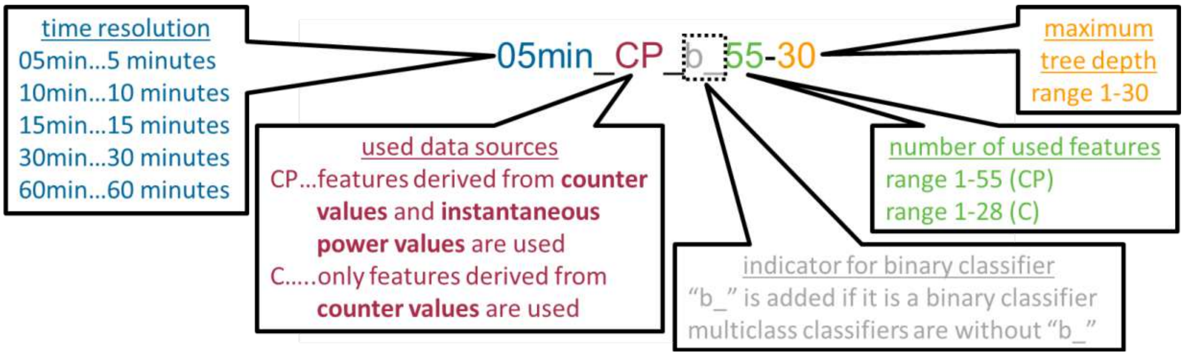


Fig. 39: Naming scheme for the random forest classifiers.

As the classifiers with 7, 9, and 18 features show a continuous mean *CMCR* value of exactly 0 over almost all evaluated maximum tree depths and have high mean *CACC* values, it was decided that the candidates should have those specific sets of features. The first three candidates that were chosen have these three numbers of features and a maximum tree depth of 30, which could be interpreted as "no limitation of the tree depth" in the case at hand.

As smaller, simpler classifier models require less computational power, and to counteract overfit, the maximum depth of the trees should be as small as possible. Therefore, the next three candidates also have 7, 9, and 18 features, but their maximum depth was set to the points where the mean *CACC* values do not increase further with increasing maximum depth.

Moreover, the random forest model with all 55 features and a maximum tree depth of 30 was also selected, namely as the reference model. The seven random forest models of the example case that were selected as candidates for the best classifier are thereby:

- 05min_CP_55-30 – model trained with all 55 features and a maximum tree depth of 30
- 05min_CP_18-30 – model trained with 18 features and a maximum tree depth of 30
- 05min_CP_18-10 – model trained with 18 features and a maximum tree depth of 10
- 05min_CP_9-30 –model trained with 9 features and a maximum tree depth of 30
- 05min_CP_ 9-14 –model trained with 9 features and a maximum tree depth of 14
- 05min_CP_7-30 –model trained with 7 features and a maximum tree depth of 30
- 05min_CP_7-14 –model trained with 7 features and a maximum tree depth of 14.

All seven selected models are especially highlighted in Fig. 37 and Fig. 38. The actual features that were used for the training of the models are shown in Tab. 12. In the same manner, as in appendix B, the features are ordered according to their feature importance, i.e., the feature in the first row is the one that was kept until the end of the backward selection process. As can be seen, the most important feature is "c_d2_slope", the slope of the scatterplot from the second-order difference quotient calculated out of the counter values. The following most important features come from all various sources. Derived features are underrepresented, indicating that the primary features seem to carry most of the relevant classification information on their own.

Tab. 12: Overview of the features used by some of the candidates for the best multiclass classifiers. The features are ordered according to their importance. Time resolution: 5 min. Used data sources: CP.

feature		random forest classifier			
order	code	05min_CP_55-30	05min_CP_18-30 & 05min_CP_18-10	05min_CP_9-30 & 05min_CP_9-14	05min_CP_7-30 & 05min_CP_7-14
1	c_d2_slope	x	x	x	x
2	p_d1_kendall	x	x	x	x
3	c_d2_spearman	x	x	x	x
4	c_d1_profile_relation_neg	x	x	x	x
5	c_d2t_spearman	x	x	x	x
6	p_d1t_kendall	x	x	x	x
7	c_d1_profile_relation	x	x	x	x
8	c_d2t_intercept_to_std_ln	x	x	x	
9	c_d2t_kendall	x	x	x	
10	p_m_profile_relation_neg	x	x		
11	c_d2_pearson	x	x		
12	p_d1_r_squared_sub	x	x		
13	p_m_profile_relation	x	x		
14	c_d2_kendall	x	x		
15	c_d2t_deviation_slope_1_ln	x	x		
16	p_d1t_spearman	x	x		
17	c_d2_deviation_slope_1	x	x		
18	p_d1_spearman	x	x		
19	p_d1_deviation_slope_1	x			
20	c_d2t_pearson	x			
21	c_d2t_intercept_to_std	x			
22	c_d2t_deviation_slope_1	x			
23	p_d1_r_squared	x			
24	c_d2_r_squared	x			
25	p_d1t_diff_sm_mm	x			
26	p_d1t_deviation_slope_1_ln	x			
27	c_d2_r_squared_sub	x			
28	c_d2t_slope	x			
29	c_d2t_diff_sm_mm	x			
30	p_d1t_pearson	x			
31	c_d2t_r_squared	x			
32	p_d1t_deviation_slope_1	x			
33	c_d2_intercept_to_std	x			
34	p_d1_pearson	x			
35	p_d1t_diff_sm_mm_ln	x			
36	c_d2t_r_squared_sub_ln	x			
37	p_d1_slope	x			
38	p_d1t_pearson_sub_ln	x			
39	p_d1t_intercept_to_std_ln	x			
40	c_d2t_r_squared_sub	x			
41	p_d1_intercept_to_std	x			
42	c_d2t_diff_sm_mm_ln	x			
43	p_d1t_threshold_to_std	x			
44	p_d1t_intercept_to_std	x			
45	p_d1t_kendall_sub_ln	x			
46	c_d2t_pearson_sub_ln	x			
47	p_d1t_r_squared_sub_ln	x			
48	c_d2t_kendall_sub_ln	x			
49	p_d1t_spearman_sub_ln	x			
50	p_d1t_slope	x			
51	c_d2t_spearman_sub_ln	x			
52	c_d2t_threshold_to_std	x			
53	p_d1t_r_squared	x			
54	p_d1t_r_squared_sub	x			
55	smoothed	x			

The heatmaps illustrating the mean *CMCR* and *CACC* results for all other of the evaluated 20 cases can be found in appendix C. They are set up in the same way as described in the example case shown in Fig. 37 and Fig. 38. The classifiers that were chosen as candidates for the best classifiers were also chosen according to the goal of a low mean *CMCR* value and a high mean *CACC* value. They are also highlighted in the heatmaps. Tab. 13 and Tab. 14 give an overview of those candidates.

Tab. 13: The candidates for the best multiclass classifiers.

designation of the classifier	time resolution of the training data	utilized data sources	number of features used	maximum tree depth
05min_CP_55-30	5 min	CP	55	30
05min_CP_18-30			18	30
05min_CP_18-10			18	10
05min_CP_9-30			9	30
05min_CP_9-14			9	14
05min_CP_7-30			7	30
05min_CP_7-14			7	14
05min_C_28-30		C	28	30
05min_C_7-30			7	30
05min_C_7-13			7	13
05min_C_5-30			5	30
05min_C_5-14			5	14
10min_CP_55-30	10 min	CP	55	30
10min_CP_40-30			40	30
10min_CP_40-10			40	10
10min_CP_17-30			17	30
10min_CP_17-15			17	15
10min_CP_12-30			12	30
10min_CP_12-9			12	9
10min_C_28-30		C	28	30
10min_C_7-22			7	22
10min_C_5-23			5	23
10min_C_3-30			3	30
10min_C_3-15			3	15
15min_CP_55-30	15 min	CP	55	30
15min_CP_53-5			53	5
15min_CP_45-5			45	5
15min_CP_25-13			25	13
15min_CP_25-4			25	4
15min_CP_13-21			13	21
15min_CP_8-8			8	8
15min_C_28-30		C	28	30
15min_C_3-22			3	22
30min_CP_55-30	30 min	CP	55	30
30min_CP_41-30			41	30
30min_CP_27-15			27	15
30min_CP_20-20			20	20
30min_CP_8-29			8	29
30min_CP_5-7			5	7
30min_C_28-28		C	28	28
30min_C_28-17			28	17
60min_CP_3-30	60 min	CP	3	30
60min_CP_3-15			3	15
60min_C_4-27		C	4	27
60min_C_4-14			4	14

Tab. 14: The candidates for the best binary classifiers.

designation of the classifier	time resolution of the training data	utilized data sources	number of features used	maximum tree depth
05min_CP_b_55-30	5 min	CP	55	30
05min_CP_b_27-30			27	30
05min_CP_b_27-13			27	13
05min_CP_b_4-30			4	30
05min_CP_b_4-13			4	13
05min_C_b_28-30		C	28	30
05min_C_b_9-30			9	30
05min_C_b_9-15			9	15
05min_C_b_7-30			7	30
05min_C_b_7-17			7	17
10min_CP_b_55-30	10 min	CP	55	30
10min_CP_b_53-8			53	8
10min_CP_b_26-5			26	5
10min_CP_b_2-11			2	11
10min_C_b_28-30		C	28	30
10min_C_b_8-16			8	16
10min_C_b_4-22			4	22
10min_C_b_3-30			3	30
15min_CP_b_55-30	15 min	CP	55	30
15min_CP_b_18-4			18	4
15min_CP_b_7-10			7	10
15min_CP_b_2-30			2	30
15min_C_b_28-30		C	28	30
15min_C_b_24-2			24	2
15min_C_b_8-19			8	19
15min_C_b_3-3			3	3
30min_CP_b_55-30	30 min	CP	55	30
30min_CP_b_45-30			45	30
30min_CP_b_45-18			45	18
30min_CP_b_34-30			34	30
30min_C_b_28-30		C	28	30
30min_C_b_15-30			15	30
30min_C_b_15-12			15	12
30min_C_b_13-12			13	12
60min_CP_b_55-30	60 min	CP	55	30
60min_CP_b_37-30			37	30
60min_CP_b_37-20			37	20
60min_CP_b_16-30			16	30
60min_CP_b_16-14			16	14
60min_C_b_28-30		C	28	30
60min_C_b_22-14			22	14
60min_C_b_6-30			6	30

A detailed study of the results in appendix C reveals that the process of selecting the candidates for the best classifiers was not as straightforward as shown in the example case. Especially in the multiclass cases with a time resolution of 30 or 60 min, the mean *CMCR* values could not be computed for many of the configurations – indicated by the “NaN” entries. This led to a very limited choice of possible candidates. Nevertheless, for each of the 20 evaluated cases, there are at least two classifier candidates for each case. In the next step, these classifier candidates were used on the validation data.

5 Validation

In this chapter, the candidates for best multiclass and binary classifiers are used on validation data from the case study and the synthetically generated data. The main purpose of this is to verify that the method works with other data than the data the classifiers were developed with. As the validation data comes from different time slots with different lengths and different time resolutions, a byproduct is the estimation of the method's prediction performance for various circumstances. Further, since the final best classifiers for each of the 20 evaluated cases were not chosen yet, but instead, only candidates were chosen, the selection of the final best classifiers is also part of this chapter.

5.1 Adjusted evaluation metrics

The classification evaluation metrics, as described in section 4.7.1, were designed to aid in the process of selecting candidates for the best classifiers. To estimate the prediction performance of the method, they must be adjusted.

One of the reasons for that is that the method automatically filters out certain possible meter combinations. That is, in the case study's network graph, see Fig. 12, the number of real main meter-submeter relations is stated as 180. In the overview of the classification problem, see Fig. 35, there are only 168 left. The missing 12 combinations were filtered out. Thus, the custom accuracy (*CACC*) metric, as stated in formula (5) for the binary case and in formula (10) for the multiclass case, only uses the 168 combinations as the denominator. To evaluate the prediction performance of the whole method, the filtered-out combinations need to be added to the denominator. The adjusted *CACC* metric is called *CACC2*. With the combinations that were filtered out during the cleaning step, see section 4.6.1, as Z, the formula for *CACC2* in the binary case is (14) and in the multiclass case is (15).

$$CACC2 = \frac{tp}{tp + fn + Z} \tag{14}$$

$$CACC2 = \frac{O + P + I}{AO + AP + AI + Z} \tag{15}$$

Through this alteration, the *CACC2* metric expresses how many of all the real main meter-submeter relations the method could correctly detect.

The *CMCR* metric also needs some adjustments to be suitable for estimating the prediction performance of the method. From the original *CMCR* metric, two new metrics, *CMCR2* and *CMCR2C*, were derived, which take those adjustments into consideration.

In the binary case, there is no need for the first adjustment, therefore, the *CMCR2* metric, see (16), is identical to the original CMCR formula of the binary case (7).

$$CMCR2 = CMCR = \frac{fp}{tp + fp} \tag{16}$$

In order that the original *CMCR* formula of the multiclass case (13) yields a value between zero and one, a classifier must predict at least one entry to the cluster “obvious relations” and at least one entry to the clusters “rare partial supply” or “inconspicuous relations”. If the classifier fails to do so, the *CMCR* is not defined, and the Python implementation of the formula returns “NaN”. To counteract this behavior, the *CMCR2* is defined as shown in (17).

$$CMCR2 = \begin{cases} \dfrac{ON}{O + OP + OI + ON}, & P + PO + PI + PN + I + IO + IP + IN = 0 \\ \dfrac{PN + IN}{P + PO + PI + PN + I + IO + IP + IN}, & O + OP + OI + ON = 0 \\ CMCR, & \text{otherwise} \end{cases} \tag{17}$$

Due to this alteration, the custom misclassification rate can also be computed for cases where the multiclass classifier does not predict at least one entry to the cluster “obvious relations” and at least one entry to the clusters “rare partial supply” or “inconspicuous relations”. The only case in which the *CMCR2* metric cannot be computed and its Python implementation would yield “NaN” is the case where all entries are predicted to the cluster “no relations”. As the intention of the custom misclassification rate is that it shall give an indication of how many of the predictions to all clusters besides “no relations” are from the cluster “no relations”, this behavior is acceptable. Such a case can be considered as “the method failed to detect any relationships between the energy meters”.

During the analysis of the predictions that the classifiers made for the validation data, some misclassifications occurred that can be considered a correct prediction and a misclassification at the same time: The classifiers sometimes identified the main meter’s superior meter, and in some rare cases they even identified the superior meter of the main meter’s superior meter. To take this special type of misclassifications into account, another metric was set up, namely *CMCR2C*. It is simply the *CMCR2* metric, but all misclassifications that were made between a meter and its main meter’s superior meter or even the superior meter of the main meter’s superior meter were excluded from the calculation of the metric.

Tab. 15 gives an overview of all evaluation metrics used in this work. All metrics were calculated for all validation cases, but to keep the presentation of the results short, primarily *CACC2* and *CMCR2C* are used to discuss the method’s performance.

Tab. 15: Overview of the original classification evaluation metrics and the adjusted evaluation metrics.

metric		description
custom accuracy	*CACC*	the number of the correctly predicted main meter-submeter relationships set in relation to the number of main meter-submeter relationships that were not filtered out
	CACC2	the number of the correctly predicted main meter-submeter relationships set in relation to the total number of main meter-submeter relationships
custom misclassification rate	*CMCR*	the weighted average of the misclassifications from the cluster "no relations" to the other clusters; in the case of the case study data, the anticipated PV and blind misclassifications are excluded
	CMCR2	*CMCR* adjusted for the case that the multiclass classifier does not predict at least one entry to the cluster "obvious relations" and at least one entry to the clusters "rare partial supply" or "inconspicuous relations"
	CMCR2C	*CMCR2*, but the misclassifications to the main meter's superior meter and the superior meter of the main meter's superior meter are excluded

Even though in the custom misclassification rate metrics, there are several exclusions of certain misclassifications, these misclassifications can be seen in alternative depictions of the results, e.g., in the network graphs. These are presented and explained at the end of the following subchapters 5.2 and 5.3.

5.2 Case study data from other time slots

As described in section 4, the method was developed, and the classifiers trained with the case study's energy monitoring data from a 5-week time slot from the fall of 2017. The case study's energy monitoring data from the entire year of 2018 is used as the basis for the first part of the validation of the method – 19 time slots with varying lengths and from varying seasons were chosen from it. Fig. 40 illustrates both the 5-week time slot used for training and the 19 time slots from 2018 that were chosen for the validation.

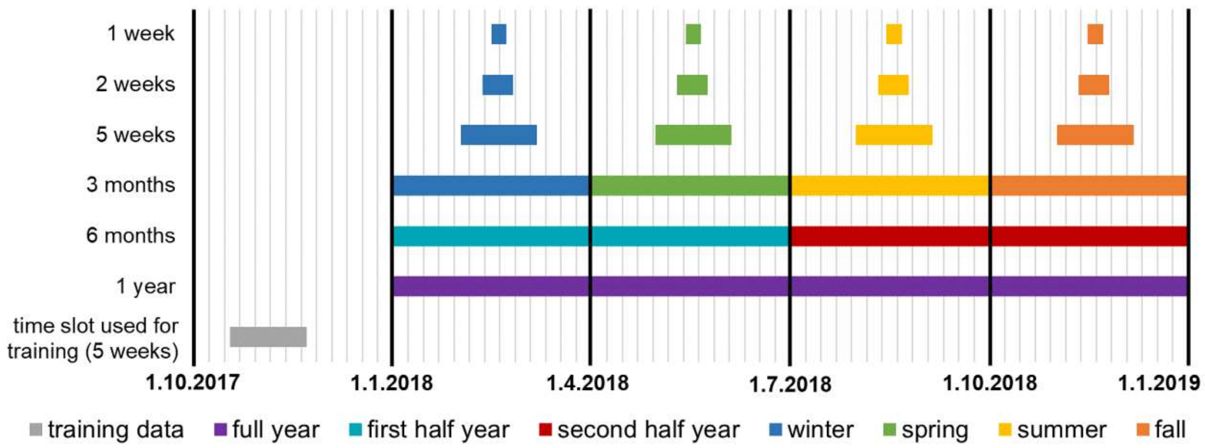


Fig. 40: The time slots of the case study data which were used for the training and the performance evaluation of the candidates for the best multiclass and binary classifiers.

As the energy measured by some of the energy meters – mostly heat meters – is dependent on the season, most of the time slots were chosen according to the four seasons. To evaluate the impact of different lengths of the time slots, the time slots have lengths of 1 week, 2 weeks, 5 weeks, 3 months, 6 months, and the entire year. The 5-week training time slot was also included in the entire validation process. As all classifiers were trained with 80% of the data from these five weeks, the results for this time slot give an indication of the peak prediction performance that could be achieved by the method.

To be able to evaluate different time resolutions, the data from each of these 20 time slots were resampled to the time resolutions the respective classifiers were trained for. Hence, the method was validated with 100 different datasets that originated from the 20 mentioned time slots and have time resolutions of 5 min, 10 min, 15 min, 30 min, and 60 min.

Smoothing, feature extraction, cleaning, and clustering steps – as described in sections 4.4 to 4.6.2 – were conducted with each of these 100 datasets. The result is 100 new datasets suitable to be used as input for the classifiers: Each of them consists of the 55 features for each possible meter combination, provided the combination was not automatically filtered out. All candidates for the best multiclass and binary classifiers were then applied to these 100 new datasets, regardless of the time resolution, the classifiers were trained with.

The results for the *CACC2* and *CMCR2C* metrics that each classifier achieved for the 100 datasets are presented in Appendix D. Due to the sheer number of results, the metrics for datasets that have the same time slot length and time resolution were averaged; e.g., the metrics for a time slot length of 5 weeks are the mean of the metrics for the time slots “5 weeks winter”, “5 weeks spring”, “5 weeks summer” and “5 weeks fall”. Fig. 41 and Fig. 42 illustrate the results by using one of the classifiers as an example.

Fig. 41 illustrates the accuracy metric *CACC2* of the classifier 05min_CP_18-30 when used on the datasets that were calculated out of the monitoring data with 5 min time resolution. As expected, the classifier achieves the highest *CACC2* value when applied to the training data: Approximately 0.63, which means that 63% of the real main meter-submeter relations of the case study were correctly detected by the method. This value is substantially higher than the values that were achieved with the other datasets with *CACC2* values between 0.18 and 0.3, meaning that approximately 25% of the real main meter-submeter relations were identified. The average values for each time slot length indicate that with decreasing time slot length, the *CACC2* also slightly decreases.

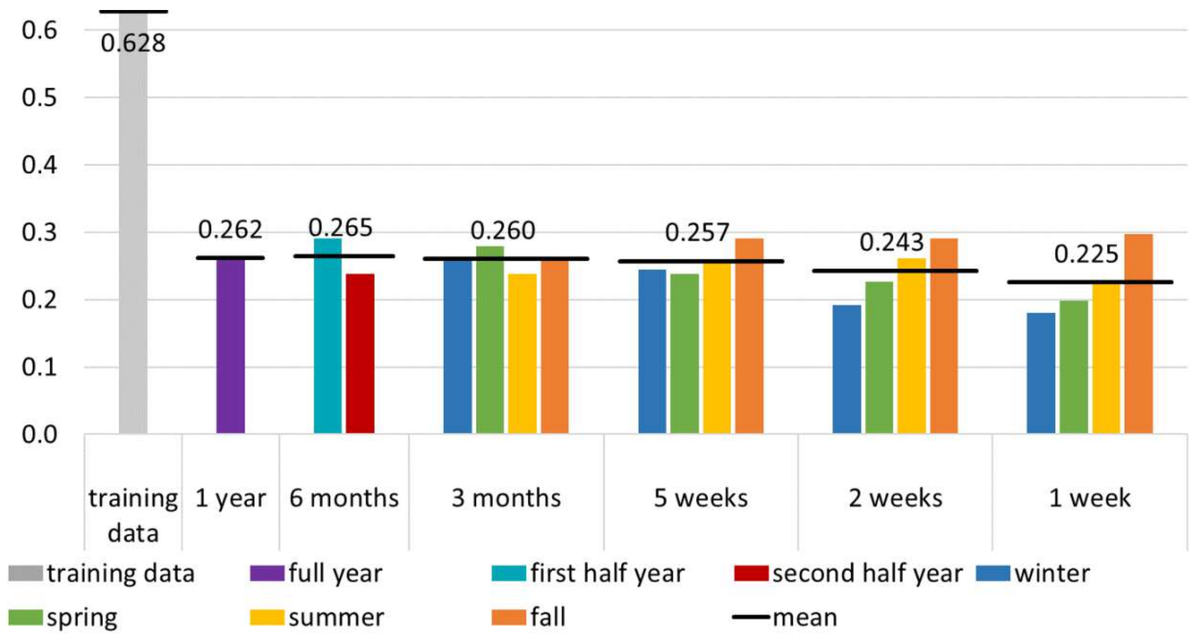


Fig. 41: CACC2 values achieved by the classifier 05min_CP_18-30 when providing it with case study data with 5 min time resolution and varying time slot length.

To correctly evaluate the method's prediction performance, the classifier's misclassification rate was also investigated. Fig. 42 shows the *CMCR2C* metric for the same case that was shown in Fig. 41.

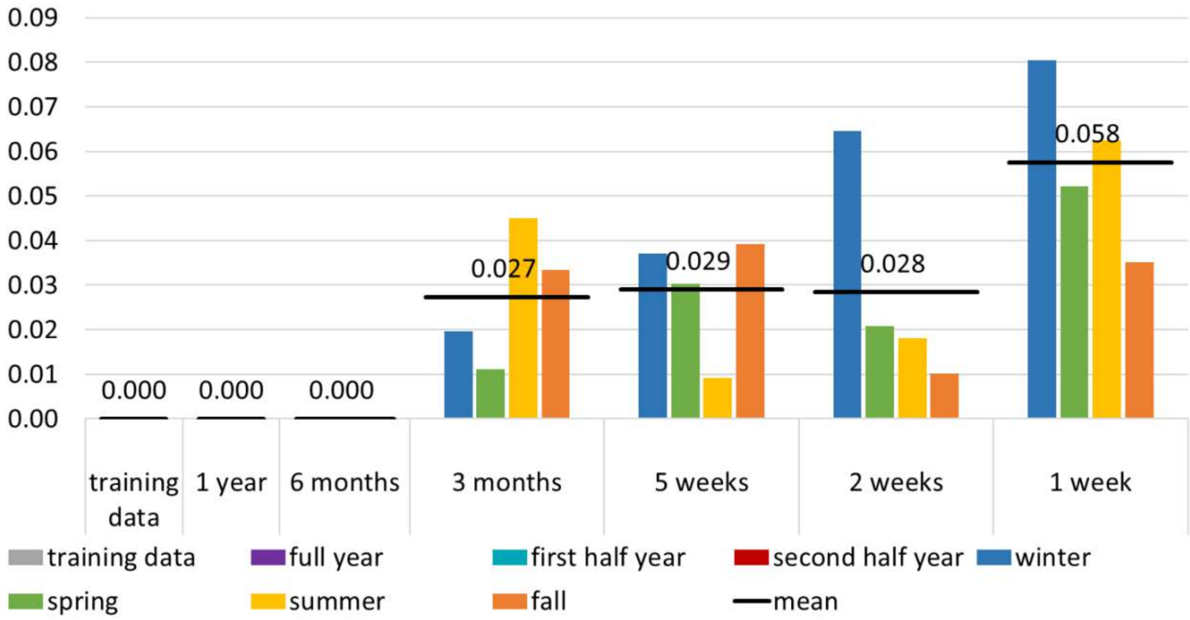


Fig. 42: CMCR2C values achieved by the classifier 05min_CP_18-30 when providing it with case study data with 5 min time resolution and varying time slot length.

As the *CMCR2C* metric excludes misclassifications between certain identical consumers (blinds and PV) and misclassifications between submeters and the meters superior to their main meters, the metric presents an indicator for real misclassifications where there is no obvious cause for the misclassification. Later in this section, the following section, and in appendix E, these excluded misclassifications are also displayed and discussed.

Fig. 42 shows that for the training dataset and the datasets based on monitoring data from time slots with a 1-year or 6-month length, there were no real misclassifications. For the datasets based on time slot lengths of 3 months, 5 weeks, or 2 weeks, the mean *CMCR2C* is approximately 0.03, which can be interpreted as 3% of the main meter-submeter relationships indicated by the classifier are misclassifications. In the case of the time slot length of 1 week, this rate is approximately twice as high. An investigation of the underlying single *CMCR2C* values that are the basis for the mean *CMCR2C* reveals that the misclassification rate can vary significantly from case to case.

A detailed analysis of the *CACC2* and *CMCR2C* metrics of all classifier candidates when used on the 100 datasets from the case study (see also appendix D), revealed the following findings:

- Regardless of the time resolution, the peak accuracy metric *CACC2* that can be observed when using the classifier candidates on the dataset from the 5-week training time slot is almost always approximately 0.6.
- A classifier should only be used on datasets that were calculated out of the monitoring data that has the same time resolution that the classifier was trained with. In some cases, classifiers with a higher time resolution than the dataset achieve higher *CACC2* values than classifiers with the same time resolution as the dataset. The downside of this is that with the higher *CACC2* values, there are also higher *CMCR2C* values. As misclassifications shall be avoided, these cases are deemed worse than the cases with fitting time resolution.
- Binary classifiers generally achieve higher *CACC2* values than multiclass classifiers. Again, the downside is that there are higher *CMCR2C* values. As misclassifications shall be avoided, the binary classifiers are deemed worse than the multiclass classifiers.
- Classifiers using features that were calculated using both available sources, the counter values and the instantaneous power values (CP), performed better than classifiers using only features calculated out of the counter values (C).
- A higher number of features used by a classifier leads to better prediction performance.
- Reducing the maximum tree depth does not improve the prediction performance. Classifiers with higher maximum tree depth generally show better prediction performance.
- Providing data from time slots with a higher time slot length leads to better prediction performance.

- Given that the dataset has the same time resolution as the classifier, the higher the time resolution, the better the prediction performance.

- For each time resolution, there is at least one classifier candidate with the same time resolution that achieves a *CMCR2C* value below 0.1 and a *CACC2* value above 0.15, given the time slot length is long enough. If a misclassification rate of approximately 10% (*CMCR2C* = 0.1) is considered an acceptable error rate, this result can be interpreted that the method basically works for all investigated time resolutions. This interpretation was later proven wrong with the second set of validation data, see the following section 5.3 or the second half of the appendix D. Considering both sets of validation data, only classifiers with a time resolution of 5 min, 10 min, or 15 min achieve acceptable prediction performances in both cases. This issue is elaborated further in section 5.3.

In appendix D, some classifiers are highlighted in red. These are the classifier candidates that were chosen as the six final best classifiers – one for each time resolution and data source setting (CP or C):

- 05min_CP_18-30
- 05min_C_7-30
- 10min_CP_55-30
- 10min_C_28-30
- 15min_CP_55-30
- 15min_C_28-30

They were chosen according to the performance that they showed when used on the validation data from the case study and the validation data from the synthetically generated data. As can be seen, by their designation, all of them are multiclass classifiers, and they all use the highest value of the allowed maximum tree depth, namely 30. Besides the classifiers that were trained with the 5 min data, all of them use the maximum number of features available for each data source setting.

Fig. 43 and Fig. 44 illustrate the prediction performance achieved by these final best classifiers when used on the validation data of the case study. The figures show the influence of the time slot length of the datasets on the *CACC2* and *CMCR2C* metrics.

Fig. 43 shows that the classifiers achieve a *CACC2* value of approximately 0.6 when used on the data that the classifiers were trained with. For all other cases, the *CACC2* stays between 0.2 and 0.27 at an almost constant level. For the CP classifiers, there is a slight drop from

5 weeks to 2 weeks and from 2 weeks to 1 week. Besides the fact that the CP classifiers have slightly higher *CACC2* metrics than the C classifiers, there is little difference between the six final classifiers. More differences can be observed when investigating the *CMCR2C* metrics, see Fig. 44.

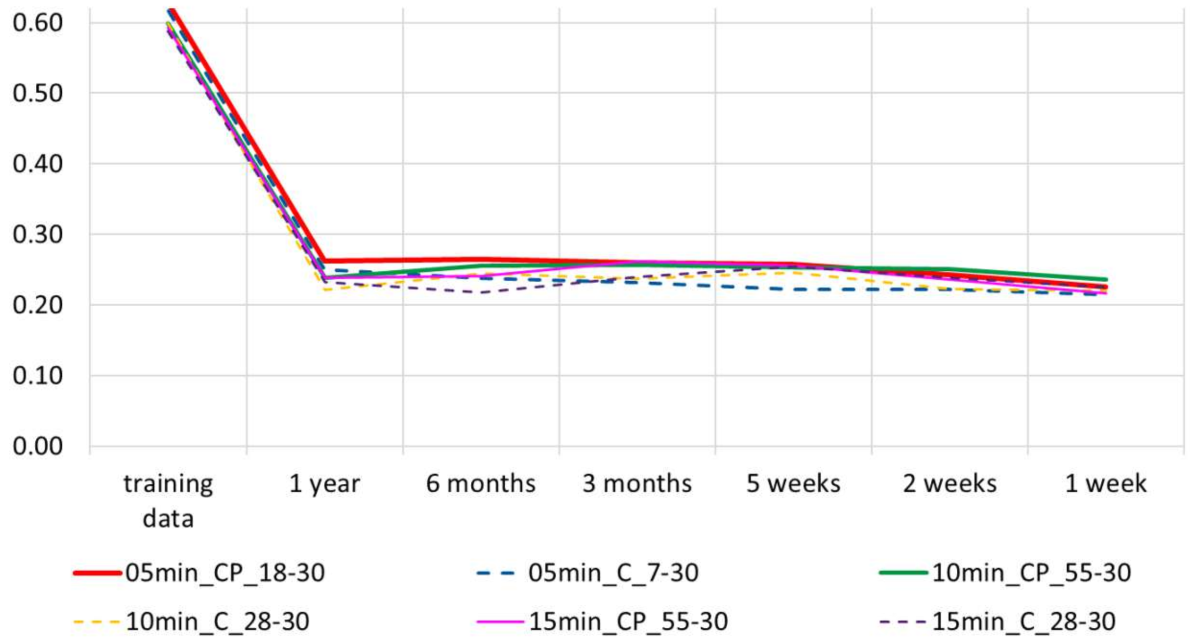


Fig. 43: Mean CACC2 values achieved by the six final classifiers when providing them with case study data with varying time slot length – the time resolution of the provided data fitting to the resolution that the models were trained for.

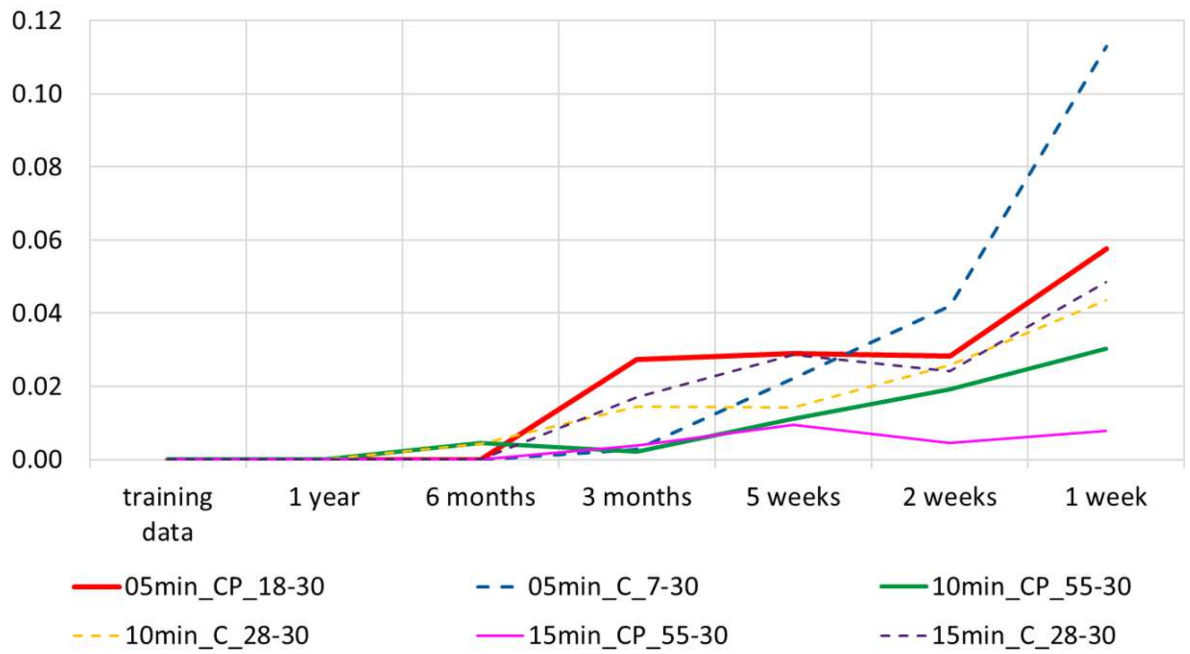


Fig. 44: Mean CMCR2C values achieved by the six final classifiers when providing them with case study data with varying time slot length – the time resolution of the provided data fitting to the resolution that the models were trained for.

Fig. 44 shows the trend that the smaller the time slot length, the worse the *CMCR2C*. While the *CMCR2C* is zero or almost zero for the cases of the training data, the time slot with a length of 1 year, or the time slots with a length of 6 months, the metric increases for data from time slots with smaller time slot lengths. For data from time slots with a length of 3 months, 5 weeks, or 2 weeks the *CMCR2C* value lies at 0.04 or below, which can be interpreted as a maximum of 4% of the cases with indicated main meter-submeter relationship are misclassifications. When the time slot length drops to 1 week, for most cases, higher misclassification rates must be expected.

Surprisingly, the classifiers with the highest time resolution of 5 min have the highest *CMCR2C* values at time slot lengths of 2 weeks or 1 week. As one week's worth of data in a higher time resolution encompasses more data points than one week worth of data with a lower time resolution, there is more information in the data with a higher time resolution. Thus, it would have been expected that the classifiers would achieve the best performance with the highest time resolution, i.e., the 5 min data.

An investigation of the underlying single *CMCR2C* values of the separate cases from each season revealed no obvious reasons for that. It is suspected that the reason for this behavior is the monitoring data from the energy meters with low meter resolutions, i.e., primarily the heat meters.

To this point, the prediction performance of the classifiers has only been discussed based on the *CACC2* and *CMCR2C* metrics. Fig. 45 and Fig. 46 display the results in another, more detailed way. Both figures illustrate the number of correctly identified main meter-submeter relationships and the misclassifications that wrongfully indicate a main meter-submeter relationship. Again, the average results that the six final classifiers achieve are presented for each of the investigated time slot lengths.

The misclassifications are split into different categories to enable a more detailed analysis. Contrary to the misclassification rate metric *CMCR2C*, the misclassifications between certain identical consumers (blinds and PV) and between submeters and the meters superior to their main meters are not excluded. Instead, they are displayed as separate categories.

As can be seen in Fig. 45, all classifiers could, on average, at least correctly identify 37 of the real main meter-submeter relationships – regardless of the time slot length. The number of correctly identified main meter-submeter relationships is slightly higher in the case of the CP classifiers than in the case of the C classifiers.

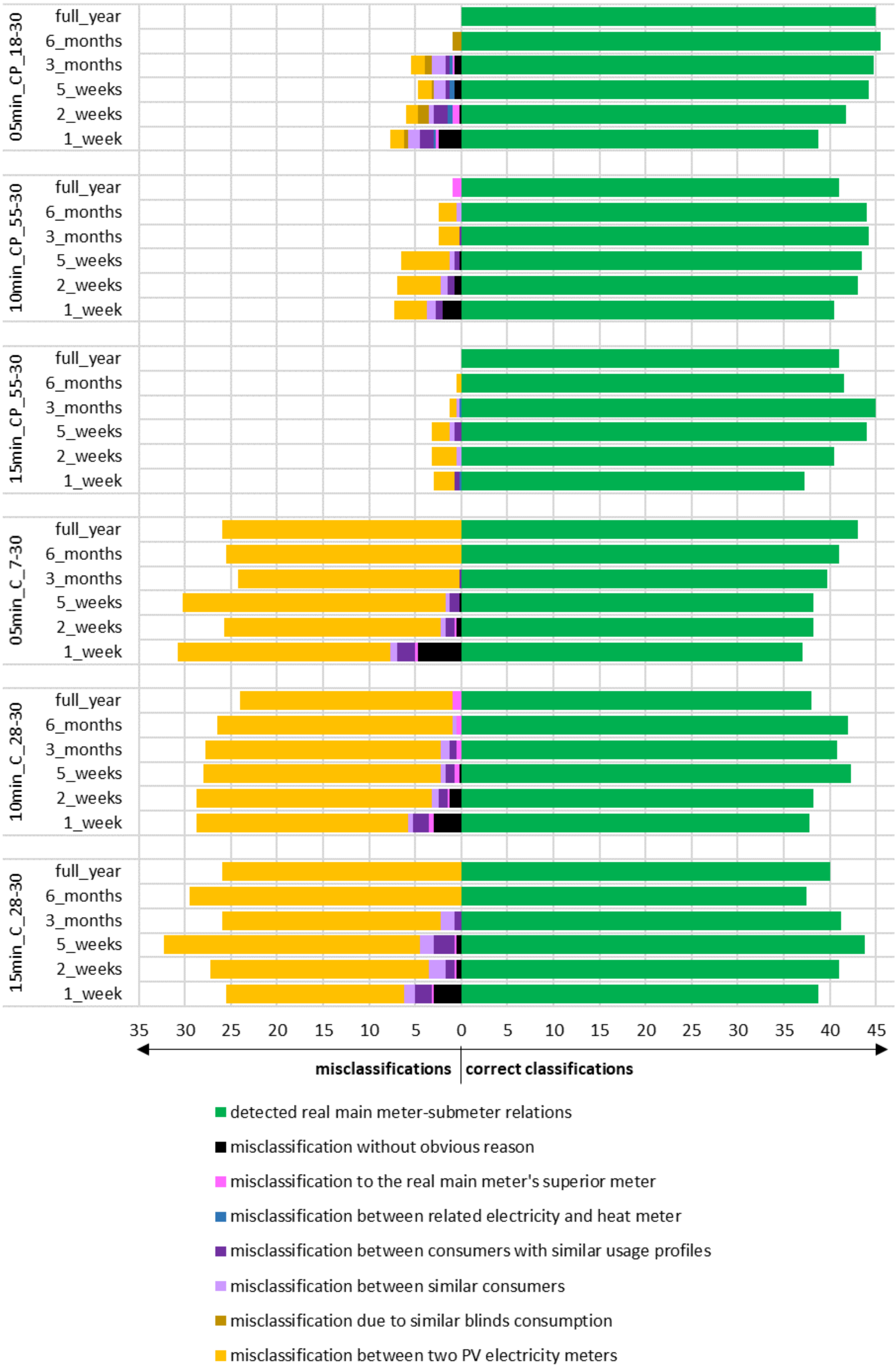


Fig. 45: Average number of cases classified to the clusters "obvious relations", "rare partial supply" and "inconspicuous relations" when using the method on case study data from other time slots.

Moreover, Fig. 45 shows that the CP classifiers lead to fewer misclassifications than the C classifiers. The difference can mostly be attributed to the misclassifications between two PV electricity meters. This observation indicates that the features calculated with the instantaneous power values load profiles carry additional information that aids in preventing misclassifications between electricity meters that measure similar PV supply.

The misclassifications that can be attributed to the similar electricity consumption of the case study's automatic blinds are surprisingly low. They only occur in the predictions of the classifier 05min_CP_18-30.

Misclassifications between similar consumers are a category of misclassifications between consumers that are of the same type and have the same size in terms of energy consumption. E.g., the case study's two positive pressure ventilation systems or the two parts of the safety lighting system. Naturally, in this category, there are only misclassifications between energy meters of the same type, i.e., between two electricity meters or two heat meters.

In the category "misclassifications between consumers with similar usage profiles", the consumers have different types or different sizes, but they are used at the same time due to user presence or interaction; e.g., the heat consumption of the air curtains at the ground floor and the electricity consumption of the lecture hall, which is accessed by walking through the air curtains. In this category, misclassifications between any energy meter – regardless of the type – are possible.

Each of the final classifiers has, at least in some cases, misclassifications that fall into these two categories. Interestingly, in the case of the C classifiers, the number of these misclassifications seems to decrease with increasing time resolution. Still, in the case of the CP classifiers, the number seems to increase with increasing time resolution. An investigation of the underlying cases did not reveal any obvious reason for that.

Another category with very few cases is the category "misclassifications between related electricity and heat meter". E.g., the electricity consumption of a laboratory room and the thermal power of the room's cooling system. Misclassifications of this category occurred almost exclusively in the predictions of the classifier 05min_CP_18-30.

Misclassifications of the category "misclassifications to the real main meter's superior meter" can primarily be observed in the results of the classifier 10min_C_28-30. There they occur at every time slot length. In the results of the other classifiers, they mostly occur at time slot lengths of 2 weeks or 1 week. The classifier 10min_CP_55-30 only has misclassifications of this category at the time slot length of a full year, and the classifier 15min_CP_55-30 does not have them at all.

While there are some causal relationships in the misclassification categories that were discussed until now, this is not the case for the category “misclassification without obvious reason”. All misclassifications where there seems to be no causal relationship are put into this category. Hence, this category can be considered as some sort of general indicator for the quality of the method. If the number of entries in this category is too high, the method might be unsuitable for the given case. As the goal of the method is to detect real main meter-submeter relationships by comparing the meter’s load profiles, misclassifications without obvious reason are a danger to the usefulness of the method. If there are misclassifications that cannot be attributed to similarities in the load profiles, the classifiers seem to be trained to detect something besides the similarities. Considering that most of the misclassifications without obvious reason occurred in the cases where there is a time slot length of 1 week, this would be an argument against the use of the method in such cases. It can be concluded that the method should only be used when there is monitoring data from at least 2 weeks.

In the case of the 15min_CP_55-30 classifier, there is not even one misclassification of this category. Overall, this classifier shows extremely low numbers of misclassifications while the number of detected real main meter-submeter relations remains relatively high.

Fig. 45 illustrates the prediction results according to the nature of the underlying classification problem, namely in a binary way. Classifications of detected real main meter-submeter relations to the clusters “obvious relations”, “rare partial supply”, and “inconspicuous relations” are considered correct predictions – even if they were misclassified in between those three clusters. Only misclassifications from the cluster “no relations” to these three other clusters are considered real misclassifications. As the goal of using multiclass classifiers for the method was to have a cluster where the predictions to this cluster can be considered especially trustworthy, it must be evaluated whether it was reached or not. Fig. 46 is an adjusted version of Fig. 45, where the results are only presented for the cluster “obvious relations”.

Fig. 46 clearly shows that the goal was reached: There are almost no misclassifications in the cluster “obvious relations”. Most of the few misclassifications are from the categories of the expected misclassifications between PV electricity meters or due to similar blinds consumptions. As these misclassifications are caused by the fundamental issue that consumers with identical or almost identical load profiles are practically impossible to distinguish from each other by their load profiles only, these misclassifications must be expected within the method’s results.

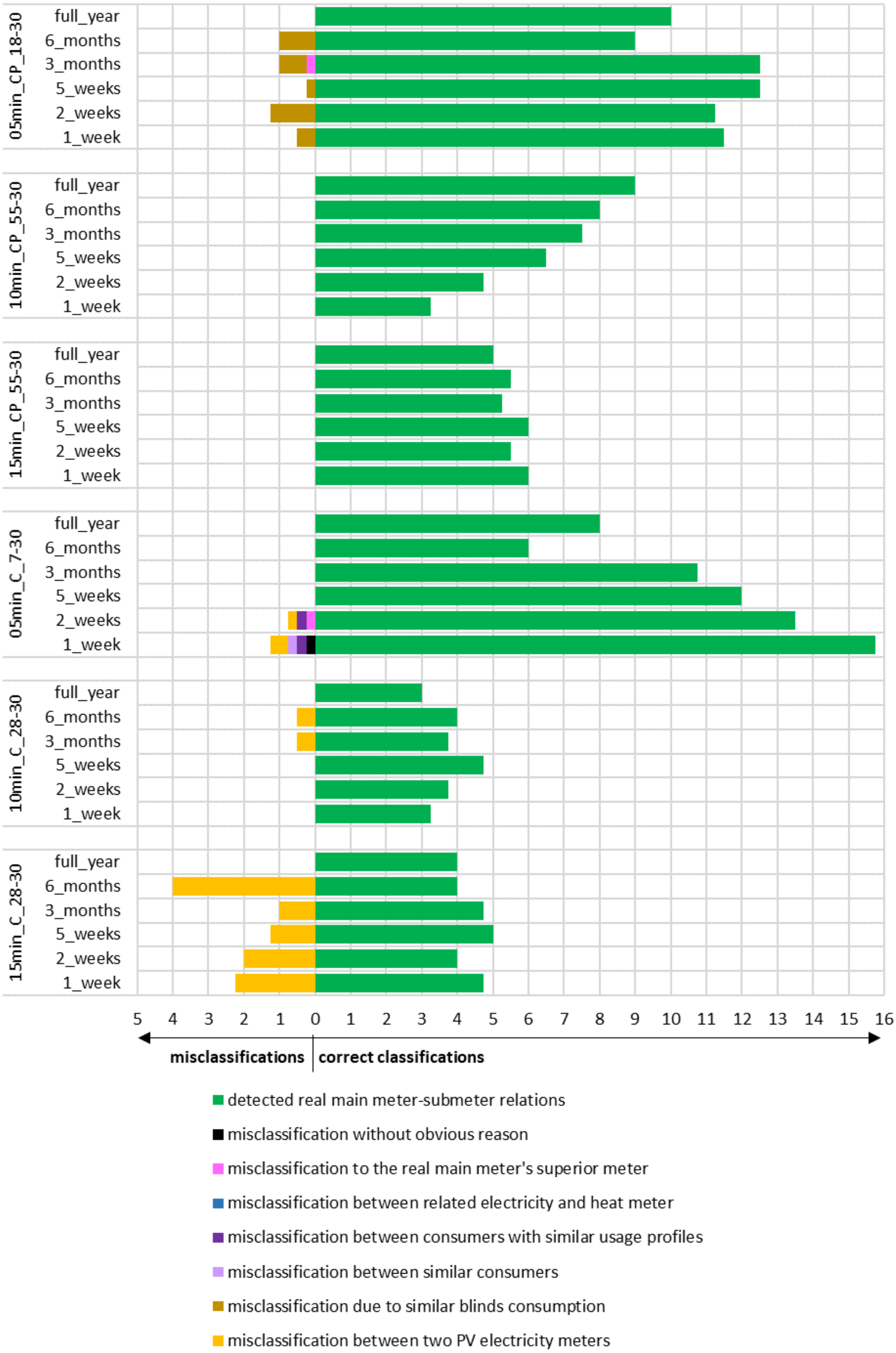


Fig. 46: Average number of cases classified to the cluster "obvious relations" when using the method on case study data from other time slots.

Regarding the influence of the time slot length, there is no clear pattern that can be observed in the results depicted in Fig. 46. In the case of the classifier 10min_CP_55-30, the number of detected real main meter-submeter relations increases with increasing time slot length, but in the case of the classifier 05min_C_7-30, the number decreases with increasing time slot length. The results of the other classifiers do not show a clear trend.

Except for the classifier 10min_C_28-30, there is a general trend that with increasing time resolution, the number of detected real main meter-submeter relations increases too.

Until now, the prediction results of the classifiers have been presented in a condensed way: Either the average values of several cases were shown, or the results were illustrated by metrics. In the network diagram in Fig. 47, the prediction results that one single classifier achieves in one single case are presented. In appendix E, there are further examples of the results that the six final best classifiers achieved in different cases. These examples are approximately one-third of all the evaluated cases. They visualize the prediction performance of the method that can be expected when it is provided with monitoring data, which is like the monitoring data of the case study.

The basis for the network graph in Fig. 47 is the network graph that was already presented in section 3.1 in Fig. 12: The nodes represent the energy meters on different levels of the hierarchy, and the connecting grey lines represent real main meter-submeter relationships.

The colored lines represent the correctly identified main meter-submeter relationships and the misclassifications between two meters without main meter-submeter relationship. The coloring scheme and the categories are identical to those used in Fig. 45 and Fig. 46. By using three different line styles, it is indicated in which cluster the classifier predicted each meter combination pair. The visualization cannot depict which of the two meters is the assumed main meter and which is the assumed submeter. In the case of correct predictions and misclassifications over two different meter hierarchy levels, the assumed main meter is generally the one on the higher level, and thus, the assumed submeter is the one on the lower level.

In the top area of the network graph, all relevant information about the classifier and the underlying data is presented. Further, all classification evaluation metrics, as described in sections 4.7.1 and 5.1, are shown.

The correct predictions to the cluster “no relations” are not illustrated. Their number can be estimated by subtracting the number of the filtered-out cases (7,609) and the number of real main meter-submeter relations (180) from the number of possible meter combinations (24,492).

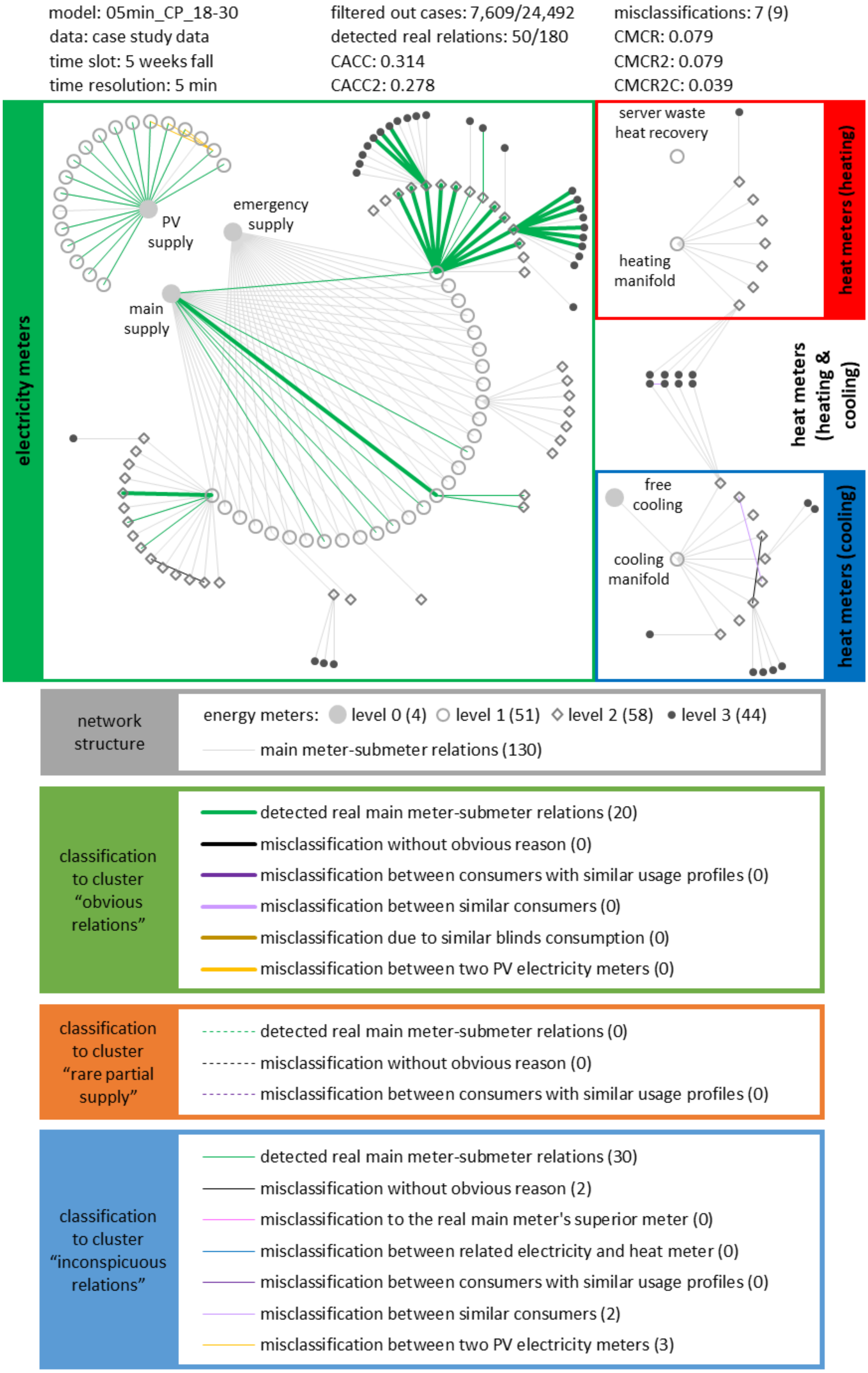


Fig. 47: Network diagram illustrating the prediction performance of the classifier 05min_CP_18-30 when providing it with case study data from the time slot “5 weeks fall” with 5 min time resolution.

Misclassifications of a submeter to its main meter's superior meter, provided the real main meter-submeter relationship was detected as well, are not directly illustrated in the network graph in Fig. 47. Such misclassifications can easily be identified by comparing the average power level of the two meters that were indicated as main meters. The meter with a lower power level will be the real main meter, and the one with a higher power level will be the main meter's superior meter. The number of the misclassifications that were not visualized is indirectly shown in the top area under the parameter "misclassifications": It is the difference between the visualized misclassifications – the number outside of the brackets – and the total number of misclassifications – the number in the brackets.

An analysis of the results of the case presented in Fig. 47 shows that approximately 28% of the real main meter-submeter relations were correctly detected by the method. All of these correctly detected relations were between electricity meters. The method failed to detect real relations between heat meters. The relatively low counter resolution of the heat meters and the thermal inertia of the heating and cooling system explains this. Further, the method also failed to detect real relations between the electrical emergency supply and its submeters. This is not surprising, as the emergency supply is only active for a few hours during the year. If the relations between the emergency supply and its submeters (29 relations) and the relations between heat meters (33 relations) are excluded, the rate of correctly identified relations increases to 42%.

Setting up the method with multiclass classifiers partially works as intended: The predictions to the cluster "obvious relations" are free of misclassifications – the misclassifications occurred only in the cluster "inconspicuous relations". Meanwhile, there is not one single prediction to the cluster "rare partial supply". As all entries of this cluster are meter combinations between the electrical emergency supply and its submeters, the reason for that is, again, that the emergency supply is only active for a few hours during the whole year.

An investigation of the misclassifications reveals that, besides the three misclassifications between two PV electricity meters, there is only one misclassification between the electricity meters. Contrasting these misclassifications with the fact that the method could correctly identify 50 real main meter-submeter relations between electricity meters, the method can be considered successful – at least in the case of electricity meters.

All other misclassifications occurred between heat meters. As the method failed to predict any real main meter-submeter relations between the heat meters, misclassifications between them do neither contribute to nor endanger the method's usefulness for the identification of the hierarchy of heat meters.

The case presented in Fig. 47 is a good example of the average prediction performance of the method. Appendix E shows the full range of cases – some cases with slightly better prediction performance and some with slightly worse prediction performance.

A detailed analysis of all example cases of the case study in appendix E leads to similar findings that can be summarized as follows:

- The method basically works for electricity meters.
- The method basically fails for heat meters. Only in very few cases the method correctly predicted a main meter-submeter relationship between heat meters.
- Predictions to the cluster “obvious relations” are more trustworthy than predictions to the cluster “inconspicuous relations”, as there are fewer misclassifications in the cluster “obvious relations” than in the other one.
- Only when provided with the original training data the method can detect real main meter-submeter relations that belong to the cluster “rare partial supply”. Otherwise, there is not one single case where a real main meter-submeter relation was predicted to that cluster. As in only one single case, there is a misclassification to this cluster, the total removal of this cluster from the method would not endanger the prediction performance of the whole method.
- Classifiers that utilize features derived from both data sources, counter values, and instantaneous power values (CP), perform better than classifiers that utilize only features derived from the counter values (C). In the case of the C classifiers, there are significantly more misclassifications between two PV electricity meters.
- With increasing time slot length, the prediction performance of the method increases slightly. For the case of a time slot length of 1 week, the number of misclassifications significantly rises.

5.3 Synthetically generated data

The second part of the method is conducted by using all candidates for the best classifiers on data that is derived from the synthetically generated data. As described in section 3.2, the synthetical data was generated by combining random pieces from measured load profiles. Thus, the synthetically generated data lacks seasonal differences. Fig. 48 depicts how the validation data time slots were chosen from the one year of synthetically generated data.

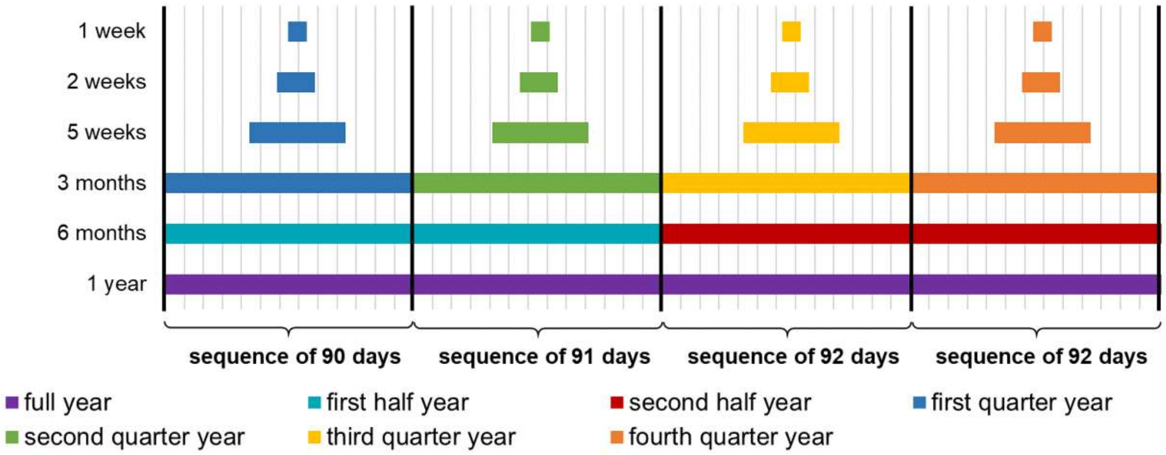


Fig. 48: The time slots of the synthetically generated data which were used for the performance evaluation of the candidates for the best multiclass and binary classifiers.

The selection of the time slots was conducted in the same manner as with the validation data chosen from the case study's monitoring data. The only differences between the case study (Fig. 40) and the synthetically generated data (Fig. 48) are that the time slots are not named according to seasons and that there is no training data time slot. Thus, in the case of the synthetically generated data, there are only 19 time slots instead of 20.

The data from these 19 time slots were also resampled to the time resolutions that the classifiers were trained for. Hence, the method was validated with 95 different datasets that come from the 19 mentioned time slots and have a time resolution of 5 min, 10 min, 15 min, 30 min, and 60 min.

The method's steps for smoothing, feature extraction, cleaning, and clustering – as described in sections 4.4 to 4.6.2 – were conducted with each of these 95 different datasets. The result is 95 new datasets suitable to be used as input for the classifiers: Each of them consists of the 55 features for each possible meter combination, provided the combination was not automatically filtered out. All candidates for the best multiclass and binary classifiers were then applied to these 95 new datasets, regardless of the time resolution, the classifiers were trained with.

The results for the *CACC2* and *CMCR2C* metrics that each classifier achieved for the 95 datasets are presented after the results of the case study in appendix D, i.e., in the second half. Fig. 49 and Fig. 50 illustrate the results of one classifier as an example.

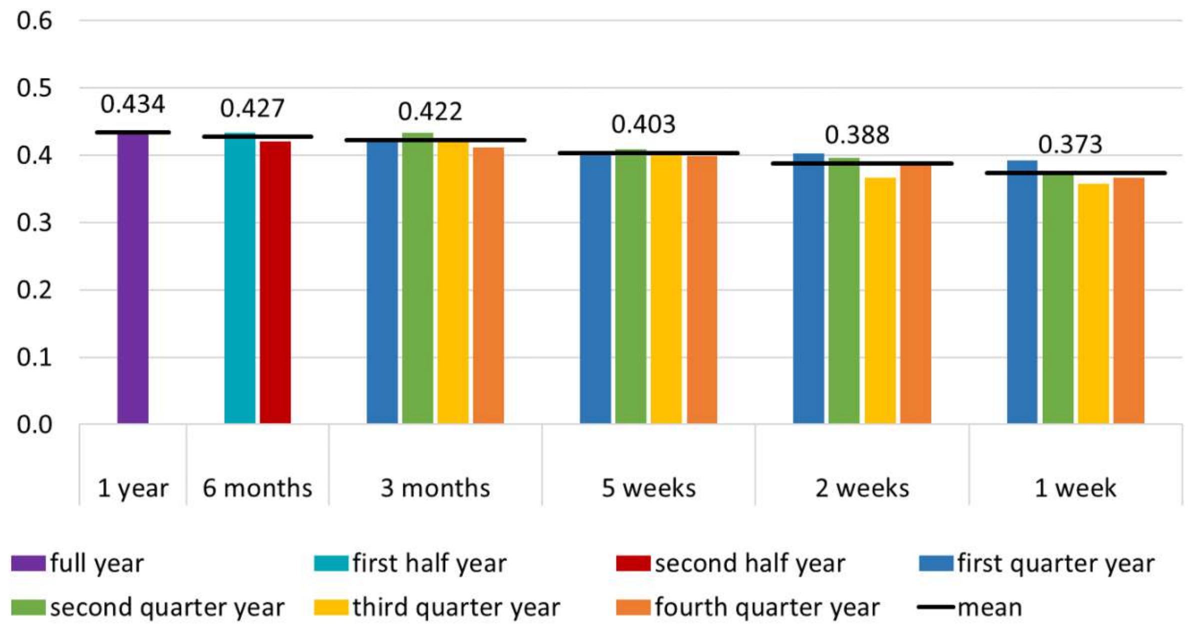


Fig. 49: CACC2 values achieved by the classifier 05min_CP_18-30 when providing it with synthetically generated data with 5 min time resolution and varying time slot length.

Fig. 49 shows *CACC2* metrics that are quite like the *CACC2* metrics that the same classifier achieved when applied to the case study data – see Fig. 41 – but on a higher level and in a narrower band. The *CACC2* values are in a band between 0.35 and 0.44. Similar to the case of the case study data, the values seem to slightly decrease with decreasing time slot length. The higher level is because the synthetically generated data is exclusively from electricity meters. The lower *CACC2* level in the graph of the case study, Fig. 41, can be explained by the main meter-submeter relations of heat meters that the method failed to identify. If the heat meter combinations would be removed from the calculation of the *CACC2*, the *CACC2* values achieved with the case study data would increase to around 0.4, which means that approximately 40% of the real main meter-submeter relations would have been correctly detected.

Regarding the misclassification metric *CMCR2C*, the graph of the synthetically generated data, see Fig. 50, looks quite different from the case study graph displayed in Fig. 42. For datasets derived from monitoring data with a time slot length of 3 months, 5 weeks, or 2 weeks there were practically no misclassifications – the *CMCR2C* value is either exactly zero or at least near zero. For the case of the case study data, the value was around 0.03. On the other hand, for time slot lengths of 1 week, the mean *CMCR2C* value of around 0.18 is much higher than the value achieved with the case study data, which was around 0.06.

When analyzing the results in the second half of appendix D, the same can be observed with the other cases: The method achieves relatively low *CMCR2C* values at time slot lengths larger than 1 week and relatively high *CMCR2C* values at time slot lengths of 1 week.

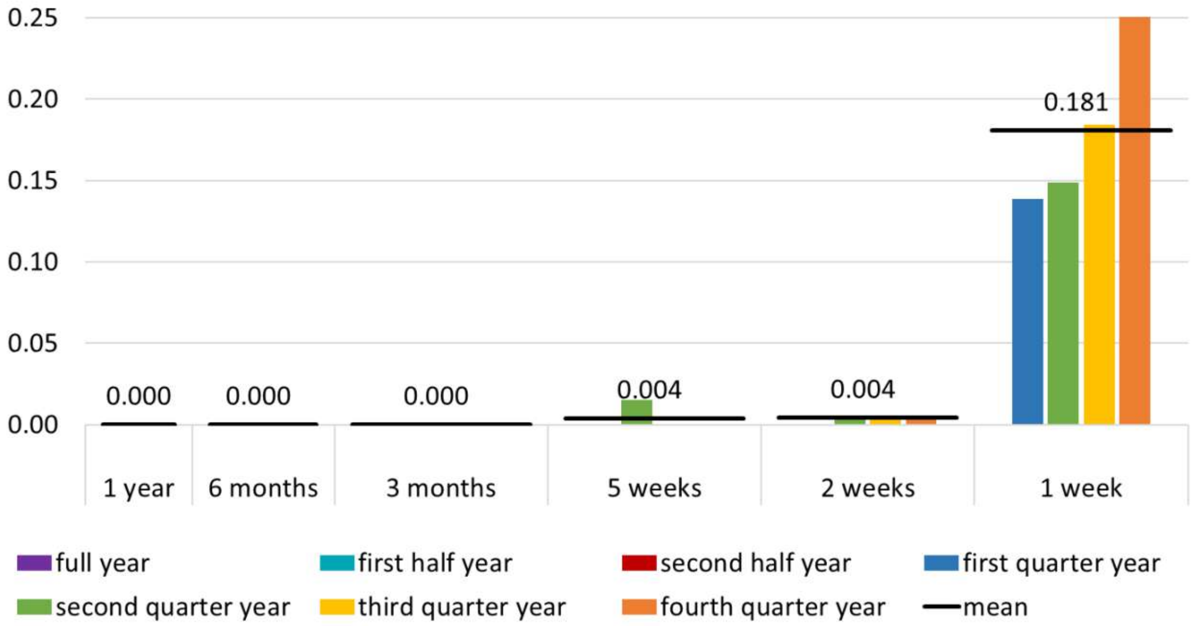


Fig. 50: CMCR2C values achieved by the classifier 05min_CP_18-30 when providing it with synthetically generated data with 5 min time resolution and varying time slot length.

Further, as already indicated during the selection of the six final best classifiers in section 5.2, the prediction results that the classifier candidates achieved with the synthetically generated validation data with a time resolution of 30 min and 60 min are relatively poor. In most of the cases, the accuracy metric *CACC2* is below 0.1, i.e., less than 10% of the real main meter-submeter relations were detected by the method.

Even though some of the 30 min and 60 min classifiers achieved *CMCR2C* values like those displayed in Fig. 50, the low *CACC2* metrics are deemed unacceptable. Thus, the six final best classifiers were selected from those that were trained with data that has a time resolution of 5 min, 10 min, or 15 min.

Fig. 51 and Fig. 52 illustrate the prediction performance achieved by these final best classifiers when used on the validation data derived from the synthetically generated data. The figures show the influence of the time slot length of the datasets on the *CACC2* and *CMCR2C* metrics.

A comparison of Fig. 51 with its counterpart, the graph of the *CACC2* metrics achieved with the case study data, Fig. 43, reveals that there are significant differences: For the case study data, the *CACC2* values are on a similar level, in case of the synthetically generated data, they are not.

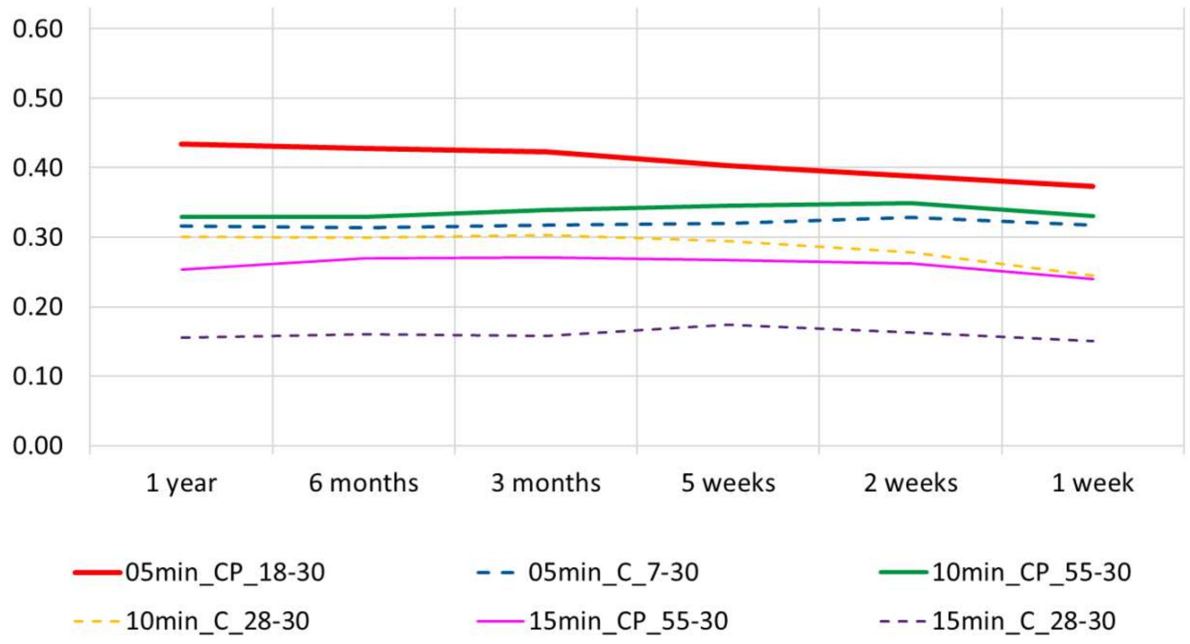


Fig. 51: Mean CACC2 values achieved by the six final classifiers when providing them with synthetically generated data with varying time slot length – the time resolution of the provided data fitting to the resolution that the models were trained for.

While the *CACC2* values of all six classifiers stayed almost constant in a band between 0.2 and 0.27 in the case of the case study, Fig. 51 shows that the values stay almost constant at very distinct levels in the case of the synthetically generated data. The classifier 05min_CP_18-30 shows values around 0.4, while 15min_C_28-30 stays slightly above values of 0.15. All other classifiers have almost constant values on levels in between those two extremes.

This very significant difference between the two sources of validation data indicates that the classifiers are somewhat biased towards the case study data. An investigation of the *CMCR2C* metrics, displayed in Fig. 52 indicates that the bias does not render the method useless for other data besides the case study data. Except for the case of data from time slots with a length of 1 week, the six best classifiers achieved even lower and thus better *CMCR2C* metrics with the synthetically generated validation data than with the case study validation data.

Fig. 52 presents that the *CMCR2C* metrics of the two 5 min classifiers and classifier 15min_C_28-30 show a similar pattern as in Fig. 50: Low or near zero *CMCR2C* values at time slot lengths larger than 1 week and relatively high *CMCR2C* values at time slot lengths of 1 week. The other three classifiers also achieved relatively low CMCR2C values for the 1-week time slot.

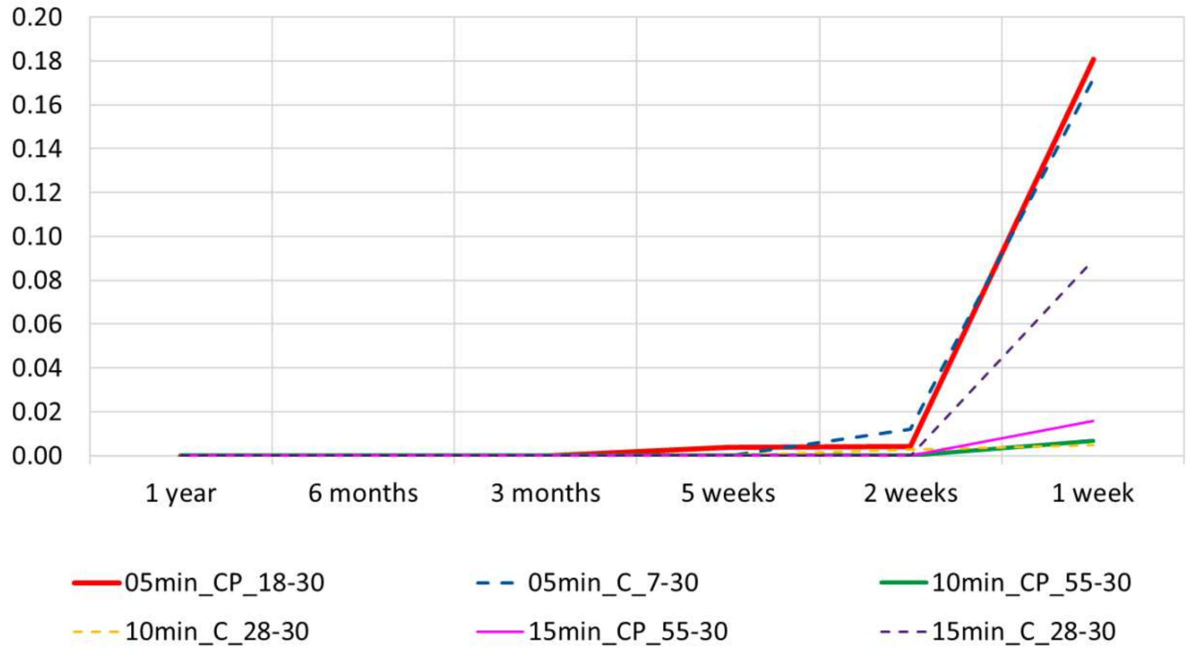


Fig. 52: Mean CMCR2C values achieved by the six final classifiers when providing them with synthetically generated data with varying time slot length – the time resolution of the provided data fitting to the resolution that the models were trained for

That the method can generally achieve *CACC2* values above zero while the *CMCR2C* metrics remain practically zero for most of the cases can be interpreted as proof that the method basically works. This means that the method correctly predicts main meter-submeter relations with a reasonably low chance that the predictions are misclassifications.

Similar to the presentation of the classifier performance in the case of the case study data in Fig. 45 and Fig. 46, the classifier performance in case of the synthetically generated data is shown in Fig. 53 and Fig. 54. Apart from the cases where the time slot length is 1 week, there are almost no major misclassifications, as was also indicated by the near-zero *CMCR2C* metric. The only misclassifications that occurred in a noteworthy number of cases are the misclassifications to the main meter's superior meter. In the case of the 5 min classifiers there are several of them – independent of the time slot length.

Nevertheless, the number of misclassifications is much lower than the correctly identified real main meter-submeter relations. A general tendency can be observed that with increasing time resolution, the number of correctly identified real main meter-submeter relations increases too. This observation is different than the one that was made during the analysis of the results of the case study that were presented in Fig. 45 in the same manner. In Fig. 45, the number of identified real relations stayed quite constant for all cases of time resolution and time slot length.

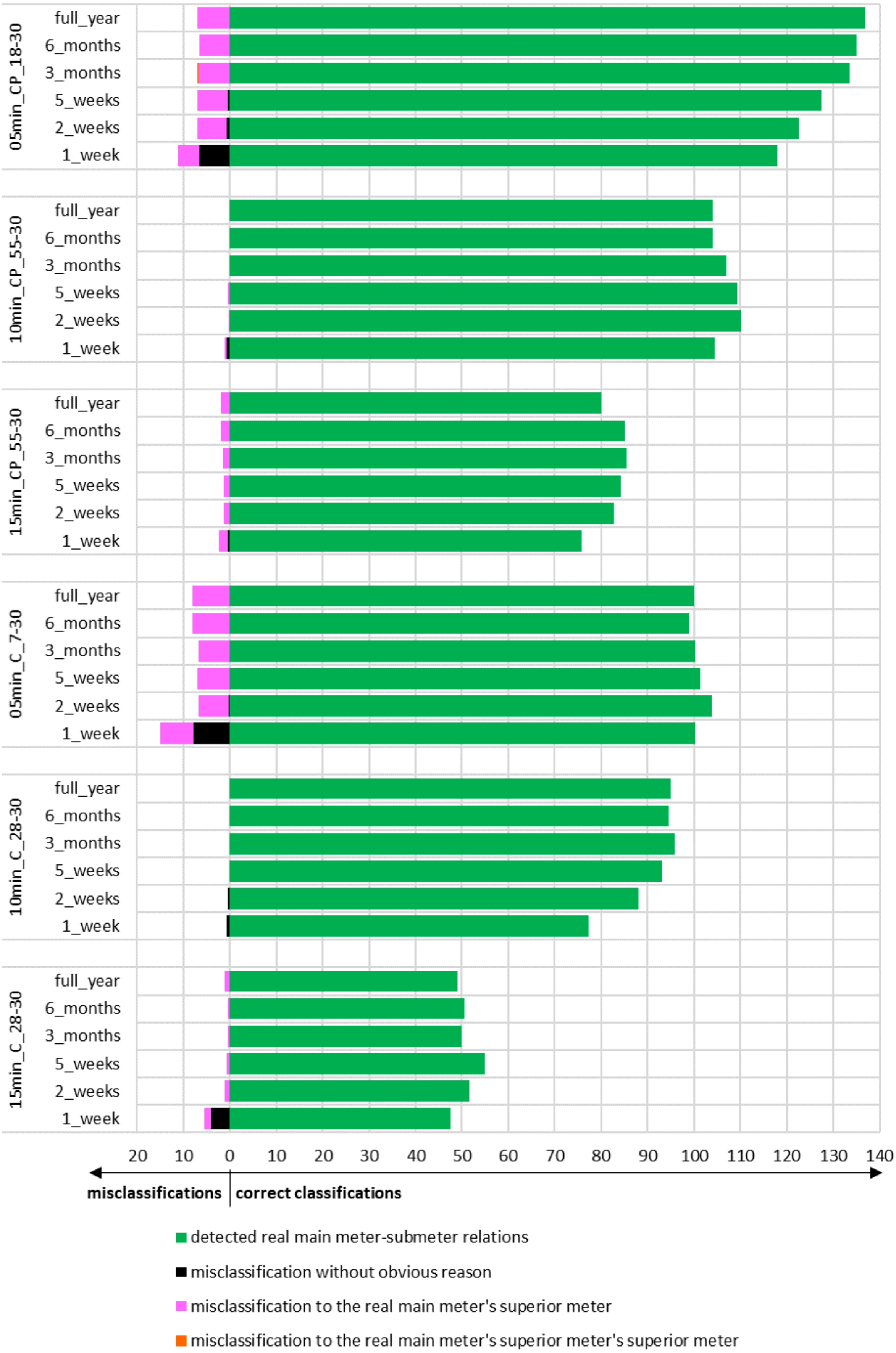


Fig. 53: Average number of cases classified to the clusters "obvious relations", "rare partial supply" and "inconspicuous relations" when using the method on the synthetically generated data.

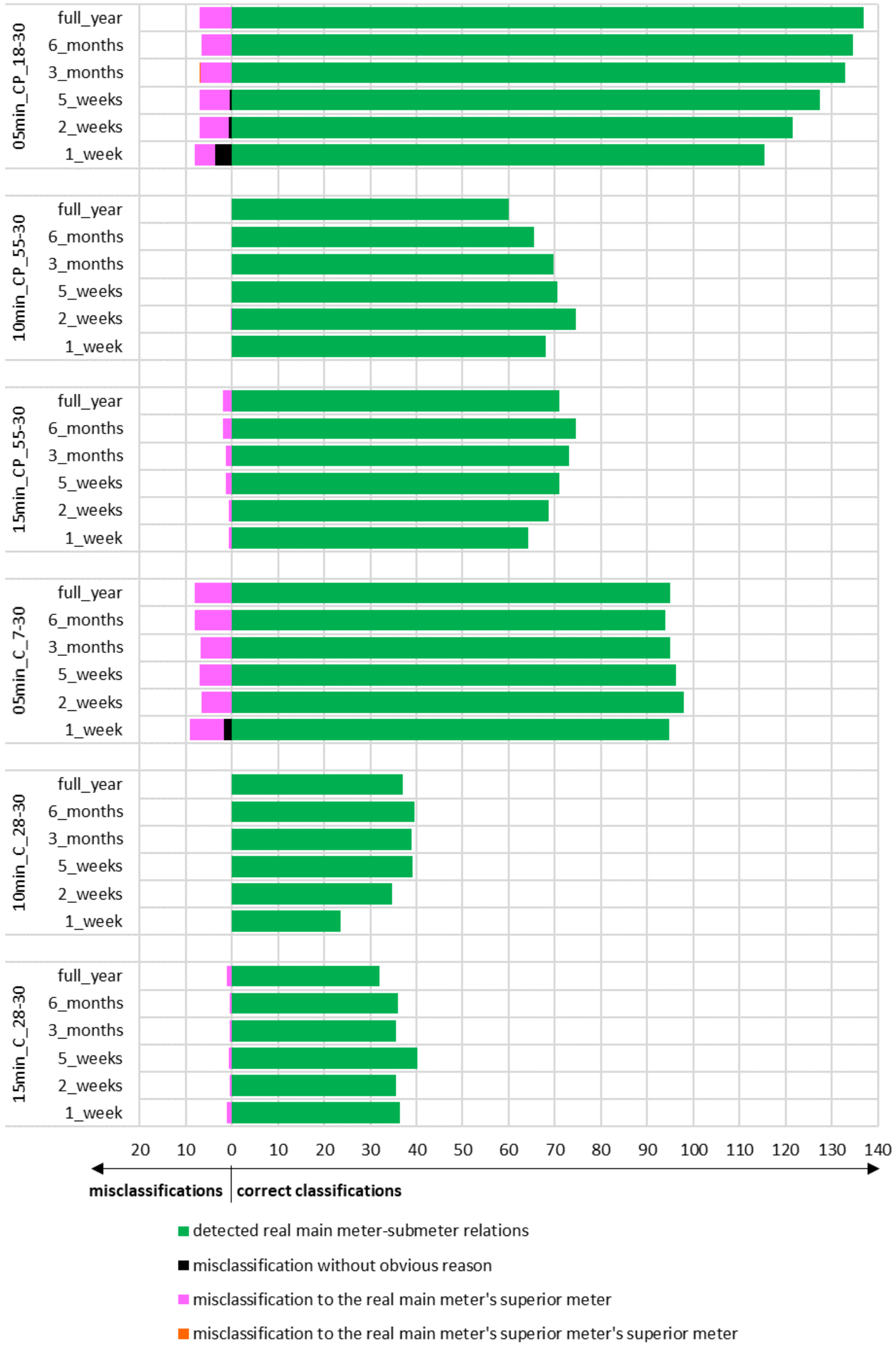


Fig. 54: Average number of cases classified to the cluster "obvious relations" when using the method on the synthetically generated data.

The reason for this difference appears to be the bias towards the case study data. Especially the classifiers with lower time resolution – 10 min and 15 min – seem to benefit from the bias. In the case of the case study data, it enables them to detect more real main meter-submeter relations and thus helps to push their *CACC2* metrics towards the level of the 5 min classifiers.

Fig. 54 is an adjusted version of Fig. 53, where the results are only presented for the cluster “obvious relations”. When comparing those two graphs, it becomes obvious that most of the real main meter-submeter relations detected by the 5 min classifiers were predicted to the cluster “obvious relations”. This observation is valid for all different time slot lengths.

In the case of the 10 min and 15 min classifiers, there is a notable difference between the numbers of the detected real main meter-submeter relations. With the knowledge that there are practically no correct classifications to cluster “rare partial supply”, this means that some of them were classified to cluster “inconspicuous relations”.

As the misclassifications in Fig. 54 are almost the same as in Fig. 53, the intended increased trustworthiness of cluster “obvious relations” was not observed for the case of the synthetically generated validation data. Nevertheless, the number of correctly detected real main meter-submeter relations is relatively high compared to the number of misclassifications – especially when considering that the misclassifications to the main meter’s superior meter are somewhat debatable misclassifications.

Fig. 55 presents the prediction results that one classifier achieved in one case. It is based upon the network graph of the synthetically generated data that was presented in Fig. 17 in section 3.2. The formatting of Fig. 55 is identical to the one of Fig. 47 – the results of the same classifier when used on one case of the case study validation data.

It can be seen in Fig. 55 that the classifier could correctly predict approximately 39% of the real main meter-submeter relations. Besides several misclassifications to a main meter’s superior meter, there is not one single misclassification. 10 of the 33 misclassifications to the main meter’s superior meter are visible in the network graph. The other 23 misclassifications are not displayed as the corresponding real main meter-submeter relationship was detected as well, and thus these misclassifications can easily be identified and filtered out.

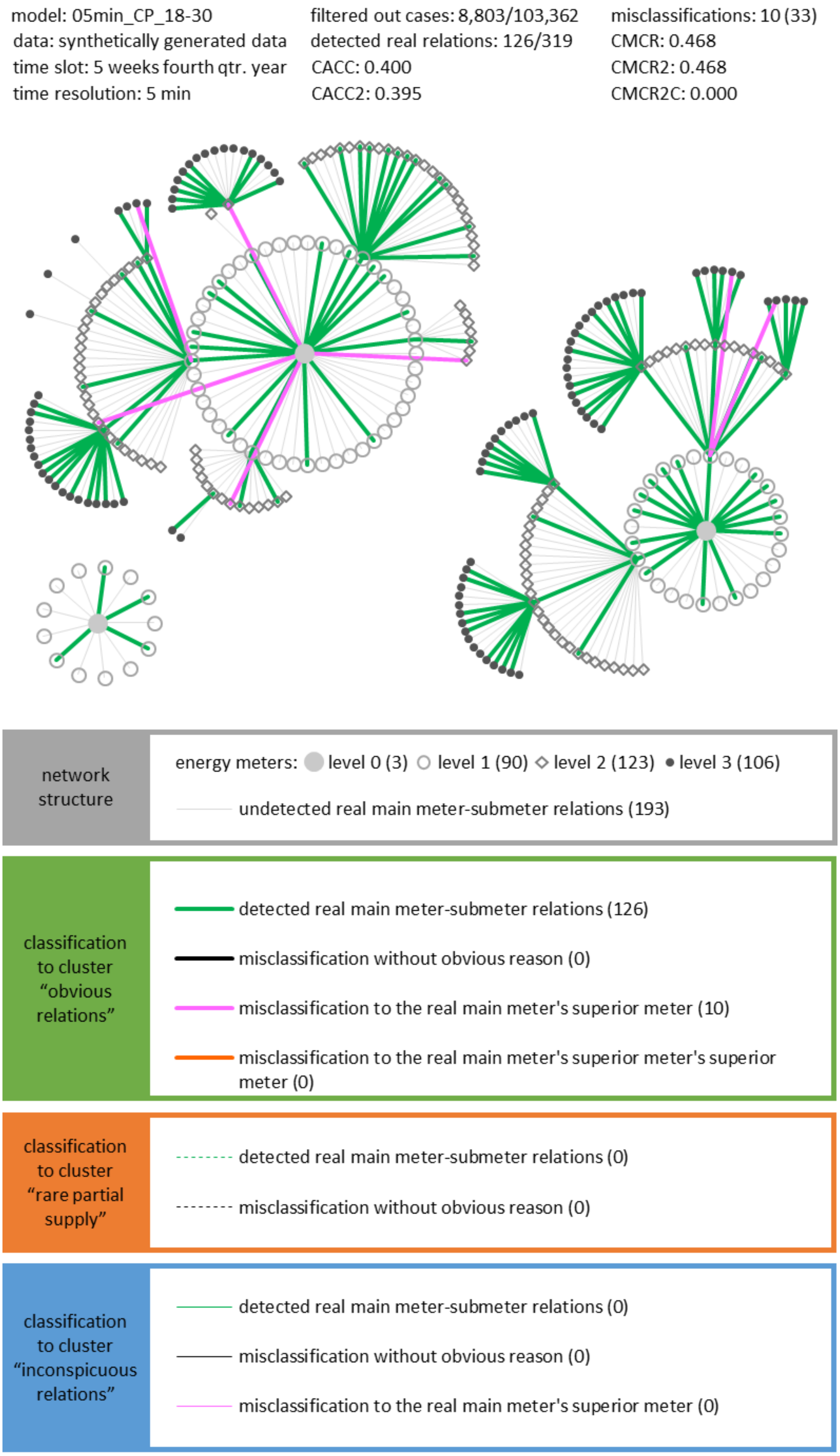


Fig. 55: Network diagram illustrating the prediction performance of the classifier 05min_CP_18-30 when providing it with synthetically generated data from the time slot “5 weeks fourth quarter year” (equivalent to time slot “5 weeks fall” from the case study data) with 5 min time resolution.

Even though these cases can easily be dealt with, they impact certain classification evaluation metrics. In the top area of Fig. 55, the metrics *CMCR* and *CMCR2* are relatively high, while the *CMCR2C* metric is exactly zero. The reason for that is that some of the misclassifications between a submeter and its main meter's superior meter were classified to the cluster "inconspicuous relations". As in the presented case, there are no detected real main meter-submeter relations classified to this cluster, the impact of a cluster full of misclassifications on the *CMCR* and *CMCR2* metrics is significant. Respectively, the misclassification metrics should have been set up in another way. Due to their actual setup, they are prone to such numerical issues.

In the second half of appendix E, there are further examples of the prediction performance of the six final best classifiers for some cases of the synthetically generated validation data. The prediction performance of the classifiers generally behaves as presented in Fig. 51 and Fig. 52: There are almost no misclassifications for cases where the time slot length is larger than 1 week, and the number of detected real main meter-submeter relations decreases with decreasing time resolution. Further, C classifiers perform worse than CP classifiers.

6 Conclusion

This thesis presents a method that can detect main meter-submeter relationships solely by analyzing the main energy monitoring data from energy meters: Measured counter values and instantaneous power values. The method is based on the machine learning method "random forest". It is shown how the energy monitoring data must be processed to be suitable to be used as input for random forests. Moreover, it is shown how the random forests must be trained to fulfill the objective of detecting main meter-submeter relationships.

There are already methods that fulfill a similar purpose, but they differ from the method developed in this thesis in various aspects: Either the methods use voltage measurements to infer main meter-submeter relationships, or they require that there are no unmetered consumers in the meter hierarchy.

The method presented in this thesis uses the time series of counter values and instantaneous power values to infer main meter-submeter relationships, and it allows for some consumers to be unmetered. It does not directly evaluate the load profiles of the energy meters. Instead, it derives characterization indicators, so-called features, from the load profiles of two energy meters, where one is assumed to be the main meter and the other one to be the submeter. These features are then used as input for random forest classification.

The method was developed based on the monitoring data of a case study, the Plus-Energy Office High-Rise Building – a highly efficient office building used by TU Wien. Additional, previously unused data from this case study and synthetically generated data were used to validate the method and estimate its application limits. The findings of the development and the validation of the method can be summarized as follows:

- The method's main application is the detection of main meter-submeter relations between electricity meters. It is not suitable for the detection of main meter-submeter relations between heat meters.
- As the method focuses on detecting similar changes in load profiles, the prediction performance heavily depends on the situation in which the method is applied. If there are not enough changes in the load profiles, the method is bound to fail.
- As the method is built to infer main meter-submeter relationships solely from monitoring data without additional knowledge about the meter hierarchy, it considers all possible meter combinations. For the case of n energy meters, there are $n \cdot (n-1)$ possible combinations, and features are calculated for all of them. Therefore, the calculation of features gets relatively compute-intensive in the case of large numbers of energy meters.

- Random forests for multiclass classification are the method's main component. They classify to four clusters: "obvious relations", "rare partial supply", "inconspicuous relations", and "no relations". Cluster "no relations" encompasses all meter combinations where there is no real main meter-submeter relationship. As this is the case for most combinations, this cluster is by far the largest cluster. The other three clusters encompass the cases where there are real main meter-submeter relationships. Cluster "rare partial supply" includes entries where there is such a relationship, but only during short periods of time, e.g., a secondary power supply that is not always active. The method failed to correctly classify entries to this cluster. Respectively, the cluster could have been left out and its entries assigned to the other clusters. The other two clusters, "obvious relations" and "inconspicuous relations", contain all other cases where there is a real main meter-submeter relationship. Predictions to the cluster "obvious relations" can be considered more trustworthy than those to the cluster "inconspicuous relations", as there are generally fewer misclassifications to cluster "obvious relations".

- The method was developed for two different cases:

    1. Time series data from both the counter values as well as the instantaneous power values of energy meters are available.

    2. Only time series data of the counter values of the energy meters are available.

- It was found that the method's prediction performance is better in the first case – the one where more data is available.

- Monitoring data in five different time resolutions were investigated: 5 min, 10 min, 15 min, 30 min, and 60 min. Generally, it can be stated that the higher the time resolution, the better the prediction performance.

- Monitoring data from time slots with six different time slot lengths were investigated: 1 year, 6 months, 3 months, 5 weeks, 2 weeks, and 1 week. The larger the time slot length, the better the prediction performance.

- Given an electricity meter hierarchy and monitoring data similar to the one of the case study, it can be expected that approximately 40% of the hierarchy can be inferred by the method while there are almost zero real misclassifications. The prerequisite for this is that both time series data from the counter values and the instantaneous power values are available, the time resolution is 5 min, and the time slot length is at least 2 weeks. In the cases of 10 min and 15 min time resolution, the rate drops to approximately 34% and 26%. With the same prerequisites fulfilled, the method detected only a few parts of the case study's heat meter hierarchy – in most cases, only one main meter-submeter relationship out of 33.

- The method showed a bias towards the case study's monitoring data. This bias became obvious while analyzing the impact of different time resolutions for both validation data sources. Specifically, the results achieved with the validation data from other time slots of the case study's monitoring data indicated that different time resolutions have almost no impact on the prediction performance. The results achieved with the synthetically generated validation data indicated that different time resolutions do significantly influence the prediction performance. As the validation results mainly differ in having no classifications to the cluster "inconspicuous relations" in the case of the synthetically generated data, it was concluded that the bias is primarily present in this cluster. For both validation data sources, practically all predictions to the cluster "obvious relations" were correct, and there were almost no misclassifications to this cluster.
- Considering the bias towards monitoring data from the case study, the impact of different time resolutions on the prediction performance must be evaluated by analyzing only the results achieved with the synthetically generated validation data. These results indicate that the method should only be used for time resolutions equal to or higher than 15 min – specifically, 5 min, 10 min, and 15 min. For the cases of 30 min and 60 min time resolution, less than 10% of the electricity meter hierarchy could be correctly identified, which is deemed too low to be useful.
- The final output of the development of the method constituted a guide on how to develop classifiers for monitoring data with 5 min, 10 min, and 15 min time resolution out of the case study's monitoring data from a specific 5-week time slot. As these classifiers are developed for two cases, (i) the case where both the counter values as well as the instantaneous power values of energy meters are available, and (ii) the case where only the meter's counter values are available, there are six final best classifiers.

Compared to the prediction performance of other, already existing, similar methods [26–30], the prediction performance of the method developed in this thesis is low. The other methods can detect between 75% and 100% of the electricity meter hierarchy, while the developed method achieves approximately 40% in the case of monitoring data with a 5 min time resolution. As the other methods were developed and evaluated under different circumstances, the lower prediction performance of the new method does not void its usefulness. As explained in section 2.5, none of the existing methods works with energy measurement data instead of voltage measurement data *and* allows for unmetered consumers. That this new method can do so is its unique feature. This unique feature enables the method to be the key element for a process to accelerate the validation of energy monitoring systems. It can be the basic building block for a set of methods that aids in the tasks of generating an overview of the real energy meter hierarchy and identifying errors.

7 Outlook

In this work it was shown that the method, which uses random forest classifiers to infer main meter-submeter relationships out of energy monitoring data, basically works. The prediction performance that can be expected from it has been evaluated for several different settings and cases. Besides using the method, as intended, as the key element for the envisioned process of accelerating the validation of energy monitoring systems, there are several further development steps that can be taken. Most of them aim to improve the method or extend its application field.

As a first step, a more detailed analysis of the bias towards the monitoring data from the case study should be conducted. As a bias is generally something that might cause unexpected behavior, it should be avoided. It could lead to the situation that the method produces misclassifications because some calculated features coincidentally match the features of some cases of the training data.

Another improvement step could be the revision of the decision process for the best classifier candidates. Especially the classification evaluation metric "custom misclassification rate" (*CMCR*) used in this process should be reworked. As this metric led to several cases where it was undefined due to a division by zero, the resulting heat maps have some blind spots. During the model selection, these blind spots appeared to be reasonable as they indicated models that did not fully replicate the desired multiclass classification, but during the validation, it became apparent that this does not necessarily mean that these models produce worse prediction results than other models. Maybe models hidden in these blind spots would have shown a better prediction performance.

The method proved unsuitable for the classification of heat meters. One possible explanation is the thermal inertia and losses of the piping systems, which lead to the situation that a consumption change measured by a submeter will not be immediately registered by the main meter or that it is registered with a different magnitude. Dynamic time warping [43] might be an approach to tackle this issue.

One general issue is that the computation time of the method increases exponentially with the number of energy meters. To reduce the computation time, an algorithm to filter out meter combinations that are very unlikely to have a main meter-submeter relationship could be implemented. E.g., the general power levels of the meters could be used to decide whether the meter combination should be filtered out or not.

Even though the suggested further development steps can be considered optional, they would increase the method's usefulness. Especially the reduction of the computation time would be beneficial when the method is used, as intended, as the key element for the envisioned process of accelerating the validation of energy monitoring systems. It would allow an automatic search for wrongfully parameterized current transformer values in electricity meters, which was one of the main errors encountered in the Plus-Energy Office High-Rise Building’s energy monitoring system.

8 Literature

[1] Eurostat, Energy balance sheets 2017 data. 2019 Edition, Publications Office of the European Union, 2019. https://doi.org/10.2785/10223.

[2] European Parliament, Directive 2018/844/EU, 2018. http://data.europa.eu/eli/dir/2018/844/oj.

[3] A. Esser, A. Dunne, T. Meeusen, S. Quaschning, W. Denis, Comprehensive study of building energy renovation activities and the uptake of nearly zero-energy buildings in the EU. Final report, European Commission, 2019. https://doi.org/10.2833/14675.

[4] O. Hernandez, Elithis Tower in Dijon, France, REHVA Eur. HVAC J. 48 (2011) 53–57. http://www.rehva.eu/fileadmin/hvac-dictio/03-2011/Elithis_Tower_in_Dijon__France.pdf.

[5] A. David, M. Leeb, T. Bednar, Comparison of the planned and the real energy consumption of the world's first (Plus-)Plus-Energy Office High-Rise Building, in: Energy Procedia, 2017. https://doi.org/10.1016/j.egypro.2017.09.726.

[6] Y. Geng, W. Ji, Z. Wang, B. Lin, Y. Zhu, Energy & Buildings A review of operating performance in green buildings : Energy use , indoor environmental quality and occupant satisfaction, Energy Build. 183 (2019) 500–514. https://doi.org/10.1016/j.enbuild.2018.11.017.

[7] S. Cozza, J. Chambers, A. Brambilla, M.K. Patel, In search of optimal consumption: A review of causes and solutions to the Energy Performance Gap in residential buildings, Energy Build. 249 (2021) 111253. https://doi.org/10.1016/j.enbuild.2021.111253.

[8] Z. (John) Zhai, A. Salazar, Assessing the implications of submetering with energy analytics to building energy savings, Energy Built Environ. (2020). https://doi.org/10.1016/j.enbenv.2019.08.002.

[9] A. David, T. Bednar, Transition towards an energy-efficient scientific Office Setup in Austria's largest (Plus-)Plus-Energy Office Building - An Analysis, in: E3S Web Conf., 2020. https://doi.org/10.1051/e3sconf/202017222008.

[10] E. Mills, P. Mathew, Monitoring-Based Commissioning: Benchmarking Analysis of 24 UC/CSU/IOU Projects, 2009. https://doi.org/10.1080/01998595.2014.10844605.

[11] O. Guerra-santin, C. Aidan, In-use monitoring of buildings : An overview of data collection methods, Energy Build. 93 (2015) 189–207. https://doi.org/10.1016/j.enbuild.2015.02.042.

[12] U. Habib, Fault Detection and Diagnosis in Building Energy Systems, TU Wien, 2016.

[13] W. Kim, S. Katipamula, A review of fault detection and diagnostics methods for building systems, Sci. Technol. Built Environ. (2018). https://doi.org/10.1080/23744731.2017.1318008.

[14] Y. Zhao, T. Li, X. Zhang, C. Zhang, Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future, Renew. Sustain. Energy Rev. (2019). https://doi.org/10.1016/j.rser.2019.04.021.

[15] S.C. Yip, K.S. Wong, W.P. Hew, M.T. Gan, R.C.W. Phan, S.W. Tan, Detection of energy theft and defective smart meters in smart grids using linear regression, Int. J. Electr. Power Energy Syst. 91 (2017) 230–240. https://doi.org/10.1016/j.ijepes.2017.04.005.

[16] S.C. Yip, W.N. Tan, C.K. Tan, M.T. Gan, K.S. Wong, An anomaly detection framework for identifying energy theft and defective meters in smart grids, Int. J. Electr. Power Energy Syst. (2018). https://doi.org/10.1016/j.ijepes.2018.03.025.

[17] M. Liu, D. Liu, G. Sun, Y. Zhao, D. Wang, F. Liu, X. Fang, Q. He, D. Xu, Detection of Malfunctioning Smart Electricity Meter, CoRR. (2019).

[18] J. Stevens, P. Smith, Guidance for Electricity and Gas Meter Installation Consumer Facing Issues, (2015).

[19] ADDC Guidelines for Energy Meter Installation, (2018).

[20] P. Concannon, An Investigation into Heat Meter Measurement Errors Final Report, 2014.

[21] European Union, Directive 2009/72/EC, Off. J. Eur. Union. L211 (2009) L 211/55-L 211/93.

[22] Bericht zur Einführung von intelligenten Messgeräten in Österreich 2021, 2021. https://www.e-control.at/documents/1785851/1811582/Monitoringbericht_Smart+Meter_2021.pdf.

[23] S.S.S.R. Depuru, L. Wang, V. Devabhaktuni, N. Gudi, Smart meters for power grid - Challenges, issues, advantages and status, 2011 IEEE/PES Power Syst. Conf. Expo. PSCE 2011. (2011) 1–7. https://doi.org/10.1109/PSCE.2011.5772451.

[24] U. Greveler, B. Justus, D. Loehr, Multimedia content identification through smart meter power usage profiles, Comput. Priv. Data Prot. (2012).

[25] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, R. Steinmetz, On the accuracy of appliance identification based on distributed load metering data, 2012 Sustain. Internet ICT Sustain. Sustain. 2012. (2012).

[26] M. Zhang, W. Luan, S. Guo, P. Wang, Topology identification method of distribution network based on smart meter measurements, China Int. Conf. Electr. Distrib. CICED. (2018) 372–376. https://doi.org/10.1109/CICED.2018.8592228.

[27] L. Chao, Z. Lei, L. Yuhang, Topology Checking Method for Low Voltage Distribution Network Based on Fuzzy C-Means Clustering Algorithm, Proc. 2020 IEEE Int. Conf. Artif. Intell. Comput. Appl. ICAICA 2020. (2020) 1077–1080. https://doi.org/10.1109/ICAICA50127.2020.9182696.

[28] C. Qian, W. Meiyan, G. Ding, H. Fei-Hu, S. Sheila Jepchirchir, Topology Identification Method for Primary Distribution Network with Limited Smart Meter Data, Proc. - 2021 6th Asia Conf. Power Electr. Eng. ACPEE 2021. (2021) 1611–1616. https://doi.org/10.1109/ACPEE51499.2021.9436892.

[29] S. Liu, D. Zhou, K. Li, S. Hou, Topology Identification Method of Low-Voltage Distribution Network Based on Regression Analysis of Voltage Characteristic Parameters, Proc. - 2021 6th Int. Conf. Smart Grid Electr. Autom. ICSGEA 2021. (2021) 75–79. https://doi.org/10.1109/ICSGEA53208.2021.00023.

[30] S.J. Pappu, N. Bhatt, R. Pasumarthy, A. Rajeswaran, Identifying Topology of Low Voltage Distribution Networks Based on Smart Meter Data, IEEE Trans. Smart Grid. 9 (2018) 5113–5122. https://doi.org/10.1109/TSG.2017.2680542.

[31] T.O. Ayodele, Types of Machine Learning Algorithms, New Adv. Mach. Learn. 3 (2010) 19–48. https://doi.org/10.5772/9385.

[32] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning. Data Mining, Inference, and Prediction., Second Edi, Springer, 2017.

[33] K.P. Murphy, Machine Learning: A Probabilistic Perspective, in: The MIT Press, 2012.

[34] G.A. Susto, A. Cenedese, M. Terzi, Time-Series Classification Methods: Review and Applications to Power Systems Data, Elsevier Inc., 2018. https://doi.org/10.1016/B978-0-12-811968-6.00009-7.

[35] L. Breiman, Random Forests, Mach. Learn. 45 (2001) 5–32.

[36] A. David, T. Bednar, H. Schöberl, Comparison of the designed and the measured performance of TU Wien ’ s refurbished high-rise, in: 21. Int. Passiv. 2017, n.d.: pp. 463–468.

[37] H. Schöberl, R. Hofer, M. Leeb, T. Bednar, G. Kratochwil, Schriftenreihe 47/2014: Österreichs größtes Plus- Energie-Bürogebäude am Standort Getreidemarkt der TU Wien, Bundesministerium für Verkehr, Innovation und Technologie, Vienna, 2014.

[38] Open Power System Data, Data Package Household Data, Version 2020-04-15. (2020). https://data.open-power-system-data.org/household_data/2020-04-15. (Primary data from various sources, for a complete list see URL).

[39] W. McKinney, Data Structures for Statistical Computing in Python, in: Proc. 9th Python Sci. Conf., 2010. https://doi.org/10.25080/majora-92bf1922-00a.

[40] M. Hossin, M.N. Sulaiman, A Review on Evaluation Metrics for Data Classification Evaluations, Int. J. Data Min. Knowl. Manag. Process. 5 (2015) 01–11. https://doi.org/10.5121/ijdkp.2015.5201.

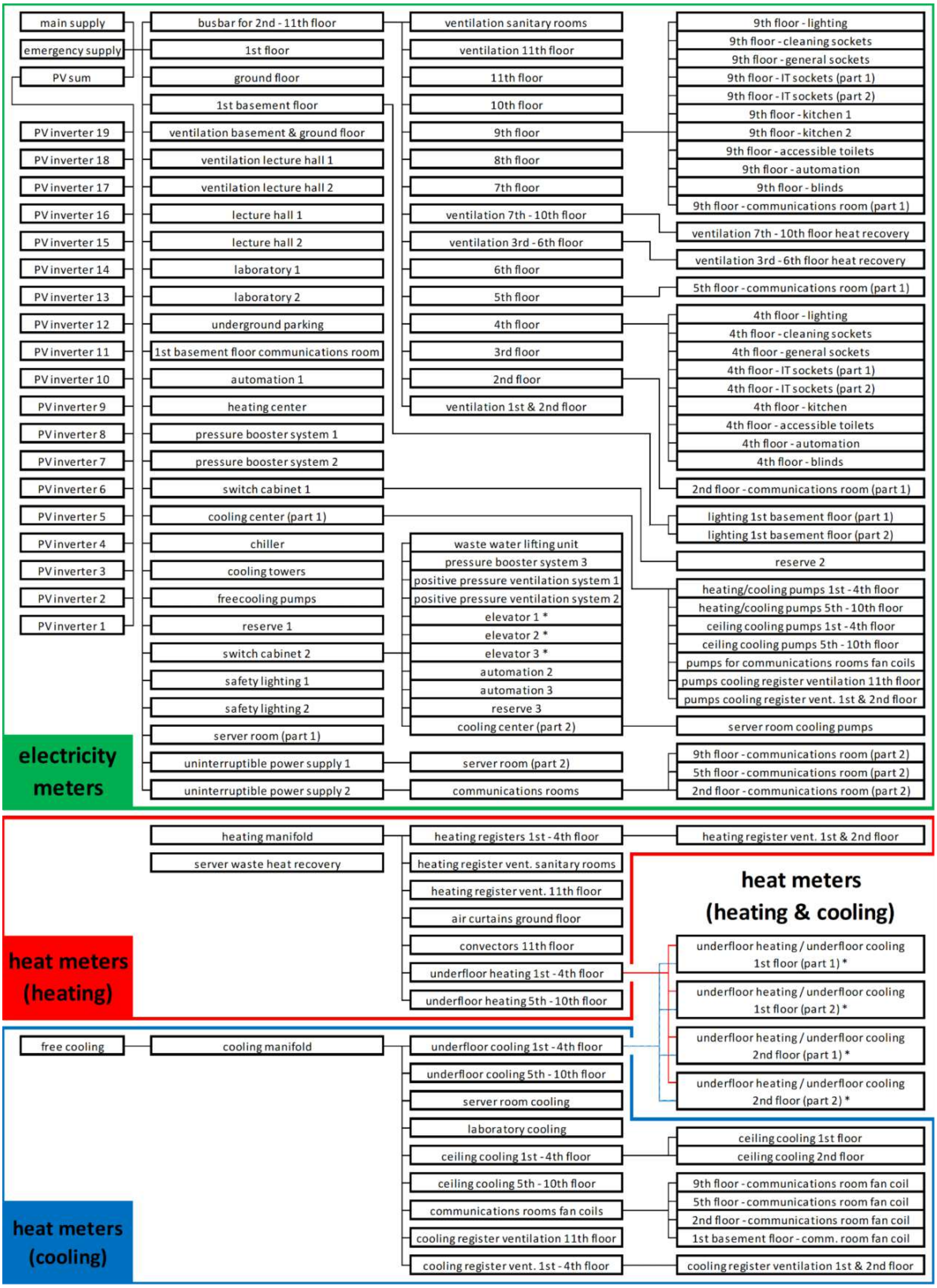
[41] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. (2011) 2825–2830.

[42] C. Strobl, A.L. Boulesteix, A. Zeileis, T. Hothorn, Bias in random forest variable importance measures: Illustrations, sources and a solution, BMC Bioinformatics. 8 (2007). https://doi.org/10.1186/1471-2105-8-25.

[43] E. Keogh, C.A. Ratanamahatana, Exact indexing of dynamic time warping, Knowl. Inf. Syst. 7 (2005) 358–386. https://doi.org/10.1007/s10115-004-0154-9.

9 Appendix A: Structure and topology of the case study's energy meters

Overview of the structure and topology of the (Plus-)Plus-Energy Office High-Rise Building's energy meters. The connecting lines indicate the main meter-submeter relationships. Except for the meters of the PV inverters, the meters attached to the right side of a connecting line can be considered as submeters of the meter attached to the left side of the connecting line. Meters with two counters are marked by *.



10 Appendix B: Order of features as determined during backward selection

order	data resolution: **5 min** classifier type: **multiclass** used data sources: **CP**	data resolution: **5 min** classifier type: **multiclass** used data sources: **C**	data resolution: **5 min** classifier type: **binary** used data sources: **CP**	data resolution: **5 min** classifier type: **binary** used data sources: **C**
1	c_d2_slope	c_d2_slope	c_d2_slope	c_d2_spearman
2	p_d1_kendall	c_d2_spearman	p_d1t_kendall	c_d2_slope
3	c_d2_spearman	c_d1_profile_relation_neg	c_d2_spearman	c_d2t_spearman
4	c_d1_profile_relation_neg	c_d2t_spearman	c_d1_profile_relation_neg	c_d1_profile_relation_neg
5	c_d2t_spearman	c_d1_profile_relation	c_d2t_spearman	c_d2t_kendall
6	p_d1t_kendall	c_d2t_kendall	p_m_profile_relation	c_d1_profile_relation
7	c_d1_profile_relation	c_d2t_pearson	p_d1_kendall	c_d2t_pearson
8	c_d2t_intercept_to_std_ln	c_d2t_intercept_to_std_ln	c_d2t_intercept_to_std_ln	c_d2t_intercept_to_std_ln
9	c_d2t_kendall	c_d2t_deviation_slope_1_ln	c_d2t_pearson	c_d2_kendall
10	p_m_profile_relation_neg	c_d2_kendall	c_d2t_kendall	c_d2_pearson
11	c_d2_pearson	c_d2_pearson	c_d1_profile_relation	c_d2t_deviation_slope_1_ln
12	p_d1_r_squared_sub	c_d2_deviation_slope_1	p_d1t_spearman	c_d2_deviation_slope_1
13	p_m_profile_relation	c_d2t_intercept_to_std	c_d2_kendall	c_d2_r_squared
14	c_d2_kendall	c_d2_r_squared	p_d1_r_squared	c_d2t_intercept_to_std
15	c_d2t_deviation_slope_1_ln	c_d2t_deviation_slope_1	p_m_profile_relation_neg	c_d2t_r_squared_sub
16	p_d1t_spearman	c_d2t_r_squared_sub_ln	c_d2_pearson	c_d2t_r_squared_sub_ln
17	c_d2_deviation_slope_1	c_d2_r_squared_sub	p_d1_deviation_slope_1	c_d2_r_squared_sub
18	p_d1_spearman	c_d2t_slope	p_d1_spearman	c_d2t_slope
19	p_d1_deviation_slope_1	c_d2t_r_squared	c_d2t_diff_sm_mm	c_d2t_diff_sm_mm
20	c_d2t_pearson	c_d2_intercept_to_std	p_d1t_diff_sm_mm	c_d2_intercept_to_std
21	c_d2t_intercept_to_std	c_d2t_r_squared_sub	c_d2t_intercept_to_std	c_d2t_r_squared
22	c_d2t_deviation_slope_1	c_d2t_diff_sm_mm	c_d2_r_squared_sub	c_d2t_deviation_slope_1
23	p_d1_r_squared	c_d2t_pearson_sub_ln	p_d1_r_squared_sub	c_d2t_pearson_sub_ln
24	c_d2_r_squared	c_d2t_kendall_sub_ln	c_d2_deviation_slope_1	c_d2t_kendall_sub_ln
25	p_d1t_diff_sm_mm	c_d2t_spearman_sub_ln	p_d1t_deviation_slope_1_ln	c_d2t_diff_sm_mm_ln
26	p_d1t_deviation_slope_1_ln	c_d2t_diff_sm_mm_ln	c_d2_r_squared	c_d2t_spearman_sub_ln
27	c_d2_r_squared_sub	c_d2t_threshold_to_std	c_d2t_r_squared	c_d2t_threshold_to_std
28	c_d2t_slope	smoothed	c_d2t_slope	smoothed
29	c_d2t_diff_sm_mm		c_d2t_deviation_slope_1_ln	
30	p_d1t_pearson		p_d1_pearson	
31	c_d2t_r_squared		c_d2_intercept_to_std	
32	p_d1t_deviation_slope_1		p_d1t_deviation_slope_1	
33	c_d2_intercept_to_std		c_d2t_diff_sm_mm_ln	
34	p_d1_pearson		p_d1t_diff_sm_mm_ln	
35	p_d1t_diff_sm_mm_ln		c_d2t_deviation_slope_1	
36	c_d2t_r_squared_sub_ln		p_d1t_pearson	
37	p_d1_slope		c_d2t_r_squared_sub_ln	
38	p_d1t_pearson_sub_ln		p_d1_slope	
39	p_d1t_intercept_to_std_ln		c_d2t_r_squared_sub	
40	c_d2t_r_squared_sub		p_d1t_pearson_sub_ln	
41	p_d1_intercept_to_std		p_d1t_threshold_to_std	
42	c_d2t_diff_sm_mm_ln		c_d2t_pearson_sub_ln	
43	p_d1t_threshold_to_std		p_d1_intercept_to_std	
44	p_d1t_intercept_to_std		p_d1t_intercept_to_std_ln	
45	p_d1t_kendall_sub_ln		p_d1t_spearman_sub_ln	
46	c_d2t_pearson_sub_ln		c_d2t_kendall_sub_ln	
47	p_d1t_r_squared_sub_ln		p_d1t_intercept_to_std	
48	c_d2t_kendall_sub_ln		p_d1t_r_squared_sub_ln	
49	p_d1t_spearman_sub_ln		p_d1t_kendall_sub_ln	
50	p_d1t_slope		c_d2t_spearman_sub_ln	
51	c_d2t_spearman_sub_ln		p_d1t_slope	
52	c_d2t_threshold_to_std		c_d2t_threshold_to_std	
53	p_d1t_r_squared		p_d1t_r_squared	
54	p_d1t_r_squared_sub		p_d1t_r_squared_sub	
55	smoothed		smoothed	

order	data resolution: **10 min** classifier type: **multiclass** used data sources: **CP**	data resolution: **10 min** classifier type: **multiclass** used data sources: **C**	data resolution: **10 min** classifier type: **binary** used data sources: **CP**	data resolution: **10 min** classifier type: **binary** used data sources: **C**
1	c_d2_slope	c_d2_slope	c_d2_slope	c_d2_slope
2	p_d1_spearman	c_d2_pearson	p_d1_spearman	c_d2t_pearson
3	p_d1t_spearman	c_d2t_slope	p_d1t_spearman	c_d2t_slope
4	c_d2t_slope	c_d2_spearman	c_d2t_slope	c_d2_pearson
5	c_d1_profile_relation_neg	c_d1_profile_relation_neg	p_d1t_kendall	c_d2_spearman
6	p_d1t_kendall	c_d2t_spearman	c_d2_pearson	c_d2t_spearman
7	p_d1_kendall	c_d1_profile_relation	p_m_profile_relation_neg	c_d1_profile_relation_neg
8	p_m_profile_relation_neg	c_d2t_intercept_to_std_ln	p_d1_kendall	c_d1_profile_relation
9	c_d2t_intercept_to_std_ln	c_d2t_pearson	c_d1_profile_relation_neg	c_d2t_intercept_to_std_ln
10	c_d2t_pearson	c_d2_kendall	c_d2t_intercept_to_std_ln	c_d2_kendall
11	p_m_profile_relation	c_d2t_kendall	c_d2t_pearson	c_d2t_kendall
12	c_d2t_spearman	c_d2t_intercept_to_std	p_m_profile_relation	c_d2_r_squared
13	p_d1_r_squared_sub	c_d2_r_squared	c_d2_spearman	c_d2t_intercept_to_std
14	c_d1_profile_relation	c_d2t_deviation_slope_1_ln	p_d1_r_squared	c_d2t_deviation_slope_1_ln
15	p_d1_slope	c_d2_intercept_to_std	c_d1_profile_relation	c_d2_intercept_to_std
16	c_d2_pearson	c_d2t_r_squared	p_d1_slope	c_d2_r_squared_sub
17	c_d2_spearman	c_d2_r_squared_sub	c_d2t_spearman	c_d2t_r_squared
18	p_d1_r_squared	c_d2_deviation_slope_1	p_d1_r_squared_sub	c_d2t_pearson_sub_ln
19	p_d1t_kendall_sub_ln	c_d2t_r_squared_sub	p_d1t_kendall_sub_ln	c_d2_deviation_slope_1
20	c_d2t_intercept_to_std	c_d2t_deviation_slope_1	c_d2_r_squared	c_d2t_r_squared_sub
21	c_d2t_kendall	c_d2t_pearson_sub_ln	c_d2t_intercept_to_std	c_d2t_spearman_sub_ln
22	p_d1_deviation_slope_1	c_d2t_spearman_sub_ln	c_d2_kendall	c_d2t_diff_sm_mm
23	c_d2_r_squared	c_d2t_kendall_sub_ln	p_d1_deviation_slope_1	c_d2t_deviation_slope_1
24	p_d1t_r_squared_sub_ln	c_d2t_r_squared_sub_ln	p_d1t_r_squared_sub_ln	c_d2t_kendall_sub_ln
25	c_d2_kendall	c_d2t_diff_sm_mm	c_d2_r_squared_sub	c_d2t_r_squared_sub_ln
26	p_d1t_pearson	c_d2t_threshold_to_std	c_d2t_kendall	c_d2t_diff_sm_mm_ln
27	c_d2_r_squared_sub	c_d2t_diff_sm_mm_ln	p_d1_pearson	c_d2t_threshold_to_std
28	c_d2t_deviation_slope_1_ln	smoothed	p_d1t_spearman_sub_ln	smoothed
29	c_d2_intercept_to_std		c_d2_intercept_to_std	
30	p_d1_pearson		c_d2t_diff_sm_mm	
31	p_d1t_spearman_sub_ln		p_d1t_pearson	
32	p_d1t_slope		p_d1t_slope	
33	p_d1_intercept_to_std		p_d1t_deviation_slope_1	
34	p_d1t_deviation_slope_1		c_d2t_deviation_slope_1_ln	
35	p_d1t_threshold_to_std		p_d1t_threshold_to_std	
36	c_d2t_diff_sm_mm		c_d2t_diff_sm_mm_ln	
37	c_d2_deviation_slope_1		c_d2t_r_squared	
38	c_d2t_r_squared		p_d1_intercept_to_std	
39	p_d1t_intercept_to_std_ln		c_d2_deviation_slope_1	
40	p_d1t_pearson_sub_ln		p_d1t_pearson_sub_ln	
41	c_d2t_r_squared_sub		c_d2t_r_squared_sub	
42	c_d2t_deviation_slope_1		p_d1t_intercept_to_std	
43	p_d1t_intercept_to_std		c_d2t_pearson_sub_ln	
44	c_d2t_pearson_sub_ln		p_d1t_intercept_to_std_ln	
45	c_d2t_diff_sm_mm_ln		p_d1t_r_squared_sub	
46	p_d1t_r_squared_sub		p_d1t_deviation_slope_1_ln	
47	p_d1t_deviation_slope_1_ln		c_d2t_r_squared_sub_ln	
48	c_d2t_spearman_sub_ln		p_d1t_r_squared	
49	p_d1t_r_squared		c_d2t_deviation_slope_1	
50	c_d2t_r_squared_sub_ln		c_d2t_spearman_sub_ln	
51	c_d2t_kendall_sub_ln		c_d2t_kendall_sub_ln	
52	p_d1t_diff_sm_mm		p_d1t_diff_sm_mm	
53	c_d2t_threshold_to_std		c_d2t_threshold_to_std	
54	p_d1t_diff_sm_mm_ln		p_d1t_diff_sm_mm_ln	
55	smoothed		smoothed	

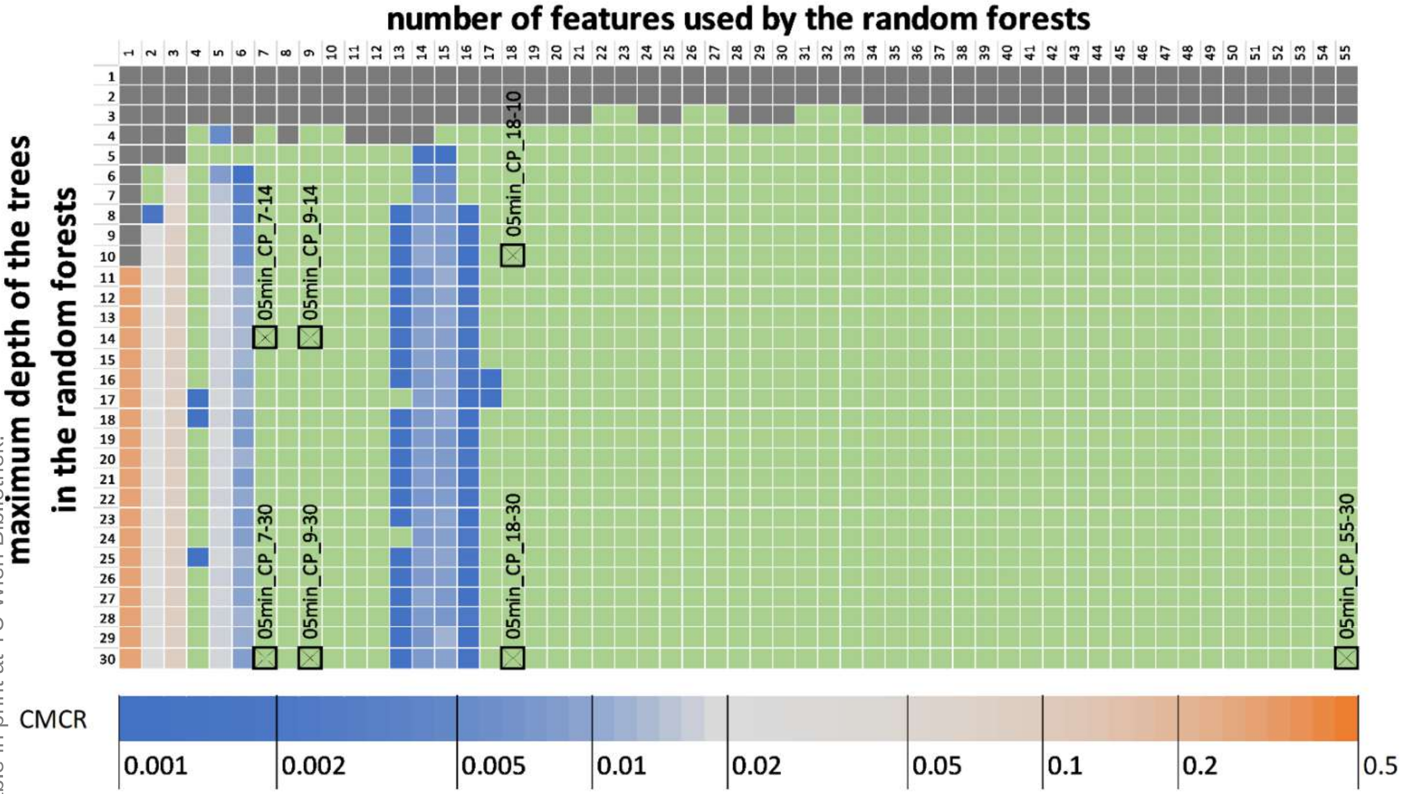
order	data resolution: **15 min** classifier type: **multiclass** used data sources: **CP**	data resolution: **15 min** classifier type: **multiclass** used data sources: **C**	data resolution: **15 min** classifier type: **binary** used data sources: **CP**	data resolution: **15 min** classifier type: **binary** used data sources: **C**
1	p_d1_spearman	c_d2_slope	p_d1_spearman	c_d2_slope
2	c_d2_slope	c_d2t_kendall	c_d2_slope	c_d2t_kendall
3	p_d1t_spearman	c_d1_profile_relation	p_d1t_spearman	c_d1_profile_relation
4	p_m_profile_relation	c_d2t_slope	p_m_profile_relation	c_d2_kendall
5	p_d1t_kendall	c_d2_kendall	p_d1t_kendall	c_d2t_slope
6	c_d2t_slope	c_d1_profile_relation_neg	p_d1_kendall	c_d2_spearman
7	p_d1_kendall	c_d2_spearman	c_d2t_slope	c_d1_profile_relation_neg
8	c_d1_profile_relation_neg	c_d2t_intercept_to_std_ln	c_d1_profile_relation_neg	c_d2t_intercept_to_std_ln
9	p_m_profile_relation_neg	c_d2t_spearman	c_d2t_kendall	c_d2t_spearman
10	c_d2_kendall	c_d2_intercept_to_std	c_d2t_intercept_to_std_ln	c_d2t_r_squared
11	c_d1_profile_relation	c_d2_pearson	c_d1_profile_relation	c_d2_pearson
12	c_d2t_intercept_to_std_ln	c_d2t_deviation_slope_1_ln	c_d2_kendall	c_d2t_intercept_to_std
13	c_d2t_kendall	c_d2t_pearson	p_d1_deviation_slope_1	c_d2t_pearson
14	p_d1_deviation_slope_1	c_d2t_intercept_to_std	p_d1_r_squared	c_d2t_pearson_sub_ln
15	p_d1_r_squared	c_d2t_r_squared	p_m_profile_relation_neg	c_d2_intercept_to_std
16	c_d2_spearman	c_d2_r_squared	c_d2_spearman	c_d2t_r_squared_sub
17	c_d2t_intercept_to_std	c_d2t_r_squared_sub	c_d2t_intercept_to_std	c_d2_r_squared
18	p_d1_r_squared_sub	c_d2t_deviation_slope_1	p_d1_slope	c_d2t_deviation_slope_1_ln
19	p_d1_slope	c_d2_r_squared_sub	c_d2t_spearman	c_d2_r_squared_sub
20	c_d2t_spearman	c_d2t_pearson_sub_ln	p_d1_r_squared_sub	c_d2t_deviation_slope_1
21	c_d2t_r_squared_sub	c_d2_deviation_slope_1	p_d1t_kendall_sub_ln	c_d2t_r_squared_sub_ln
22	c_d2_intercept_to_std	c_d2t_r_squared_sub_ln	c_d2t_r_squared	c_d2_deviation_slope_1
23	c_d2_deviation_slope_1	c_d2t_kendall_sub_ln	c_d2_intercept_to_std	c_d2t_kendall_sub_ln
24	p_d1t_kendall_sub_ln	c_d2t_spearman_sub_ln	c_d2t_r_squared_sub_ln	c_d2t_spearman_sub_ln
25	p_d1t_deviation_slope_1	c_d2t_diff_sm_mm	p_d1t_pearson	c_d2t_threshold_to_std
26	p_d1t_pearson	c_d2t_threshold_to_std	p_d1t_spearman_sub_ln	c_d2t_diff_sm_mm
27	c_d2_pearson	c_d2t_diff_sm_mm_ln	c_d2_pearson	c_d2t_diff_sm_mm_ln
28	c_d2t_deviation_slope_1_ln	smoothed	p_d1_pearson	smoothed
29	c_d2t_pearson		c_d2t_r_squared_sub	
30	p_d1_pearson		p_d1t_intercept_to_std	
31	p_d1_intercept_to_std		c_d2t_pearson	
32	p_d1t_spearman_sub_ln		p_d1t_deviation_slope_1	
33	c_d2t_deviation_slope_1		c_d2_deviation_slope_1	
34	c_d2_r_squared		c_d2_r_squared	
35	c_d2t_r_squared		p_d1t_r_squared	
36	p_d1t_slope		p_d1t_slope	
37	p_d1t_intercept_to_std		p_d1_intercept_to_std	
38	c_d2_r_squared_sub		c_d2_r_squared_sub	
39	p_d1t_r_squared_sub		c_d2t_pearson_sub_ln	
40	c_d2t_r_squared_sub_ln		c_d2t_deviation_slope_1	
41	p_d1t_intercept_to_std_ln		p_d1t_intercept_to_std_ln	
42	p_d1t_deviation_slope_1_ln		p_d1t_r_squared_sub	
43	p_d1t_pearson_sub_ln		c_d2t_deviation_slope_1_ln	
44	c_d2t_pearson_sub_ln		p_d1t_pearson_sub_ln	
45	p_d1t_r_squared		c_d2t_spearman_sub_ln	
46	p_d1t_threshold_to_std		p_d1t_deviation_slope_1_ln	
47	c_d2t_kendall_sub_ln		c_d2t_kendall_sub_ln	
48	p_d1t_r_squared_sub_ln		p_d1t_threshold_to_std	
49	c_d2t_spearman_sub_ln		p_d1t_r_squared_sub_ln	
50	c_d2t_diff_sm_mm		c_d2t_threshold_to_std	
51	c_d2t_threshold_to_std		p_d1t_diff_sm_mm	
52	p_d1t_diff_sm_mm		c_d2t_diff_sm_mm	
53	smoothed		smoothed	
54	c_d2t_diff_sm_mm_ln		p_d1t_diff_sm_mm_ln	
55	p_d1t_diff_sm_mm_ln		c_d2t_diff_sm_mm_ln	

order	data resolution: **30 min** classifier type: **multiclass** used data sources: **CP**	data resolution: **30 min** classifier type: **multiclass** used data sources: **C**	data resolution: **30 min** classifier type: **binary** used data sources: **CP**	data resolution: **30 min** classifier type: **binary** used data sources: **C**
1	c_d2_slope	c_d2_slope	p_d1_kendall	c_d2_spearman
2	p_d1_kendall	c_d2_spearman	c_d2_slope	c_d2_slope
3	p_m_profile_relation	c_d2t_spearman	p_m_profile_relation	c_d1_profile_relation
4	p_d1t_kendall	c_d1_profile_relation_neg	p_d1t_kendall	c_d2t_spearman
5	c_d2t_kendall	c_d2t_intercept_to_std_ln	c_d2t_kendall	c_d2_pearson
6	c_d1_profile_relation_neg	c_d2t_slope	p_d1t_spearman	c_d2t_intercept_to_std_ln
7	p_d1t_spearman	c_d2t_kendall	c_d1_profile_relation_neg	c_d2t_kendall
8	c_d2t_slope	c_d1_profile_relation	c_d2t_intercept_to_std_ln	c_d1_profile_relation_neg
9	c_d2t_intercept_to_std_ln	c_d2_pearson	c_d2t_slope	c_d2t_slope
10	p_d1_spearman	c_d2_kendall	p_d1_slope	c_d2_kendall
11	c_d1_profile_relation	c_d2t_intercept_to_std	c_d2_kendall	c_d2t_pearson
12	c_d2_kendall	c_d2t_pearson	c_d1_profile_relation	c_d2t_intercept_to_std
13	p_d1_slope	c_d2_intercept_to_std	p_d1_spearman	c_d2t_r_squared_sub_ln
14	p_m_profile_relation_neg	c_d2t_spearman_sub_ln	c_d2t_pearson	c_d2_intercept_to_std
15	c_d2_pearson	c_d2t_deviation_slope_1	p_m_profile_relation_neg	c_d2t_r_squared_sub
16	c_d2t_intercept_to_std	c_d2_r_squared	c_d2t_intercept_to_std	c_d2_r_squared
17	c_d2t_spearman	c_d2t_r_squared_sub	c_d2_spearman	c_d2t_spearman_sub_ln
18	p_d1t_kendall_sub_ln	c_d2t_deviation_slope_1_ln	p_d1t_kendall_sub_ln	c_d2_r_squared_sub
19	p_d1t_slope	c_d2_r_squared_sub	p_d1t_slope	c_d2_deviation_slope_1
20	c_d2t_pearson	c_d2_deviation_slope_1	c_d2_pearson	c_d2t_r_squared
21	p_d1t_intercept_to_std	c_d2t_kendall_sub_ln	c_d2t_spearman	c_d2t_kendall_sub_ln
22	c_d2_spearman	c_d2t_r_squared	p_d1t_intercept_to_std	c_d2t_deviation_slope_1
23	c_d2_intercept_to_std	c_d2t_r_squared_sub_ln	c_d2t_r_squared_sub_ln	c_d2t_pearson_sub_ln
24	p_d1t_spearman_sub_ln	c_d2t_pearson_sub_ln	c_d2_intercept_to_std	c_d2t_deviation_slope_1_ln
25	p_d1t_intercept_to_std_ln	c_d2t_threshold_to_std	p_d1t_spearman_sub_ln	c_d2t_threshold_to_std
26	c_d2t_r_squared_sub_ln	c_d2t_diff_sm_mm	p_d1_pearson	c_d2t_diff_sm_mm
27	p_d1t_pearson	c_d2t_diff_sm_mm_ln	p_d1t_intercept_to_std_ln	c_d2t_diff_sm_mm_ln
28	c_d2t_deviation_slope_1	smoothed	p_d1t_pearson	smoothed
29	p_d1t_pearson_sub_ln		c_d2_r_squared	
30	p_d1_pearson		p_d1t_pearson_sub_ln	
31	c_d2_deviation_slope_1		c_d2_deviation_slope_1	
32	c_d2_r_squared		c_d2t_deviation_slope_1	
33	c_d2t_deviation_slope_1_ln		c_d2_r_squared_sub	
34	c_d2_r_squared_sub		c_d2t_spearman_sub_ln	
35	p_d1_r_squared		p_d1t_deviation_slope_1_ln	
36	p_d1t_threshold_to_std		c_d2t_r_squared_sub	
37	p_d1_intercept_to_std		p_d1_r_squared_sub	
38	p_d1_r_squared_sub		c_d2t_deviation_slope_1_ln	
39	c_d2t_spearman_sub_ln		p_d1_r_squared	
40	c_d2t_r_squared		p_d1t_r_squared_sub_ln	
41	p_d1t_deviation_slope_1		p_d1t_deviation_slope_1	
42	c_d2t_r_squared_sub		p_d1_intercept_to_std	
43	p_d1t_deviation_slope_1_ln		c_d2t_r_squared	
44	p_d1t_r_squared_sub_ln		c_d2t_pearson_sub_ln	
45	c_d2t_kendall_sub_ln		p_d1_deviation_slope_1	
46	p_d1_deviation_slope_1		c_d2t_kendall_sub_ln	
47	c_d2t_pearson_sub_ln		p_d1t_r_squared_sub	
48	c_d2t_threshold_to_std		p_d1t_threshold_to_std	
49	p_d1t_r_squared		p_d1t_r_squared	
50	c_d2t_diff_sm_mm		smoothed	
51	p_d1t_r_squared_sub		c_d2t_threshold_to_std	
52	smoothed		c_d2t_diff_sm_mm	
53	c_d2t_diff_sm_mm_ln		c_d2t_diff_sm_mm_ln	
54	p_d1t_diff_sm_mm_ln		p_d1t_diff_sm_mm_ln	
55	p_d1t_diff_sm_mm		p_d1t_diff_sm_mm	

order	data resolution: **60 min** classifier type: **multiclass** used data sources: **CP**	data resolution: **60 min** classifier type: **multiclass** used data sources: **C**	data resolution: **60 min** classifier type: **binary** used data sources: **CP**	data resolution: **60 min** classifier type: **binary** used data sources: **C**
1	p_d1_kendall	c_d2_spearman	c_d2_spearman	c_d2_spearman
2	p_m_profile_relation	c_d1_profile_relation	p_d1_kendall	c_d2t_pearson
3	c_d2_spearman	c_d2t_pearson	c_d1_profile_relation	c_d1_profile_relation
4	p_d1_slope	c_d2_slope	c_d2_pearson	c_d2t_spearman
5	c_d2t_spearman	c_d2t_spearman	p_d1_slope	c_d2_pearson
6	p_d1t_kendall	c_d2_pearson	c_d2t_spearman	c_d2t_r_squared_sub_ln
7	c_d1_profile_relation_neg	c_d1_profile_relation_neg	p_d1t_kendall	c_d2_kendall
8	c_d2t_pearson	c_d2_kendall	p_m_profile_relation	c_d1_profile_relation_neg
9	c_d1_profile_relation	c_d2t_intercept_to_std_ln	c_d2t_pearson	c_d2_r_squared
10	c_d2_r_squared	c_d2_r_squared	p_d1t_slope	c_d2_slope
11	p_d1t_slope	c_d2t_kendall	c_d2t_intercept_to_std_ln	c_d2t_kendall
12	c_d2_kendall	c_d2t_r_squared_sub_ln	c_d2t_kendall	c_d2t_intercept_to_std_ln
13	p_d1t_spearman	c_d2_r_squared_sub	c_d2_r_squared	c_d2_r_squared_sub
14	c_d2t_intercept_to_std_ln	c_d2t_slope	p_d1_spearman	c_d2t_r_squared
15	c_d2t_kendall	c_d2t_r_squared	c_d1_profile_relation_neg	c_d2t_r_squared_sub
16	p_m_profile_relation_neg	c_d2t_intercept_to_std	p_d1t_spearman_sub_ln	c_d2t_intercept_to_std
17	c_d2_pearson	c_d2t_pearson_sub_ln	c_d2_kendall	c_d2t_slope
18	p_d1_spearman	c_d2_intercept_to_std	c_d2_r_squared_sub	c_d2t_pearson_sub_ln
19	c_d2_slope	c_d2t_spearman_sub_ln	p_d1t_spearman	c_d2_deviation_slope_1
20	p_d1t_spearman_sub_ln	c_d2t_r_squared_sub	c_d2_slope	c_d2t_spearman_sub_ln
21	p_d1t_intercept_to_std_ln	c_d2_deviation_slope_1	c_d2t_r_squared_sub	c_d2t_kendall_sub_ln
22	c_d2_r_squared_sub	c_d2t_kendall_sub_ln	p_d1t_deviation_slope_1_ln	c_d2_intercept_to_std
23	c_d2t_r_squared	c_d2t_deviation_slope_1	c_d2t_intercept_to_std	c_d2t_deviation_slope_1
24	p_d1t_kendall_sub_ln	c_d2t_deviation_slope_1_ln	p_d1t_r_squared	c_d2t_deviation_slope_1_ln
25	c_d2_deviation_slope_1	c_d2t_threshold_to_std	p_m_profile_relation_neg	c_d2t_threshold_to_std
26	p_d1t_pearson	smoothed	c_d2_deviation_slope_1	c_d2t_diff_sm_mm
27	p_d1t_intercept_to_std	c_d2t_diff_sm_mm	c_d2t_r_squared	smoothed
28	c_d2t_intercept_to_std	c_d2t_diff_sm_mm_ln	p_d1t_intercept_to_std_ln	c_d2t_diff_sm_mm_ln
29	c_d2t_r_squared_sub		c_d2t_pearson_sub_ln	
30	p_d1t_deviation_slope_1_ln		p_d1t_pearson	
31	c_d2t_spearman_sub_ln		p_d1_intercept_to_std	
32	c_d2_intercept_to_std		c_d2t_spearman_sub_ln	
33	p_d1_pearson		p_d1t_deviation_slope_1	
34	p_d1t_pearson_sub_ln		p_d1t_kendall_sub_ln	
35	p_d1_intercept_to_std		p_d1t_intercept_to_std	
36	c_d2t_pearson_sub_ln		p_d1t_r_squared_sub	
37	c_d2t_slope		p_d1_pearson	
38	p_d1t_r_squared_sub_ln		p_d1_deviation_slope_1	
39	p_d1t_deviation_slope_1		p_d1t_pearson_sub_ln	
40	c_d2t_r_squared_sub_ln		c_d2t_slope	
41	p_d1_r_squared		c_d2t_r_squared_sub_ln	
42	c_d2t_kendall_sub_ln		c_d2_intercept_to_std	
43	p_d1_deviation_slope_1		c_d2t_kendall_sub_ln	
44	p_d1t_r_squared		p_d1_r_squared	
45	p_d1_r_squared_sub		p_d1t_r_squared_sub_ln	
46	p_d1t_r_squared_sub		p_d1_r_squared_sub	
47	p_d1t_threshold_to_std		c_d2t_deviation_slope_1	
48	c_d2t_deviation_slope_1		p_d1t_threshold_to_std	
49	c_d2t_threshold_to_std		c_d2t_deviation_slope_1_ln	
50	c_d2t_deviation_slope_1_ln		c_d2t_threshold_to_std	
51	smoothed		c_d2t_diff_sm_mm_ln	
52	c_d2t_diff_sm_mm		smoothed	
53	c_d2t_diff_sm_mm_ln		c_d2t_diff_sm_mm	
54	p_d1t_diff_sm_mm_ln		p_d1t_diff_sm_mm_ln	
55	p_d1t_diff_sm_mm		p_d1t_diff_sm_mm	

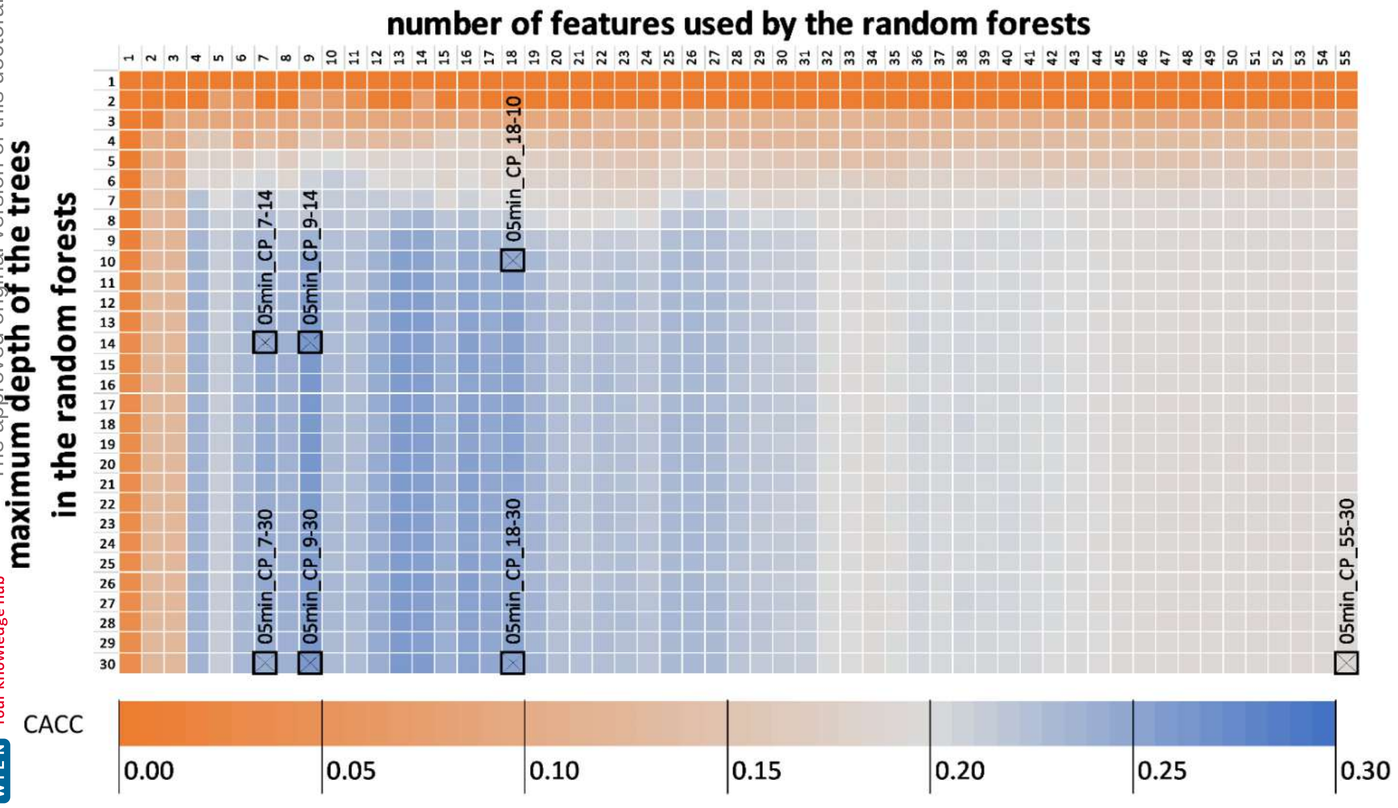
11 Appendix C: Selection of the best classifier candidates

data resolution: **5 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **multiclass**



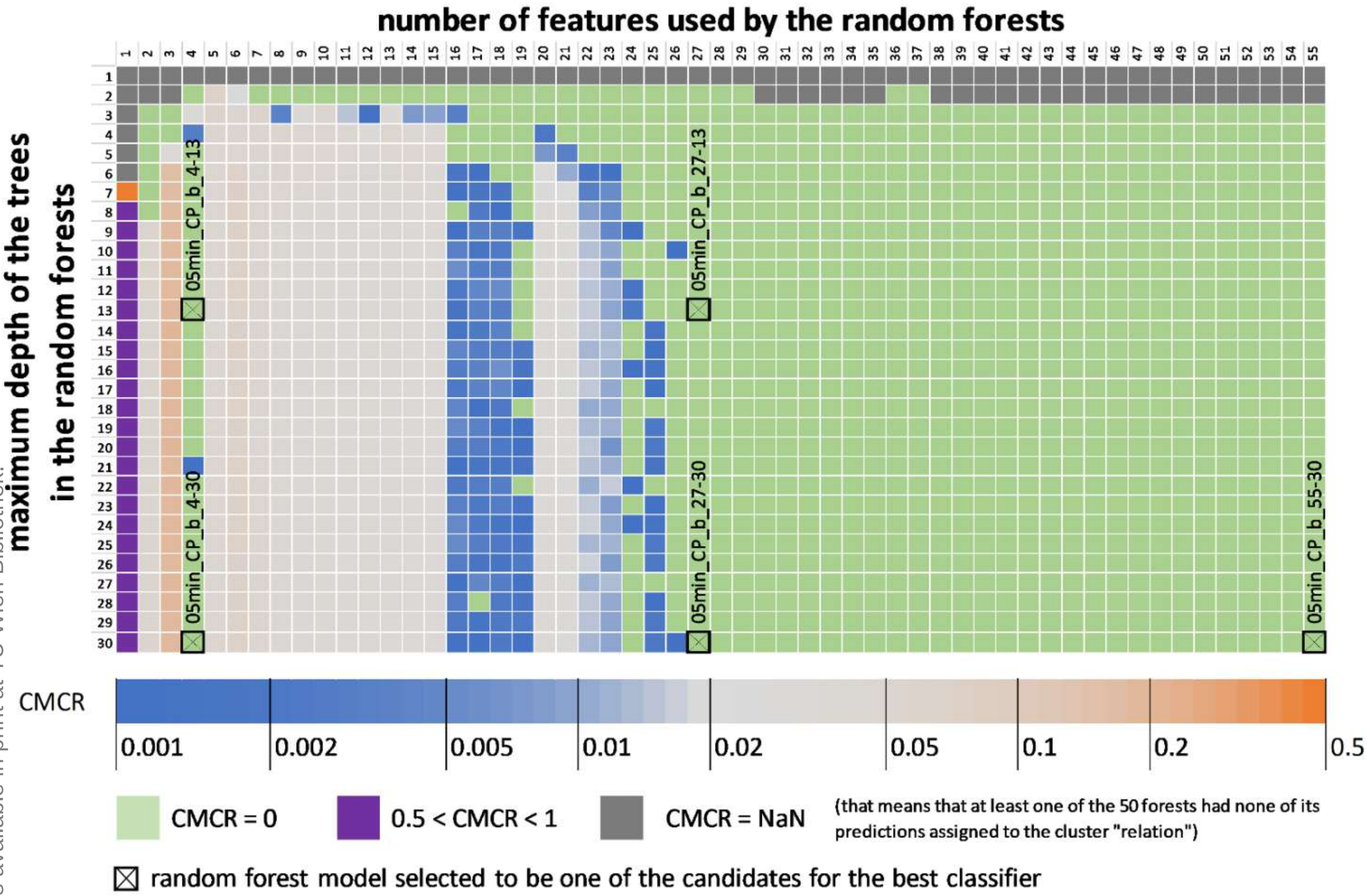
CMCR = 0 0.5 < CMCR < 1 CMCR = NaN (that means that at least one of the 50 forests had none of its predictions assigned to cluster "obvious relations" or none assigned to the clusters "rare partial supply" and "inconspicuous relations")

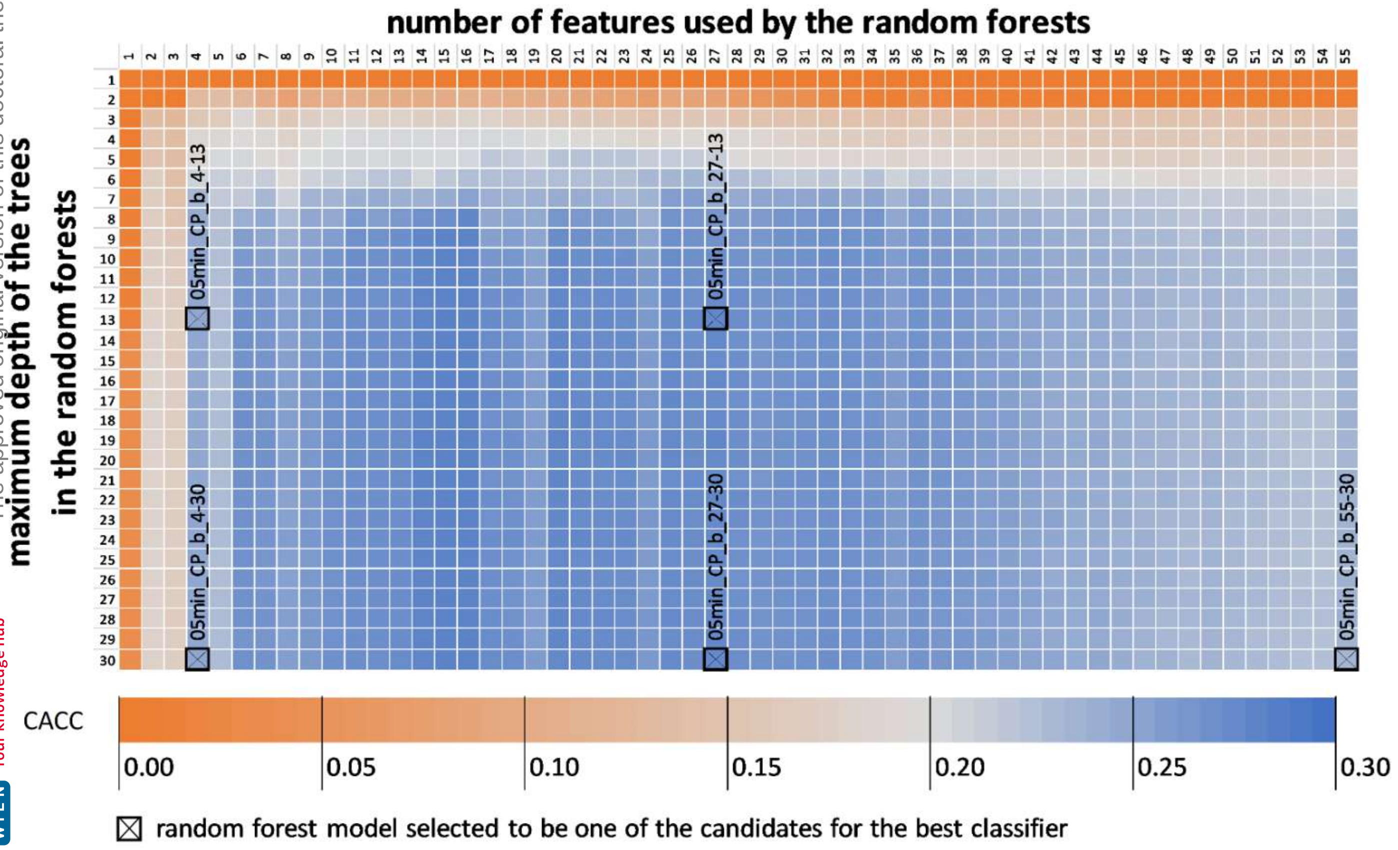
☒ random forest model selected to be one of the candidates for the best classifier



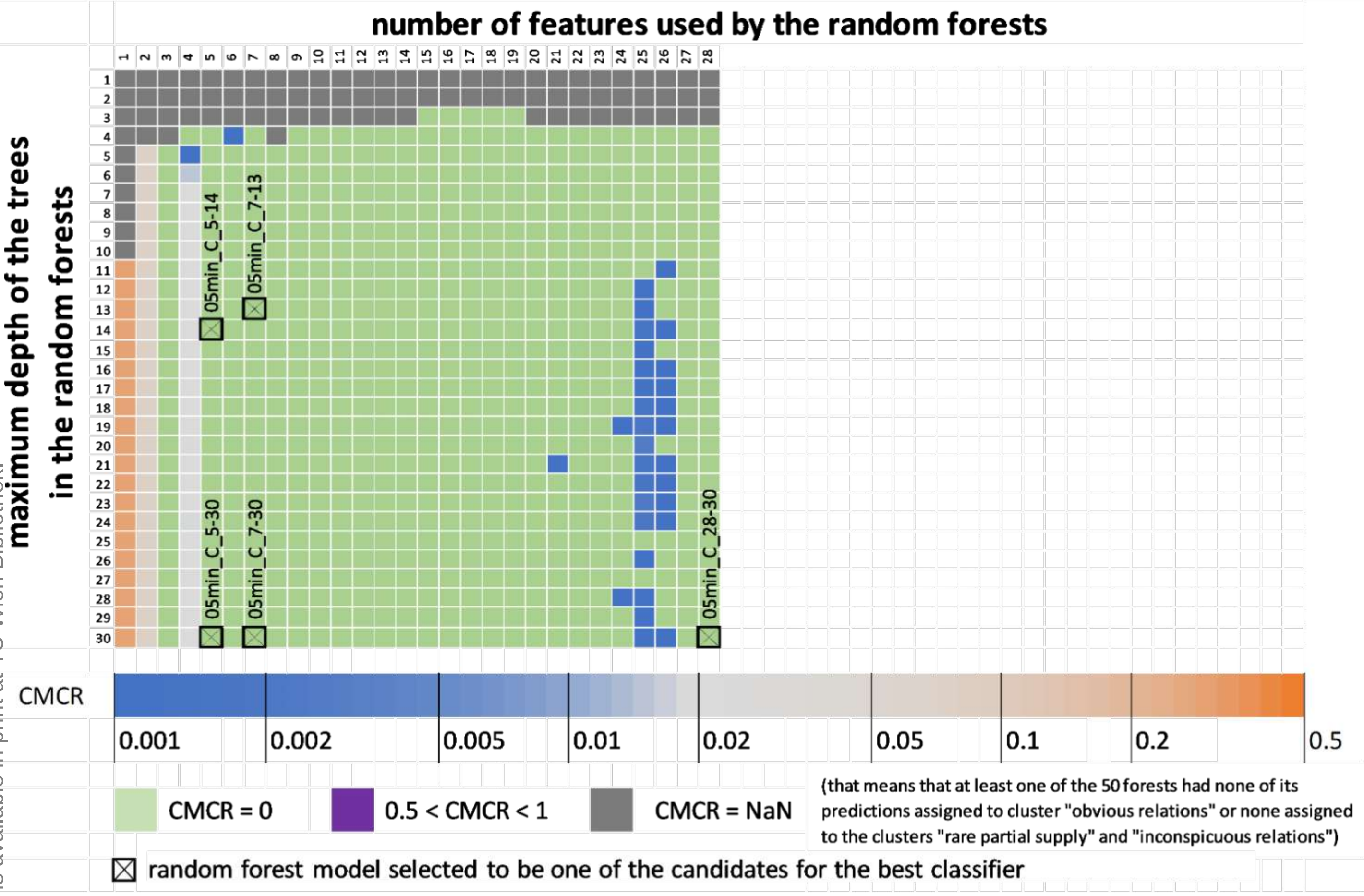
☒ random forest model selected to be one of the candidates for the best classifier

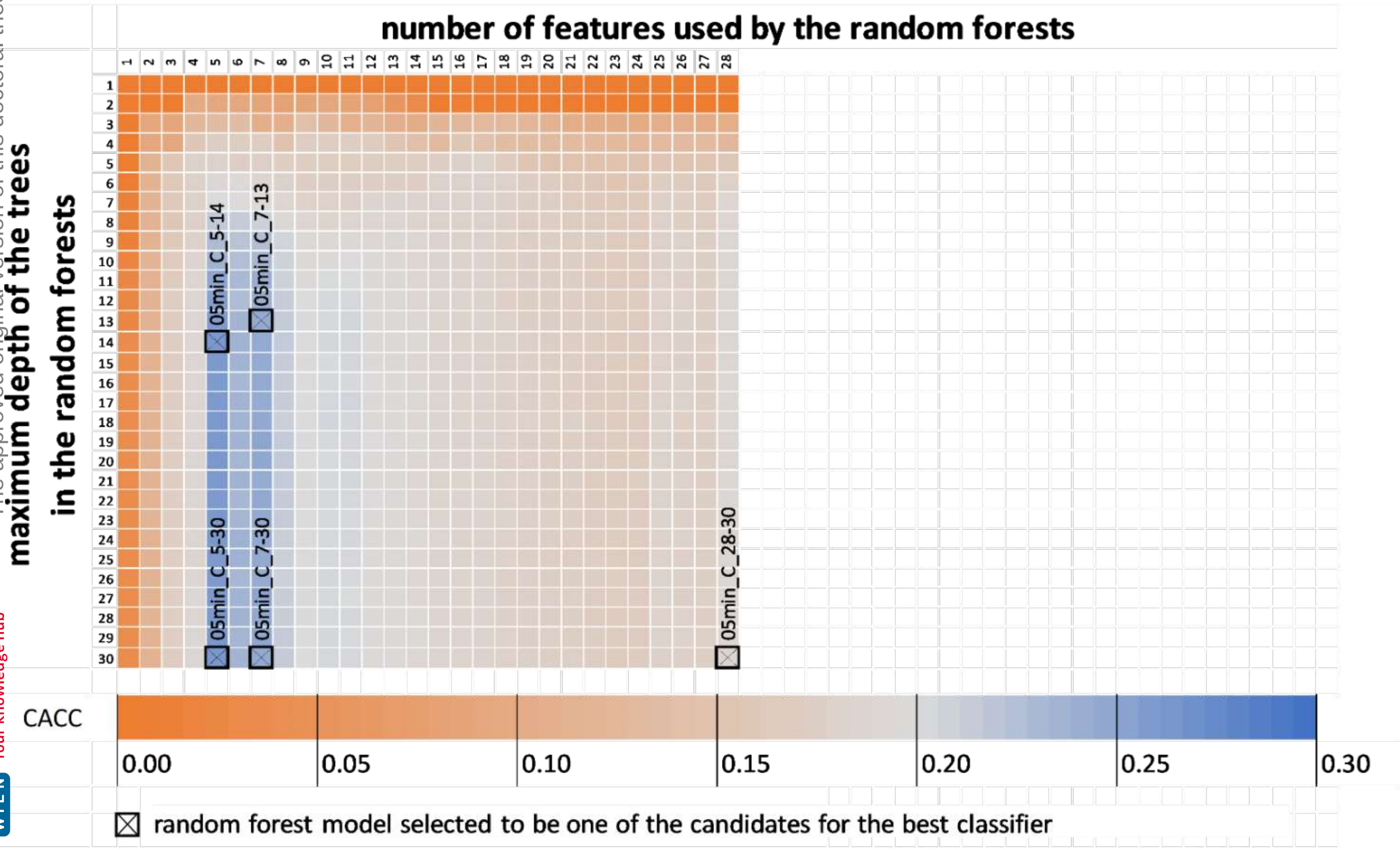
data resolution: **5 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **binary**



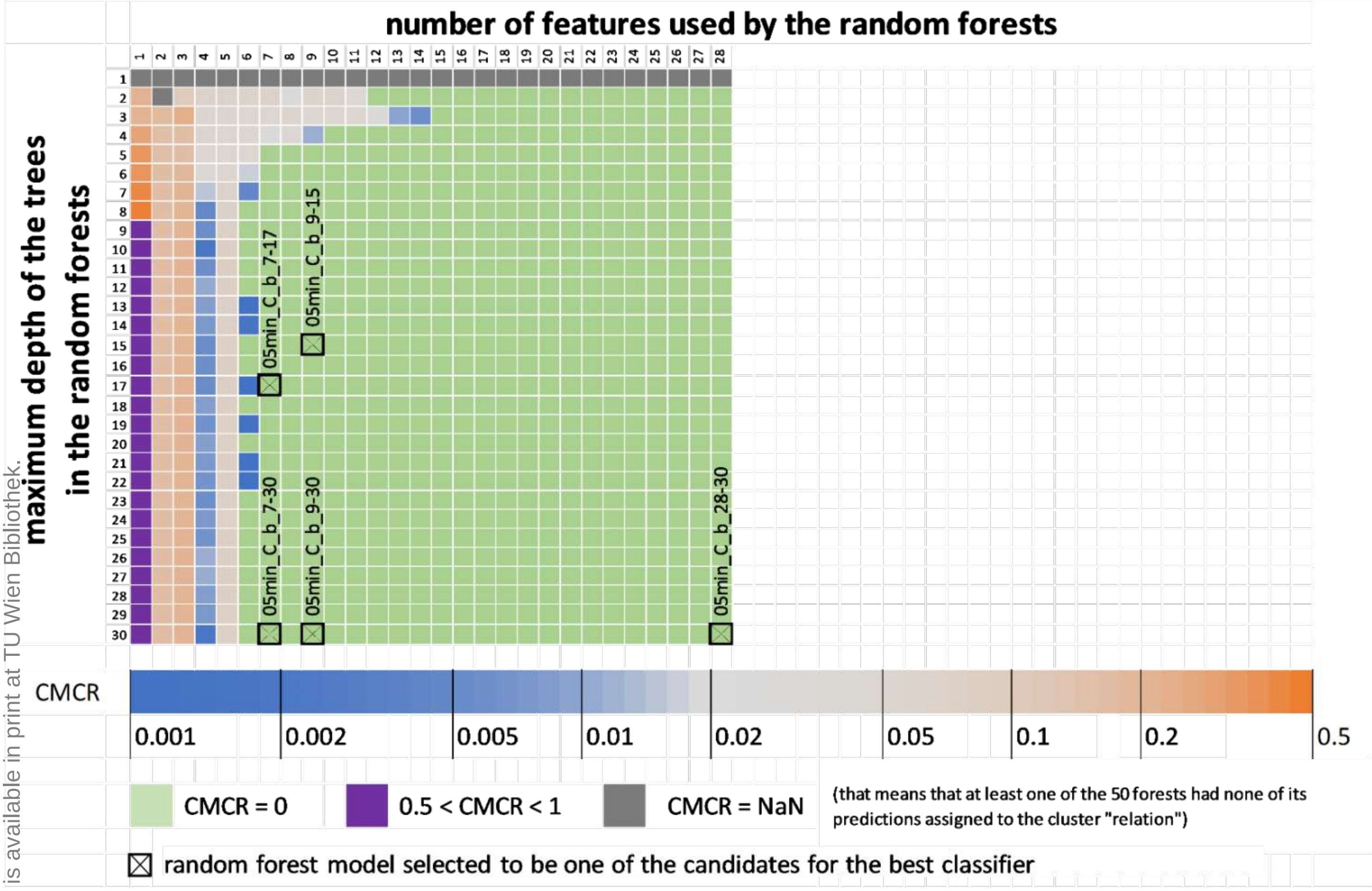


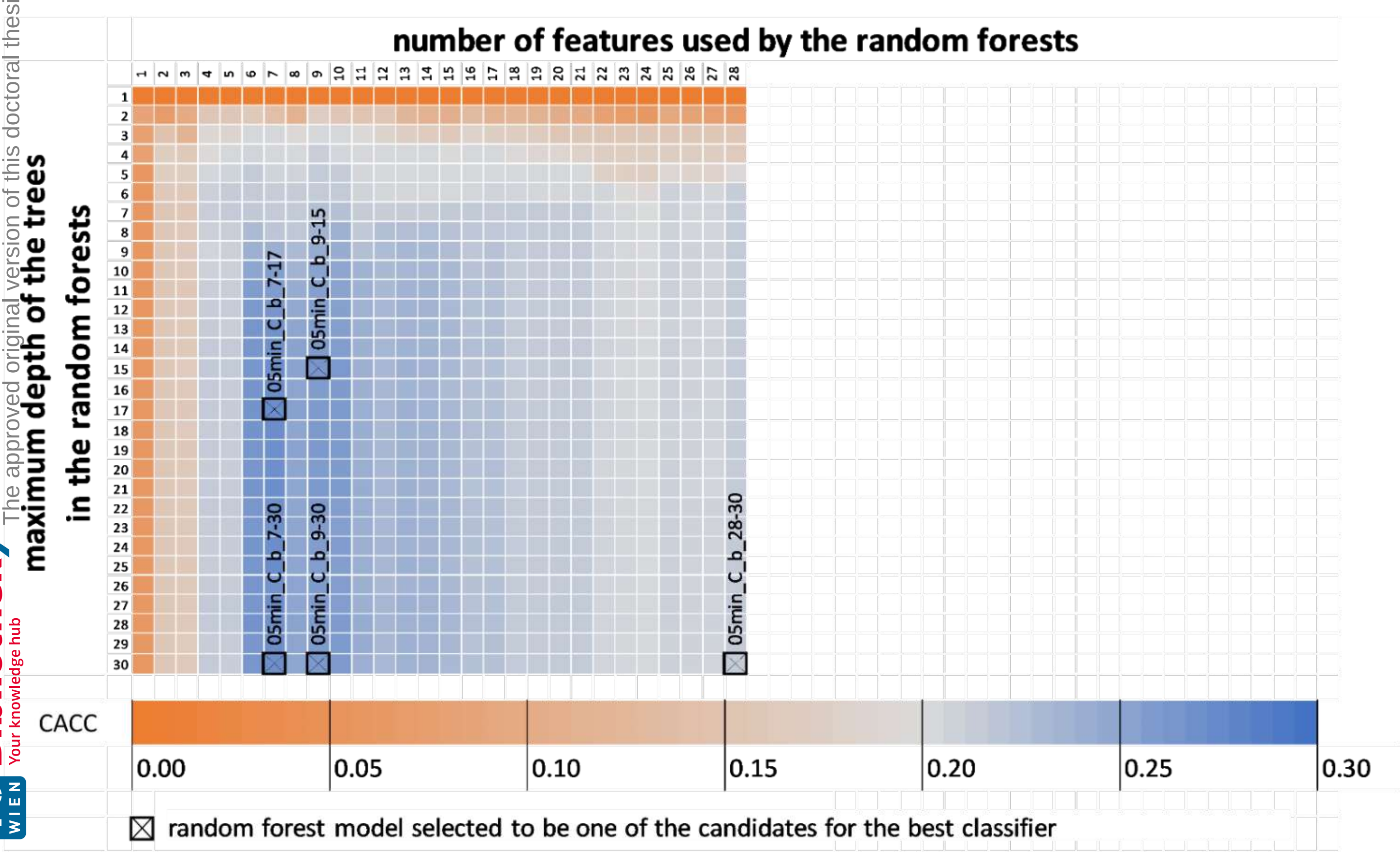
data resolution: **5 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **multiclass**



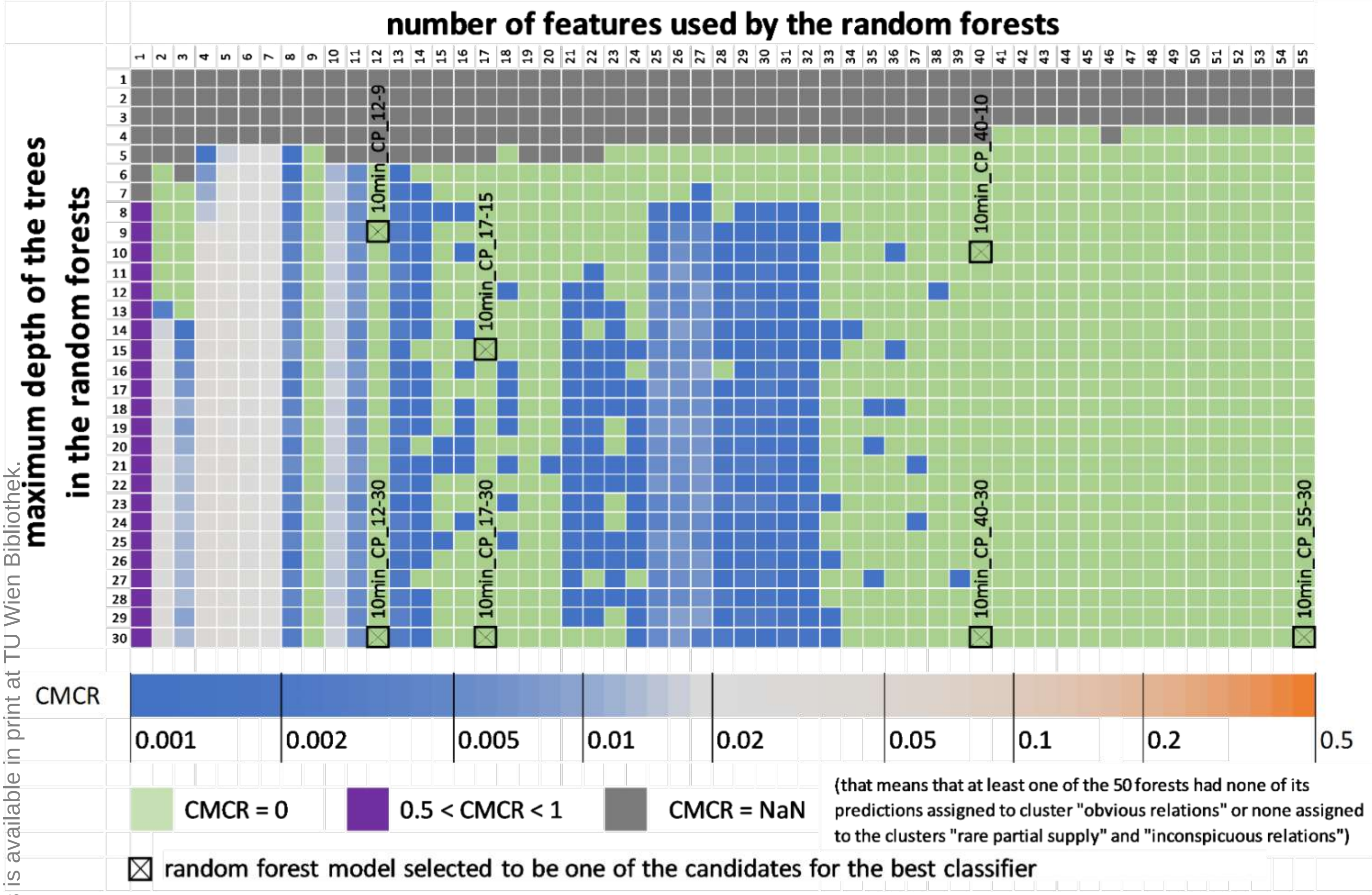


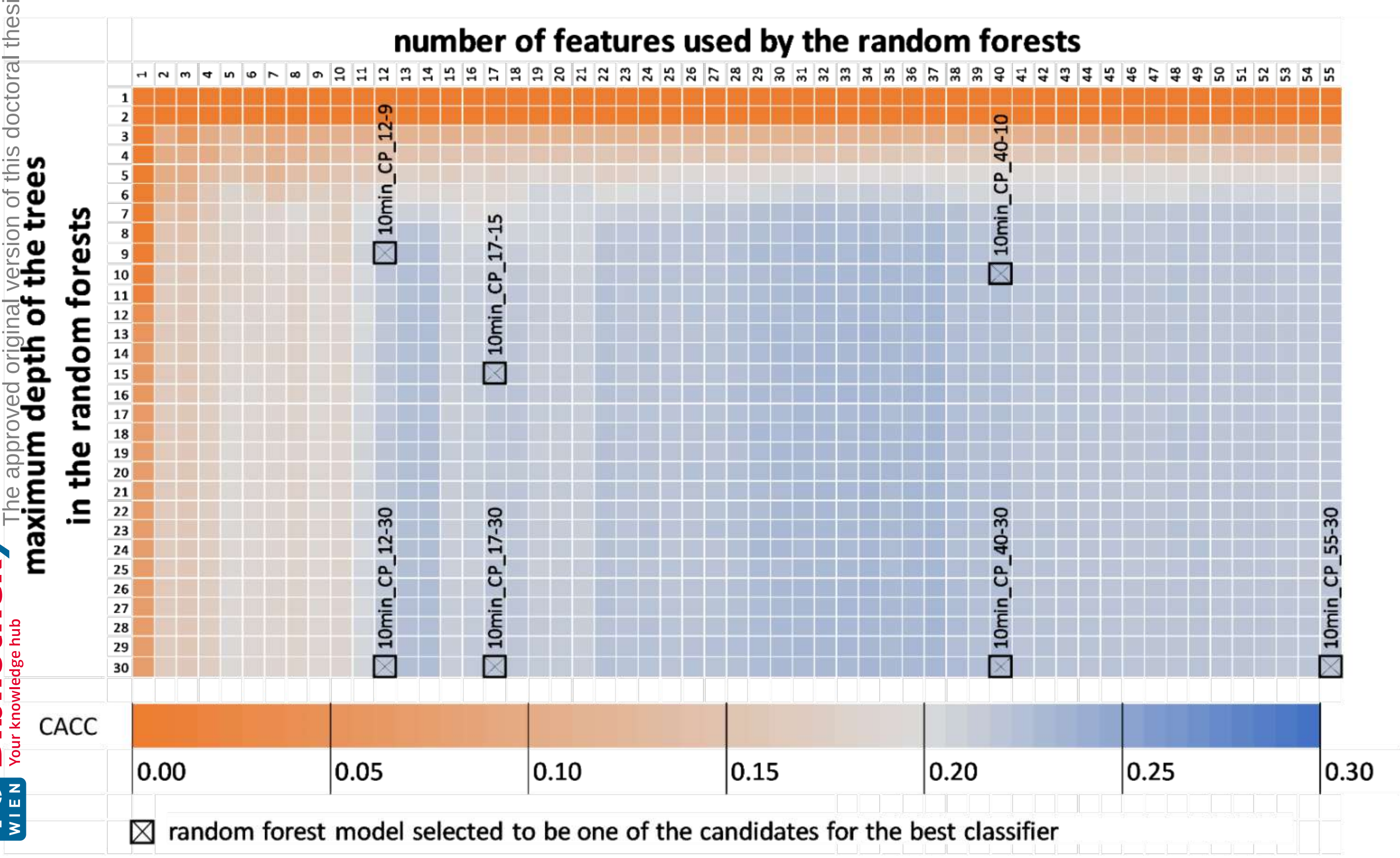
data resolution: **5 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **binary**



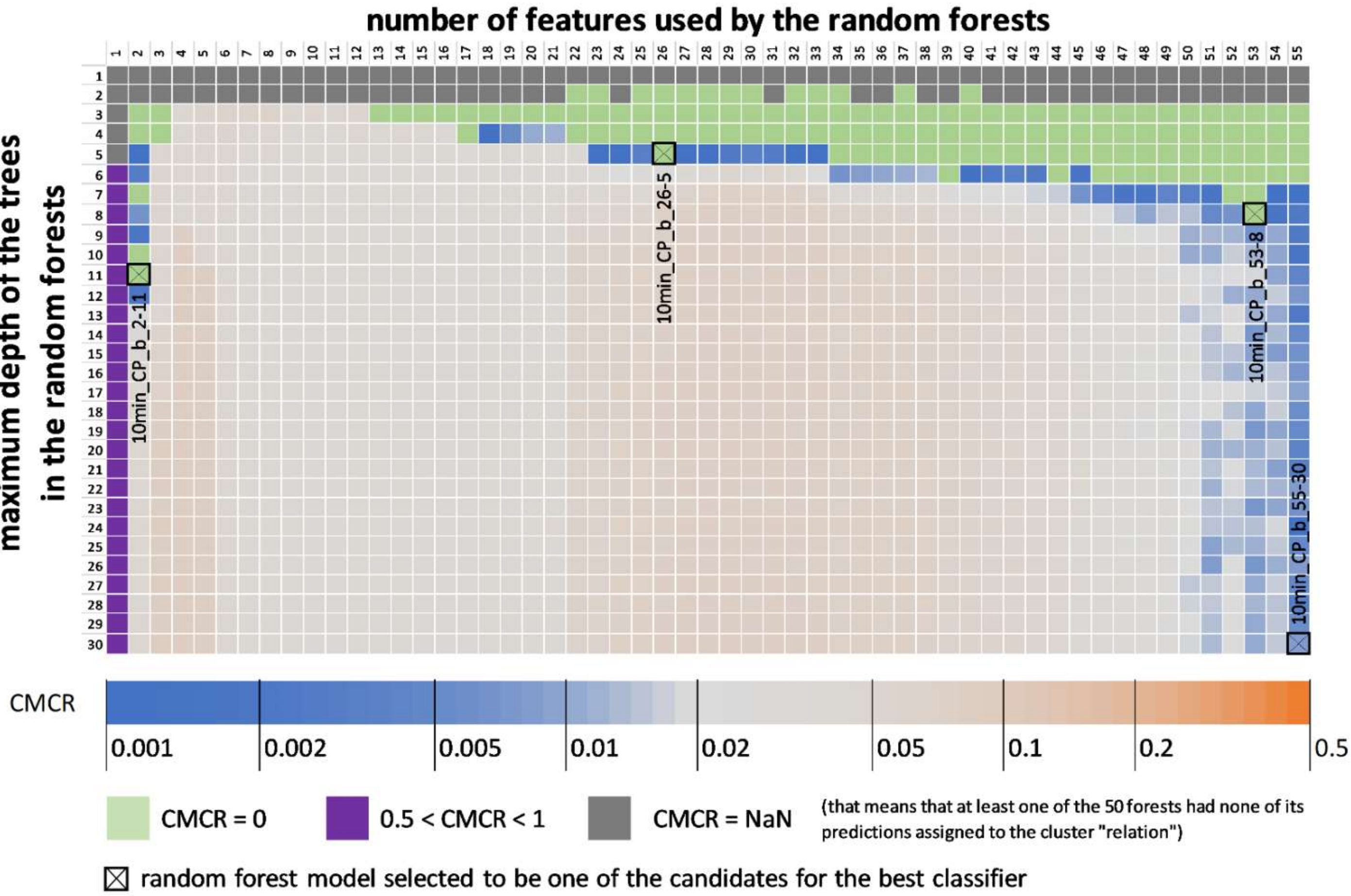


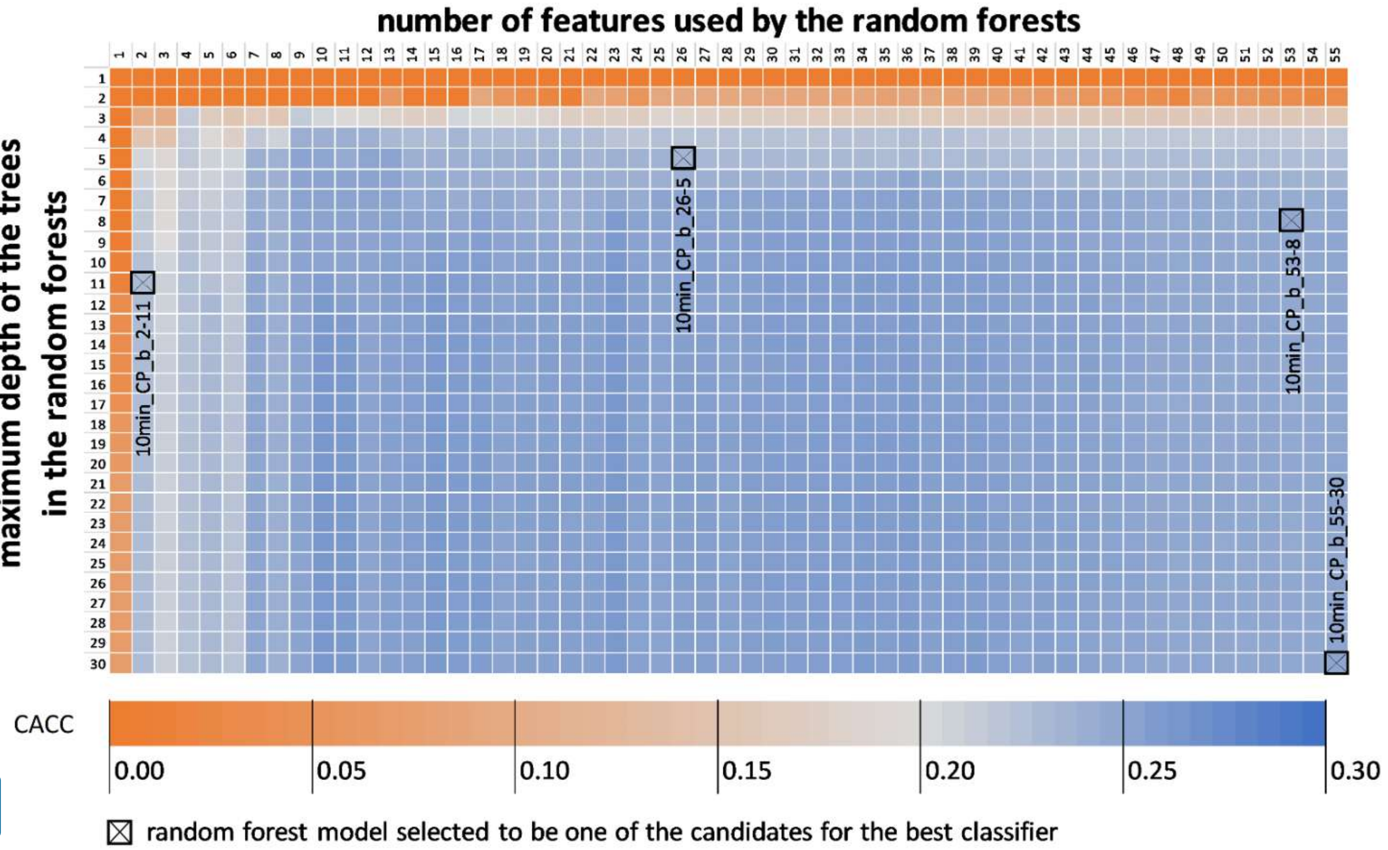
data resolution: **10 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **multiclass**



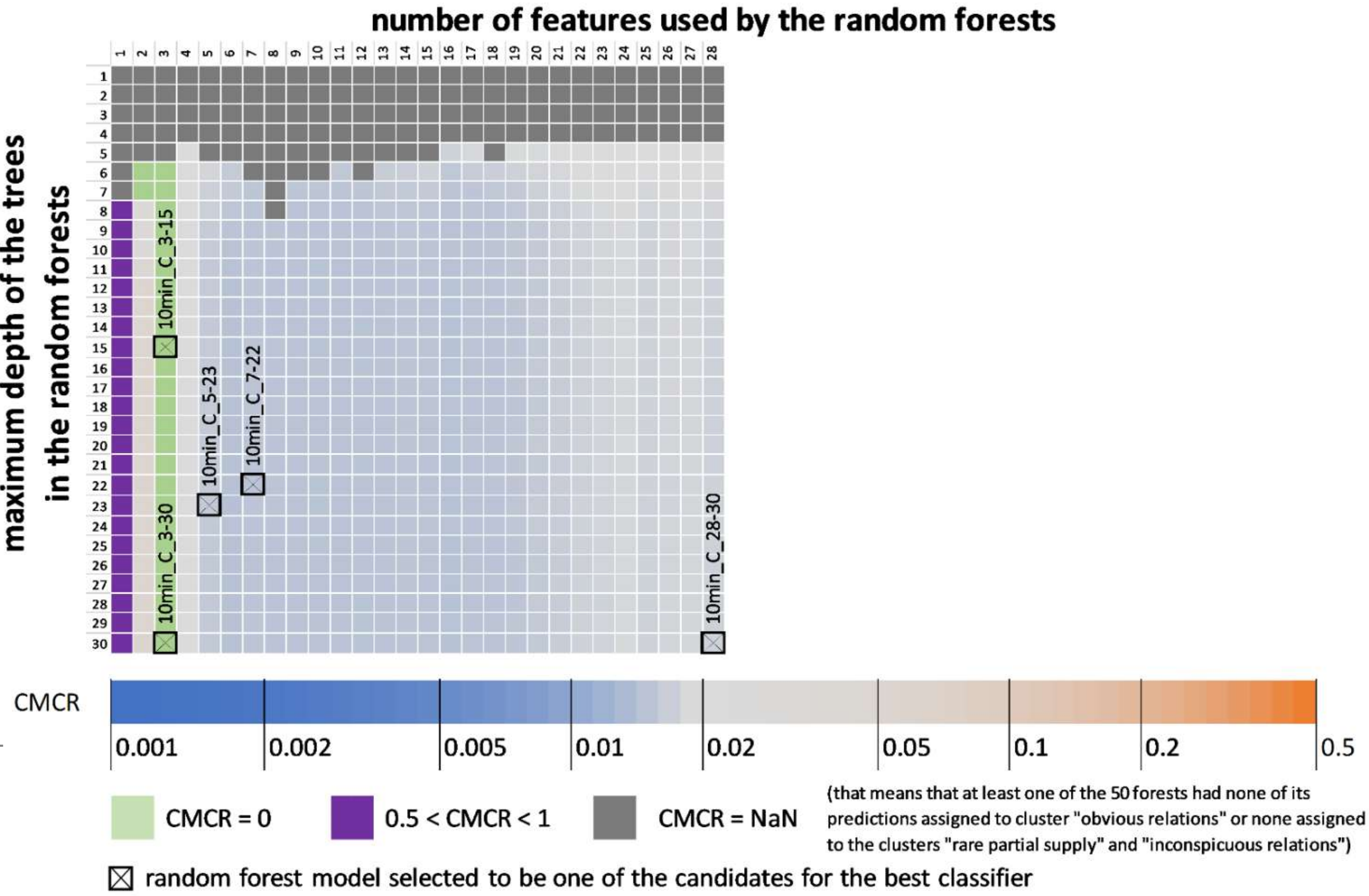


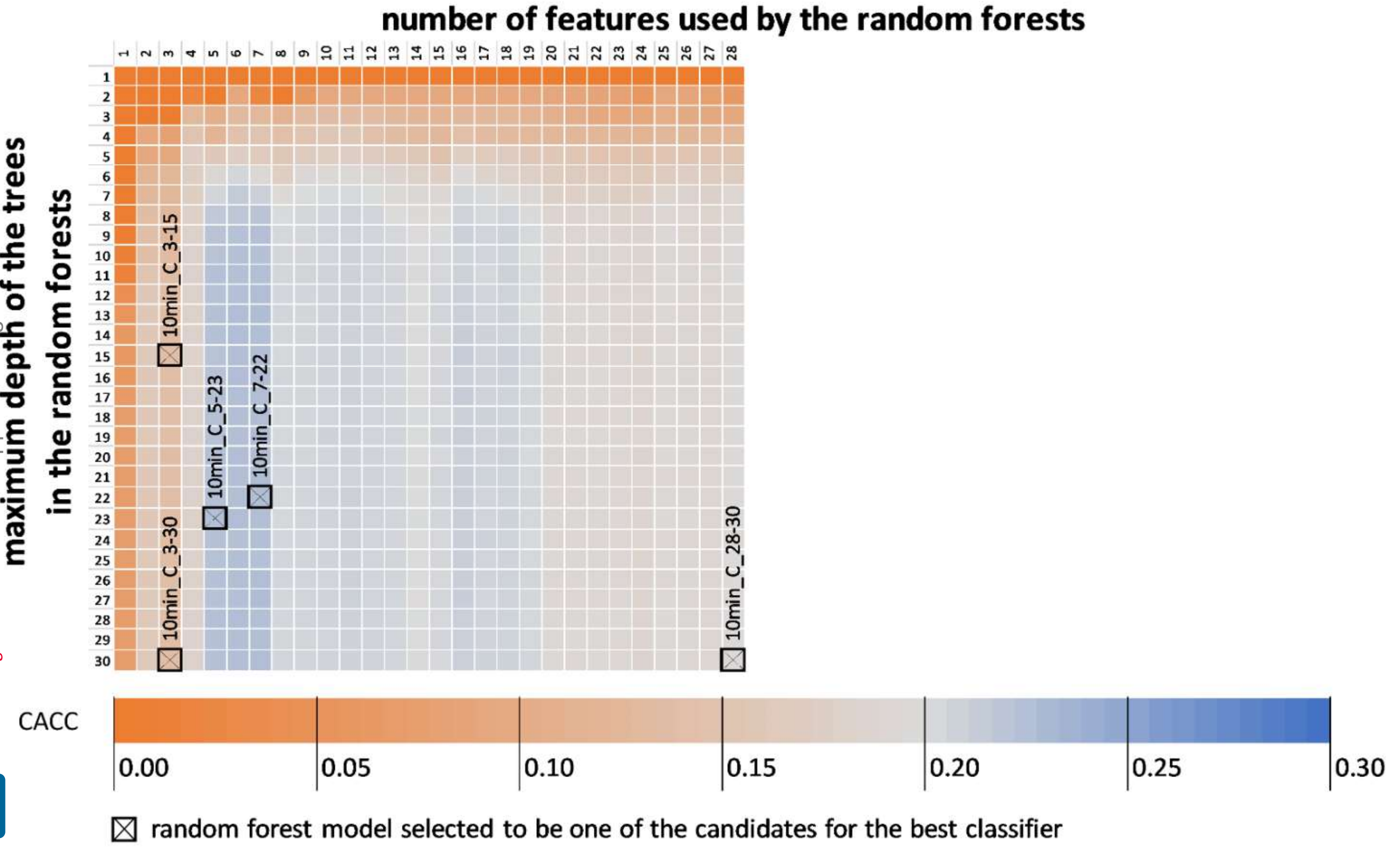
data resolution: **10 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **binary**



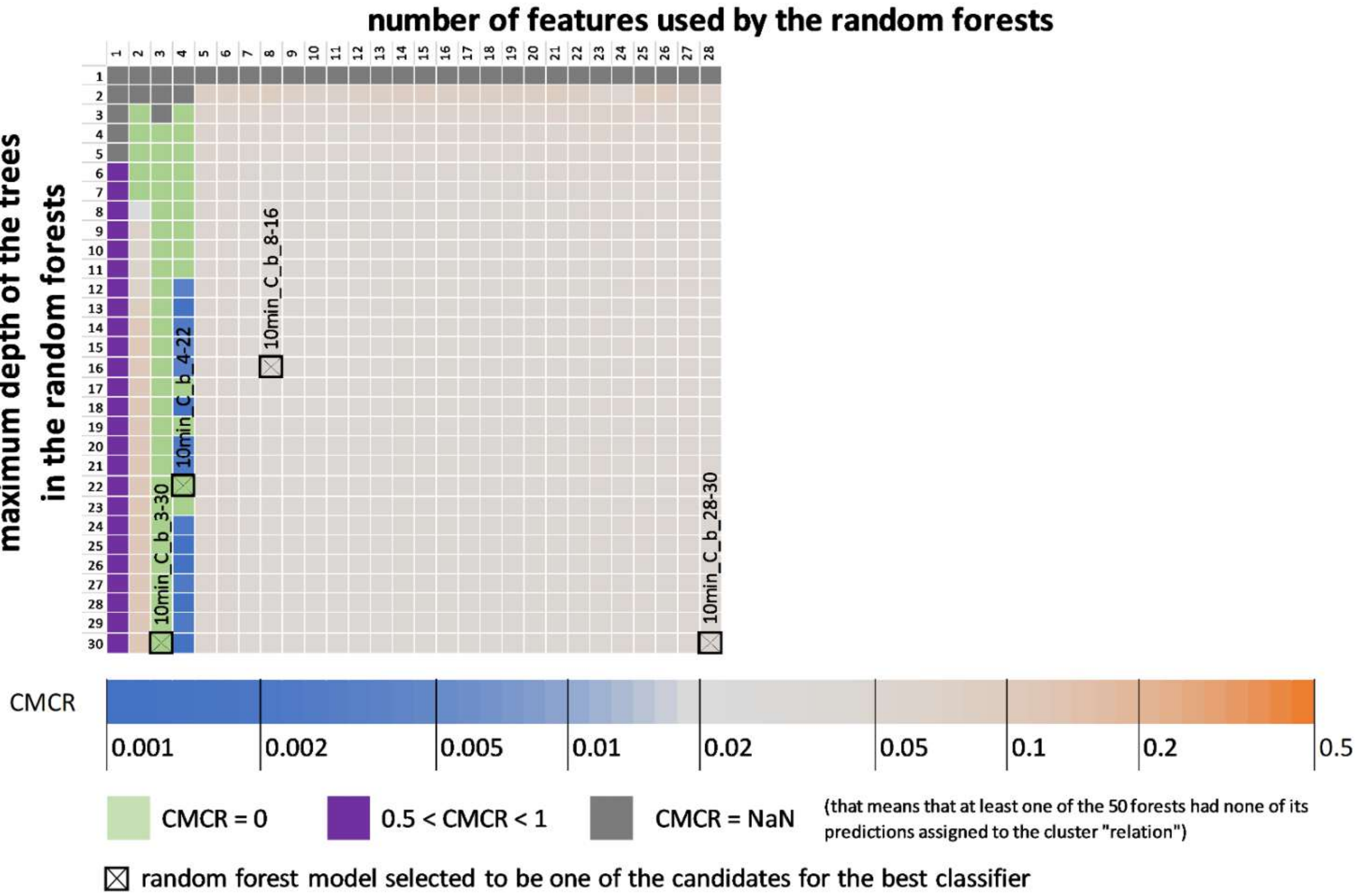


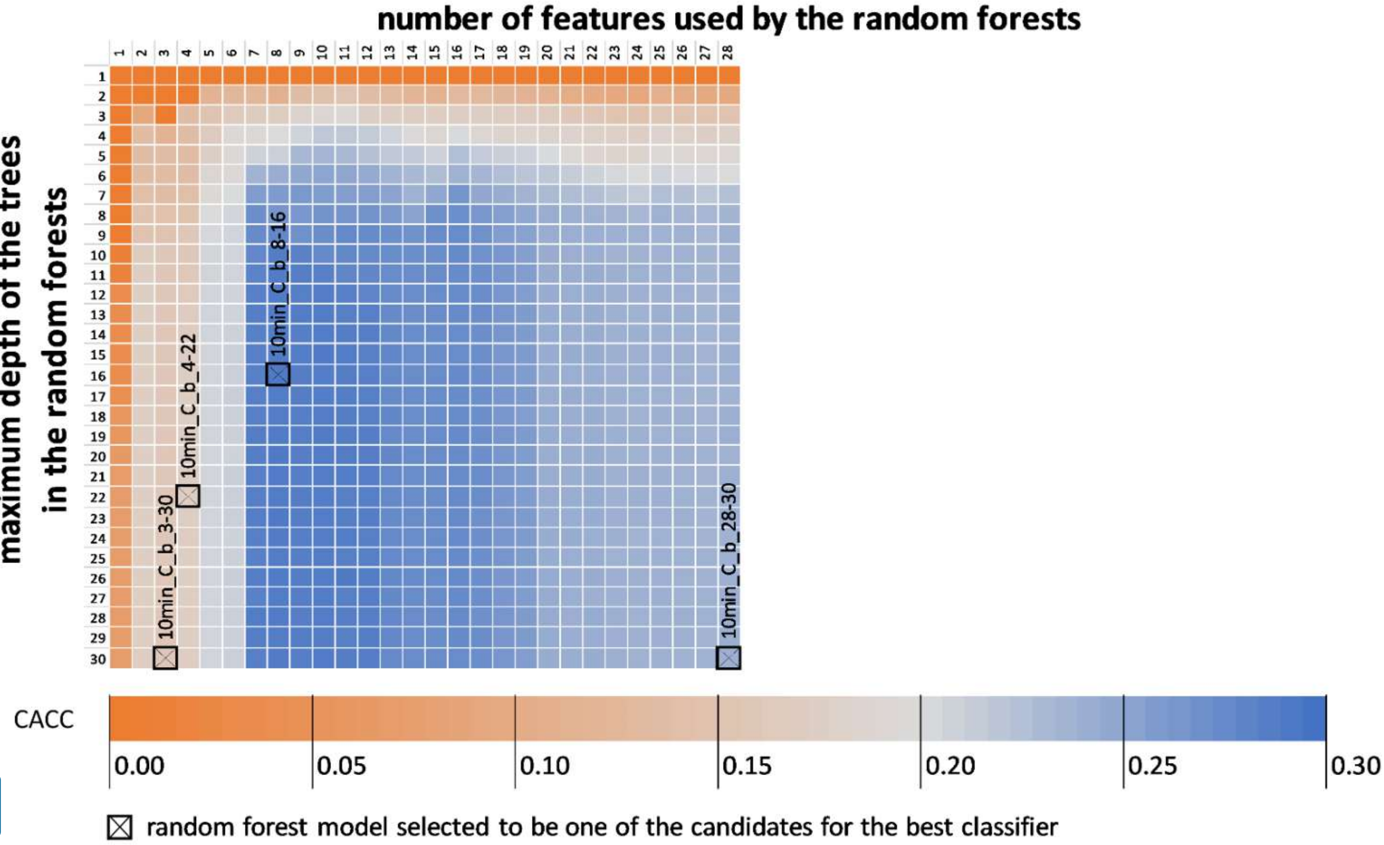
data resolution: **10 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **multiclass**



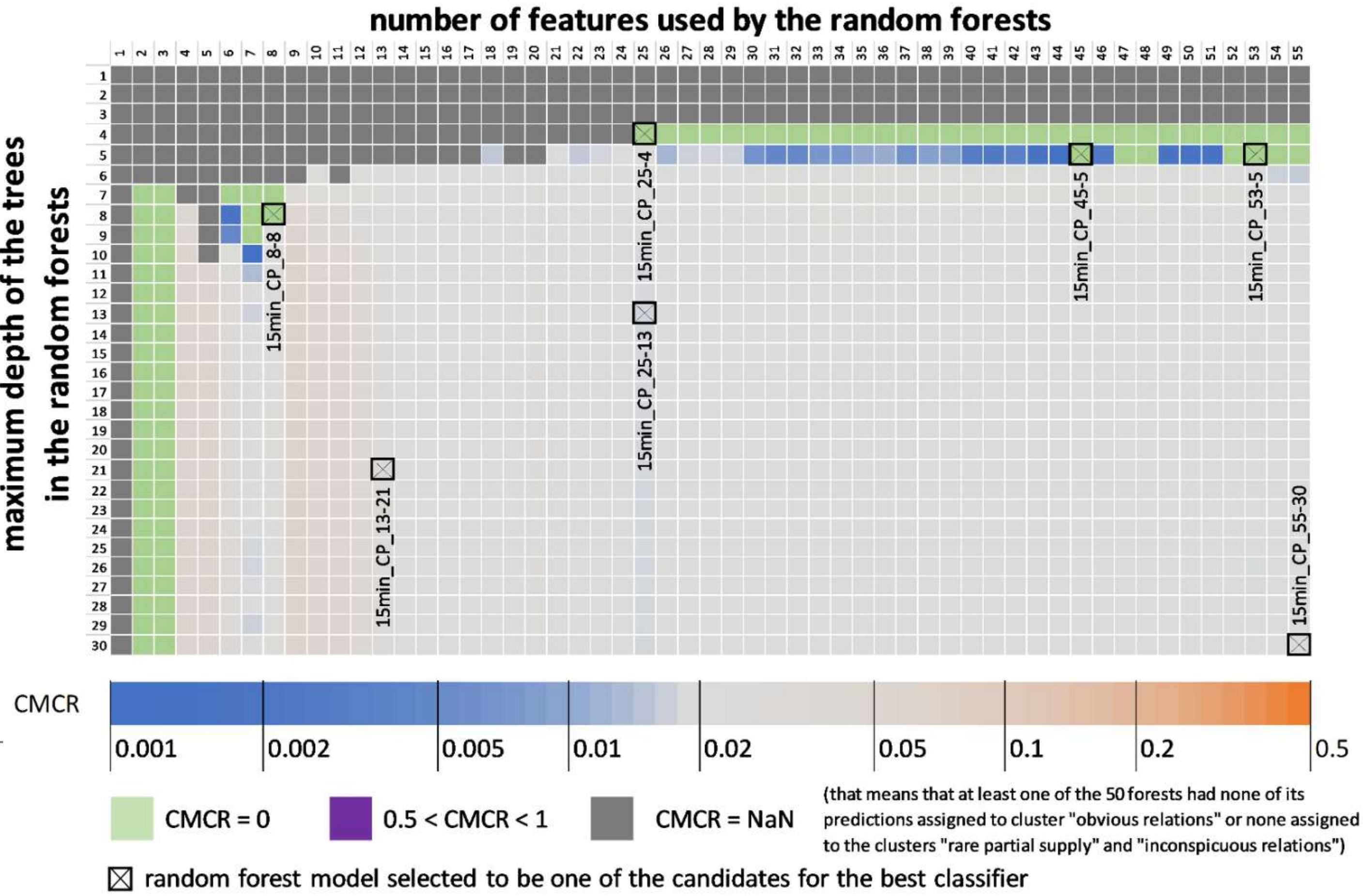


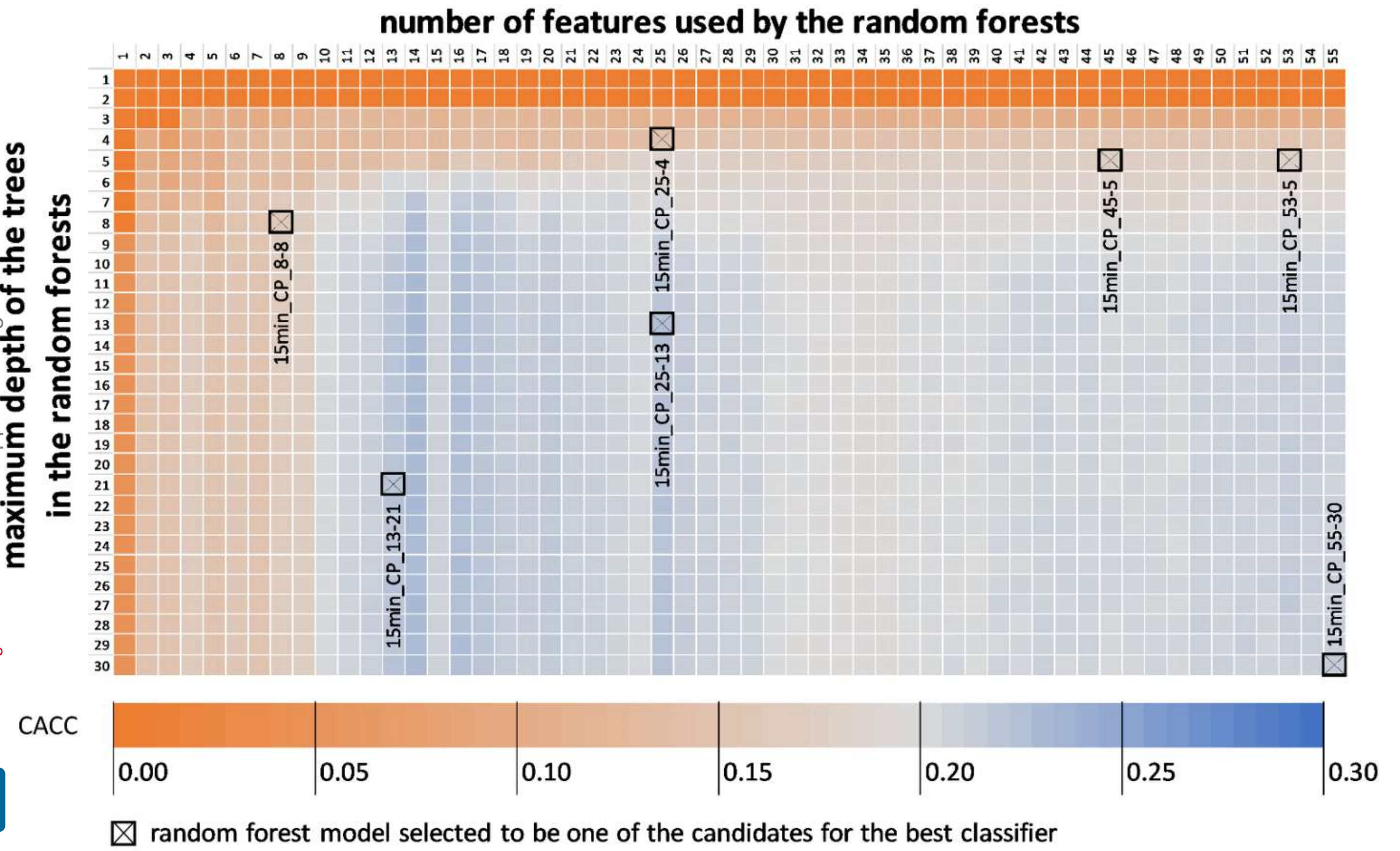
data resolution: **10 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **binary**



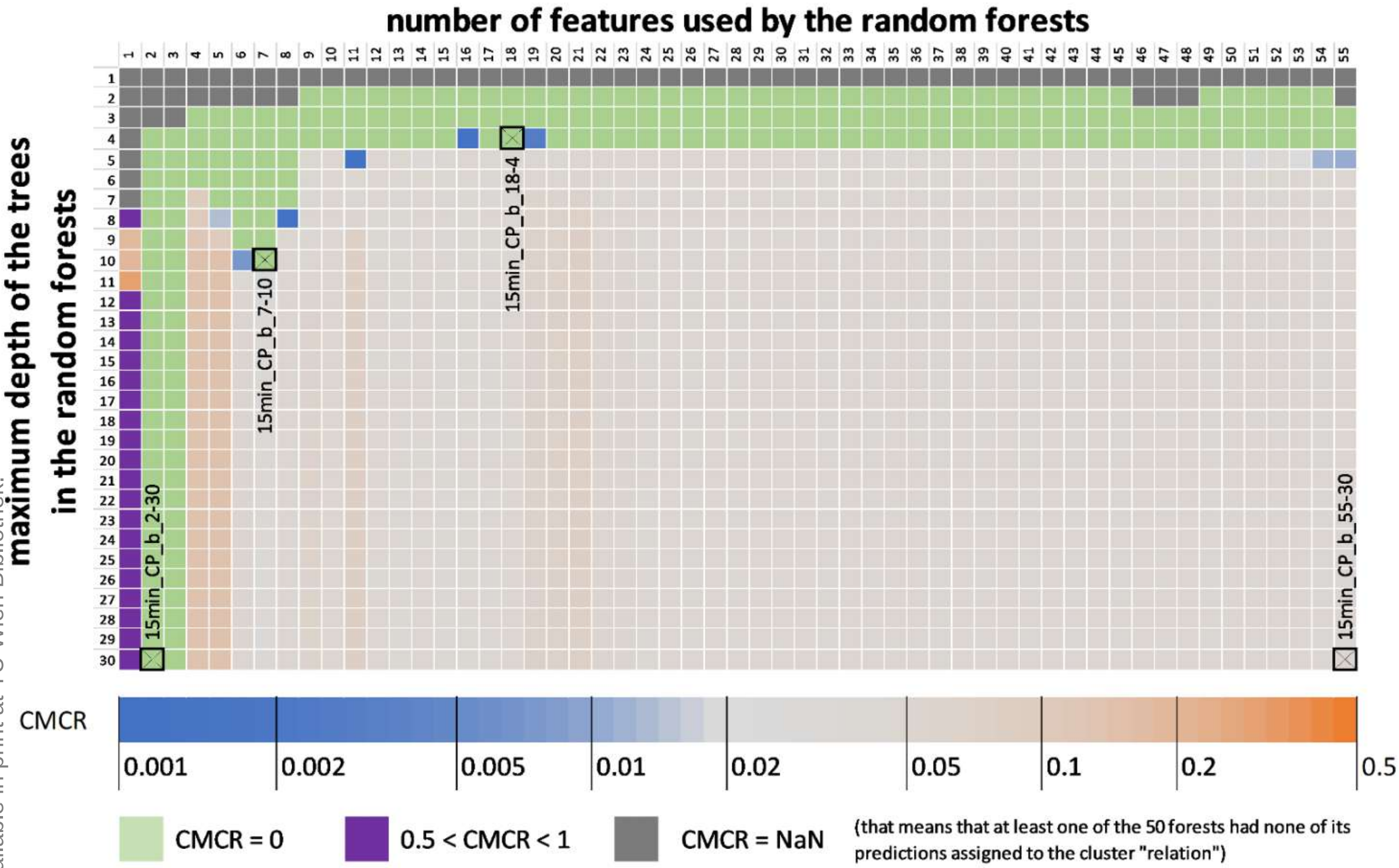


data resolution: **15 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **multiclass**

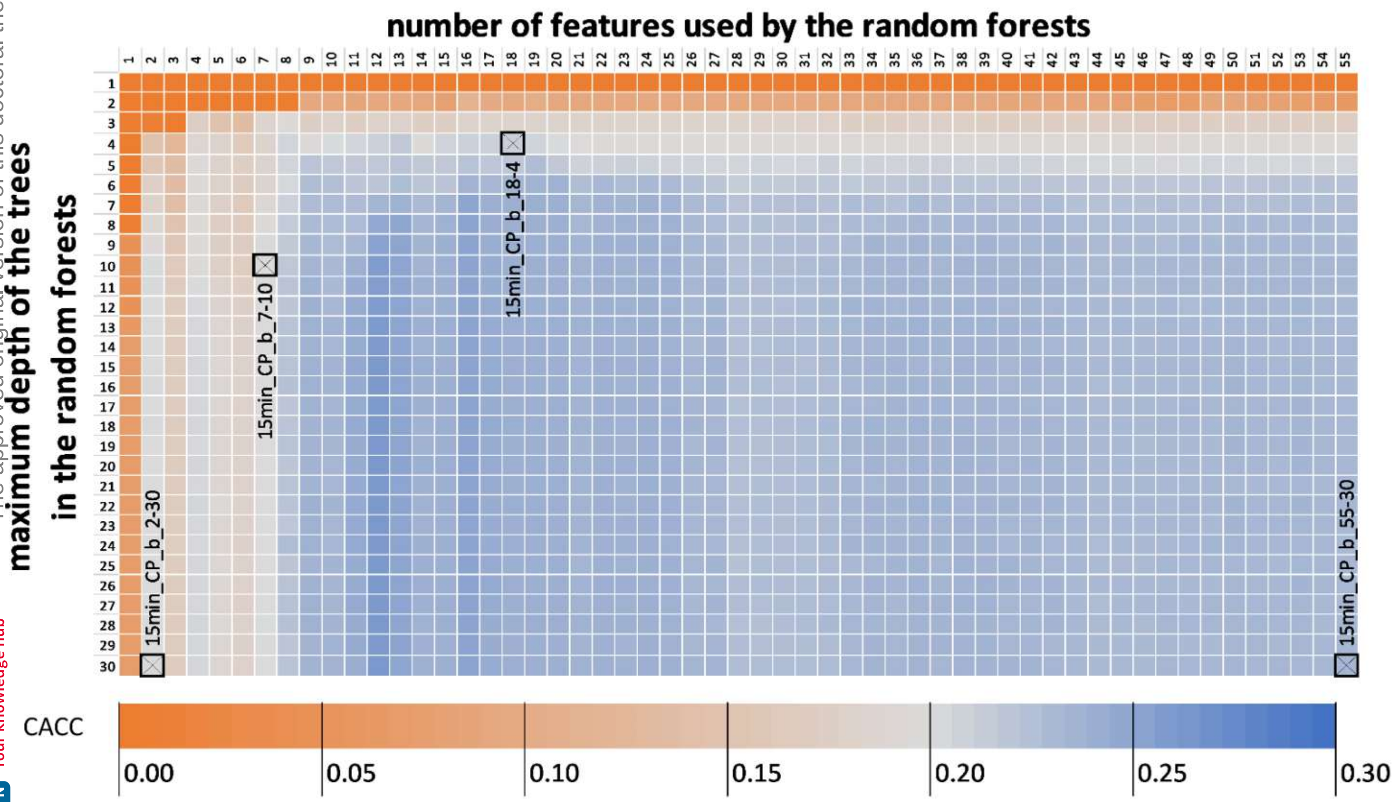




data resolution: **15 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **binary**

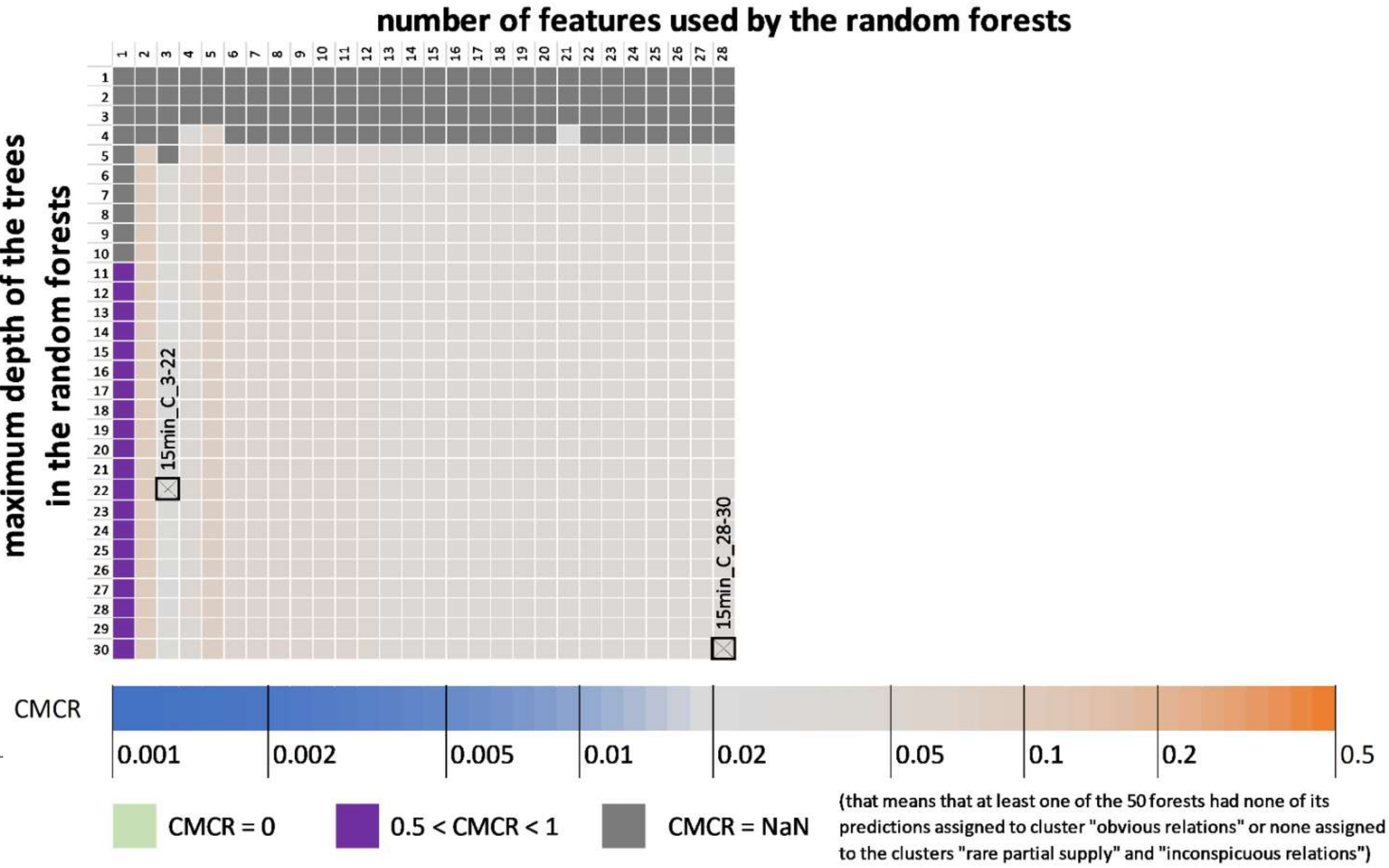


☒ random forest model selected to be one of the candidates for the best classifier

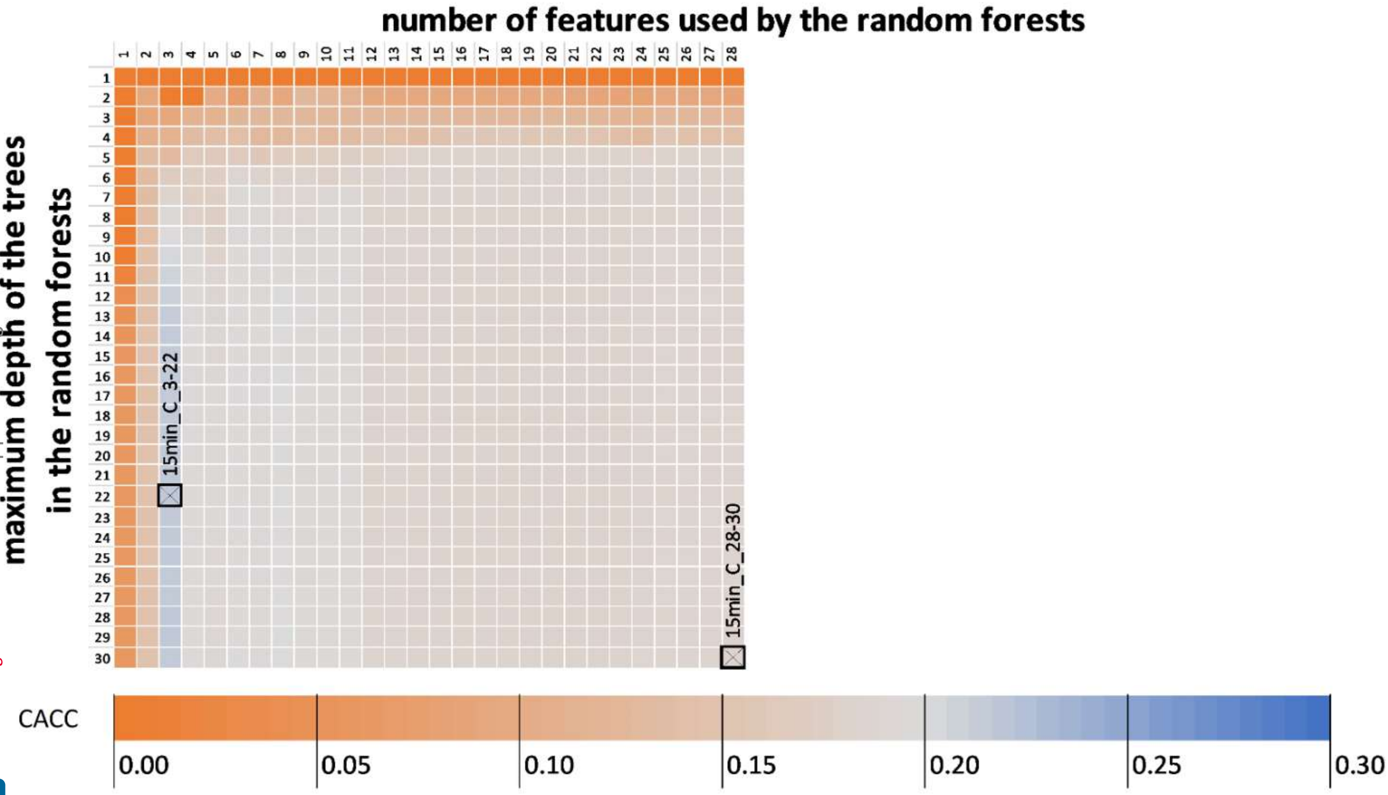


☒ random forest model selected to be one of the candidates for the best classifier

data resolution: **15 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **multiclass**

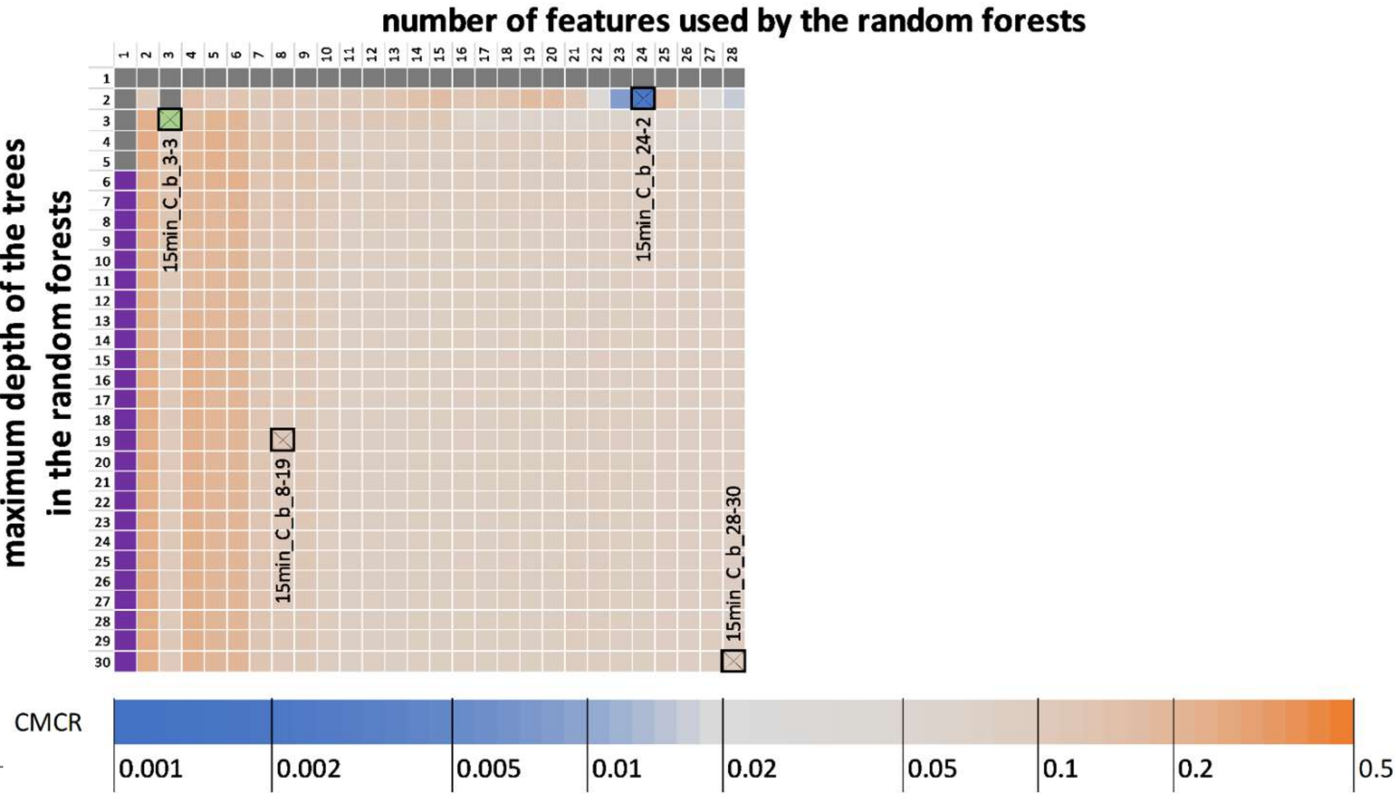


☒ random forest model selected to be one of the candidates for the best classifier



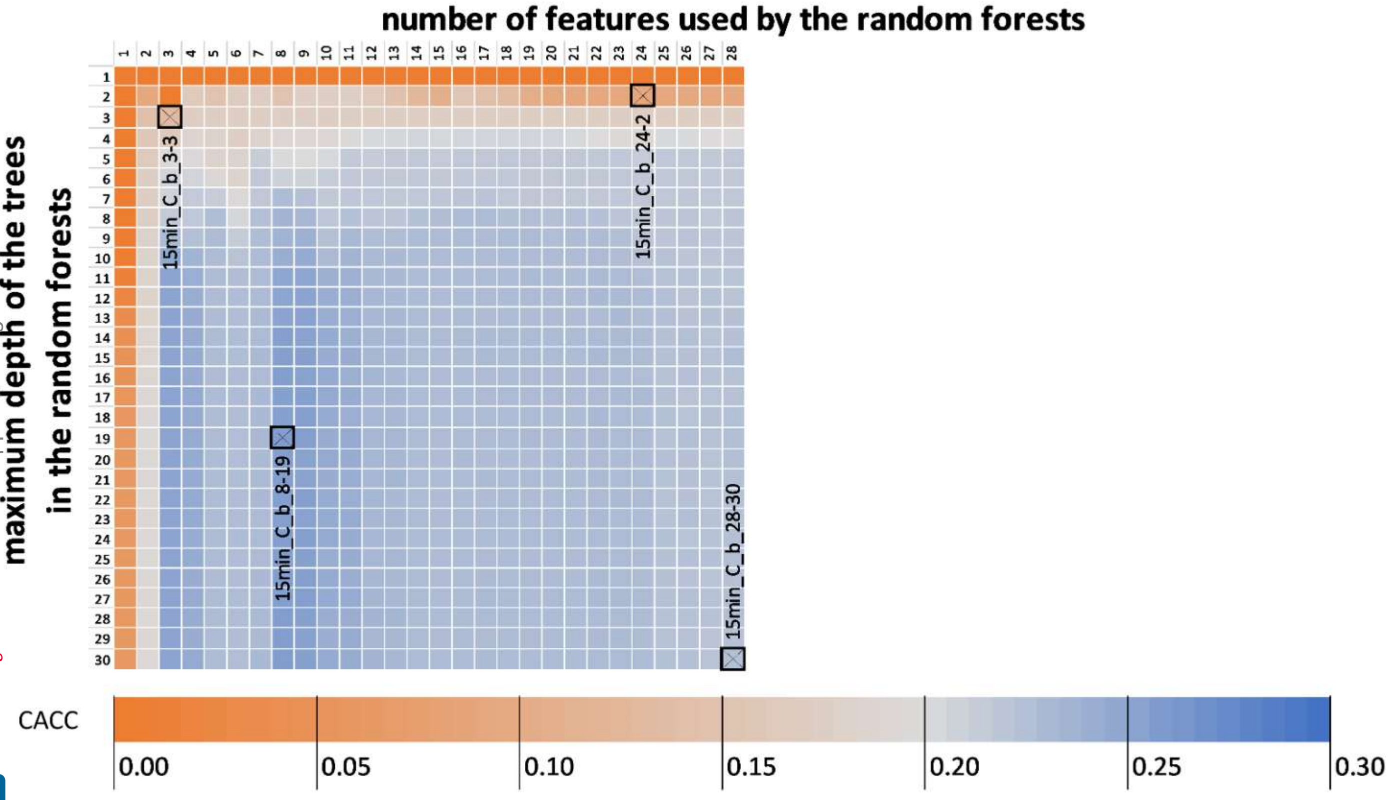
☒ random forest model selected to be one of the candidates for the best classifier

data resolution: **15 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **binary**



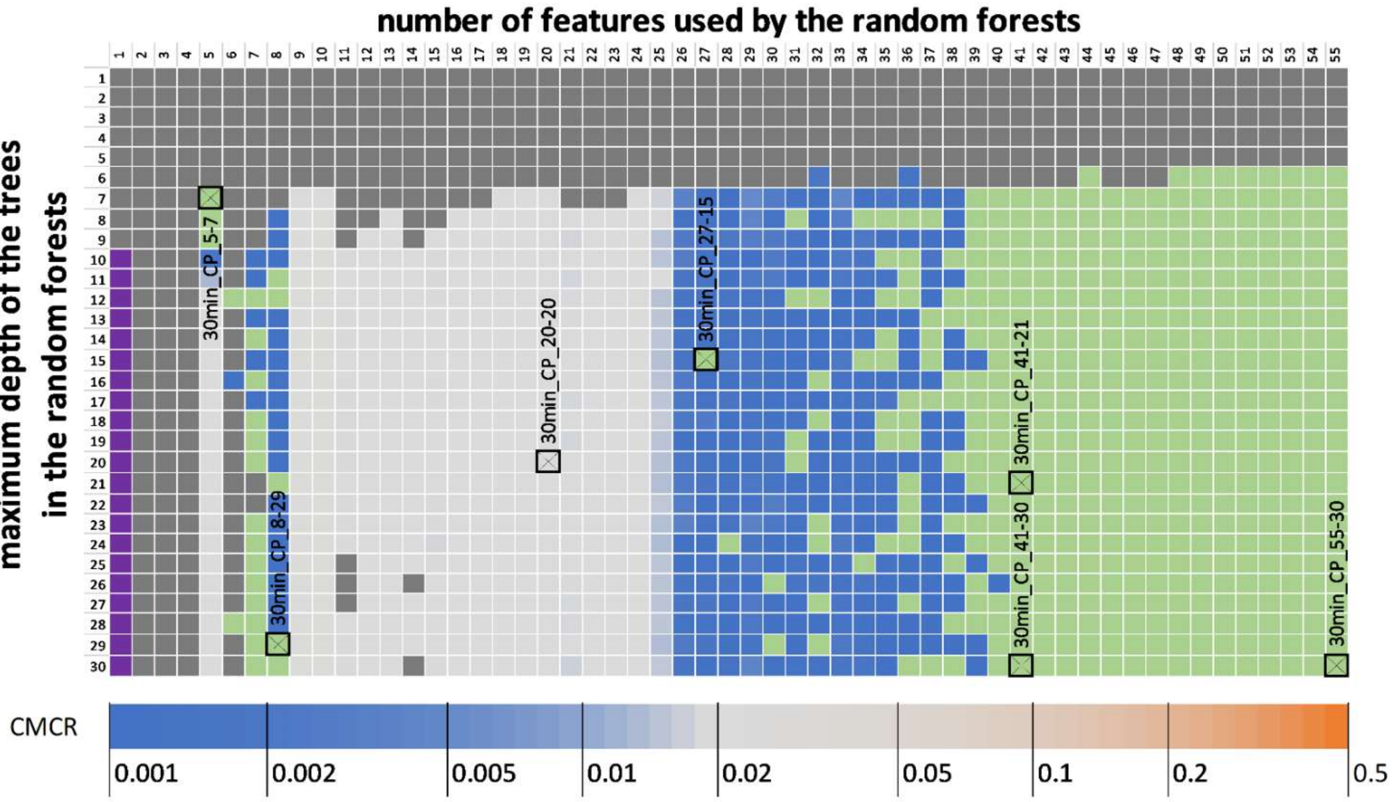
CMCR = 0 0.5 < CMCR < 1 CMCR = NaN (that means that at least one of the 50 forests had none of its predictions assigned to the cluster "relation")

☒ random forest model selected to be one of the candidates for the best classifier



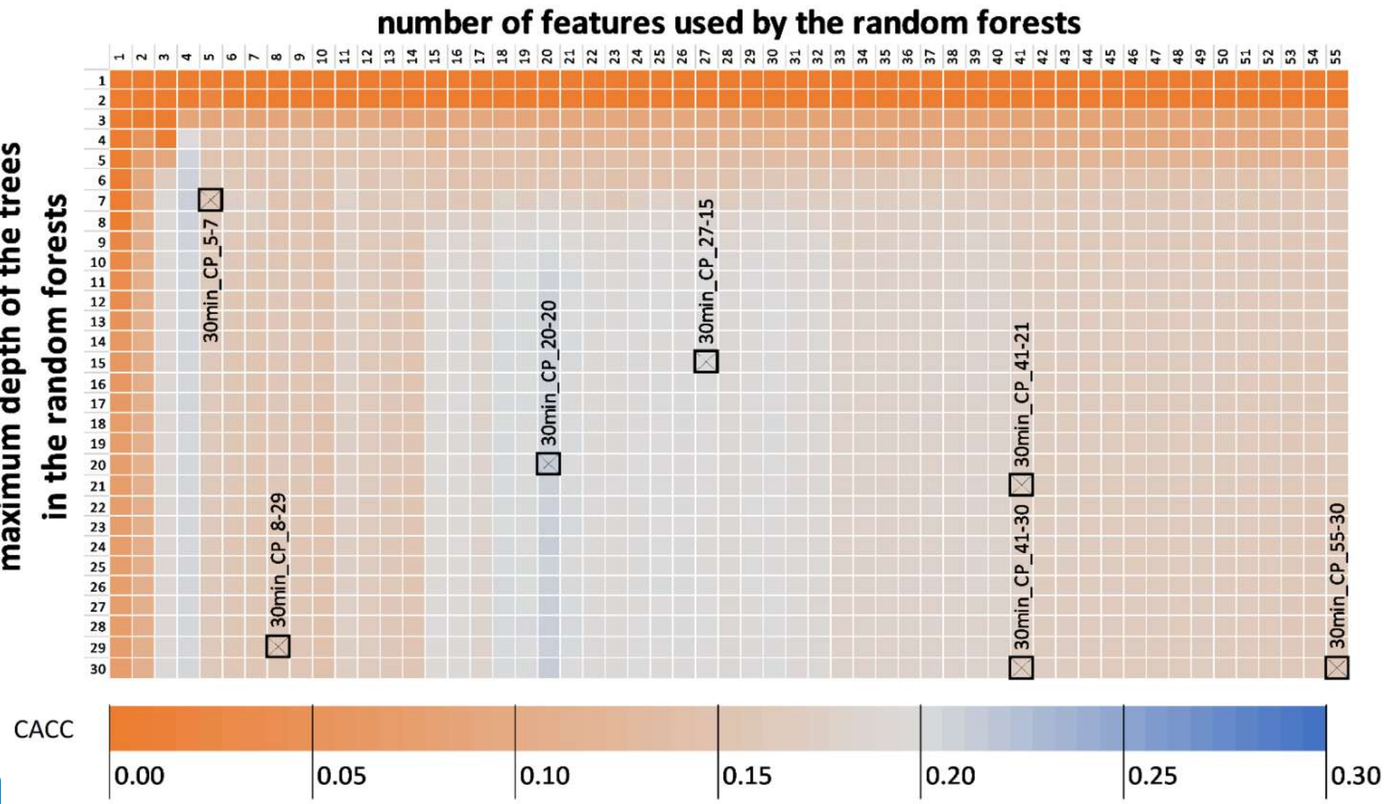
☒ random forest model selected to be one of the candidates for the best classifier

data resolution: **30 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **multiclass**



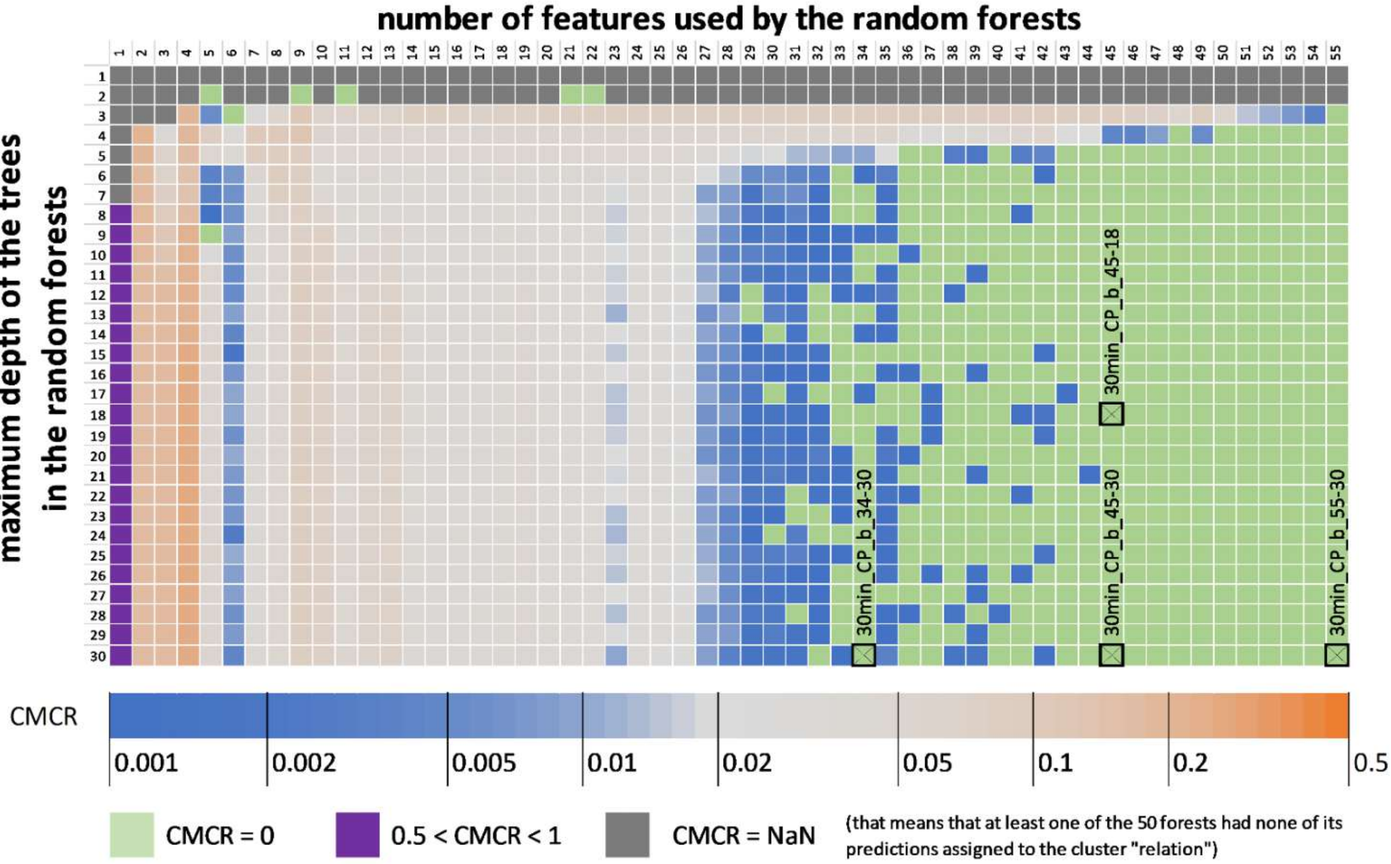
CMCR = 0 0.5 < CMCR < 1 CMCR = NaN (that means that at least one of the 50 forests had none of its predictions assigned to cluster "obvious relations" or none assigned to the clusters "rare partial supply" and "inconspicuous relations")

☒ random forest model selected to be one of the candidates for the best classifier

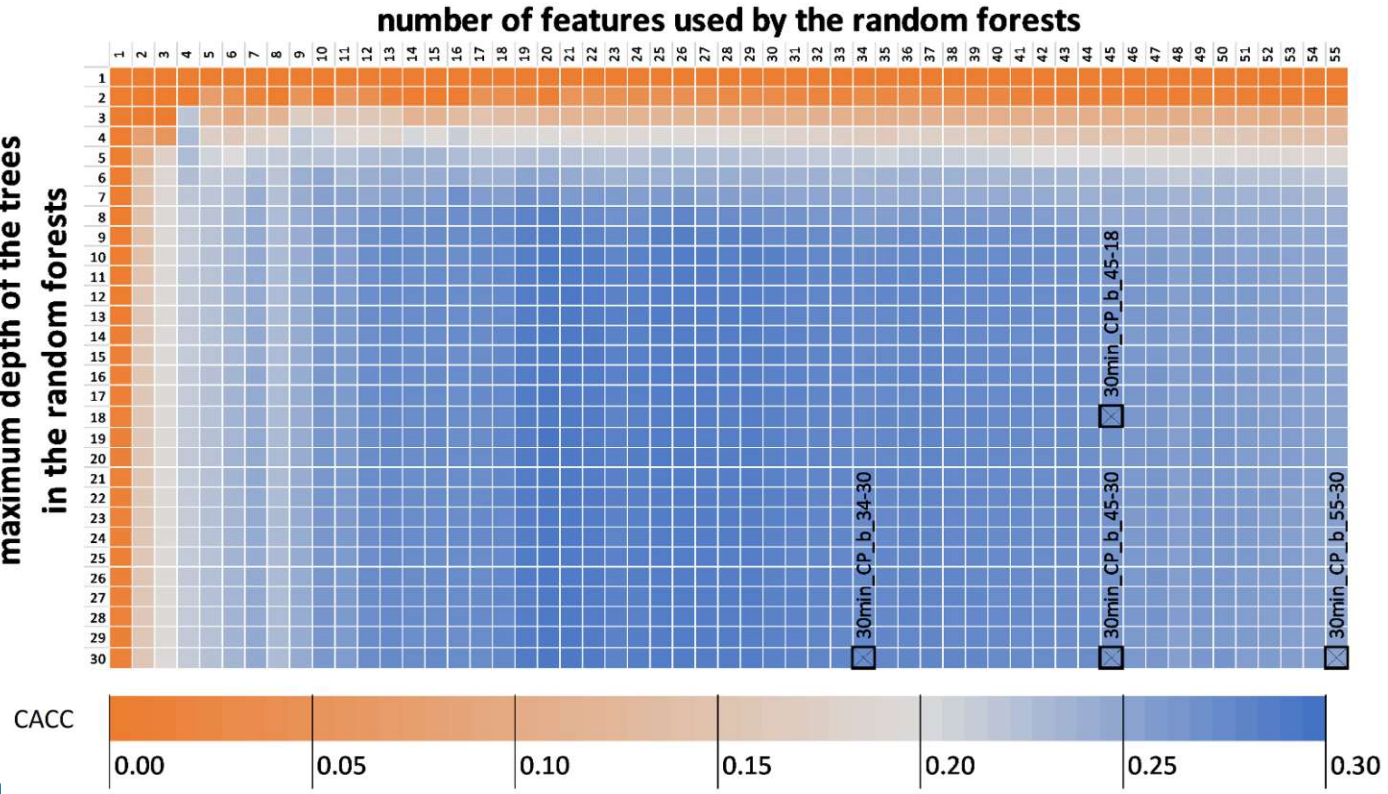


☒ random forest model selected to be one of the candidates for the best classifier

data resolution: **30 min** used data sources: **CP** (features derived from both counter values and
classifier type: **binary** instantaneous power values are used)

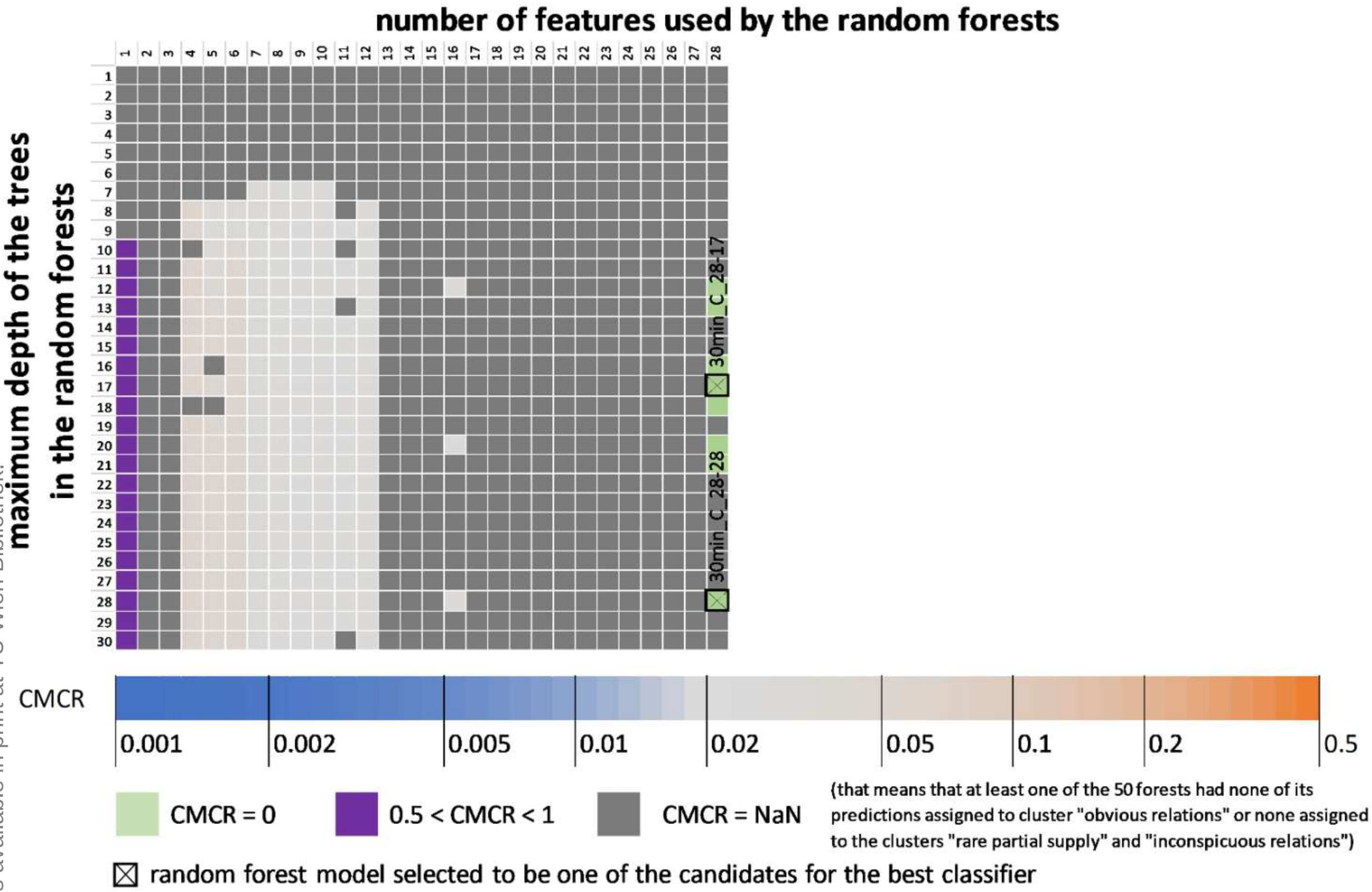


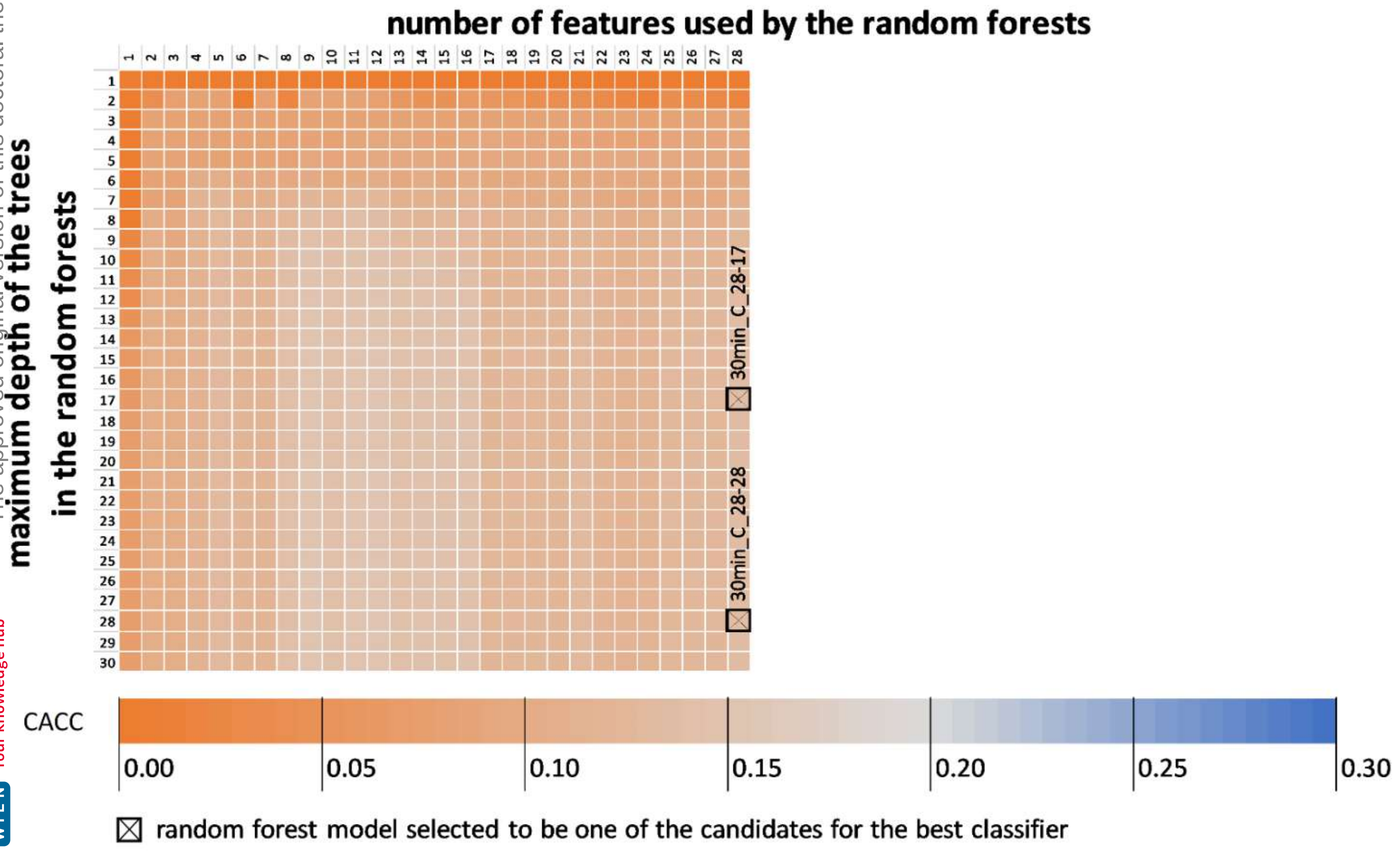
☒ random forest model selected to be one of the candidates for the best classifier



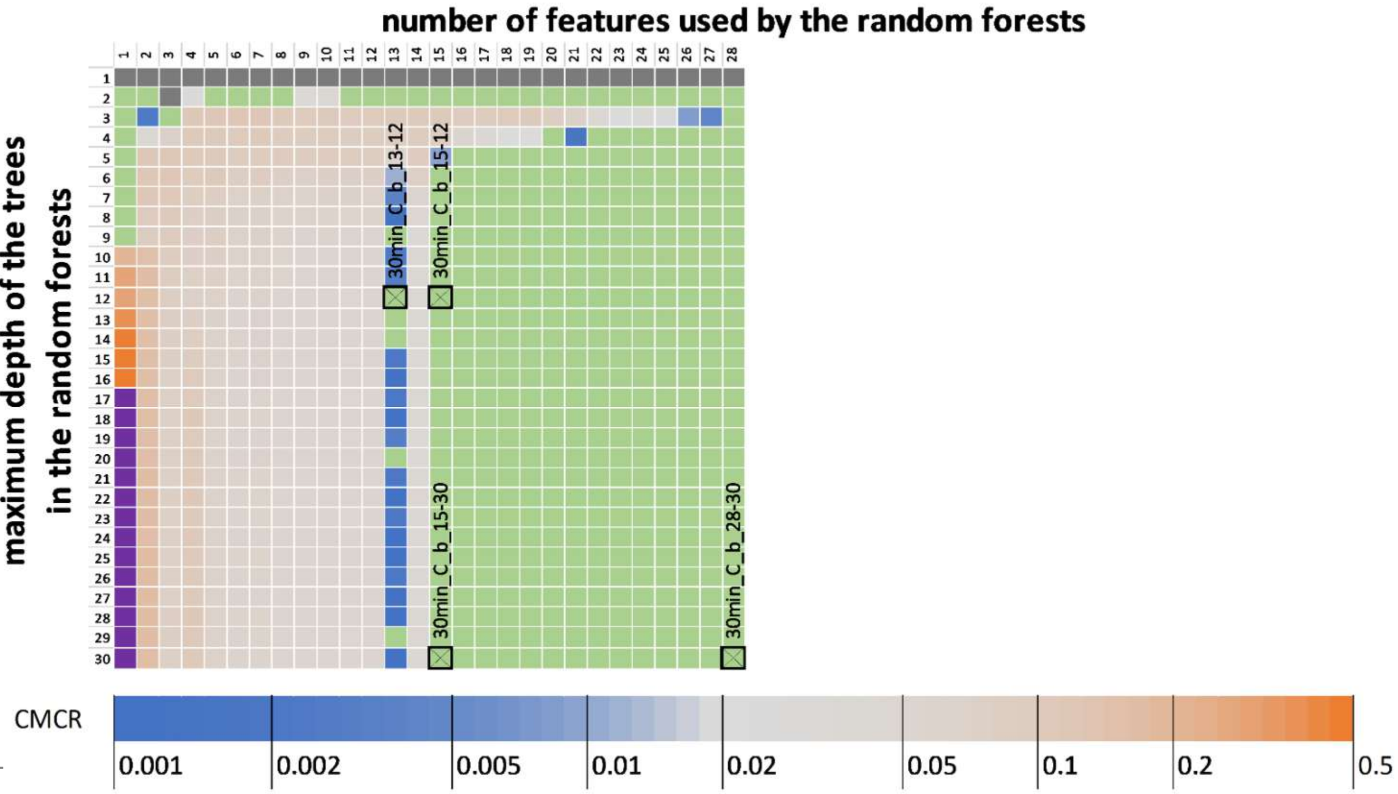
☒ random forest model selected to be one of the candidates for the best classifier

data resolution: **30 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **multiclass**



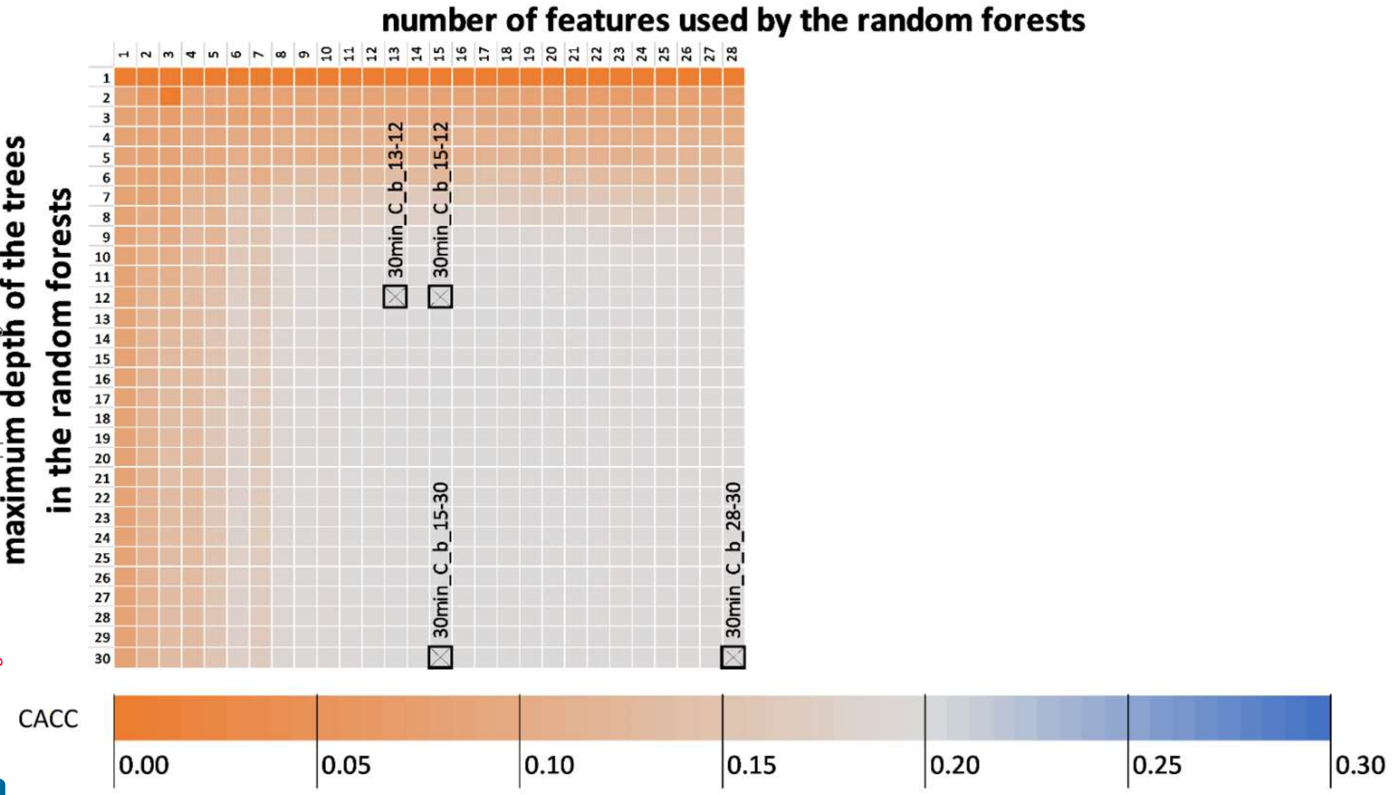


data resolution: **30 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **binary**



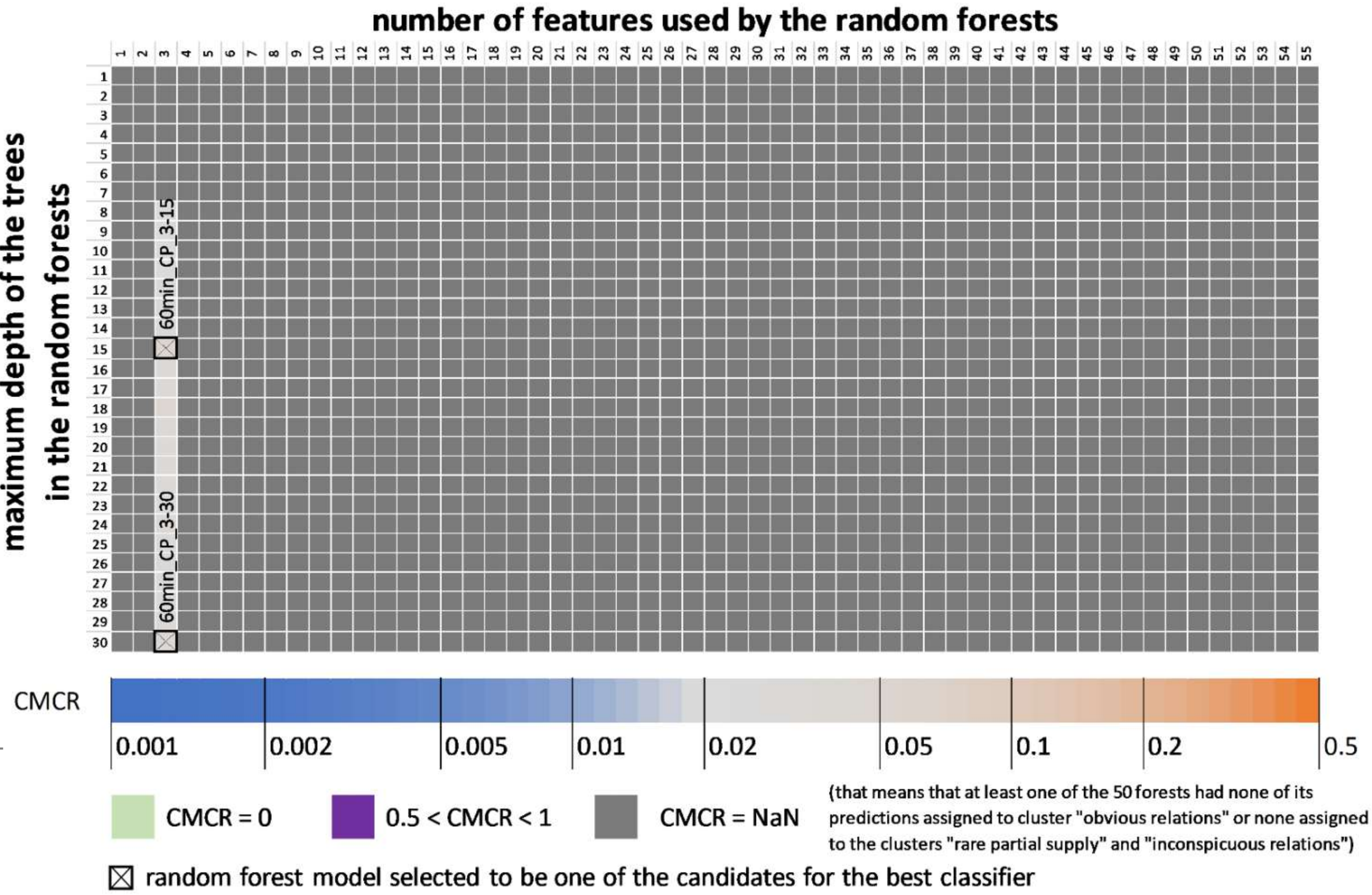
CMCR = 0 0.5 < CMCR < 1 CMCR = NaN (that means that at least one of the 50 forests had none of its predictions assigned to the cluster "relation")

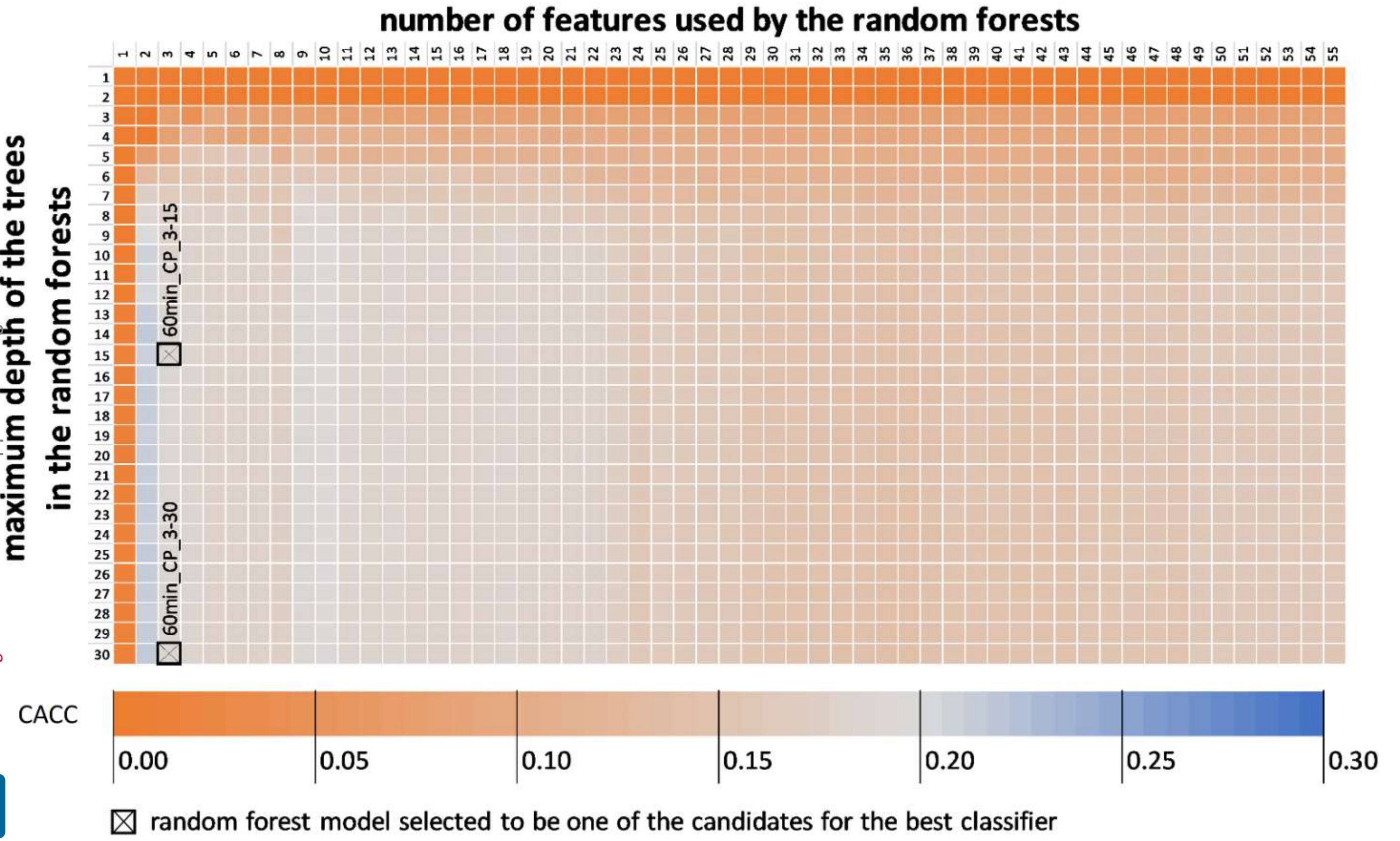
☒ random forest model selected to be one of the candidates for the best classifier



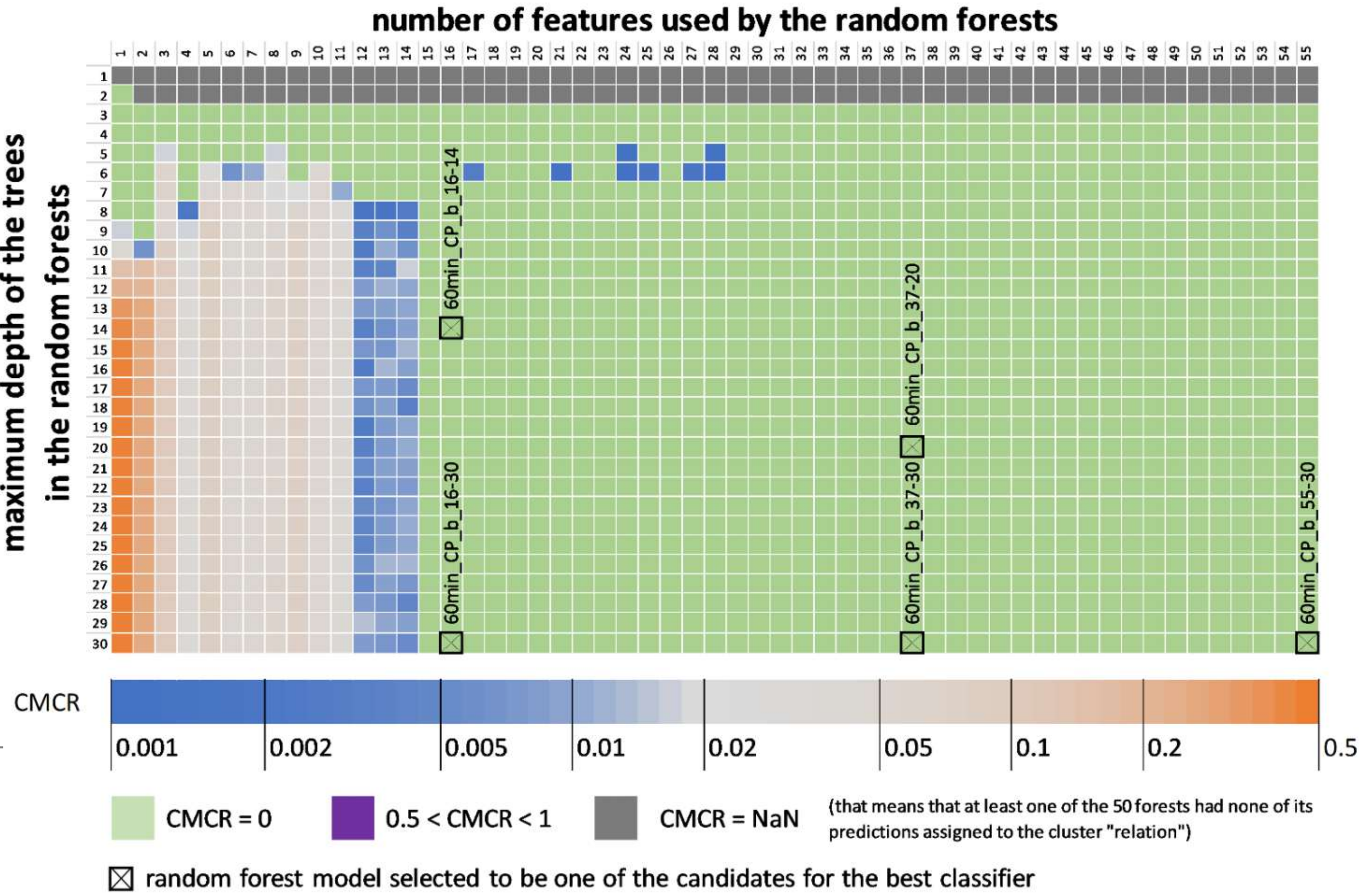
☒ random forest model selected to be one of the candidates for the best classifier

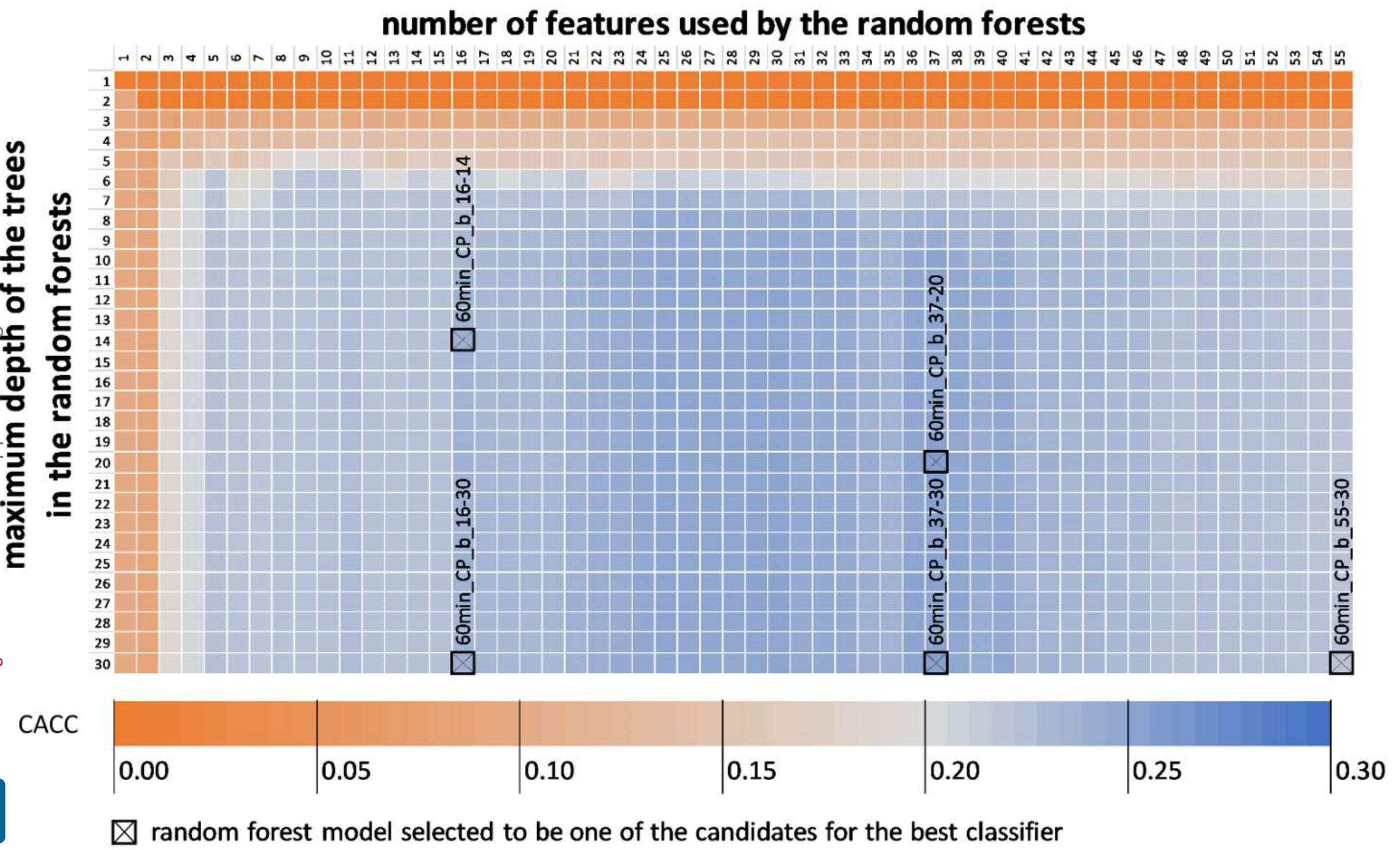
data resolution: **60 min** used data sources: **CP** (features derived from both counter values
classifier type: **multiclass** and instantaneous power values are used)



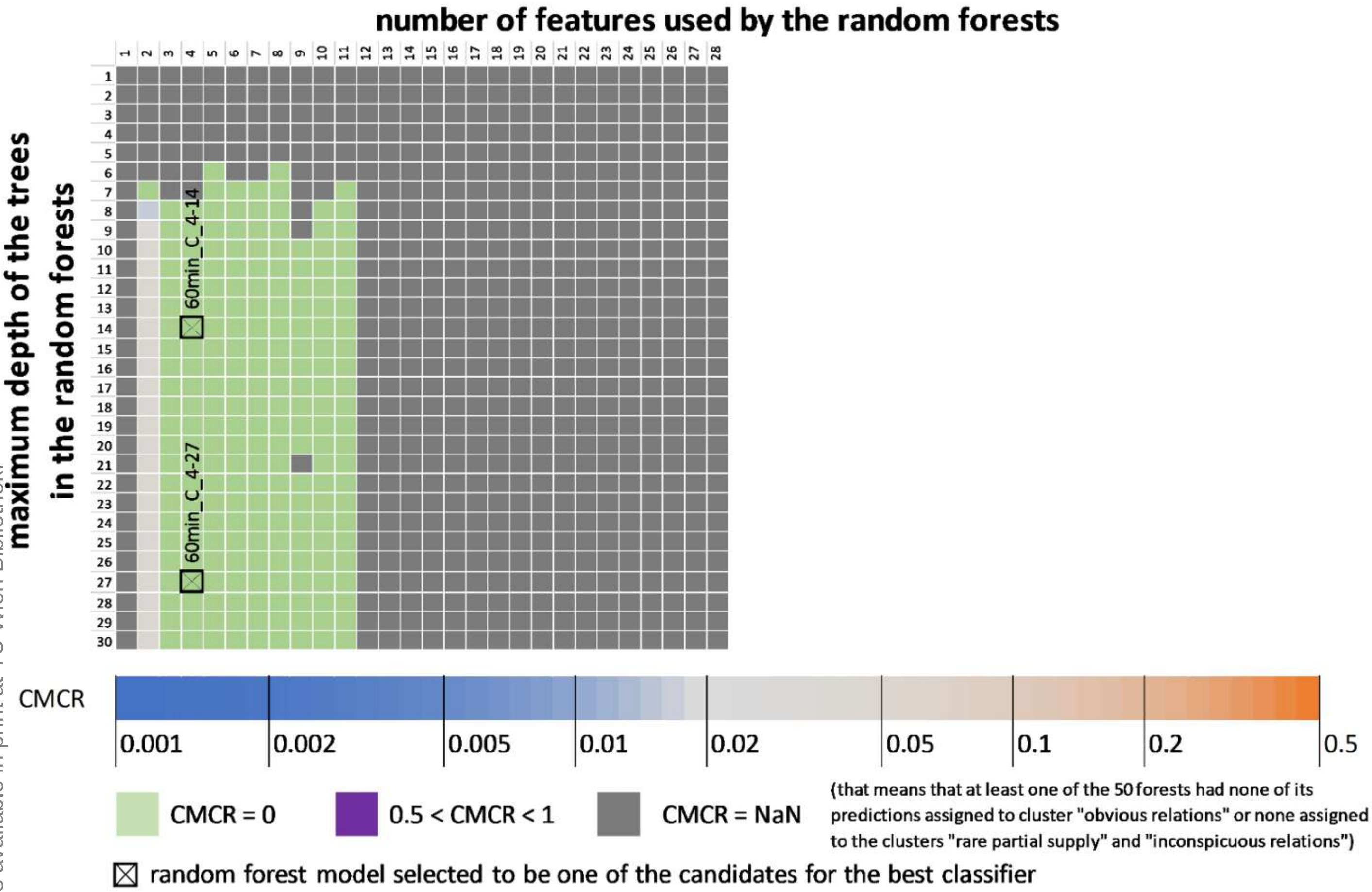


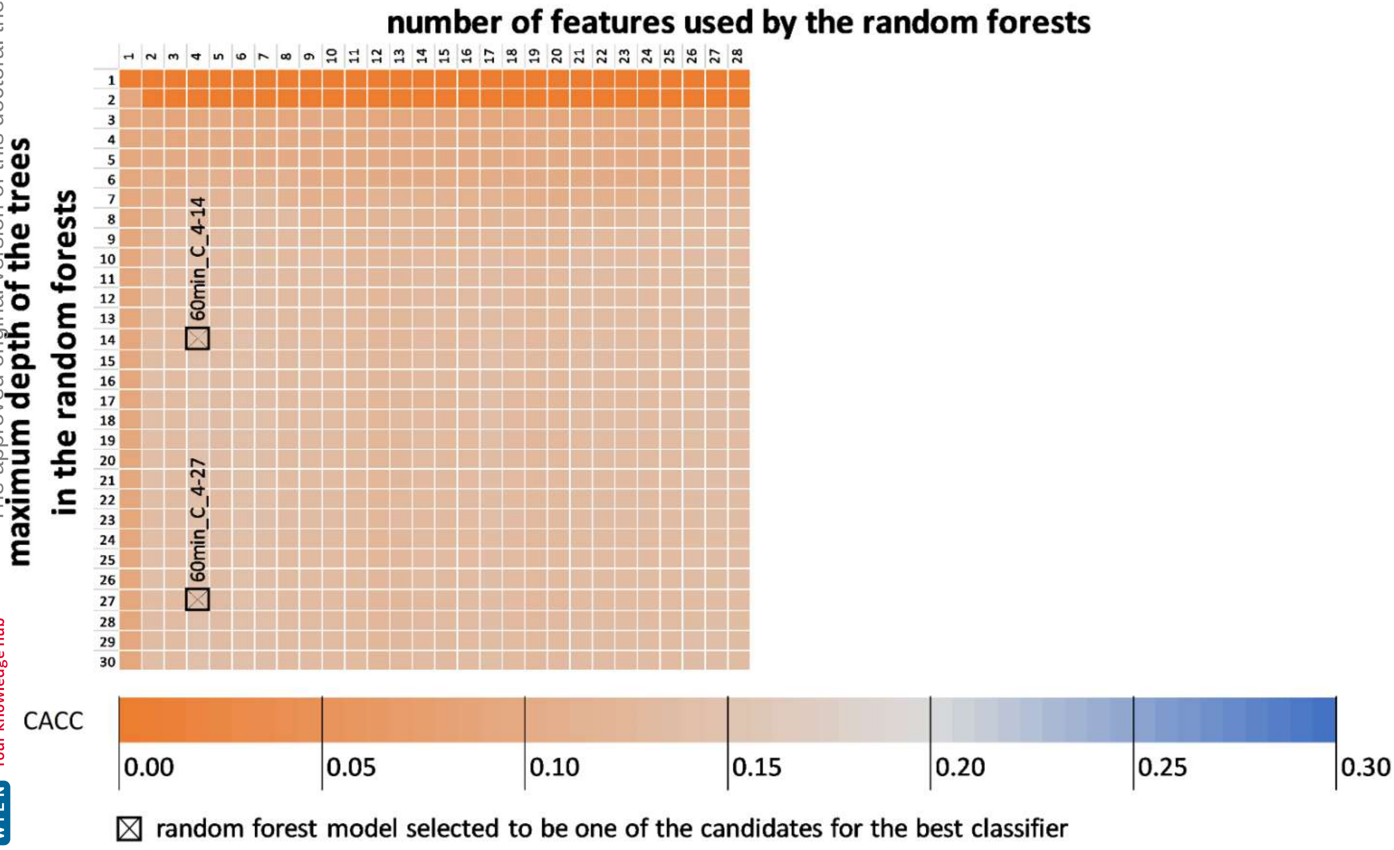
data resolution: **60 min** used data sources: **CP** (features derived from both counter values and instantaneous power values are used)
classifier type: **binary**



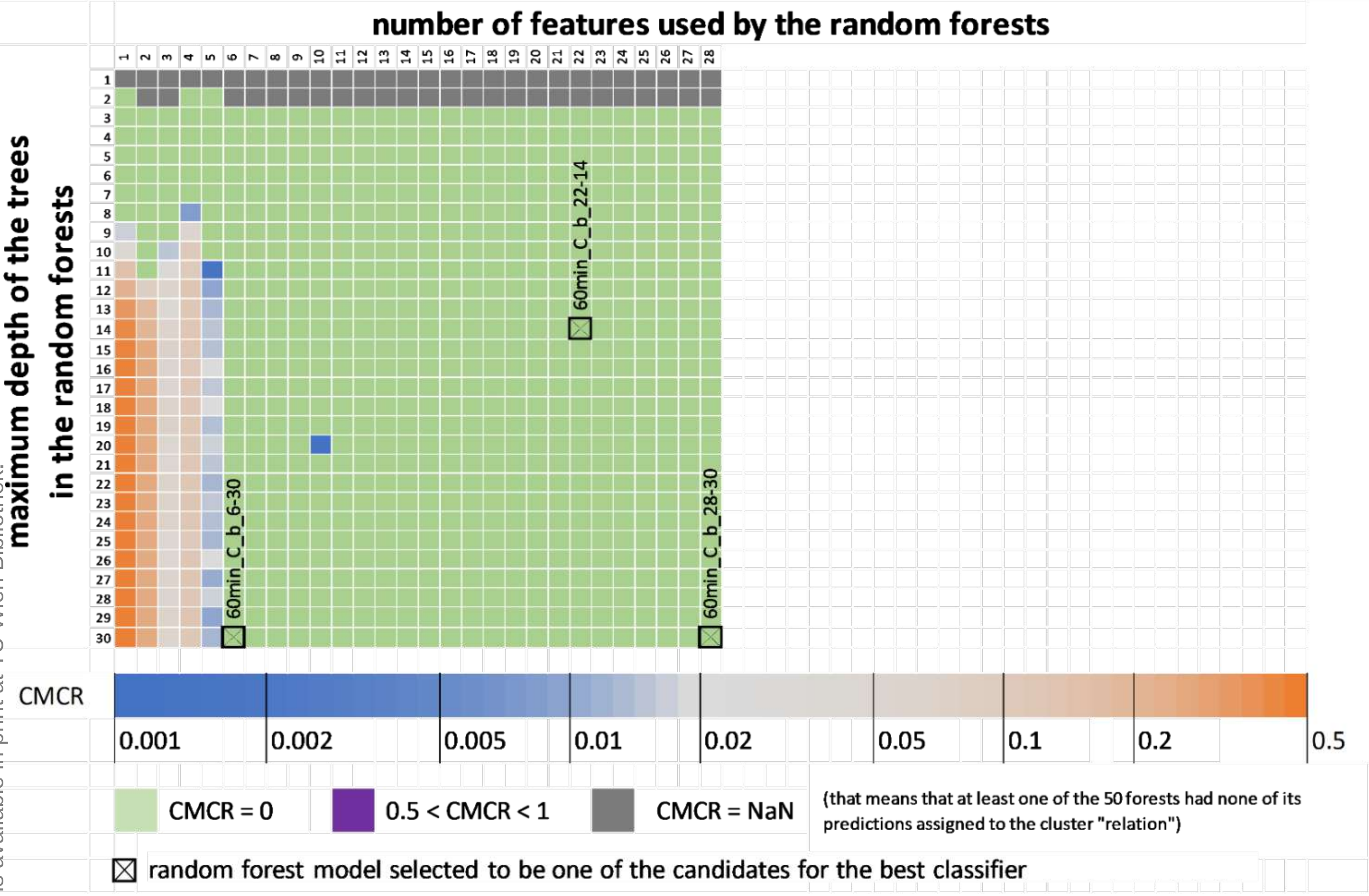


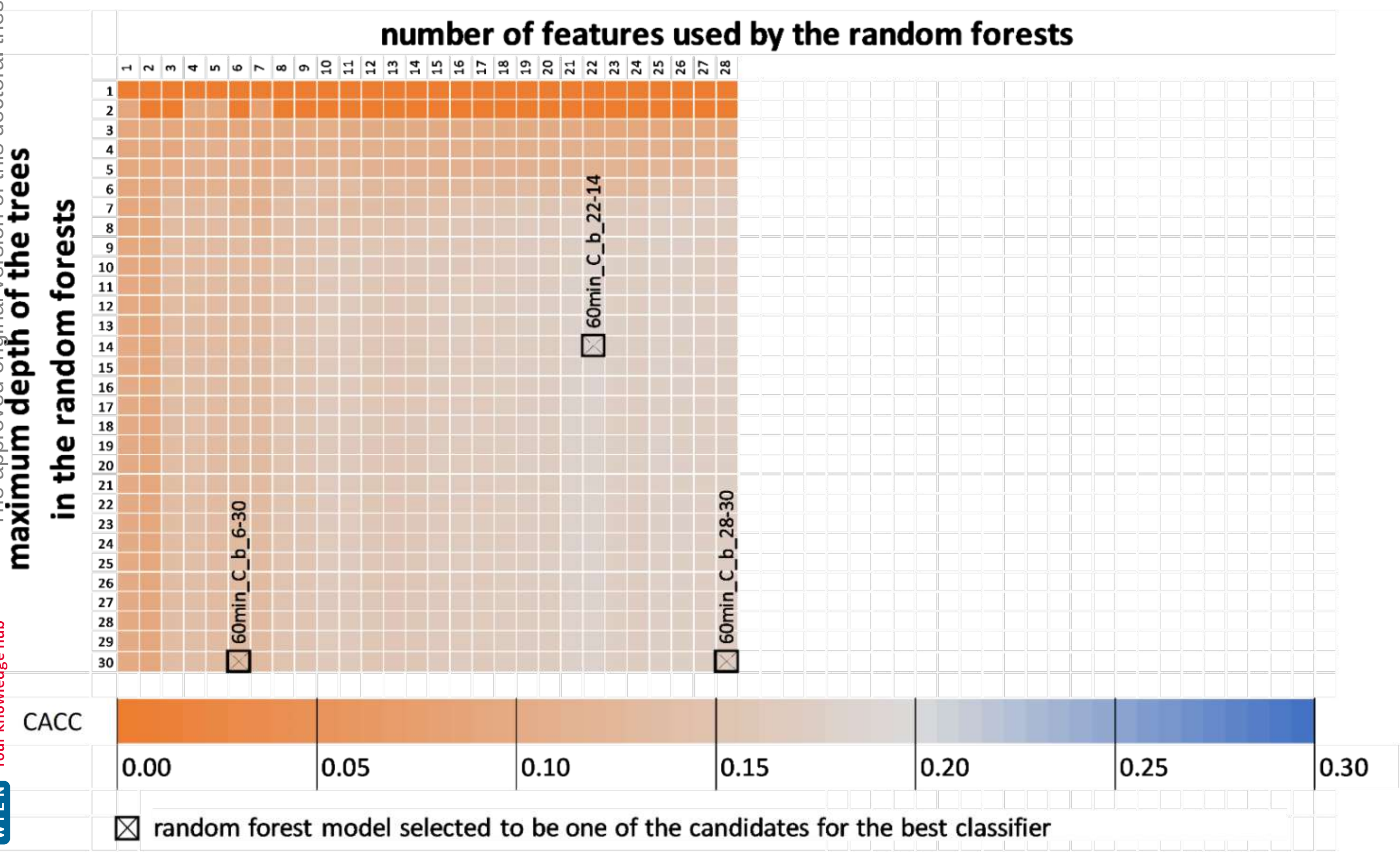
data resolution: **60 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **multiclass**





data resolution: **60 min** used data sources: **C** (only features derived from counter values are used)
classifier type: **binary**





12 Appendix D: Classification performance of the best classifier candidates

The tables in this appendix show the values for the *CMCR2C* and *CACC2* metrics that were achieved by the best classifier candidates when provided with data from different time slots or with different time resolutions. For the time slots with the same time slot length, their *CMCR2C* and *CACC2* metrics are averaged (e.g., the metrics for a time slot length of 5 weeks are the mean of the time slots “5 weeks winter”, “5 weeks spring”, “5 weeks summer”, and “5 weeks fall”). Cells without value indicate that at least one of the averaged *CMCR2C* values was “NaN”, which was caused by zero correctly identified main meter-submeter relations.

data: **case study data** time slot: **original training data**

		mean CMCR2C					mean CACC2				
time resolution		5 min	10 min	15 min	30 min	60 min	5 min	10 min	15 min	30 min	60 min
best classifier candidates	05min_CP_55-30	0.000	0.075	0.221	0.595	0.853	0.593	0.343	0.320	0.355	0.401
	05min_CP_18-30	0.000	0.082	0.295	0.711	0.877	0.628	0.297	0.302	0.349	0.430
	05min_CP_18-10	0.000	0.124	0.314	0.713	0.875	0.535	0.302	0.302	0.343	0.453
	05min_CP_9-30	0.000	0.145	0.328	0.689	0.873	0.622	0.308	0.326	0.390	0.471
	05min_CP_9-14	0.000	0.135	0.321	0.677	0.863	0.616	0.308	0.314	0.390	0.471
	05min_CP_7-30	0.000	0.098	0.252	0.684	0.858	0.616	0.279	0.326	0.378	0.465
	05min_CP_7-14	0.000	0.098	0.260	0.691	0.864	0.610	0.279	0.314	0.378	0.459
	05min_CP_b_55-30	0.000	0.123	0.319	0.681	0.873	0.610	0.372	0.372	0.430	0.483
	05min_CP_b_27-30	0.000	0.167	0.374	0.744	0.884	0.628	0.378	0.360	0.453	0.529
	05min_CP_b_27-13	0.000	0.167	0.384	0.741	0.886	0.622	0.378	0.355	0.448	0.523
	05min_CP_b_4-30	0.000	0.228	0.429	0.747	0.844	0.610	0.256	0.279	0.343	0.436
	05min_CP_b_4-13	0.000	0.279	0.442	0.751	0.846	0.599	0.256	0.279	0.337	0.442
	05min_C_28-30	0.000	0.057	0.323	0.768	0.894	0.593	0.331	0.349	0.395	0.483
	05min_C_7-30	0.000	0.112	0.256	0.751	0.891	0.616	0.279	0.326	0.395	0.471
	05min_C_7-13	0.000	0.112	0.243	0.745	0.893	0.587	0.279	0.331	0.395	0.459
	05min_C_5-30	0.000	0.122	0.308	0.767	0.888	0.622	0.256	0.302	0.331	0.395
	05min_C_5-14	0.000	0.122	0.311	0.775	0.892	0.622	0.256	0.302	0.320	0.401
	05min_C_b_28-30	0.000	0.197	0.318	0.786	0.885	0.593	0.331	0.349	0.424	0.512
	05min_C_b_9-30	0.000	0.192	0.479	0.819	0.908	0.610	0.343	0.360	0.453	0.529
	05min_C_b_9-15	0.000	0.194	0.475	0.822	0.909	0.610	0.337	0.360	0.453	0.529
	05min_C_b_7-30	0.000	0.183	0.380	0.797	0.895	0.622	0.337	0.360	0.442	0.523
	05min_C_b_7-17	0.000	0.183	0.376	0.795	0.892	0.628	0.337	0.366	0.442	0.523
	10min_CP_55-30	0.012	0.000	0.035	0.245	0.678	0.180	0.599	0.297	0.349	0.413
	10min_CP_40-30	0.111	0.000	0.033	0.283	0.661	0.174	0.599	0.314	0.343	0.424
	10min_CP_40-10	0.013	0.000	0.032	0.289	0.699	0.174	0.581	0.320	0.360	0.436
	10min_CP_17-30	0.043	0.000	0.228	0.595	0.755	0.174	0.593	0.355	0.419	0.459
	10min_CP_17-15	0.043	0.000	0.228	0.593	0.755	0.174	0.593	0.349	0.419	0.459
	10min_CP_12-30	0.299	0.000	0.159	0.612	0.805	0.180	0.605	0.343	0.384	0.465
	10min_CP_12-9	0.367	0.000	0.145	0.617	0.826	0.169	0.523	0.331	0.384	0.436
	10min_CP_b_55-30	0.030	0.000	0.131	0.492	0.769	0.186	0.599	0.308	0.378	0.465
	10min_CP_b_53-8	0.030	0.000	0.121	0.475	0.761	0.186	0.453	0.297	0.366	0.459
	10min_CP_b_26-5	0.063	0.000	0.123	0.459	0.772	0.174	0.256	0.291	0.343	0.453
	10min_CP_b_2-11	0.143	0.000	0.386	0.723	0.900	0.140	0.512	0.250	0.227	0.262
	10min_C_28-30	0.023	0.004	0.052	0.452	0.904	0.180	0.599	0.279	0.331	0.413
	10min_C_7-22	0.165	0.004	0.194	0.555	0.841	0.192	0.599	0.308	0.355	0.430
	10min_C_5-23	0.254	0.004	0.200	0.647	0.852	0.174	0.593	0.302	0.343	0.436
	10min_C_3-30	0.051	0.000	0.147	0.481	0.845	0.076	0.552	0.192	0.145	0.180
	10min_C_3-15	0.051	0.000	0.163	0.486	0.814	0.076	0.552	0.186	0.145	0.180
	10min_C_b_28-30	0.086	0.010	0.136	0.422	0.717	0.186	0.593	0.297	0.343	0.436
	10min_C_b_8-16	0.108	0.009	0.182	0.523	0.762	0.192	0.622	0.314	0.360	0.448
	10min_C_b_4-22	0.050	0.000	0.319	0.667	0.832	0.110	0.564	0.186	0.192	0.267
	10min_C_b_3-30	0.059	0.000	0.163	0.522	0.808	0.093	0.558	0.209	0.192	0.198

data: **case study data**		time slot: **original training data**								
	mean CMCR2C					**mean CACC2**				
time resolution	**5 min**	**10 min**	**15 min**	**30 min**	**60 min**	**5 min**	**10 min**	**15 min**	**30 min**	**60 min**
best classifier candidates										
15min_CP_55-30	0.000	0.000	0.004	0.169	0.693	0.163	0.169	0.593	0.308	0.349
15min_CP_53-5	0.000	0.000	0.000	0.043	0.590	0.151	0.163	0.262	0.297	0.314
15min_CP_45-5	0.000	0.000	0.000	0.038	0.587	0.151	0.163	0.256	0.297	0.320
15min_CP_25-13	0.000	0.000	0.004	0.174	0.667	0.163	0.169	0.599	0.331	0.372
15min_CP_25-4	0.000	0.000	0.000	0.050	0.869	0.151	0.163	0.209	0.273	0.302
15min_CP_13-21	0.000	0.000	0.004	0.321	0.821	0.169	0.186	0.581	0.308	0.337
15min_CP_8-8	0.000	0.000	0.000	0.233	0.719	0.140	0.169	0.349	0.227	0.198
15min_CP_b_55-30	0.000	0.000	0.010	0.263	0.703	0.169	0.180	0.593	0.326	0.407
15min_CP_b_18-4	0.033	0.000	0.045	0.238	0.672	0.169	0.163	0.244	0.279	0.349
15min_CP_b_7-10	0.069	0.061	0.000	0.203	0.636	0.157	0.180	0.483	0.273	0.250
15min_CP_b_2-30	0.143	0.156	0.000	0.391	0.837	0.105	0.157	0.581	0.227	0.192
15min_C_28-30	0.000	0.000	0.007	0.200	0.808	0.157	0.180	0.587	0.326	0.360
15min_C_3-22	0.000	0.000	0.004	0.343	0.861	0.145	0.186	0.587	0.273	0.343
15min_C_b_28-30	0.000	0.031	0.019	0.296	0.738	0.169	0.180	0.593	0.331	0.384
15min_C_b_24-2	0.000	0.000	0.000	0.048	0.333	0.029	0.105	0.105	0.116	0.140
15min_C_b_8-19	0.063	0.000	0.028	0.378	0.677	0.174	0.186	0.610	0.326	0.378
15min_C_b_3-3	0.000	0.000	0.000	0.163	0.441	0.116	0.134	0.157	0.209	0.221
30min_CP_55-30	0.000	0.000	0.000	0.000	0.368	0.081	0.140	0.157	0.570	0.291
30min_CP_41-30	0.000	0.000	0.000	0.000	0.419	0.099	0.134	0.157	0.587	0.285
30min_CP_27-15	0.000	0.000	0.000	0.000	0.517	0.099	0.128	0.145	0.605	0.302
30min_CP_20-20	0.000	0.042	0.013	0.004	0.603	0.116	0.134	0.151	0.610	0.320
30min_CP_8-29	0.000	0.042	0.013	0.000	0.758	0.081	0.134	0.151	0.587	0.308
30min_CP_5-7	0.000	0.000	0.000	0.000	0.294	0.099	0.128	0.145	0.320	0.279
30min_CP_b_55-30	0.000	0.000	0.000	0.000	0.396	0.099	0.151	0.163	0.616	0.320
30min_CP_b_45-30	0.000	0.000	0.000	0.000	0.417	0.099	0.151	0.157	0.622	0.326
30min_CP_b_45-18	0.000	0.000	0.000	0.000	0.417	0.099	0.151	0.157	0.622	0.326
30min_CP_b_34-30	0.000	0.000	0.000	0.000	0.438	0.099	0.145	0.157	0.628	0.314
30min_C_28-28	0.000	0.000	0.000	0.000	0.757	0.058	0.122	0.151	0.564	0.267
30min_C_28-17	0.000	0.000	0.000	0.000	0.757	0.058	0.122	0.151	0.564	0.267
30min_C_b_28-30	0.000	0.000	0.000	0.000	0.383	0.070	0.134	0.157	0.587	0.291
30min_C_b_15-30	0.000	0.000	0.000	0.000	0.344	0.041	0.128	0.157	0.587	0.244
30min_C_b_15-12	0.000	0.000	0.036	0.000	0.364	0.047	0.134	0.157	0.570	0.244
30min_C_b_13-12	0.000	0.000	0.069	0.000	0.384	0.035	0.134	0.157	0.576	0.262
60min_CP_3-30	0.000	0.000	0.000	0.000	0.000	0.023	0.081	0.081	0.099	0.581
60min_CP_3-15	0.000	0.000	0.000	0.000	0.000	0.023	0.081	0.081	0.099	0.576
60min_CP_b_55-30		0.000	0.000	0.000	0.000	0.000	0.093	0.116	0.122	0.599
60min_CP_b_37-30		0.000	0.000	0.000	0.000	0.000	0.087	0.122	0.122	0.605
60min_CP_b_37-20		0.000	0.000	0.000	0.000	0.000	0.087	0.122	0.122	0.605
60min_CP_b_16-30		0.000	0.000	0.000	0.000	0.000	0.070	0.128	0.134	0.610
60min_CP_b_16-14		0.000	0.000	0.000	0.000	0.000	0.076	0.122	0.134	0.610
60min_C_4-27	0.000	0.000	0.000	0.000	0.000	0.052	0.076	0.099	0.116	0.558
60min_C_4-14	0.000	0.067	0.000	0.000	0.000	0.052	0.081	0.099	0.116	0.558
60min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.047	0.099	0.110	0.134	0.570
60min_C_b_22-14	0.000	0.000	0.000	0.000	0.000	0.047	0.099	0.110	0.134	0.581
60min_C_b_6-30	0.000	0.000	0.000	0.000	0.000	0.052	0.064	0.110	0.110	0.552

data: **case study data**		time resolution: **5 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**05min_CP_55-30**	0.000	0.004	0.010	0.011	0.005	0.007	0.308	0.302	0.289	0.283	0.251	0.234
	05min_CP_18-30	0.000	0.000	0.027	0.029	0.028	0.058	0.262	0.265	0.260	0.257	0.243	0.225
	05min_CP_18-10	0.009	0.005	0.026	0.026	0.027	0.054	0.267	0.265	0.257	0.257	0.244	0.224
	05min_CP_9-30	0.000	0.013	0.015	0.024	0.047	0.113	0.262	0.256	0.250	0.240	0.241	0.219
	05min_CP_9-14	0.000	0.017	0.015	0.021	0.048	0.106	0.256	0.259	0.249	0.240	0.237	0.221
	05min_CP_7-30	0.000	0.013	0.015	0.020	0.045	0.086	0.267	0.267	0.247	0.247	0.240	0.230
	05min_CP_7-14	0.000	0.013	0.015	0.022	0.050	0.093	0.267	0.267	0.249	0.249	0.237	0.227
	05min_CP_b_55-30	0.000	0.009	0.023	0.029	0.015	0.029	0.314	0.326	0.310	0.307	0.286	0.269
	05min_CP_b_27-30	0.000	0.009	0.031	0.065	0.050	0.076	0.308	0.328	0.323	0.313	0.291	0.279
	05min_CP_b_27-13	0.000	0.009	0.031	0.070	0.046	0.087	0.308	0.326	0.320	0.311	0.291	0.275
	05min_CP_b_4-30	0.021	0.031	0.049	0.086	0.081	0.191	0.267	0.267	0.253	0.249	0.257	0.227
	05min_CP_b_4-13	0.021	0.032	0.044	0.086	0.080	0.205	0.267	0.262	0.253	0.249	0.257	0.225
	05min_C_28-30	0.000	0.033	0.005	0.027	0.019	0.057	0.267	0.259	0.243	0.249	0.238	0.230
	05min_C_7-30	0.000	0.000	0.003	0.022	0.042	0.113	0.250	0.238	0.231	0.222	0.222	0.215
	05min_C_7-13	0.000	0.000	0.007	0.022	0.042	0.106	0.256	0.238	0.234	0.224	0.227	0.215
	05min_C_5-30	0.000	0.000	0.003	0.012	0.045	0.106	0.262	0.244	0.238	0.233	0.225	0.215
	05min_C_5-14	0.000	0.000	0.003	0.012	0.045	0.104	0.262	0.244	0.237	0.235	0.224	0.218
	05min_C_b_28-30	0.021	0.012	0.018	0.032	0.053	0.160	0.273	0.265	0.256	0.263	0.246	0.243
	05min_C_b_9-30	0.000	0.000	0.035	0.060	0.088	0.203	0.267	0.247	0.247	0.253	0.249	0.237
	05min_C_b_9-15	0.000	0.000	0.040	0.060	0.093	0.211	0.267	0.247	0.247	0.254	0.249	0.237
	05min_C_b_7-30	0.000	0.000	0.024	0.049	0.097	0.192	0.262	0.247	0.246	0.257	0.251	0.237
	05min_C_b_7-17	0.000	0.000	0.024	0.049	0.092	0.192	0.262	0.244	0.246	0.257	0.251	0.237
	10min_CP_55-30	0.000	0.000	0.033	0.015	0.019	0.024	0.081	0.110	0.132	0.158	0.170	0.156
	10min_CP_40-30	0.000	0.000	0.042	0.079	0.027	0.032	0.116	0.128	0.135	0.161	0.167	0.157
	10min_CP_40-10	0.000	0.000	0.033	0.046	0.027	0.032	0.110	0.122	0.134	0.160	0.166	0.157
	10min_CP_17-30	0.000	0.000	0.045	0.019	0.042	0.050	0.169	0.174	0.164	0.172	0.182	0.177
	10min_CP_17-15	0.000	0.000	0.045	0.019	0.042	0.050	0.169	0.169	0.164	0.172	0.182	0.177
	10min_CP_12-30	0.012	0.006	0.009	0.057	0.062	0.191	0.169	0.172	0.166	0.163	0.173	0.172
	10min_CP_12-9	0.012	0.006	0.009	0.057	0.063	0.166	0.169	0.169	0.166	0.161	0.170	0.167
	10min_CP_b_55-30	0.000	0.000	0.008	0.036	0.060	0.074	0.081	0.105	0.145	0.174	0.182	0.174
	10min_CP_b_53-8	0.000	0.000	0.008	0.037	0.068	0.075	0.076	0.096	0.138	0.166	0.177	0.169
	10min_CP_b_26-5	0.000	0.000	0.017	0.063	0.076	0.087	0.169	0.172	0.163	0.167	0.174	0.163
	10min_CP_b_2-11	0.067	0.301	0.282	0.225	0.328	0.333	0.081	0.084	0.090	0.121	0.096	0.105
	10min_C_28-30	0.000	0.000	0.016	0.018	0.023	0.031	0.163	0.148	0.156	0.166	0.173	0.169
	10min_C_7-22	0.000	0.000	0.085	0.162	0.344	0.478	0.174	0.172	0.169	0.173	0.173	0.167
	10min_C_5-23	0.000	0.000	0.085	0.156	0.350	0.417	0.163	0.163	0.164	0.167	0.172	0.161
	10min_C_3-30	0.000	0.000	0.163	0.227	0.235	0.181	0.035	0.023	0.038	0.049	0.054	0.044
	10min_C_3-15	0.000	0.000	0.087	0.220	0.239	0.171	0.035	0.023	0.039	0.049	0.052	0.045
	10min_C_b_28-30	0.000	0.000	0.048	0.067	0.095	0.104	0.128	0.154	0.164	0.179	0.185	0.180
	10min_C_b_8-16	0.000	0.000	0.061	0.094	0.187	0.207	0.174	0.183	0.179	0.190	0.190	0.179
	10min_C_b_4-22	0.250	0.258	0.354	0.358	0.431	0.406	0.035	0.044	0.077	0.086	0.106	0.100
	10min_C_b_3-30	0.000	0.056	0.354	0.308	0.350	0.377	0.035	0.041	0.051	0.073	0.080	0.065

data: **case study data** time resolution: **5 min**

	time slot	mean CMCR2C full_year	6_months	3_months	5_weeks	2_weeks	1_week	mean CACC2 full_year	6_months	3_months	5_weeks	2_weeks	1_week
best classifier candidates	15min_CP_55-30	0.000	0.000	0.000	0.000	0.004	0.004	0.134	0.125	0.118	0.132	0.138	0.138
	15min_CP_53-5	0.000	0.000	0.000	0.000	0.004	0.004	0.064	0.064	0.084	0.112	0.125	0.118
	15min_CP_45-5	0.000	0.000	0.000	0.000	0.005	0.004	0.093	0.087	0.093	0.116	0.121	0.115
	15min_CP_25-13	0.000	0.000	0.000	0.010	0.012	0.008	0.163	0.128	0.119	0.129	0.138	0.140
	15min_CP_25-4	0.000	0.000	0.000	0.000	0.008	0.016	0.145	0.110	0.109	0.116	0.121	0.113
	15min_CP_13-21	0.000	0.000	0.000	0.007	0.011	0.025	0.163	0.148	0.144	0.145	0.150	0.141
	15min_CP_8-8	0.000	0.021	0.019	0.005	0.106	0.013	0.047	0.052	0.068	0.086	0.090	0.092
	15min_CP_b_55-30	0.000	0.000	0.000	0.000	0.011	0.028	0.093	0.110	0.118	0.140	0.142	0.145
	15min_CP_b_18-4	0.000	0.000	0.009	0.019	0.047	0.062	0.163	0.163	0.157	0.150	0.148	0.145
	15min_CP_b_7-10	0.059	0.125	0.151	0.104	0.121	0.128	0.093	0.081	0.089	0.099	0.121	0.119
	15min_CP_b_2-30	0.167	0.204	0.195	0.197	0.203	0.229	0.058	0.073	0.081	0.073	0.065	0.070
	15min_C_28-30	0.000	0.000	0.000	0.005	0.007	0.011	0.163	0.134	0.118	0.141	0.148	0.145
	15min_C_3-22	0.000	0.000	0.010	0.021	0.049	0.084	0.151	0.145	0.141	0.141	0.137	0.124
	15min_C_b_28-30	0.000	0.000	0.010	0.011	0.028	0.028	0.163	0.140	0.135	0.147	0.156	0.150
	15min_C_b_24-2			0.000	0.000	0.167	0.000	0.000	0.000	0.004	0.013	0.007	0.016
	15min_C_b_8-19	0.000	0.000	0.034	0.066	0.106	0.107	0.180	0.163	0.164	0.166	0.161	0.156
	15min_C_b_3-3	0.000	0.000	0.000	0.000	0.014	0.000	0.116	0.116	0.106	0.099	0.102	0.092
	30min_CP_55-30	0.000	0.000	0.000	0.000	0.025	0.000	0.006	0.012	0.025	0.042	0.061	0.058
	30min_CP_41-30	0.000	0.000	0.000	0.000	0.021	0.000	0.012	0.020	0.033	0.061	0.070	0.068
	30min_CP_27-15	0.000	0.000	0.000	0.000	0.021	0.000	0.035	0.035	0.057	0.070	0.070	0.070
	30min_CP_20-20	0.000	0.000	0.000	0.016	0.185	0.031	0.081	0.067	0.076	0.081	0.083	0.089
	30min_CP_8-29	0.000	0.000	0.000	0.019	0.046	0.044	0.099	0.093	0.090	0.086	0.094	0.086
	30min_CP_5-7	0.000	0.000	0.000	0.000	0.026	0.015	0.105	0.105	0.097	0.084	0.102	0.083
	30min_CP_b_55-30	0.000	0.000	0.000	0.000	0.019	0.000	0.017	0.023	0.039	0.064	0.073	0.070
	30min_CP_b_45-30	0.000	0.000	0.000	0.000	0.019	0.000	0.017	0.023	0.036	0.062	0.071	0.070
	30min_CP_b_45-18	0.000	0.000	0.000	0.000	0.019	0.000	0.017	0.023	0.036	0.062	0.071	0.070
	30min_CP_b_34-30	0.000	0.000	0.000	0.000	0.019	0.025	0.023	0.035	0.057	0.074	0.074	0.073
	30min_C_28-28		0.000	0.000	0.000	0.036	0.000	0.000	0.015	0.023	0.042	0.036	0.035
	30min_C_28-17		0.000	0.000	0.000	0.036	0.000	0.000	0.015	0.023	0.044	0.036	0.035
	30min_C_b_28-30	0.000	0.000	0.000	0.000	0.031	0.000	0.006	0.023	0.028	0.057	0.047	0.042
	30min_C_b_15-30	0.000	0.000	0.000	0.000	0.042	0.000	0.006	0.017	0.033	0.041	0.035	0.029
	30min_C_b_15-12	0.000	0.000	0.000	0.000	0.036	0.000	0.006	0.017	0.041	0.044	0.035	0.031
	30min_C_b_13-12	0.000	0.000	0.000	0.000	0.028	0.000	0.006	0.017	0.033	0.042	0.045	0.035
	60min_CP_3-30	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.009	0.022	0.038	0.032	0.020
	60min_CP_3-15	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.012	0.023	0.036	0.031	0.022
	60min_CP_b_55-30					0.083	0.000	0.000	0.000	0.000	0.000	0.012	0.013
	60min_CP_b_37-30					0.056	0.000	0.000	0.000	0.000	0.000	0.015	0.016
	60min_CP_b_37-20					0.056	0.000	0.000	0.000	0.000	0.000	0.015	0.016
	60min_CP_b_16-30			0.000	0.000	0.092	0.000	0.000	0.000	0.006	0.013	0.019	0.019
	60min_CP_b_16-14			0.000	0.000	0.092	0.000	0.000	0.000	0.006	0.012	0.019	0.019
	60min_C_4-27		0.000	0.125	0.000	0.077	0.167	0.000	0.003	0.019	0.033	0.035	0.022
	60min_C_4-14			0.125	0.000	0.073	0.167	0.000	0.000	0.017	0.033	0.036	0.022
	60min_C_b_28-30			0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.019	0.039	0.028
	60min_C_b_22-14			0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.020	0.042	0.029
	60min_C_b_6-30	0.000		0.000	0.000	0.000	0.000	0.006	0.000	0.010	0.019	0.031	0.023

data: **case study data**		time resolution: **10 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**05min_CP_55-30**	0.008	0.030	0.022	0.052	0.098	0.057	0.302	0.328	0.328	0.320	0.304	0.291
	05min_CP_18-30	0.016	0.059	0.074	0.095	0.139	0.187	0.297	0.311	0.304	0.294	0.283	0.270
	05min_CP_18-10	0.015	0.060	0.068	0.101	0.143	0.180	0.308	0.311	0.310	0.294	0.285	0.281
	05min_CP_9-30	0.043	0.053	0.100	0.125	0.200	0.260	0.297	0.308	0.307	0.291	0.297	0.283
	05min_CP_9-14	0.028	0.050	0.094	0.135	0.201	0.246	0.291	0.305	0.305	0.289	0.297	0.283
	05min_CP_7-30	0.035	0.046	0.087	0.118	0.170	0.224	0.291	0.305	0.304	0.292	0.283	0.281
	05min_CP_7-14	0.028	0.047	0.090	0.115	0.156	0.222	0.291	0.308	0.302	0.294	0.288	0.282
	05min_CP_b_55-30	0.047	0.066	0.095	0.122	0.176	0.217	0.355	0.369	0.374	0.359	0.352	0.350
	05min_CP_b_27-30	0.057	0.085	0.134	0.196	0.256	0.344	0.384	0.375	0.387	0.369	0.369	0.352
	05min_CP_b_27-13	0.058	0.078	0.131	0.194	0.260	0.337	0.378	0.378	0.385	0.366	0.368	0.352
	05min_CP_b_4-30	0.080	0.082	0.189	0.197	0.309	0.379	0.267	0.262	0.273	0.267	0.263	0.249
	05min_CP_b_4-13	0.060	0.089	0.192	0.197	0.310	0.381	0.273	0.267	0.273	0.266	0.265	0.250
	05min_C_28-30	0.000	0.030	0.071	0.117	0.178	0.251	0.308	0.328	0.324	0.302	0.308	0.307
	05min_C_7-30	0.018	0.034	0.089	0.119	0.236	0.306	0.285	0.288	0.297	0.285	0.289	0.275
	05min_C_7-13	0.018	0.026	0.088	0.112	0.241	0.277	0.285	0.291	0.297	0.286	0.289	0.279
	05min_C_5-30	0.019	0.022	0.101	0.144	0.273	0.306	0.279	0.288	0.283	0.269	0.269	0.263
	05min_C_5-14	0.019	0.017	0.100	0.144	0.278	0.324	0.273	0.288	0.286	0.270	0.269	0.266
	05min_C_b_28-30	0.018	0.065	0.136	0.169	0.295	0.381	0.314	0.334	0.337	0.321	0.321	0.324
	05min_C_b_9-30	0.068	0.105	0.202	0.296	0.404	0.520	0.320	0.317	0.327	0.307	0.318	0.323
	05min_C_b_9-15	0.068	0.118	0.216	0.297	0.407	0.518	0.320	0.317	0.328	0.310	0.320	0.323
	05min_C_b_7-30	0.037	0.093	0.184	0.261	0.378	0.496	0.302	0.311	0.330	0.315	0.321	0.326
	05min_C_b_7-17	0.037	0.093	0.180	0.249	0.374	0.496	0.302	0.311	0.330	0.314	0.321	0.324
	10min_CP_55-30	0.000	0.005	0.002	0.011	0.019	0.030	0.238	0.256	0.257	0.253	0.250	0.235
	10min_CP_40-30	0.000	0.005	0.002	0.014	0.020	0.044	0.238	0.262	0.263	0.257	0.250	0.247
	10min_CP_40-10	0.000	0.005	0.002	0.010	0.024	0.046	0.238	0.259	0.265	0.266	0.247	0.251
	10min_CP_17-30	0.009	0.011	0.015	0.027	0.053	0.069	0.250	0.282	0.281	0.283	0.263	0.260
	10min_CP_17-15	0.009	0.011	0.015	0.029	0.050	0.109	0.250	0.282	0.281	0.282	0.263	0.260
	10min_CP_12-30	0.010	0.013	0.011	0.040	0.148	0.187	0.227	0.259	0.267	0.270	0.253	0.259
	10min_CP_12-9	0.010	0.009	0.012	0.033	0.129	0.165	0.227	0.253	0.251	0.263	0.241	0.250
	10min_CP_b_55-30	0.000	0.010	0.020	0.043	0.071	0.117	0.250	0.267	0.278	0.279	0.266	0.270
	10min_CP_b_53-8	0.000	0.011	0.027	0.046	0.081	0.134	0.244	0.259	0.265	0.263	0.259	0.253
	10min_CP_b_26-5	0.000	0.011	0.049	0.061	0.099	0.146	0.233	0.253	0.249	0.253	0.238	0.230
	10min_CP_b_2-11	0.208	0.250	0.209	0.295	0.422	0.413	0.244	0.209	0.221	0.209	0.189	0.199
	10min_C_28-30	0.000	0.004	0.014	0.014	0.026	0.044	0.221	0.244	0.237	0.246	0.222	0.219
	10min_C_7-22	0.000	0.004	0.024	0.031	0.128	0.167	0.238	0.256	0.251	0.244	0.237	0.234
	10min_C_5-23	0.000	0.013	0.048	0.041	0.126	0.188	0.244	0.253	0.244	0.240	0.235	0.234
	10min_C_3-30	0.014	0.028	0.104	0.064	0.270	0.345	0.157	0.154	0.153	0.140	0.135	0.141
	10min_C_3-15	0.017	0.031	0.081	0.068	0.277	0.305	0.140	0.148	0.153	0.138	0.134	0.141
	10min_C_b_28-30	0.000	0.013	0.044	0.056	0.074	0.123	0.221	0.244	0.247	0.259	0.241	0.234
	10min_C_b_8-16	0.000	0.022	0.080	0.094	0.132	0.221	0.244	0.265	0.262	0.263	0.256	0.250
	10min_C_b_4-22	0.143	0.138	0.261	0.280	0.566	0.759	0.174	0.180	0.189	0.179	0.170	0.164
	10min_C_b_3-30	0.114	0.080	0.161	0.205	0.386	0.553	0.180	0.166	0.172	0.161	0.157	0.154

data: **case study data**		time resolution: **10 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**15min_CP_55-30**	0.000	0.000	0.000	0.003	0.003	0.003	0.192	0.203	0.188	0.193	0.180	0.180
	15min_CP_53-5	0.000	0.000	0.000	0.000	0.004	0.000	0.169	0.172	0.164	0.164	0.157	0.157
	15min_CP_45-5	0.000	0.000	0.000	0.000	0.004	0.003	0.169	0.169	0.164	0.161	0.157	0.153
	15min_CP_25-13	0.000	0.000	0.003	0.012	0.011	0.006	0.186	0.206	0.190	0.190	0.185	0.189
	15min_CP_25-4	0.000	0.000	0.000	0.003	0.006	0.007	0.163	0.166	0.163	0.161	0.160	0.151
	15min_CP_13-21	0.000	0.000	0.003	0.017	0.029	0.044	0.186	0.201	0.195	0.198	0.190	0.177
	15min_CP_8-8	0.000	0.000	0.003	0.038	0.030	0.022	0.174	0.169	0.164	0.161	0.154	0.141
	15min_CP_b_55-30	0.000	0.000	0.000	0.027	0.030	0.022	0.180	0.203	0.196	0.205	0.189	0.198
	15min_CP_b_18-4	0.000	0.000	0.000	0.031	0.023	0.055	0.180	0.186	0.176	0.182	0.182	0.176
	15min_CP_b_7-10	0.147	0.041	0.100	0.162	0.208	0.141	0.169	0.183	0.182	0.174	0.158	0.154
	15min_CP_b_2-30	0.152	0.159	0.130	0.229	0.236	0.260	0.163	0.169	0.161	0.153	0.140	0.147
	15min_C_28-30	0.000	0.000	0.021	0.015	0.017	0.015	0.198	0.192	0.193	0.195	0.192	0.185
	15min_C_3-22	0.000	0.006	0.019	0.020	0.043	0.066	0.186	0.177	0.180	0.185	0.183	0.176
	15min_C_b_28-30	0.000	0.000	0.075	0.049	0.047	0.046	0.203	0.203	0.205	0.211	0.208	0.199
	15min_C_b_24-2	0.000	0.000	0.000	0.015	0.031	0.021	0.105	0.099	0.087	0.083	0.092	0.081
	15min_C_b_8-19	0.000	0.000	0.032	0.071	0.110	0.096	0.203	0.218	0.218	0.208	0.198	0.199
	15min_C_b_3-3	0.000	0.000	0.041	0.019	0.038	0.010	0.140	0.134	0.131	0.128	0.128	0.128
	30min_CP_55-30	0.000	0.000	0.000	0.011	0.007	0.000	0.128	0.134	0.126	0.126	0.134	0.131
	30min_CP_41-30	0.000	0.000	0.000	0.011	0.008	0.004	0.128	0.131	0.129	0.126	0.129	0.131
	30min_CP_27-15	0.000	0.000	0.000	0.021	0.021	0.007	0.134	0.137	0.134	0.124	0.125	0.129
	30min_CP_20-20	0.000	0.000	0.000	0.020	0.007	0.004	0.140	0.142	0.135	0.137	0.131	0.131
	30min_CP_8-29	0.000	0.000	0.000	0.029	0.022	0.015	0.140	0.140	0.138	0.135	0.125	0.125
	30min_CP_5-7	0.000	0.000	0.000	0.010	0.043	0.035	0.122	0.122	0.121	0.128	0.118	0.116
	30min_CP_b_55-30	0.000	0.000	0.000	0.010	0.020	0.000	0.145	0.142	0.142	0.137	0.137	0.140
	30min_CP_b_45-30	0.000	0.000	0.000	0.010	0.021	0.000	0.140	0.142	0.140	0.137	0.131	0.132
	30min_CP_b_45-18	0.000	0.000	0.000	0.010	0.021	0.000	0.140	0.142	0.140	0.137	0.131	0.131
	30min_CP_b_34-30	0.000	0.000	0.000	0.020	0.021	0.010	0.134	0.145	0.141	0.137	0.131	0.142
	30min_C_28-28	0.000	0.000	0.000	0.000	0.014	0.000	0.122	0.116	0.119	0.116	0.112	0.113
	30min_C_28-17	0.000	0.000	0.000	0.000	0.014	0.000	0.122	0.116	0.121	0.119	0.113	0.115
	30min_C_b_28-30	0.000	0.000	0.000	0.000	0.031	0.019	0.134	0.137	0.129	0.128	0.124	0.128
	30min_C_b_15-30	0.000	0.000	0.000	0.011	0.024	0.000	0.134	0.128	0.126	0.119	0.116	0.115
	30min_C_b_15-12	0.000	0.000	0.000	0.011	0.023	0.000	0.145	0.134	0.128	0.121	0.119	0.113
	30min_C_b_13-12	0.000	0.000	0.000	0.020	0.034	0.023	0.145	0.140	0.135	0.126	0.122	0.121
	60min_CP_3-30	0.000	0.000	0.042	0.042	0.190	0.056	0.041	0.041	0.060	0.062	0.047	0.042
	60min_CP_3-15	0.000	0.000	0.042	0.042	0.190	0.056	0.041	0.041	0.060	0.061	0.048	0.044
	60min_CP_b_55-30	0.000	0.000	0.000	0.000	0.025	0.000	0.029	0.035	0.054	0.052	0.055	0.055
	60min_CP_b_37-30	0.000	0.000	0.000	0.031	0.042	0.021	0.035	0.032	0.058	0.051	0.058	0.054
	60min_CP_b_37-20	0.000	0.000	0.000	0.031	0.042	0.021	0.035	0.032	0.058	0.051	0.058	0.054
	60min_CP_b_16-30	0.000	0.000	0.000	0.021	0.042	0.021	0.064	0.067	0.068	0.068	0.063	0.054
	60min_CP_b_16-14	0.000	0.000	0.000	0.021	0.042	0.019	0.058	0.067	0.068	0.070	0.064	0.055
	60min_C_4-27	0.000	0.000	0.021	0.106	0.076	0.021	0.099	0.087	0.073	0.052	0.070	0.067
	60min_C_4-14	0.000	0.000	0.021	0.064	0.076	0.021	0.105	0.093	0.073	0.057	0.071	0.070
	60min_C_b_28-30	0.000	0.000	0.000	0.000	0.016	0.000	0.099	0.105	0.081	0.083	0.087	0.080
	60min_C_b_22-14	0.000	0.000	0.000	0.015	0.028	0.000	0.105	0.102	0.084	0.084	0.087	0.074
	60min_C_b_6-30	0.000	0.000	0.000	0.023	0.056	0.021	0.099	0.096	0.074	0.055	0.057	0.060

data: **case study data**		time resolution: **15 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**05min_CP_55-30**	0.042	0.084	0.121	0.209	0.227	0.270	0.326	0.340	0.344	0.331	0.318	0.311
	05min_CP_18-30	0.108	0.192	0.199	0.265	0.329	0.432	0.314	0.294	0.311	0.317	0.301	0.278
	05min_CP_18-10	0.095	0.205	0.193	0.250	0.322	0.407	0.337	0.311	0.313	0.320	0.307	0.298
	05min_CP_9-30	0.155	0.195	0.199	0.296	0.415	0.444	0.297	0.317	0.330	0.331	0.326	0.310
	05min_CP_9-14	0.148	0.191	0.197	0.281	0.392	0.438	0.297	0.323	0.326	0.331	0.331	0.308
	05min_CP_7-30	0.139	0.161	0.180	0.225	0.338	0.425	0.297	0.317	0.315	0.323	0.311	0.292
	05min_CP_7-14	0.132	0.154	0.184	0.232	0.358	0.438	0.297	0.320	0.321	0.327	0.315	0.295
	05min_CP_b_55-30	0.176	0.176	0.273	0.353	0.405	0.476	0.355	0.366	0.387	0.394	0.376	0.372
	05min_CP_b_27-30	0.291	0.334	0.402	0.428	0.516	0.588	0.355	0.387	0.398	0.398	0.392	0.387
	05min_CP_b_27-13	0.294	0.335	0.406	0.431	0.513	0.590	0.349	0.387	0.397	0.398	0.392	0.385
	05min_CP_b_4-30	0.384	0.397	0.438	0.396	0.518	0.571	0.262	0.282	0.282	0.282	0.270	0.260
	05min_CP_b_4-13	0.378	0.379	0.437	0.395	0.518	0.587	0.267	0.285	0.282	0.283	0.278	0.263
	05min_C_28-30	0.135	0.199	0.177	0.289	0.433	0.560	0.343	0.346	0.353	0.350	0.344	0.331
	05min_C_7-30	0.111	0.173	0.225	0.300	0.452	0.556	0.291	0.317	0.327	0.321	0.314	0.308
	05min_C_7-13	0.106	0.176	0.216	0.303	0.457	0.542	0.291	0.311	0.327	0.320	0.315	0.311
	05min_C_5-30	0.105	0.166	0.221	0.338	0.475	0.571	0.297	0.305	0.311	0.295	0.291	0.267
	05min_C_5-14	0.105	0.163	0.221	0.339	0.472	0.572	0.297	0.314	0.314	0.299	0.297	0.270
	05min_C_b_28-30	0.259	0.284	0.305	0.393	0.539	0.621	0.349	0.358	0.359	0.359	0.355	0.349
	05min_C_b_9-30	0.402	0.423	0.452	0.484	0.641	0.719	0.337	0.352	0.358	0.353	0.356	0.350
	05min_C_b_9-15	0.412	0.426	0.458	0.488	0.642	0.721	0.331	0.352	0.360	0.355	0.358	0.349
	05min_C_b_7-30	0.313	0.355	0.402	0.442	0.604	0.699	0.331	0.349	0.356	0.352	0.356	0.349
	05min_C_b_7-17	0.296	0.358	0.398	0.438	0.600	0.693	0.331	0.349	0.356	0.352	0.356	0.349
	10min_CP_55-30	0.000	0.004	0.019	0.035	0.045	0.062	0.262	0.291	0.298	0.299	0.288	0.279
	10min_CP_40-30	0.000	0.004	0.039	0.040	0.062	0.107	0.267	0.294	0.305	0.301	0.291	0.286
	10min_CP_40-10	0.000	0.004	0.041	0.041	0.065	0.114	0.267	0.291	0.307	0.305	0.297	0.291
	10min_CP_17-30	0.133	0.095	0.135	0.171	0.233	0.302	0.308	0.317	0.324	0.321	0.314	0.308
	10min_CP_17-15	0.133	0.095	0.135	0.171	0.231	0.301	0.308	0.317	0.324	0.321	0.312	0.308
	10min_CP_12-30	0.008	0.053	0.080	0.162	0.186	0.362	0.291	0.302	0.314	0.311	0.298	0.297
	10min_CP_12-9	0.008	0.053	0.084	0.162	0.182	0.344	0.291	0.291	0.304	0.301	0.291	0.294
	10min_CP_b_55-30	0.000	0.018	0.083	0.127	0.171	0.228	0.291	0.320	0.324	0.324	0.326	0.313
	10min_CP_b_53-8	0.000	0.019	0.077	0.119	0.158	0.229	0.279	0.302	0.314	0.314	0.318	0.295
	10min_CP_b_26-5	0.000	0.011	0.092	0.157	0.169	0.239	0.256	0.279	0.289	0.292	0.294	0.289
	10min_CP_b_2-11	0.278	0.341	0.349	0.384	0.430	0.519	0.227	0.238	0.246	0.266	0.257	0.235
	10min_C_28-30	0.000	0.004	0.017	0.040	0.101	0.104	0.238	0.244	0.265	0.269	0.262	0.259
	10min_C_7-22	0.008	0.008	0.074	0.206	0.316	0.398	0.262	0.270	0.282	0.288	0.267	0.273
	10min_C_5-23	0.008	0.103	0.063	0.098	0.353	0.432	0.256	0.267	0.283	0.288	0.266	0.279
	10min_C_3-30	0.056	0.146	0.207	0.141	0.284	0.612	0.186	0.189	0.180	0.164	0.150	0.135
	10min_C_3-15	0.057	0.130	0.212	0.137	0.281	0.607	0.180	0.192	0.179	0.161	0.150	0.134
	10min_C_b_28-30	0.022	0.031	0.073	0.139	0.211	0.319	0.262	0.270	0.279	0.288	0.273	0.267
	10min_C_b_8-16	0.039	0.048	0.133	0.216	0.301	0.418	0.285	0.288	0.302	0.305	0.285	0.288
	10min_C_b_4-22	0.213	0.283	0.345	0.307	0.673	0.759	0.215	0.212	0.205	0.180	0.160	0.180
	10min_C_b_3-30	0.179	0.208	0.232	0.205	0.490	0.636	0.186	0.203	0.209	0.199	0.158	0.160

data: **case study data**		time resolution: **15 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**15min_CP_55-30**	0.000	0.000	0.004	0.009	0.005	0.008	0.238	0.241	0.262	0.256	0.235	0.217
	15min_CP_53-5	0.000	0.000	0.002	0.009	0.006	0.003	0.192	0.201	0.208	0.221	0.203	0.195
	15min_CP_45-5	0.000	0.000	0.007	0.011	0.006	0.010	0.203	0.201	0.209	0.215	0.202	0.193
	15min_CP_25-13	0.000	0.000	0.010	0.024	0.012	0.027	0.238	0.233	0.256	0.267	0.249	0.215
	15min_CP_25-4	0.000	0.000	0.006	0.014	0.016	0.027	0.198	0.198	0.208	0.212	0.195	0.185
	15min_CP_13-21	0.000	0.000	0.013	0.051	0.068	0.115	0.244	0.247	0.249	0.249	0.230	0.218
	15min_CP_8-8	0.023	0.032	0.037	0.065	0.069	0.085	0.192	0.206	0.203	0.193	0.186	0.158
	15min_CP_b_55-30	0.000	0.000	0.025	0.035	0.021	0.051	0.250	0.259	0.269	0.278	0.250	0.238
	15min_CP_b_18-4	0.000	0.000	0.029	0.076	0.065	0.117	0.209	0.218	0.234	0.234	0.221	0.214
	15min_CP_b_7-10	0.132	0.149	0.157	0.174	0.155	0.234	0.192	0.218	0.217	0.214	0.201	0.173
	15min_CP_b_2-30	0.156	0.216	0.226	0.262	0.214	0.347	0.221	0.212	0.193	0.203	0.192	0.148
	15min_C_28-30	0.000	0.000	0.017	0.029	0.024	0.048	0.233	0.218	0.240	0.254	0.238	0.225
	15min_C_3-22	0.000	0.005	0.013	0.035	0.039	0.043	0.233	0.230	0.238	0.238	0.219	0.209
	15min_C_b_28-30	0.000	0.000	0.046	0.083	0.091	0.153	0.238	0.230	0.254	0.269	0.253	0.241
	15min_C_b_24-2	0.000	0.000	0.013	0.000	0.026	0.054	0.105	0.105	0.102	0.103	0.102	0.097
	15min_C_b_8-19	0.000	0.000	0.042	0.119	0.190	0.230	0.238	0.241	0.262	0.267	0.253	0.243
	15min_C_b_3-3	0.000	0.000	0.000	0.055	0.029	0.083	0.145	0.148	0.140	0.147	0.138	0.134
	30min_CP_55-30	0.000	0.000	0.000	0.000	0.004	0.006	0.140	0.142	0.141	0.151	0.151	0.141
	30min_CP_41-30	0.000	0.000	0.000	0.000	0.003	0.010	0.140	0.145	0.144	0.145	0.148	0.140
	30min_CP_27-15	0.000	0.000	0.000	0.012	0.007	0.017	0.145	0.142	0.137	0.141	0.147	0.144
	30min_CP_20-20	0.000	0.000	0.004	0.018	0.014	0.017	0.145	0.142	0.144	0.154	0.145	0.145
	30min_CP_8-29	0.000	0.000	0.007	0.105	0.043	0.043	0.157	0.157	0.157	0.154	0.144	0.131
	30min_CP_5-7	0.000	0.000	0.020	0.023	0.030	0.068	0.145	0.140	0.137	0.141	0.135	0.118
	30min_CP_b_55-30	0.000	0.000	0.000	0.007	0.010	0.027	0.151	0.154	0.154	0.167	0.161	0.157
	30min_CP_b_45-30	0.000	0.000	0.010	0.015	0.010	0.016	0.151	0.151	0.150	0.163	0.161	0.156
	30min_CP_b_45-18	0.000	0.000	0.010	0.015	0.010	0.016	0.151	0.151	0.150	0.163	0.161	0.156
	30min_CP_b_34-30	0.000	0.000	0.000	0.033	0.018	0.034	0.145	0.151	0.157	0.172	0.163	0.164
	30min_C_28-28	0.000	0.000	0.012	0.003	0.011	0.007	0.140	0.140	0.137	0.141	0.138	0.131
	30min_C_28-17	0.000	0.000	0.012	0.003	0.011	0.007	0.145	0.140	0.137	0.142	0.138	0.132
	30min_C_b_28-30	0.000	0.000	0.010	0.009	0.018	0.026	0.163	0.157	0.150	0.153	0.147	0.148
	30min_C_b_15-30	0.000	0.000	0.011	0.009	0.019	0.038	0.151	0.148	0.144	0.145	0.145	0.141
	30min_C_b_15-12	0.000	0.000	0.011	0.009	0.020	0.019	0.157	0.148	0.144	0.148	0.144	0.141
	30min_C_b_13-12	0.000	0.000	0.009	0.018	0.038	0.043	0.169	0.151	0.156	0.153	0.148	0.148
	60min_CP_3-30	0.000	0.000	0.000	0.051	0.174	0.064	0.076	0.087	0.076	0.067	0.067	0.057
	60min_CP_3-15	0.000	0.000	0.000	0.051	0.188	0.063	0.076	0.076	0.074	0.067	0.065	0.058
	60min_CP_b_55-30	0.000	0.000	0.000	0.000	0.021	0.000	0.076	0.078	0.074	0.076	0.080	0.074
	60min_CP_b_37-30	0.000	0.000	0.000	0.014	0.036	0.042	0.064	0.078	0.071	0.077	0.084	0.076
	60min_CP_b_37-20	0.000	0.000	0.000	0.014	0.036	0.042	0.064	0.078	0.071	0.077	0.084	0.076
	60min_CP_b_16-30	0.000	0.000	0.000	0.033	0.032	0.053	0.093	0.093	0.083	0.087	0.096	0.083
	60min_CP_b_16-14	0.000	0.000	0.000	0.033	0.032	0.053	0.093	0.096	0.081	0.087	0.094	0.084
	60min_C_4-27	0.000	0.000	0.000	0.000	0.060	0.095	0.105	0.096	0.099	0.090	0.094	0.086
	60min_C_4-14	0.000	0.000	0.000	0.000	0.060	0.095	0.093	0.096	0.102	0.092	0.094	0.086
	60min_C_b_28-30	0.000	0.000	0.000	0.000	0.013	0.015	0.110	0.110	0.105	0.105	0.103	0.096
	60min_C_b_22-14	0.000	0.000	0.013	0.000	0.013	0.030	0.110	0.110	0.108	0.106	0.103	0.092
	60min_C_b_6-30	0.000	0.000	0.000	0.000	0.034	0.076	0.099	0.093	0.097	0.094	0.093	0.087

data: **case study data**		time resolution: **30 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**05min_CP_55-30**	0.407	0.470	0.601	0.650	0.661	0.708	0.320	0.352	0.359	0.340	0.343	0.324
	05min_CP_18-30	0.634	0.627	0.652	0.715	0.727	0.760	0.355	0.363	0.343	0.328	0.344	0.327
	05min_CP_18-10	0.621	0.622	0.651	0.722	0.732	0.770	0.355	0.369	0.347	0.324	0.349	0.328
	05min_CP_9-30	0.631	0.666	0.678	0.714	0.734	0.775	0.407	0.392	0.395	0.368	0.374	0.340
	05min_CP_9-14	0.611	0.643	0.659	0.693	0.718	0.767	0.413	0.398	0.397	0.369	0.374	0.340
	05min_CP_7-30	0.582	0.589	0.643	0.688	0.719	0.752	0.407	0.395	0.387	0.362	0.363	0.342
	05min_CP_7-14	0.599	0.601	0.658	0.705	0.728	0.760	0.407	0.398	0.390	0.363	0.365	0.343
	05min_CP_b_55-30	0.712	0.697	0.706	0.743	0.763	0.781	0.390	0.413	0.427	0.413	0.419	0.410
	05min_CP_b_27-30	0.754	0.748	0.748	0.789	0.808	0.831	0.419	0.439	0.442	0.422	0.439	0.437
	05min_CP_b_27-13	0.756	0.750	0.749	0.785	0.807	0.832	0.419	0.433	0.439	0.422	0.436	0.430
	05min_CP_b_4-30	0.756	0.739	0.719	0.749	0.756	0.788	0.337	0.358	0.374	0.347	0.349	0.314
	05min_CP_b_4-13	0.752	0.741	0.721	0.750	0.758	0.794	0.349	0.355	0.374	0.350	0.353	0.320
	05min_C_28-30	0.673	0.685	0.717	0.771	0.797	0.836	0.343	0.358	0.391	0.398	0.382	0.368
	05min_C_7-30	0.653	0.694	0.694	0.756	0.802	0.826	0.384	0.390	0.395	0.378	0.375	0.353
	05min_C_7-13	0.622	0.676	0.686	0.745	0.796	0.822	0.384	0.390	0.398	0.375	0.374	0.347
	05min_C_5-30	0.626	0.690	0.710	0.763	0.795	0.825	0.308	0.314	0.330	0.313	0.311	0.294
	05min_C_5-14	0.644	0.696	0.722	0.765	0.793	0.825	0.326	0.323	0.327	0.317	0.320	0.299
	05min_C_b_28-30	0.706	0.748	0.758	0.812	0.836	0.860	0.378	0.375	0.408	0.401	0.404	0.385
	05min_C_b_9-30	0.794	0.819	0.812	0.841	0.863	0.883	0.419	0.401	0.423	0.413	0.407	0.406
	05min_C_b_9-15	0.796	0.819	0.813	0.841	0.863	0.884	0.419	0.401	0.423	0.416	0.408	0.407
	05min_C_b_7-30	0.769	0.790	0.793	0.826	0.853	0.872	0.413	0.404	0.422	0.411	0.407	0.401
	05min_C_b_7-17	0.772	0.793	0.792	0.825	0.853	0.871	0.413	0.401	0.422	0.411	0.406	0.400
	10min_CP_55-30	0.202	0.255	0.294	0.252	0.322	0.349	0.343	0.349	0.355	0.333	0.323	0.318
	10min_CP_40-30	0.351	0.270	0.350	0.269	0.384	0.411	0.337	0.358	0.356	0.331	0.330	0.321
	10min_CP_40-10	0.317	0.313	0.363	0.275	0.402	0.407	0.343	0.360	0.359	0.334	0.333	0.321
	10min_CP_17-30	0.503	0.551	0.587	0.549	0.603	0.606	0.384	0.381	0.382	0.369	0.360	0.352
	10min_CP_17-15	0.503	0.559	0.592	0.553	0.611	0.610	0.384	0.381	0.382	0.369	0.360	0.352
	10min_CP_12-30	0.507	0.603	0.620	0.559	0.640	0.636	0.378	0.366	0.375	0.360	0.349	0.349
	10min_CP_12-9	0.473	0.607	0.620	0.582	0.620	0.646	0.372	0.366	0.368	0.352	0.343	0.344
	10min_CP_b_55-30	0.421	0.391	0.452	0.457	0.524	0.571	0.360	0.363	0.371	0.358	0.356	0.340
	10min_CP_b_53-8	0.378	0.357	0.432	0.443	0.504	0.550	0.355	0.360	0.365	0.352	0.349	0.333
	10min_CP_b_26-5	0.355	0.395	0.482	0.491	0.528	0.563	0.349	0.352	0.352	0.347	0.342	0.327
	10min_CP_b_2-11	0.667	0.658	0.699	0.698	0.752	0.751	0.262	0.253	0.270	0.256	0.237	0.244
	10min_C_28-30	0.082	0.092	0.129	0.231	0.324	0.377	0.326	0.326	0.333	0.318	0.320	0.308
	10min_C_7-22	0.363	0.433	0.485	0.601	0.460	0.602	0.314	0.323	0.337	0.324	0.326	0.318
	10min_C_5-23	0.519	0.490	0.494	0.661	0.647	0.693	0.314	0.326	0.337	0.318	0.317	0.313
	10min_C_3-30	0.400	0.247	0.522	0.432	0.525	0.608	0.180	0.177	0.166	0.166	0.151	0.157
	10min_C_3-15	0.381	0.272	0.529	0.428	0.543	0.602	0.163	0.180	0.166	0.166	0.151	0.156
	10min_C_b_28-30	0.241	0.292	0.388	0.456	0.516	0.592	0.349	0.337	0.349	0.321	0.318	0.310
	10min_C_b_8-16	0.437	0.491	0.541	0.585	0.656	0.696	0.337	0.352	0.347	0.327	0.327	0.321
	10min_C_b_4-22	0.585	0.630	0.660	0.672	0.724	0.812	0.198	0.180	0.186	0.176	0.172	0.160
	10min_C_b_3-30	0.421	0.491	0.584	0.547	0.588	0.689	0.192	0.192	0.174	0.173	0.156	0.166

data: **case study data**		time resolution: **30 min**											
		mean CMCR2C						**mean CACC2**					
time slot		**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**15min_CP_55-30**	0.027	0.065	0.105	0.137	0.220	0.205	0.302	0.308	0.308	0.302	0.283	0.278
	15min_CP_53-5	0.032	0.064	0.043	0.108	0.154	0.203	0.250	0.262	0.272	0.283	0.263	0.259
	15min_CP_45-5	0.023	0.016	0.038	0.141	0.147	0.205	0.256	0.262	0.260	0.276	0.259	0.256
	15min_CP_25-13	0.019	0.082	0.118	0.167	0.229	0.240	0.302	0.302	0.298	0.302	0.286	0.285
	15min_CP_25-4	0.034	0.022	0.048	0.059	0.079	0.186	0.262	0.250	0.262	0.260	0.249	0.240
	15min_CP_13-21	0.033	0.084	0.242	0.210	0.256	0.337	0.285	0.305	0.301	0.302	0.288	0.275
	15min_CP_8-8	0.020	0.129	0.149	0.156	0.138	0.341	0.198	0.201	0.186	0.193	0.183	0.187
	15min_CP_b_55-30	0.083	0.155	0.219	0.280	0.331	0.346	0.320	0.328	0.331	0.321	0.315	0.308
	15min_CP_b_18-4	0.164	0.152	0.215	0.220	0.288	0.308	0.267	0.288	0.291	0.292	0.275	0.267
	15min_CP_b_7-10	0.184	0.245	0.296	0.308	0.320	0.344	0.233	0.233	0.222	0.240	0.222	0.218
	15min_CP_b_2-30	0.478	0.449	0.494	0.494	0.541	0.607	0.203	0.209	0.195	0.190	0.180	0.160
	15min_C_28-30	0.096	0.088	0.116	0.173	0.254	0.402	0.267	0.291	0.294	0.292	0.288	0.276
	15min_C_3-22	0.376	0.432	0.288	0.243	0.322	0.372	0.262	0.270	0.266	0.256	0.251	0.249
	15min_C_b_28-30	0.275	0.253	0.372	0.412	0.495	0.555	0.291	0.302	0.310	0.307	0.304	0.292
	15min_C_b_24-2	0.000	0.000	0.059	0.106	0.126	0.204	0.110	0.110	0.115	0.118	0.124	0.118
	15min_C_b_8-19	0.286	0.289	0.379	0.414	0.477	0.549	0.291	0.299	0.305	0.304	0.292	0.288
	15min_C_b_3-3	0.108	0.142	0.264	0.242	0.241	0.250	0.192	0.192	0.190	0.185	0.174	0.174
	30min_CP_55-30	0.000	0.000	0.010	0.022	0.026	0.098	0.163	0.174	0.195	0.208	0.188	0.179
	30min_CP_41-30	0.000	0.000	0.008	0.024	0.021	0.052	0.163	0.174	0.205	0.212	0.196	0.193
	30min_CP_27-15	0.011	0.000	0.015	0.031	0.035	0.060	0.174	0.186	0.208	0.214	0.202	0.201
	30min_CP_20-20	0.000	0.000	0.014	0.024	0.030	0.092	0.203	0.198	0.218	0.221	0.212	0.202
	30min_CP_8-29	0.011	0.000	0.053	0.024	0.120	0.119	0.192	0.180	0.202	0.217	0.206	0.182
	30min_CP_5-7	0.000	0.083	0.016	0.045	0.054	0.027	0.174	0.189	0.205	0.195	0.182	0.170
	30min_CP_b_55-30	0.000	0.000	0.024	0.069	0.058	0.054	0.192	0.203	0.217	0.228	0.230	0.228
	30min_CP_b_45-30	0.000	0.000	0.052	0.063	0.061	0.059	0.186	0.203	0.215	0.227	0.237	0.241
	30min_CP_b_45-18	0.000	0.000	0.052	0.063	0.061	0.059	0.186	0.203	0.215	0.227	0.237	0.241
	30min_CP_b_34-30	0.000	0.015	0.056	0.079	0.072	0.097	0.186	0.201	0.227	0.240	0.244	0.251
	30min_C_28-28	0.000	0.000	0.010	0.047	0.017	0.043	0.151	0.166	0.177	0.188	0.167	0.157
	30min_C_28-17	0.000	0.000	0.010	0.044	0.017	0.043	0.151	0.169	0.182	0.189	0.167	0.158
	30min_C_b_28-30	0.000	0.000	0.026	0.076	0.060	0.068	0.163	0.180	0.199	0.201	0.186	0.179
	30min_C_b_15-30	0.000	0.000	0.048	0.062	0.092	0.103	0.169	0.177	0.198	0.201	0.180	0.177
	30min_C_b_15-12	0.000	0.000	0.062	0.080	0.097	0.103	0.169	0.180	0.192	0.199	0.180	0.177
	30min_C_b_13-12	0.000	0.000	0.060	0.074	0.085	0.120	0.163	0.186	0.201	0.195	0.185	0.183
	60min_CP_3-30	0.000	0.000	0.042	0.062	0.114	0.088	0.064	0.058	0.073	0.071	0.077	0.073
	60min_CP_3-15	0.000	0.000	0.040	0.062	0.113	0.103	0.064	0.055	0.074	0.071	0.077	0.073
	60min_CP_b_55-30	0.000	0.000	0.000	0.000	0.036	0.028	0.099	0.105	0.103	0.118	0.118	0.125
	60min_CP_b_37-30	0.000	0.000	0.000	0.000	0.038	0.026	0.105	0.102	0.100	0.110	0.119	0.124
	60min_CP_b_37-20	0.000	0.000	0.000	0.000	0.038	0.026	0.105	0.102	0.100	0.110	0.119	0.124
	60min_CP_b_16-30	0.000	0.000	0.010	0.016	0.054	0.046	0.093	0.105	0.109	0.115	0.128	0.126
	60min_CP_b_16-14	0.000	0.000	0.011	0.014	0.054	0.046	0.099	0.105	0.108	0.121	0.128	0.128
	60min_C_4-27	0.000	0.044	0.010	0.067	0.194	0.088	0.116	0.128	0.116	0.118	0.108	0.105
	60min_C_4-14	0.000	0.044	0.010	0.067	0.194	0.093	0.122	0.128	0.115	0.118	0.108	0.108
	60min_C_b_28-30	0.000	0.000	0.010	0.020	0.030	0.022	0.122	0.128	0.131	0.135	0.129	0.122
	60min_C_b_22-14	0.000	0.000	0.010	0.020	0.050	0.031	0.128	0.128	0.135	0.131	0.129	0.125
	60min_C_b_6-30	0.000	0.021	0.043	0.084	0.092	0.078	0.122	0.131	0.125	0.121	0.113	0.112

data: **case study data**		time resolution: **60 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**05min_CP_55-30**	0.740	0.794	0.844	0.868	0.851	0.870	0.401	0.384	0.379	0.375	0.371	0.366
	05min_CP_18-30	0.841	0.859	0.897	0.891	0.903	0.922	0.372	0.378	0.387	0.371	0.363	0.376
	05min_CP_18-10	0.827	0.855	0.893	0.897	0.901	0.926	0.384	0.375	0.382	0.372	0.374	0.385
	05min_CP_9-30	0.871	0.886	0.898	0.899	0.912	0.916	0.436	0.422	0.426	0.410	0.398	0.413
	05min_CP_9-14	0.869	0.880	0.892	0.893	0.906	0.912	0.436	0.424	0.429	0.408	0.401	0.416
	05min_CP_7-30	0.847	0.862	0.885	0.891	0.908	0.911	0.413	0.413	0.416	0.400	0.391	0.404
	05min_CP_7-14	0.858	0.865	0.888	0.892	0.910	0.912	0.401	0.413	0.413	0.400	0.387	0.407
	05min_CP_b_55-30	0.861	0.874	0.888	0.899	0.907	0.908	0.488	0.471	0.474	0.455	0.446	0.459
	05min_CP_b_27-30	0.886	0.892	0.903	0.910	0.920	0.925	0.488	0.471	0.490	0.475	0.462	0.496
	05min_CP_b_27-13	0.886	0.890	0.903	0.910	0.920	0.925	0.483	0.474	0.487	0.474	0.464	0.493
	05min_CP_b_4-30	0.863	0.860	0.879	0.878	0.885	0.895	0.395	0.407	0.401	0.395	0.387	0.384
	05min_CP_b_4-13	0.863	0.863	0.879	0.879	0.888	0.899	0.401	0.407	0.408	0.401	0.392	0.390
	05min_C_28-30	0.868	0.883	0.918	0.915	0.923	0.935	0.419	0.422	0.429	0.430	0.416	0.446
	05min_C_7-30	0.904	0.911	0.920	0.923	0.933	0.933	0.424	0.410	0.430	0.398	0.387	0.423
	05min_C_7-13	0.905	0.915	0.919	0.920	0.931	0.932	0.419	0.407	0.430	0.397	0.385	0.416
	05min_C_5-30	0.873	0.895	0.912	0.919	0.931	0.932	0.395	0.375	0.381	0.336	0.323	0.347
	05min_C_5-14	0.870	0.897	0.914	0.920	0.931	0.933	0.407	0.375	0.385	0.340	0.327	0.355
	05min_C_b_28-30	0.887	0.898	0.913	0.922	0.932	0.938	0.453	0.451	0.464	0.443	0.433	0.462
	05min_C_b_9-30	0.917	0.920	0.927	0.931	0.937	0.944	0.465	0.453	0.469	0.451	0.446	0.481
	05min_C_b_9-15	0.917	0.920	0.926	0.931	0.937	0.943	0.465	0.453	0.472	0.452	0.448	0.484
	05min_C_b_7-30	0.906	0.912	0.921	0.926	0.934	0.940	0.471	0.448	0.465	0.451	0.439	0.480
	05min_C_b_7-17	0.908	0.915	0.921	0.926	0.934	0.940	0.471	0.442	0.464	0.449	0.439	0.478
	10min_CP_55-30	0.520	0.561	0.573	0.588	0.670	0.668	0.372	0.369	0.398	0.385	0.388	0.398
	10min_CP_40-30	0.460	0.509	0.623	0.638	0.670	0.744	0.390	0.384	0.411	0.394	0.387	0.394
	10min_CP_40-10	0.504	0.563	0.622	0.659	0.680	0.724	0.390	0.384	0.417	0.395	0.397	0.403
	10min_CP_17-30	0.764	0.744	0.740	0.757	0.771	0.796	0.407	0.401	0.429	0.407	0.427	0.419
	10min_CP_17-15	0.764	0.744	0.753	0.755	0.772	0.795	0.407	0.401	0.429	0.407	0.427	0.419
	10min_CP_12-30	0.720	0.792	0.839	0.855	0.837	0.876	0.395	0.384	0.413	0.408	0.424	0.422
	10min_CP_12-9	0.718	0.816	0.852	0.838	0.837	0.885	0.378	0.381	0.406	0.404	0.420	0.419
	10min_CP_b_55-30	0.749	0.763	0.794	0.824	0.841	0.852	0.407	0.401	0.426	0.410	0.410	0.424
	10min_CP_b_53-8	0.754	0.759	0.790	0.816	0.832	0.841	0.390	0.398	0.422	0.400	0.406	0.419
	10min_CP_b_26-5	0.800	0.806	0.820	0.830	0.847	0.855	0.390	0.390	0.403	0.397	0.394	0.404
	10min_CP_b_2-11	0.910	0.914	0.905	0.914	0.921	0.925	0.273	0.259	0.286	0.275	0.279	0.272
	10min_C_28-30	0.706	0.485	0.702	0.451	0.491	0.540	0.343	0.340	0.369	0.368	0.366	0.385
	10min_C_7-22	0.711	0.745	0.817	0.811	0.806	0.795	0.355	0.349	0.390	0.397	0.379	0.400
	10min_C_5-23	0.666	0.716	0.781	0.802	0.812	0.826	0.349	0.349	0.390	0.392	0.381	0.403
	10min_C_3-30	0.736	0.740	0.857	0.839	0.859	0.884	0.169	0.172	0.173	0.176	0.160	0.167
	10min_C_3-15	0.719	0.739	0.854	0.838	0.855	0.884	0.169	0.172	0.173	0.174	0.163	0.164
	10min_C_b_28-30	0.719	0.746	0.782	0.801	0.834	0.859	0.366	0.352	0.384	0.371	0.374	0.385
	10min_C_b_8-16	0.800	0.813	0.830	0.838	0.868	0.887	0.355	0.360	0.394	0.397	0.382	0.404
	10min_C_b_4-22	0.768	0.793	0.834	0.847	0.877	0.876	0.227	0.224	0.233	0.222	0.208	0.227
	10min_C_b_3-30	0.688	0.723	0.807	0.822	0.839	0.847	0.198	0.195	0.186	0.179	0.176	0.183

data: **case study data**								time resolution: **60 min**					
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**15min_CP_55-30**	0.439	0.514	0.508	0.617	0.701	0.727	0.331	0.326	0.339	0.330	0.334	0.337
	15min_CP_53-5	0.427	0.446	0.402	0.516	0.648	0.643	0.302	0.294	0.320	0.312	0.315	0.320
	15min_CP_45-5	0.194	0.410	0.348	0.535	0.656	0.639	0.320	0.302	0.320	0.317	0.314	0.314
	15min_CP_25-13	0.380	0.487	0.595	0.648	0.695	0.706	0.337	0.328	0.347	0.337	0.350	0.356
	15min_CP_25-4	0.548	0.734	0.604	0.597	0.560	0.477	0.302	0.297	0.317	0.304	0.308	0.304
	15min_CP_13-21	0.627	0.683	0.762	0.745	0.772	0.737	0.337	0.331	0.337	0.311	0.315	0.315
	15min_CP_8-8	0.446	0.557	0.712	0.706	0.761	0.806	0.221	0.198	0.195	0.199	0.192	0.170
	15min_CP_b_55-30	0.665	0.687	0.730	0.754	0.769	0.780	0.355	0.346	0.369	0.358	0.363	0.371
	15min_CP_b_18-4	0.676	0.675	0.706	0.724	0.726	0.733	0.337	0.334	0.356	0.340	0.344	0.343
	15min_CP_b_7-10	0.669	0.624	0.688	0.701	0.746	0.792	0.273	0.265	0.247	0.256	0.240	0.211
	15min_CP_b_2-30	0.841	0.835	0.850	0.856	0.872	0.891	0.215	0.221	0.209	0.205	0.206	0.183
	15min_C_28-30	0.242	0.747	0.539	0.619	0.711	0.797	0.326	0.317	0.343	0.330	0.333	0.331
	15min_C_3-22	0.883	0.844	0.845	0.820	0.848	0.854	0.279	0.288	0.317	0.299	0.288	0.294
	15min_C_b_28-30	0.744	0.765	0.794	0.811	0.849	0.862	0.331	0.323	0.358	0.344	0.346	0.346
	15min_C_b_24-2	0.107	0.118	0.234	0.208	0.247	0.431	0.145	0.148	0.145	0.150	0.158	0.166
	15min_C_b_8-19	0.714	0.737	0.752	0.763	0.786	0.814	0.343	0.328	0.358	0.334	0.342	0.352
	15min_C_b_3-3	0.433	0.359	0.480	0.472	0.513	0.564	0.221	0.218	0.221	0.215	0.203	0.206
	30min_CP_55-30	0.047	0.202	0.170	0.190	0.369	0.390	0.256	0.259	0.270	0.259	0.243	0.244
	30min_CP_41-30	0.170	0.231	0.187	0.117	0.354	0.403	0.256	0.265	0.279	0.269	0.244	0.246
	30min_CP_27-15	0.266	0.252	0.242	0.254	0.476	0.568	0.273	0.273	0.285	0.281	0.273	0.272
	30min_CP_20-20	0.750	0.624	0.248	0.399	0.491	0.609	0.279	0.282	0.282	0.266	0.270	0.257
	30min_CP_8-29	0.606	0.558	0.642	0.558	0.701	0.672	0.256	0.262	0.257	0.247	0.237	0.238
	30min_CP_5-7	0.753	0.501	0.435	0.267	0.475	0.392	0.233	0.221	0.227	0.225	0.221	0.215
	30min_CP_b_55-30	0.226	0.261	0.355	0.415	0.486	0.466	0.279	0.279	0.297	0.291	0.286	0.282
	30min_CP_b_45-30	0.230	0.271	0.370	0.429	0.512	0.520	0.273	0.273	0.297	0.295	0.282	0.281
	30min_CP_b_45-18	0.230	0.271	0.370	0.429	0.512	0.520	0.273	0.273	0.297	0.295	0.282	0.281
	30min_CP_b_34-30	0.246	0.292	0.415	0.467	0.541	0.549	0.285	0.294	0.301	0.305	0.294	0.289
	30min_C_28-28	0.089	0.023	0.122	0.096	0.254	0.494	0.238	0.241	0.256	0.247	0.227	0.205
	30min_C_28-17	0.089	0.023	0.124	0.098	0.263	0.494	0.238	0.244	0.257	0.249	0.227	0.205
	30min_C_b_28-30	0.120	0.149	0.320	0.355	0.430	0.451	0.256	0.265	0.276	0.267	0.251	0.241
	30min_C_b_15-30	0.269	0.184	0.338	0.366	0.407	0.463	0.221	0.250	0.260	0.243	0.225	0.231
	30min_C_b_15-12	0.273	0.193	0.341	0.372	0.400	0.463	0.233	0.250	0.259	0.246	0.228	0.234
	30min_C_b_13-12	0.291	0.229	0.378	0.380	0.432	0.503	0.227	0.244	0.253	0.246	0.234	0.241
	60min_CP_3-30	0.025	0.042	0.267	0.151	0.229	0.379	0.157	0.134	0.161	0.148	0.137	0.126
	60min_CP_3-15	0.025	0.049	0.255	0.157	0.219	0.369	0.163	0.134	0.164	0.150	0.138	0.129
	60min_CP_b_55-30	0.000	0.000	0.055	0.081	0.113	0.162	0.163	0.186	0.177	0.208	0.201	0.187
	60min_CP_b_37-30	0.032	0.046	0.100	0.109	0.131	0.170	0.174	0.186	0.172	0.198	0.196	0.193
	60min_CP_b_37-20	0.032	0.046	0.100	0.109	0.131	0.170	0.174	0.186	0.172	0.198	0.196	0.193
	60min_CP_b_16-30	0.081	0.063	0.130	0.136	0.159	0.242	0.198	0.212	0.205	0.219	0.217	0.199
	60min_CP_b_16-14	0.105	0.064	0.130	0.138	0.167	0.242	0.198	0.209	0.205	0.217	0.217	0.199
	60min_C_4-27	0.000	0.013	0.079	0.144	0.110	0.269	0.157	0.148	0.161	0.158	0.151	0.141
	60min_C_4-14	0.000	0.014	0.065	0.109	0.108	0.272	0.151	0.142	0.163	0.158	0.153	0.138
	60min_C_b_28-30	0.000	0.016	0.066	0.080	0.112	0.112	0.174	0.163	0.185	0.185	0.182	0.174
	60min_C_b_22-14	0.000	0.016	0.121	0.073	0.161	0.179	0.169	0.166	0.192	0.188	0.182	0.167
	60min_C_b_6-30	0.000	0.031	0.190	0.136	0.188	0.256	0.145	0.163	0.166	0.167	0.161	0.131

data: **synthetically generated data** time resolution: **5 min**

	time slot	mean CMCR2C full_year	mean CMCR2C 6_months	mean CMCR2C 3_months	mean CMCR2C 5_weeks	mean CMCR2C 2_weeks	mean CMCR2C 1_week	mean CACC2 full_year	mean CACC2 6_months	mean CACC2 3_months	mean CACC2 5_weeks	mean CACC2 2_weeks	mean CACC2 1_week
best classifier candidates	05min_CP_55-30	0.000	0.000	0.000	0.034	0.100	0.185	0.472	0.473	0.465	0.459	0.452	0.423
	05min_CP_18-30	0.000	0.000	0.000	0.004	0.004	0.181	0.434	0.427	0.422	0.403	0.388	0.373
	05min_CP_18-10	0.000	0.000	0.000	0.004	0.004	0.142	0.430	0.424	0.422	0.403	0.386	0.376
	05min_CP_9-30	0.000	0.000	0.000	0.000	0.042	0.348	0.304	0.310	0.307	0.316	0.326	0.314
	05min_CP_9-14	0.000	0.000	0.000	0.000	0.042	0.327	0.304	0.309	0.306	0.316	0.324	0.314
	05min_CP_7-30	0.000	0.000	0.000	0.000	0.167	0.272	0.310	0.315	0.312	0.324	0.330	0.319
	05min_CP_7-14	0.000	0.000	0.000	0.000	0.181	0.282	0.313	0.313	0.311	0.324	0.331	0.321
	05min_CP_b_55-30	0.000	0.000	0.000	0.013	0.095	0.259	0.487	0.476	0.477	0.469	0.465	0.446
	05min_CP_b_27-30	0.000	0.000	0.002	0.023	0.099	0.260	0.487	0.486	0.480	0.469	0.458	0.429
	05min_CP_b_27-13	0.000	0.000	0.002	0.028	0.099	0.260	0.487	0.489	0.480	0.469	0.458	0.430
	05min_CP_b_4-30	0.000	0.000	0.002	0.008	0.044	0.189	0.383	0.369	0.373	0.372	0.361	0.356
	05min_CP_b_4-13	0.000	0.000	0.006	0.008	0.054	0.196	0.373	0.367	0.368	0.369	0.358	0.350
	05min_C_28-30	0.000	0.000	0.000	0.000	0.019	0.113	0.430	0.422	0.415	0.401	0.387	0.379
	05min_C_7-30	0.000	0.000	0.000	0.000	0.012	0.172	0.316	0.313	0.317	0.320	0.328	0.317
	05min_C_7-13	0.000	0.000	0.000	0.000	0.014	0.182	0.316	0.315	0.316	0.321	0.329	0.318
	05min_C_5-30	0.000	0.000	0.000	0.000	0.222	0.335	0.310	0.310	0.309	0.316	0.327	0.322
	05min_C_5-14	0.000	0.000	0.000	0.000	0.167	0.334	0.310	0.310	0.309	0.317	0.330	0.323
	05min_C_b_28-30	0.000	0.000	0.000	0.000	0.007	0.088	0.446	0.443	0.437	0.426	0.422	0.400
	05min_C_b_9-30	0.000	0.000	0.000	0.000	0.014	0.130	0.329	0.326	0.328	0.334	0.347	0.343
	05min_C_b_9-15	0.000	0.000	0.000	0.000	0.014	0.134	0.329	0.328	0.328	0.335	0.349	0.345
	05min_C_b_7-30	0.000	0.000	0.000	0.000	0.009	0.096	0.326	0.326	0.326	0.335	0.347	0.346
	05min_C_b_7-17	0.000	0.000	0.000	0.000	0.004	0.091	0.326	0.326	0.326	0.335	0.347	0.345
	10min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.026	0.316	0.310	0.320	0.329	0.333	0.316
	10min_CP_40-30	0.000	0.000	0.000	0.000	0.000	0.028	0.313	0.315	0.321	0.328	0.331	0.320
	10min_CP_40-10	0.000	0.000	0.000	0.000	0.000	0.023	0.304	0.309	0.320	0.328	0.328	0.315
	10min_CP_17-30	0.000	0.000	0.000	0.000	0.001	0.023	0.332	0.331	0.332	0.328	0.325	0.315
	10min_CP_17-15	0.000	0.000	0.000	0.000	0.001	0.021	0.332	0.331	0.331	0.328	0.325	0.314
	10min_CP_12-30	0.000	0.000	0.000	0.000	0.003	0.025	0.339	0.332	0.332	0.330	0.327	0.319
	10min_CP_12-9	0.000	0.000	0.000	0.000	0.001	0.022	0.332	0.326	0.328	0.328	0.324	0.316
	10min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.044	0.351	0.348	0.351	0.358	0.366	0.354
	10min_CP_b_53-8	0.000	0.000	0.000	0.000	0.000	0.033	0.348	0.347	0.345	0.351	0.356	0.343
	10min_CP_b_26-5	0.000	0.000	0.000	0.000	0.000	0.015	0.320	0.315	0.321	0.324	0.321	0.306
	10min_CP_b_2-11	0.140	0.134	0.124	0.147	0.338	0.512	0.253	0.215	0.218	0.214	0.191	0.185
	10min_C_28-30	0.000	0.000	0.000	0.000	0.000	0.006	0.291	0.293	0.288	0.282	0.270	0.245
	10min_C_7-22	0.000	0.000	0.000	0.000	0.000	0.001	0.215	0.217	0.214	0.218	0.219	0.222
	10min_C_5-23	0.000	0.000	0.002	0.009	0.027	0.032	0.256	0.245	0.251	0.251	0.240	0.240
	10min_C_3-30	0.028	0.020	0.035	0.037	0.074	0.086	0.073	0.084	0.085	0.074	0.073	0.062
	10min_C_3-15	0.020	0.000	0.029	0.012	0.037	0.050	0.085	0.090	0.088	0.078	0.074	0.065
	10min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.003	0.320	0.321	0.320	0.316	0.318	0.278
	10min_C_b_8-16	0.000	0.000	0.000	0.000	0.000	0.011	0.307	0.307	0.301	0.296	0.300	0.286
	10min_C_b_4-22	0.042	0.000	0.093	0.119	0.155	0.191	0.073	0.079	0.067	0.064	0.077	0.074
	10min_C_b_3-30	0.022	0.021	0.037	0.024	0.066	0.070	0.139	0.149	0.142	0.128	0.113	0.110

data: **synthetically generated data**	time resolution: **5 min**											
	mean CMCR2C						**mean CACC2**					
time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates												
15min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.007	0.165	0.165	0.180	0.194	0.201	0.199
15min_CP_53-5	0.000	0.000	0.000	0.000	0.000	0.000	0.142	0.141	0.148	0.147	0.155	0.158
15min_CP_45-5	0.000	0.000	0.000	0.000	0.000	0.000	0.136	0.133	0.136	0.141	0.150	0.160
15min_CP_25-13	0.000	0.000	0.000	0.000	0.000	0.015	0.155	0.153	0.169	0.191	0.210	0.208
15min_CP_25-4	0.000	0.000	0.000	0.000	0.000	0.000	0.130	0.141	0.140	0.137	0.122	0.111
15min_CP_13-21	0.015	0.008	0.000	0.003	0.006	0.046	0.171	0.168	0.166	0.172	0.178	0.176
15min_CP_8-8	0.010	0.000	0.000	0.000	0.002	0.054	0.184	0.190	0.184	0.187	0.184	0.173
15min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.022	0.187	0.207	0.233	0.243	0.241	0.244
15min_CP_b_18-4	0.000	0.000	0.000	0.000	0.000	0.000	0.171	0.166	0.171	0.173	0.175	0.170
15min_CP_b_7-10	0.019	0.010	0.005	0.004	0.014	0.125	0.168	0.160	0.161	0.166	0.165	0.153
15min_CP_b_2-30	0.026	0.032	0.045	0.047	0.066	0.154	0.241	0.236	0.218	0.211	0.203	0.179
15min_C_28-30	0.000	0.000	0.006	0.000	0.008	0.000	0.146	0.144	0.143	0.143	0.146	0.145
15min_C_3-22	0.000	0.000	0.009	0.004	0.005	0.059	0.101	0.092	0.084	0.079	0.070	0.068
15min_C_b_28-30	0.000	0.000	0.006	0.000	0.000	0.000	0.142	0.147	0.148	0.151	0.166	0.154
15min_C_b_24-2				0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.002
15min_C_b_8-19	0.000	0.000	0.000	0.000	0.000	0.074	0.165	0.165	0.164	0.157	0.150	0.135
15min_C_b_3-3				0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.004
30min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.011	0.016	0.021	0.032	0.032
30min_CP_41-30	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.027	0.028	0.033	0.040	0.040
30min_CP_27-15	0.000	0.000	0.000	0.000	0.000	0.000	0.022	0.022	0.023	0.028	0.030	0.036
30min_CP_20-20	0.000	0.000	0.000	0.000	0.000	0.000	0.044	0.044	0.044	0.054	0.059	0.065
30min_CP_8-29	0.000	0.000	0.000	0.000	0.000	0.000	0.041	0.041	0.044	0.042	0.046	0.061
30min_CP_5-7	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.016	0.016	0.017	0.016	0.017
30min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.000	0.041	0.046	0.057	0.069	0.076	0.089
30min_CP_b_45-30	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.038	0.045	0.054	0.062	0.069
30min_CP_b_45-18	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.038	0.046	0.054	0.062	0.069
30min_CP_b_34-30	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.041	0.051	0.059	0.066	0.079
30min_C_28-28	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.004	0.005	0.007	0.008
30min_C_28-17	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.004	0.005	0.007	0.008
30min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.011	0.014	0.021	0.025	0.024
30min_C_b_15-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.005	0.005	0.007	0.009	0.013
30min_C_b_15-12	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.005	0.006	0.008	0.009	0.013
30min_C_b_13-12	0.000	0.000	0.000	0.000	0.000	0.000	0.019	0.019	0.018	0.019	0.016	0.014
60min_CP_3-30	0.000	0.063	0.050	0.000	0.000	0.000	0.025	0.021	0.019	0.015	0.017	0.019
60min_CP_3-15	0.000	0.063	0.050	0.000	0.000	0.000	0.025	0.019	0.019	0.015	0.017	0.017
60min_CP_b_55-30					0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.006
60min_CP_b_37-30	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.005	0.008	0.009	0.014
60min_CP_b_37-20	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.005	0.008	0.009	0.014
60min_CP_b_16-30	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.009	0.009	0.013	0.015	0.016
60min_CP_b_16-14	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.009	0.009	0.011	0.014	0.016
60min_C_4-27	0.667	1.000	0.792	0.750	0.375	0.673	0.003	0.000	0.003	0.002	0.005	0.004
60min_C_4-14	0.500	0.667	0.550	0.526	0.333	0.664	0.006	0.005	0.008	0.007	0.006	0.005
60min_C_b_28-30		0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.004	0.005	0.006	0.006
60min_C_b_22-14		0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.005	0.006	0.006	0.004
60min_C_b_6-30	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.014	0.012	0.012	0.010	0.011

data: **synthetically generated data** time resolution: **10 min**

	mean CMCR2C						mean CACC2					
time slot	full_year	6_months	3_months	5_weeks	2_weeks	1_week	full_year	6_months	3_months	5_weeks	2_weeks	1_week
best classifier candidates												
05min_CP_55-30	0.000	0.000	0.005	0.019	0.111	0.236	0.497	0.492	0.487	0.478	0.456	0.425
05min_CP_18-30	0.167	0.000	0.250	0.278	0.201	0.346	0.449	0.430	0.423	0.414	0.390	0.369
05min_CP_18-10	0.000	0.000	0.139	0.163	0.180	0.321	0.456	0.434	0.423	0.413	0.395	0.371
05min_CP_9-30	0.333	0.333	0.333	0.333	0.329	0.355	0.323	0.323	0.324	0.332	0.335	0.329
05min_CP_9-14	0.333	0.333	0.333	0.333	0.331	0.358	0.326	0.323	0.324	0.332	0.334	0.331
05min_CP_7-30	0.333	0.333	0.333	0.306	0.307	0.337	0.339	0.337	0.339	0.341	0.341	0.334
05min_CP_7-14	0.333	0.333	0.333	0.306	0.310	0.335	0.335	0.334	0.338	0.341	0.342	0.339
05min_CP_b_55-30	0.000	0.000	0.002	0.022	0.140	0.385	0.500	0.494	0.496	0.489	0.477	0.453
05min_CP_b_27-30	0.000	0.006	0.015	0.042	0.169	0.454	0.506	0.502	0.501	0.490	0.472	0.447
05min_CP_b_27-13	0.000	0.003	0.015	0.042	0.180	0.457	0.509	0.500	0.502	0.490	0.472	0.445
05min_CP_b_4-30	0.024	0.023	0.037	0.047	0.175	0.426	0.389	0.402	0.395	0.370	0.373	0.354
05min_CP_b_4-13	0.025	0.028	0.035	0.041	0.177	0.421	0.367	0.391	0.392	0.365	0.371	0.354
05min_C_28-30	0.000	0.033	0.000	0.018	0.095	0.274	0.430	0.424	0.417	0.406	0.403	0.381
05min_C_7-30	0.200	0.219	0.205	0.171	0.263	0.335	0.323	0.324	0.325	0.330	0.334	0.334
05min_C_7-13	0.200	0.225	0.213	0.175	0.281	0.337	0.326	0.326	0.327	0.332	0.331	0.331
05min_C_5-30	0.333	0.337	0.333	0.336	0.340	0.375	0.323	0.324	0.324	0.332	0.336	0.332
05min_C_5-14	0.333	0.337	0.333	0.336	0.340	0.379	0.326	0.326	0.326	0.333	0.342	0.334
05min_C_b_28-30	0.000	0.000	0.000	0.005	0.093	0.333	0.456	0.453	0.447	0.434	0.426	0.407
05min_C_b_9-30	0.070	0.085	0.090	0.107	0.203	0.477	0.339	0.339	0.340	0.350	0.365	0.362
05min_C_b_9-15	0.070	0.085	0.092	0.107	0.202	0.477	0.339	0.337	0.341	0.350	0.366	0.361
05min_C_b_7-30	0.027	0.040	0.050	0.063	0.143	0.392	0.339	0.339	0.343	0.351	0.362	0.366
05min_C_b_7-17	0.027	0.040	0.050	0.049	0.140	0.385	0.339	0.339	0.343	0.350	0.362	0.366
10min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.007	0.329	0.329	0.339	0.346	0.349	0.331
10min_CP_40-30	0.000	0.000	0.000	0.000	0.002	0.008	0.326	0.332	0.339	0.347	0.353	0.331
10min_CP_40-10	0.000	0.000	0.000	0.000	0.000	0.007	0.326	0.331	0.339	0.348	0.354	0.331
10min_CP_17-30	0.000	0.000	0.000	0.000	0.001	0.010	0.335	0.337	0.339	0.347	0.351	0.339
10min_CP_17-15	0.000	0.000	0.000	0.000	0.001	0.010	0.335	0.337	0.339	0.347	0.351	0.339
10min_CP_12-30	0.000	0.000	0.000	0.000	0.000	0.009	0.335	0.339	0.339	0.347	0.352	0.338
10min_CP_12-9	0.000	0.000	0.000	0.000	0.000	0.003	0.335	0.337	0.336	0.343	0.350	0.332
10min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.021	0.351	0.366	0.378	0.384	0.384	0.371
10min_CP_b_53-8	0.000	0.000	0.000	0.000	0.000	0.011	0.348	0.361	0.367	0.372	0.372	0.358
10min_CP_b_26-5	0.000	0.000	0.000	0.000	0.000	0.007	0.329	0.332	0.335	0.343	0.343	0.322
10min_CP_b_2-11	0.096	0.113	0.100	0.151	0.254	0.414	0.269	0.263	0.263	0.244	0.226	0.194
10min_C_28-30	0.000	0.000	0.000	0.000	0.003	0.005	0.301	0.299	0.303	0.294	0.278	0.244
10min_C_7-22	0.000	0.000	0.000	0.000	0.000	0.003	0.234	0.247	0.248	0.254	0.251	0.234
10min_C_5-23	0.000	0.003	0.001	0.006	0.023	0.029	0.278	0.275	0.276	0.280	0.273	0.253
10min_C_3-30	0.000	0.027	0.000	0.010	0.047	0.051	0.092	0.078	0.075	0.067	0.064	0.063
10min_C_3-15	0.000	0.018	0.000	0.000	0.019	0.045	0.095	0.095	0.079	0.071	0.065	0.061
10min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.000	0.339	0.337	0.338	0.329	0.316	0.271
10min_C_b_8-16	0.000	0.000	0.000	0.000	0.000	0.017	0.329	0.328	0.320	0.324	0.318	0.294
10min_C_b_4-22	0.000	0.054	0.034	0.072	0.097	0.149	0.085	0.084	0.083	0.073	0.083	0.082
10min_C_b_3-30	0.000	0.020	0.000	0.000	0.029	0.072	0.149	0.152	0.135	0.115	0.111	0.098

data: **synthetically generated data**		time resolution: **10 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**15min_CP_55-30**	0.000	0.000	0.000	0.000	0.000	0.005	0.218	0.258	0.263	0.260	0.250	0.231
	15min_CP_53-5	0.000	0.000	0.000	0.000	0.000	0.000	0.174	0.174	0.183	0.208	0.204	0.182
	15min_CP_45-5	0.000	0.000	0.000	0.000	0.000	0.000	0.155	0.161	0.168	0.191	0.196	0.180
	15min_CP_25-13	0.000	0.000	0.000	0.000	0.000	0.006	0.193	0.222	0.251	0.271	0.259	0.244
	15min_CP_25-4	0.000	0.000	0.000	0.000	0.000	0.000	0.123	0.127	0.133	0.134	0.127	0.119
	15min_CP_13-21	0.000	0.000	0.005	0.008	0.020	0.044	0.199	0.196	0.197	0.203	0.215	0.203
	15min_CP_8-8	0.000	0.000	0.000	0.000	0.000	0.007	0.215	0.220	0.206	0.210	0.214	0.185
	15min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.006	0.275	0.286	0.282	0.293	0.282	0.263
	15min_CP_b_18-4	0.000	0.000	0.000	0.000	0.000	0.000	0.212	0.212	0.202	0.204	0.204	0.191
	15min_CP_b_7-10	0.000	0.008	0.016	0.000	0.027	0.046	0.190	0.199	0.192	0.189	0.196	0.161
	15min_CP_b_2-30	0.034	0.024	0.024	0.028	0.046	0.105	0.269	0.255	0.253	0.245	0.224	0.188
	15min_C_28-30	0.000	0.000	0.007	0.005	0.000	0.019	0.152	0.150	0.157	0.165	0.161	0.142
	15min_C_3-22	0.000	0.000	0.000	0.000	0.070	0.117	0.082	0.082	0.089	0.079	0.069	0.070
	15min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.011	0.161	0.172	0.169	0.187	0.172	0.157
	15min_C_b_24-2					0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002
	15min_C_b_8-19	0.000	0.000	0.000	0.000	0.024	0.183	0.168	0.165	0.161	0.165	0.159	0.147
	15min_C_b_3-3						0.000	0.000	0.000	0.000	0.000	0.000	0.002
	30min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.009	0.021	0.030	0.034	0.040
	30min_CP_41-30	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.024	0.036	0.040	0.044	0.042
	30min_CP_27-15	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.024	0.025	0.030	0.029	0.038
	30min_CP_20-20	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.038	0.047	0.047	0.051	0.055
	30min_CP_8-29	0.000	0.000	0.000	0.000	0.000	0.000	0.038	0.038	0.037	0.038	0.039	0.042
	30min_CP_5-7	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.014	0.013	0.011	0.010	0.012
	30min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.000	0.028	0.043	0.052	0.068	0.081	0.094
	30min_CP_b_45-30	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.036	0.043	0.047	0.057	0.074
	30min_CP_b_45-18	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.036	0.043	0.047	0.057	0.074
	30min_CP_b_34-30	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.040	0.047	0.052	0.060	0.077
	30min_C_28-28	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.005	0.007	0.007
	30min_C_28-17	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.005	0.007	0.008
	30min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.006	0.009	0.017	0.021	0.023
	30min_C_b_15-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.006	0.008	0.009	0.010	0.011
	30min_C_b_15-12	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.006	0.007	0.009	0.009	0.013
	30min_C_b_13-12	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.016	0.016	0.016	0.017	0.016
	60min_CP_3-30	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.016	0.014	0.017	0.014	0.015
	60min_CP_3-15	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.016	0.014	0.016	0.014	0.015
	60min_CP_b_55-30			0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.004	0.005
	60min_CP_b_37-30	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.007	0.007	0.009	0.013
	60min_CP_b_37-20	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.007	0.007	0.009	0.013
	60min_CP_b_16-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.008	0.010	0.012	0.011	0.014
	60min_CP_b_16-14	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.006	0.009	0.010	0.010	0.012
	60min_C_4-27	1.000	1.000	0.889	0.500	0.667	0.833	0.000	0.000	0.001	0.002	0.002	0.002
	60min_C_4-14	0.500	0.550	0.417	0.250	0.667	0.467	0.006	0.005	0.005	0.006	0.002	0.004
	60min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.002	0.005	0.004	0.005
	60min_C_b_22-14	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.002	0.005	0.004	0.005
	60min_C_b_6-30	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.013	0.012	0.009	0.009	0.009

data: **synthetically generated data** time resolution: **15 min**

		mean CMCR2C						mean CACC2					
	time slot	full_year	6_months	3_months	5_weeks	2_weeks	1_week	full_year	6_months	3_months	5_weeks	2_weeks	1_week
best classifier candidates	05min_CP_55-30	0.236	0.213	0.221	0.212	0.268	0.368	0.491	0.483	0.483	0.472	0.463	0.445
	05min_CP_18-30	0.323	0.323	0.306	0.334	0.365	0.463	0.437	0.430	0.422	0.402	0.391	0.378
	05min_CP_18-10	0.322	0.309	0.294	0.310	0.335	0.453	0.434	0.422	0.423	0.401	0.403	0.385
	05min_CP_9-30	0.358	0.371	0.375	0.386	0.425	0.512	0.332	0.337	0.339	0.343	0.354	0.343
	05min_CP_9-14	0.358	0.372	0.375	0.388	0.426	0.507	0.332	0.332	0.339	0.342	0.353	0.344
	05min_CP_7-30	0.351	0.356	0.368	0.375	0.402	0.475	0.342	0.347	0.348	0.353	0.358	0.347
	05min_CP_7-14	0.357	0.350	0.370	0.381	0.409	0.489	0.342	0.347	0.350	0.351	0.358	0.349
	05min_CP_b_55-30	0.180	0.190	0.223	0.283	0.424	0.630	0.503	0.498	0.499	0.500	0.487	0.476
	05min_CP_b_27-30	0.258	0.255	0.269	0.309	0.475	0.707	0.509	0.503	0.502	0.496	0.481	0.472
	05min_CP_b_27-13	0.241	0.255	0.263	0.310	0.476	0.710	0.509	0.503	0.502	0.497	0.481	0.472
	05min_CP_b_4-30	0.256	0.254	0.254	0.272	0.462	0.663	0.405	0.402	0.390	0.386	0.379	0.352
	05min_CP_b_4-13	0.240	0.245	0.237	0.261	0.459	0.664	0.402	0.399	0.385	0.381	0.375	0.350
	05min_C_28-30	0.288	0.266	0.249	0.209	0.274	0.442	0.427	0.415	0.407	0.402	0.412	0.400
	05min_C_7-30	0.314	0.328	0.329	0.333	0.396	0.494	0.335	0.329	0.334	0.334	0.348	0.338
	05min_C_7-13	0.315	0.329	0.334	0.333	0.388	0.475	0.335	0.332	0.335	0.336	0.350	0.329
	05min_C_5-30	0.346	0.362	0.361	0.384	0.412	0.489	0.335	0.334	0.336	0.340	0.350	0.341
	05min_C_5-14	0.346	0.362	0.361	0.384	0.415	0.498	0.339	0.340	0.339	0.344	0.353	0.346
	05min_C_b_28-30	0.193	0.156	0.153	0.185	0.371	0.669	0.449	0.449	0.434	0.432	0.429	0.416
	05min_C_b_9-30	0.329	0.322	0.335	0.346	0.514	0.749	0.348	0.353	0.353	0.362	0.381	0.384
	05min_C_b_9-15	0.329	0.318	0.334	0.348	0.514	0.750	0.348	0.353	0.354	0.362	0.381	0.385
	05min_C_b_7-30	0.262	0.255	0.247	0.264	0.446	0.711	0.348	0.356	0.356	0.363	0.384	0.385
	05min_C_b_7-17	0.262	0.246	0.244	0.258	0.442	0.707	0.348	0.353	0.355	0.362	0.384	0.385
	10min_CP_55-30	0.000	0.000	0.000	0.000	0.008	0.037	0.345	0.342	0.345	0.345	0.349	0.332
	10min_CP_40-30	0.000	0.000	0.000	0.002	0.012	0.036	0.345	0.345	0.347	0.347	0.353	0.335
	10min_CP_40-10	0.000	0.000	0.000	0.005	0.012	0.036	0.345	0.345	0.347	0.349	0.353	0.339
	10min_CP_17-30	0.000	0.000	0.000	0.004	0.014	0.067	0.348	0.347	0.347	0.346	0.354	0.348
	10min_CP_17-15	0.000	0.000	0.000	0.005	0.014	0.067	0.348	0.347	0.346	0.345	0.353	0.348
	10min_CP_12-30	0.000	0.000	0.001	0.015	0.027	0.086	0.348	0.350	0.349	0.347	0.355	0.348
	10min_CP_12-9	0.000	0.000	0.001	0.002	0.011	0.061	0.348	0.350	0.347	0.344	0.351	0.343
	10min_CP_b_55-30	0.000	0.000	0.000	0.002	0.024	0.095	0.370	0.373	0.381	0.388	0.388	0.366
	10min_CP_b_53-8	0.000	0.000	0.000	0.000	0.018	0.075	0.358	0.361	0.365	0.374	0.376	0.359
	10min_CP_b_26-5	0.000	0.000	0.000	0.000	0.013	0.060	0.345	0.343	0.345	0.339	0.343	0.331
	10min_CP_b_2-11	0.082	0.097	0.106	0.175	0.318	0.547	0.285	0.278	0.255	0.226	0.233	0.206
	10min_C_28-30	0.000	0.000	0.000	0.000	0.007	0.035	0.316	0.323	0.315	0.301	0.284	0.237
	10min_C_7-22	0.000	0.000	0.001	0.000	0.010	0.047	0.285	0.271	0.277	0.276	0.267	0.252
	10min_C_5-23	0.000	0.000	0.003	0.005	0.029	0.100	0.316	0.307	0.294	0.291	0.284	0.260
	10min_C_3-30	0.013	0.000	0.026	0.008	0.040	0.196	0.104	0.089	0.084	0.071	0.061	0.066
	10min_C_3-15	0.000	0.000	0.019	0.010	0.022	0.173	0.108	0.098	0.088	0.073	0.066	0.068
	10min_C_b_28-30	0.000	0.000	0.000	0.000	0.015	0.078	0.348	0.339	0.338	0.336	0.316	0.277
	10min_C_b_8-16	0.000	0.000	0.002	0.000	0.029	0.113	0.335	0.335	0.328	0.330	0.325	0.296
	10min_C_b_4-22	0.031	0.000	0.040	0.061	0.095	0.403	0.098	0.108	0.091	0.085	0.090	0.083
	10min_C_b_3-30	0.038	0.010	0.031	0.013	0.051	0.254	0.158	0.160	0.148	0.121	0.109	0.104

data: **synthetically generated data** time resolution: **15 min**

		mean CMCR2C						mean CACC2					
	time slot	full_year	6_months	3_months	5_weeks	2_weeks	1_week	full_year	6_months	3_months	5_weeks	2_weeks	1_week
best classifier candidates	15min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.016	0.253	0.269	0.271	0.267	0.262	0.240
	15min_CP_53-5	0.000	0.000	0.000	0.000	0.000	0.024	0.171	0.188	0.209	0.217	0.213	0.195
	15min_CP_45-5	0.000	0.000	0.000	0.000	0.000	0.026	0.155	0.172	0.188	0.203	0.205	0.198
	15min_CP_25-13	0.000	0.000	0.000	0.000	0.000	0.024	0.203	0.236	0.267	0.278	0.276	0.252
	15min_CP_25-4	0.000	0.000	0.000	0.000	0.000	0.000	0.136	0.141	0.145	0.146	0.133	0.126
	15min_CP_13-21	0.000	0.005	0.002	0.008	0.025	0.061	0.209	0.214	0.218	0.232	0.218	0.206
	15min_CP_8-8	0.000	0.000	0.002	0.000	0.000	0.002	0.215	0.229	0.227	0.227	0.207	0.191
	15min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.014	0.291	0.293	0.297	0.291	0.286	0.266
	15min_CP_b_18-4	0.000	0.000	0.000	0.000	0.000	0.022	0.199	0.204	0.206	0.219	0.218	0.209
	15min_CP_b_7-10	0.000	0.014	0.011	0.018	0.020	0.061	0.212	0.220	0.215	0.206	0.188	0.171
	15min_CP_b_2-30	0.033	0.028	0.033	0.029	0.076	0.206	0.275	0.271	0.258	0.238	0.216	0.191
	15min_C_28-30	0.000	0.000	0.000	0.000	0.000	0.089	0.155	0.160	0.158	0.174	0.163	0.150
	15min_C_3-22	0.000	0.000	0.003	0.022	0.084	0.345	0.092	0.097	0.093	0.081	0.078	0.070
	15min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.071	0.168	0.177	0.177	0.193	0.187	0.172
	15min_C_b_24-2						0.000	0.000	0.000	0.000	0.000	0.000	0.001
	15min_C_b_8-19	0.000	0.000	0.000	0.014	0.139	0.380	0.171	0.168	0.171	0.168	0.161	0.146
	15min_C_b_3-3			0.000	0.000	0.000	0.438	0.000	0.000	0.001	0.002	0.002	0.004
	30min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.021	0.023	0.034	0.047	0.042
	30min_CP_41-30	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.028	0.030	0.036	0.050	0.036
	30min_CP_27-15	0.000	0.000	0.000	0.000	0.000	0.000	0.022	0.024	0.028	0.028	0.036	0.034
	30min_CP_20-20	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.038	0.045	0.046	0.057	0.052
	30min_CP_8-29	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.035	0.039	0.043	0.042	0.045
	30min_CP_5-7	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.013	0.013	0.011	0.013	0.015
	30min_CP_b_55-30	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.043	0.050	0.070	0.099	0.105
	30min_CP_b_45-30	0.000	0.000	0.000	0.000	0.000	0.000	0.028	0.030	0.043	0.047	0.072	0.089
	30min_CP_b_45-18	0.000	0.000	0.000	0.000	0.000	0.000	0.028	0.030	0.043	0.047	0.072	0.088
	30min_CP_b_34-30	0.000	0.000	0.000	0.000	0.000	0.000	0.035	0.036	0.044	0.053	0.073	0.084
	30min_C_28-28		0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.003	0.002	0.004
	30min_C_28-17		0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.003	0.003	0.004
	30min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.006	0.010	0.015	0.015	0.020
	30min_C_b_15-30	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.003	0.010	0.007	0.009
	30min_C_b_15-12	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.009	0.007	0.008
	30min_C_b_13-12	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.013	0.013	0.017	0.013	0.014
	60min_CP_3-30	0.000	0.000	0.000	0.083	0.000	0.000	0.009	0.011	0.010	0.009	0.008	0.009
	60min_CP_3-15	0.000	0.000	0.000	0.083	0.000	0.000	0.009	0.011	0.010	0.010	0.009	0.010
	60min_CP_b_55-30				0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.004
	60min_CP_b_37-30			0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.004	0.005	0.007
	60min_CP_b_37-20			0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.004	0.005	0.007
	60min_CP_b_16-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.005	0.006	0.006	0.006	0.008
	60min_CP_b_16-14	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.003	0.006	0.006	0.006	0.007
	60min_C_4-27	1.000	1.000	0.917	0.667	0.889	0.833	0.000	0.000	0.001	0.002	0.001	0.001
	60min_C_4-14	0.750	0.708	0.625	0.583	0.556	0.611	0.003	0.003	0.003	0.003	0.002	0.002
	60min_C_b_28-30		0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.002	0.003	0.004
	60min_C_b_22-14		0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.002	0.003	0.002
	60min_C_b_6-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.005	0.006	0.009	0.008	0.009

data: **synthetically generated data**		time resolution: **30 min**											
		mean CMCR2C						**mean CACC2**					
time slot		**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**05min_CP_55-30**	0.304	0.316	0.347	0.444	0.561	0.688	0.497	0.495	0.487	0.466	0.459	0.422
	05min_CP_18-30	0.432	0.431	0.476	0.531	0.640	0.738	0.421	0.403	0.392	0.381	0.371	0.344
	05min_CP_18-10	0.420	0.424	0.468	0.522	0.630	0.733	0.430	0.415	0.403	0.388	0.385	0.358
	05min_CP_9-30	0.647	0.672	0.675	0.698	0.738	0.835	0.348	0.353	0.354	0.359	0.364	0.355
	05min_CP_9-14	0.654	0.674	0.678	0.699	0.734	0.833	0.345	0.353	0.353	0.361	0.368	0.354
	05min_CP_7-30	0.611	0.613	0.627	0.651	0.695	0.795	0.348	0.356	0.354	0.354	0.358	0.344
	05min_CP_7-14	0.619	0.627	0.641	0.654	0.706	0.804	0.348	0.354	0.354	0.359	0.363	0.350
	05min_CP_b_55-30	0.613	0.640	0.670	0.709	0.799	0.880	0.491	0.495	0.505	0.489	0.483	0.475
	05min_CP_b_27-30	0.672	0.677	0.681	0.722	0.824	0.902	0.491	0.498	0.506	0.484	0.479	0.479
	05min_CP_b_27-13	0.669	0.674	0.682	0.722	0.825	0.903	0.491	0.498	0.503	0.484	0.479	0.475
	05min_CP_b_4-30	0.581	0.595	0.605	0.645	0.774	0.866	0.408	0.397	0.390	0.377	0.363	0.331
	05min_CP_b_4-13	0.584	0.598	0.597	0.638	0.773	0.866	0.389	0.384	0.389	0.374	0.363	0.336
	05min_C_28-30	0.318	0.347	0.380	0.433	0.577	0.730	0.440	0.426	0.422	0.410	0.426	0.427
	05min_C_7-30	0.596	0.602	0.618	0.638	0.701	0.818	0.335	0.331	0.338	0.342	0.350	0.325
	05min_C_7-13	0.569	0.579	0.594	0.618	0.689	0.801	0.329	0.328	0.337	0.339	0.343	0.320
	05min_C_5-30	0.582	0.590	0.607	0.627	0.667	0.786	0.335	0.332	0.333	0.337	0.340	0.316
	05min_C_5-14	0.587	0.601	0.613	0.632	0.676	0.792	0.335	0.334	0.336	0.339	0.342	0.324
	05min_C_b_28-30	0.528	0.548	0.579	0.649	0.790	0.902	0.449	0.448	0.445	0.432	0.444	0.441
	05min_C_b_9-30	0.668	0.666	0.675	0.709	0.839	0.920	0.373	0.381	0.379	0.384	0.399	0.411
	05min_C_b_9-15	0.671	0.669	0.675	0.707	0.839	0.920	0.373	0.380	0.378	0.386	0.399	0.413
	05min_C_b_7-30	0.649	0.646	0.657	0.690	0.824	0.912	0.377	0.384	0.380	0.388	0.400	0.411
	05min_C_b_7-17	0.646	0.644	0.654	0.687	0.821	0.911	0.377	0.383	0.380	0.386	0.400	0.411
	10min_CP_55-30	0.068	0.108	0.132	0.165	0.182	0.231	0.354	0.356	0.356	0.352	0.347	0.328
	10min_CP_40-30	0.104	0.123	0.145	0.175	0.193	0.267	0.370	0.361	0.359	0.355	0.353	0.331
	10min_CP_40-10	0.110	0.125	0.142	0.172	0.184	0.265	0.367	0.361	0.362	0.358	0.355	0.339
	10min_CP_17-30	0.130	0.132	0.140	0.150	0.216	0.380	0.370	0.362	0.361	0.364	0.364	0.366
	10min_CP_17-15	0.130	0.132	0.141	0.150	0.216	0.380	0.370	0.362	0.360	0.363	0.364	0.365
	10min_CP_12-30	0.136	0.141	0.146	0.165	0.241	0.397	0.370	0.361	0.362	0.364	0.362	0.368
	10min_CP_12-9	0.136	0.141	0.147	0.166	0.217	0.361	0.367	0.358	0.361	0.362	0.359	0.366
	10min_CP_b_55-30	0.123	0.199	0.241	0.300	0.408	0.590	0.383	0.384	0.384	0.395	0.392	0.373
	10min_CP_b_53-8	0.126	0.200	0.234	0.287	0.373	0.544	0.373	0.377	0.373	0.377	0.370	0.359
	10min_CP_b_26-5	0.278	0.294	0.306	0.341	0.423	0.581	0.345	0.347	0.351	0.348	0.350	0.343
	10min_CP_b_2-11	0.367	0.388	0.439	0.534	0.649	0.820	0.278	0.255	0.242	0.231	0.231	0.216
	10min_C_28-30	0.000	0.004	0.023	0.061	0.128	0.233	0.307	0.307	0.300	0.290	0.286	0.263
	10min_C_7-22	0.109	0.106	0.108	0.124	0.178	0.301	0.297	0.288	0.281	0.280	0.272	0.267
	10min_C_5-23	0.106	0.110	0.107	0.144	0.244	0.416	0.326	0.316	0.306	0.298	0.298	0.287
	10min_C_3-30	0.014	0.023	0.043	0.043	0.147	0.381	0.123	0.098	0.097	0.089	0.081	0.082
	10min_C_3-15	0.013	0.021	0.028	0.013	0.108	0.318	0.133	0.101	0.103	0.089	0.085	0.080
	10min_C_b_28-30	0.000	0.013	0.055	0.144	0.248	0.511	0.348	0.350	0.335	0.322	0.311	0.282
	10min_C_b_8-16	0.227	0.232	0.232	0.276	0.424	0.655	0.345	0.348	0.350	0.331	0.324	0.315
	10min_C_b_4-22	0.104	0.148	0.171	0.240	0.404	0.632	0.136	0.119	0.123	0.117	0.123	0.121
	10min_C_b_3-30	0.015	0.030	0.048	0.071	0.208	0.487	0.206	0.157	0.143	0.130	0.128	0.127

data: **synthetically generated data** time resolution: **30 min**

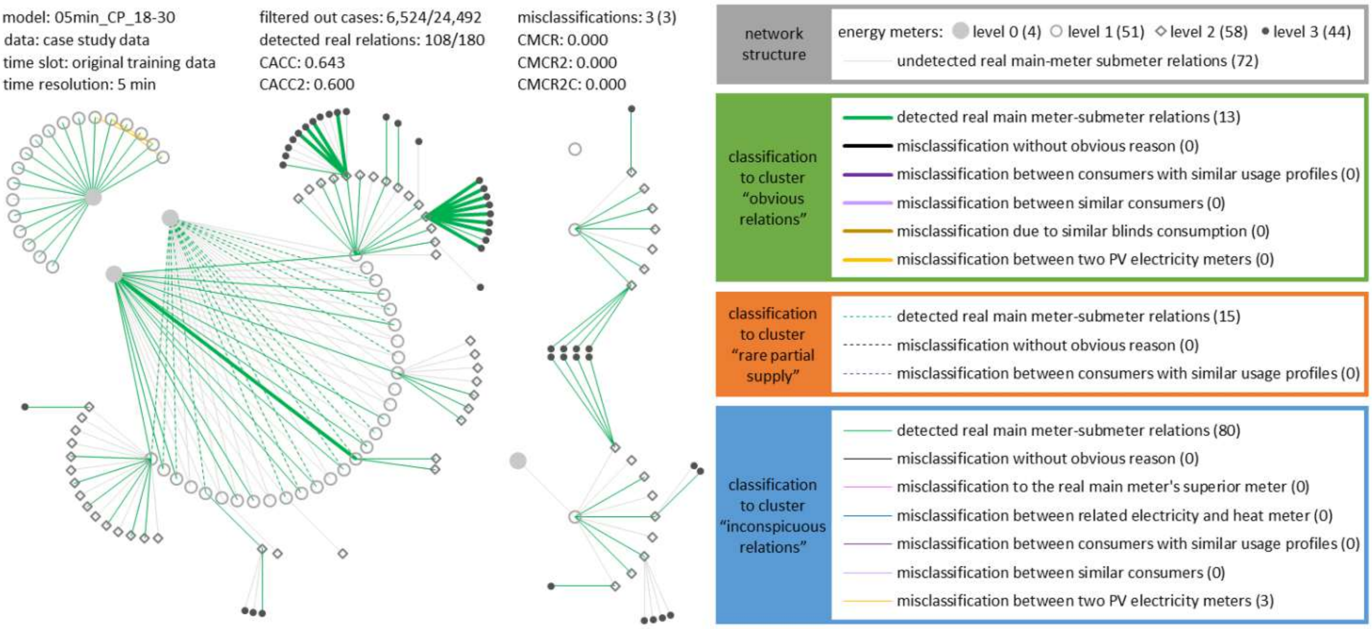
	mean CMCR2C						mean CACC2					
time slot	full_year	6_months	3_months	5_weeks	2_weeks	1_week	full_year	6_months	3_months	5_weeks	2_weeks	1_week
best classifier candidates												
15min_CP_55-30	0.000	0.018	0.038	0.091	0.144	0.231	0.259	0.266	0.270	0.263	0.260	0.244
15min_CP_53-5	0.000	0.000	0.013	0.036	0.101	0.194	0.187	0.217	0.227	0.228	0.217	0.204
15min_CP_45-5	0.000	0.008	0.009	0.030	0.105	0.201	0.177	0.206	0.220	0.229	0.216	0.202
15min_CP_25-13	0.044	0.074	0.075	0.119	0.141	0.231	0.256	0.272	0.282	0.278	0.282	0.263
15min_CP_25-4	0.000	0.007	0.009	0.017	0.054	0.114	0.177	0.180	0.176	0.172	0.162	0.158
15min_CP_13-21	0.102	0.112	0.119	0.136	0.156	0.238	0.259	0.258	0.252	0.258	0.245	0.232
15min_CP_8-8	0.006	0.016	0.016	0.049	0.079	0.150	0.272	0.259	0.248	0.242	0.218	0.199
15min_CP_b_55-30	0.000	0.010	0.027	0.086	0.169	0.384	0.285	0.291	0.285	0.285	0.289	0.278
15min_CP_b_18-4	0.132	0.146	0.179	0.203	0.226	0.372	0.250	0.250	0.239	0.243	0.235	0.235
15min_CP_b_7-10	0.000	0.050	0.076	0.142	0.228	0.449	0.259	0.242	0.230	0.222	0.206	0.191
15min_CP_b_2-30	0.191	0.201	0.205	0.238	0.331	0.534	0.282	0.258	0.248	0.227	0.222	0.203
15min_C_28-30	0.000	0.006	0.012	0.048	0.180	0.284	0.165	0.169	0.171	0.187	0.184	0.169
15min_C_3-22	0.070	0.092	0.107	0.175	0.361	0.470	0.114	0.103	0.102	0.098	0.106	0.110
15min_C_b_28-30	0.000	0.017	0.023	0.094	0.347	0.616	0.177	0.185	0.192	0.202	0.200	0.192
15min_C_b_24-2				0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.002
15min_C_b_8-19	0.127	0.165	0.187	0.332	0.571	0.707	0.196	0.195	0.191	0.191	0.191	0.184
15min_C_b_3-3	0.000	0.000	0.000	0.000	0.593	0.849	0.025	0.025	0.020	0.018	0.015	0.013
30min_CP_55-30	0.000	0.000	0.000	0.000	0.000	0.033	0.028	0.036	0.052	0.069	0.066	0.066
30min_CP_41-30	0.000	0.000	0.000	0.000	0.000	0.028	0.041	0.046	0.066	0.074	0.068	0.060
30min_CP_27-15	0.000	0.000	0.000	0.000	0.012	0.056	0.041	0.049	0.059	0.063	0.066	0.065
30min_CP_20-20	0.000	0.000	0.000	0.000	0.004	0.055	0.066	0.070	0.077	0.085	0.088	0.085
30min_CP_8-29	0.000	0.000	0.065	0.130	0.025	0.189	0.070	0.071	0.073	0.078	0.078	0.078
30min_CP_5-7	0.000	0.000	0.000	0.175	0.019	0.309	0.025	0.027	0.031	0.031	0.035	0.028
30min_CP_b_55-30	0.000	0.000	0.000	0.000	0.036	0.062	0.073	0.098	0.111	0.137	0.131	0.130
30min_CP_b_45-30	0.000	0.000	0.000	0.000	0.049	0.108	0.060	0.073	0.088	0.115	0.122	0.122
30min_CP_b_45-18	0.000	0.000	0.000	0.000	0.049	0.108	0.060	0.073	0.088	0.115	0.122	0.122
30min_CP_b_34-30	0.000	0.000	0.000	0.008	0.063	0.121	0.066	0.070	0.083	0.110	0.119	0.121
30min_C_28-28		0.000	0.000	0.000	0.000	0.063	0.000	0.005	0.006	0.013	0.011	0.007
30min_C_28-17	0.000	0.000	0.000	0.000	0.000	0.063	0.006	0.005	0.007	0.014	0.012	0.007
30min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.000	0.025	0.025	0.030	0.036	0.036	0.028
30min_C_b_15-30	0.000	0.000	0.000	0.000	0.000	0.062	0.013	0.013	0.013	0.024	0.026	0.027
30min_C_b_15-12	0.000	0.000	0.000	0.000	0.000	0.067	0.013	0.009	0.015	0.023	0.025	0.022
30min_C_b_13-12	0.000	0.000	0.000	0.000	0.000	0.051	0.022	0.024	0.025	0.029	0.032	0.026
60min_CP_3-30	0.000	0.000	0.000	0.000	0.000	0.056	0.009	0.006	0.009	0.008	0.008	0.010
60min_CP_3-15	0.000	0.000	0.000	0.000	0.000	0.056	0.009	0.006	0.009	0.008	0.008	0.010
60min_CP_b_55-30				0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.008	0.008
60min_CP_b_37-30	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.003	0.006	0.009	0.010
60min_CP_b_37-20	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.003	0.006	0.009	0.010
60min_CP_b_16-30	0.000	0.250	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.006	0.009	0.010
60min_CP_b_16-14	0.000	0.250	0.000	0.000	0.000	0.000	0.003	0.003	0.002	0.006	0.009	0.009
60min_C_4-27	1.000	1.000	1.000	0.792	1.000	0.833	0.000	0.000	0.000	0.002	0.000	0.002
60min_C_4-14	1.000	1.000	1.000	0.760	0.556	0.708	0.000	0.000	0.000	0.002	0.002	0.002
60min_C_b_28-30	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.002	0.002	0.004	0.004
60min_C_b_22-14	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.002	0.002	0.002	0.003	0.003
60min_C_b_6-30	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.005	0.006	0.005	0.005	0.006

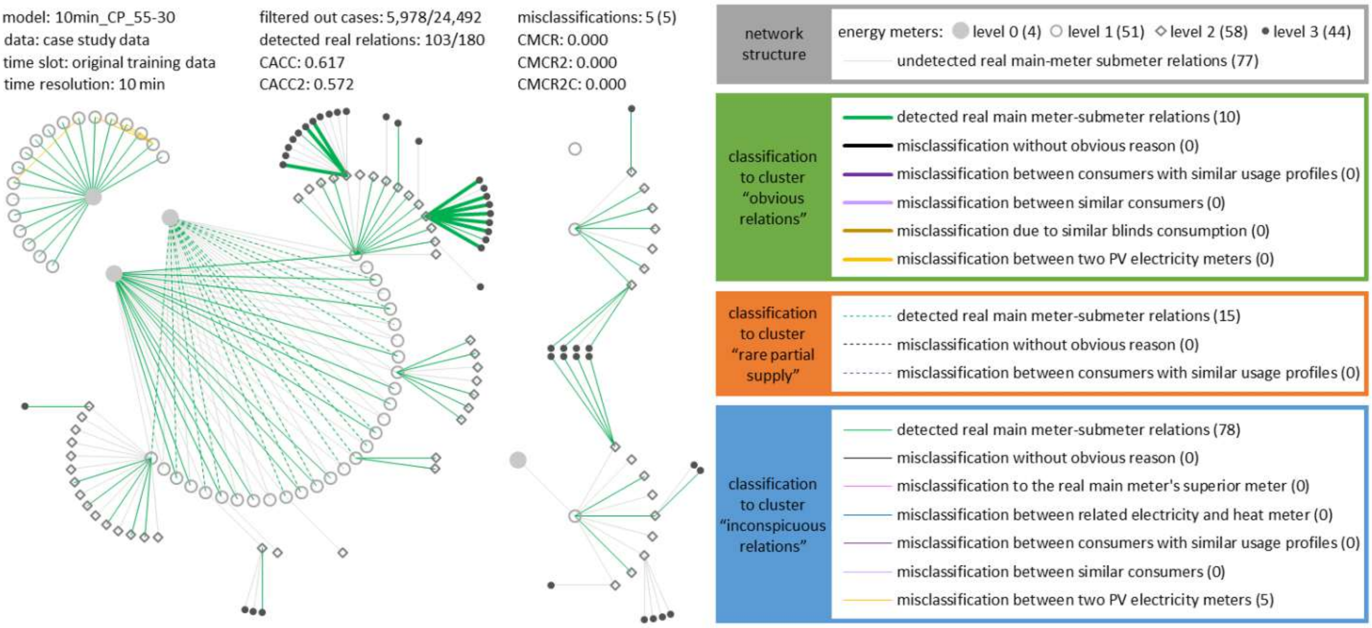
data: **synthetically generated data**		time resolution: **60 min**											
		mean CMCR2C						**mean CACC2**					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**05min_CP_55-30**	0.431	0.486	0.574	0.677	0.785	0.863	0.475	0.481	0.479	0.458	0.457	0.437
	05min_CP_18-30	0.630	0.664	0.693	0.733	0.806	0.881	0.383	0.372	0.366	0.347	0.347	0.332
	05min_CP_18-10	0.635	0.660	0.683	0.721	0.804	0.878	0.389	0.375	0.378	0.364	0.371	0.365
	05min_CP_9-30	0.820	0.823	0.835	0.845	0.893	0.934	0.373	0.381	0.376	0.381	0.377	0.388
	05min_CP_9-14	0.822	0.825	0.838	0.849	0.892	0.934	0.370	0.380	0.376	0.381	0.377	0.384
	05min_CP_7-30	0.780	0.783	0.793	0.810	0.863	0.915	0.367	0.367	0.365	0.362	0.356	0.356
	05min_CP_7-14	0.783	0.788	0.799	0.816	0.868	0.919	0.364	0.367	0.366	0.368	0.361	0.360
	05min_CP_b_55-30	0.822	0.827	0.839	0.877	0.919	0.947	0.478	0.479	0.485	0.475	0.488	0.491
	05min_CP_b_27-30	0.833	0.838	0.843	0.882	0.927	0.952	0.487	0.481	0.490	0.479	0.489	0.506
	05min_CP_b_27-13	0.831	0.837	0.843	0.883	0.927	0.952	0.487	0.484	0.488	0.478	0.488	0.506
	05min_CP_b_4-30	0.787	0.785	0.796	0.825	0.881	0.923	0.380	0.384	0.377	0.373	0.361	0.353
	05min_CP_b_4-13	0.791	0.792	0.802	0.830	0.886	0.925	0.383	0.380	0.373	0.369	0.358	0.355
	05min_C_28-30	0.522	0.581	0.620	0.703	0.811	0.880	0.421	0.416	0.424	0.422	0.445	0.472
	05min_C_7-30	0.781	0.787	0.799	0.825	0.873	0.922	0.358	0.362	0.359	0.350	0.357	0.358
	05min_C_7-13	0.781	0.789	0.799	0.823	0.868	0.918	0.361	0.359	0.357	0.346	0.351	0.353
	05min_C_5-30	0.764	0.772	0.786	0.802	0.847	0.903	0.326	0.334	0.326	0.323	0.323	0.324
	05min_C_5-14	0.769	0.778	0.786	0.808	0.852	0.906	0.339	0.337	0.339	0.329	0.334	0.328
	05min_C_b_28-30	0.809	0.818	0.828	0.870	0.925	0.954	0.440	0.438	0.442	0.438	0.460	0.488
	05min_C_b_9-30	0.850	0.851	0.855	0.877	0.932	0.957	0.399	0.402	0.403	0.414	0.434	0.474
	05min_C_b_9-15	0.851	0.852	0.856	0.877	0.932	0.958	0.399	0.402	0.404	0.415	0.436	0.475
	05min_C_b_7-30	0.854	0.852	0.855	0.874	0.929	0.954	0.392	0.402	0.402	0.415	0.431	0.474
	05min_C_b_7-17	0.852	0.850	0.852	0.872	0.928	0.954	0.392	0.402	0.402	0.415	0.431	0.473
	10min_CP_55-30	0.243	0.251	0.266	0.284	0.299	0.408	0.373	0.375	0.377	0.366	0.362	0.345
	10min_CP_40-30	0.245	0.265	0.275	0.298	0.337	0.485	0.380	0.378	0.381	0.369	0.362	0.346
	10min_CP_40-10	0.250	0.255	0.274	0.300	0.338	0.473	0.377	0.381	0.385	0.373	0.369	0.352
	10min_CP_17-30	0.375	0.406	0.411	0.488	0.556	0.662	0.392	0.389	0.390	0.382	0.388	0.384
	10min_CP_17-15	0.378	0.409	0.409	0.487	0.562	0.659	0.392	0.389	0.389	0.382	0.388	0.384
	10min_CP_12-30	0.371	0.406	0.414	0.482	0.568	0.686	0.396	0.394	0.391	0.384	0.388	0.388
	10min_CP_12-9	0.321	0.355	0.383	0.457	0.562	0.641	0.396	0.389	0.390	0.381	0.385	0.385
	10min_CP_b_55-30	0.668	0.685	0.707	0.741	0.800	0.880	0.383	0.394	0.396	0.388	0.396	0.382
	10min_CP_b_53-8	0.665	0.676	0.694	0.727	0.786	0.866	0.377	0.383	0.388	0.380	0.383	0.367
	10min_CP_b_26-5	0.692	0.704	0.719	0.746	0.803	0.877	0.373	0.378	0.382	0.369	0.372	0.366
	10min_CP_b_2-11	0.797	0.813	0.820	0.847	0.891	0.943	0.259	0.248	0.252	0.239	0.239	0.229
	10min_C_28-30	0.169	0.179	0.189	0.219	0.257	0.310	0.326	0.326	0.325	0.308	0.300	0.297
	10min_C_7-22	0.225	0.226	0.230	0.259	0.333	0.443	0.335	0.342	0.333	0.323	0.309	0.316
	10min_C_5-23	0.255	0.392	0.451	0.545	0.553	0.688	0.361	0.358	0.348	0.331	0.335	0.324
	10min_C_3-30	0.083	0.152	0.103	0.266	0.498	0.698	0.117	0.109	0.104	0.093	0.094	0.107
	10min_C_3-15	0.067	0.088	0.087	0.251	0.496	0.690	0.120	0.114	0.108	0.095	0.093	0.107
	10min_C_b_28-30	0.451	0.494	0.542	0.592	0.706	0.819	0.354	0.351	0.335	0.327	0.319	0.307
	10min_C_b_8-16	0.662	0.675	0.682	0.720	0.791	0.869	0.370	0.370	0.366	0.354	0.344	0.339
	10min_C_b_4-22	0.607	0.604	0.591	0.606	0.684	0.808	0.133	0.134	0.129	0.135	0.141	0.150
	10min_C_b_3-30	0.227	0.317	0.354	0.446	0.621	0.764	0.161	0.163	0.151	0.138	0.144	0.142

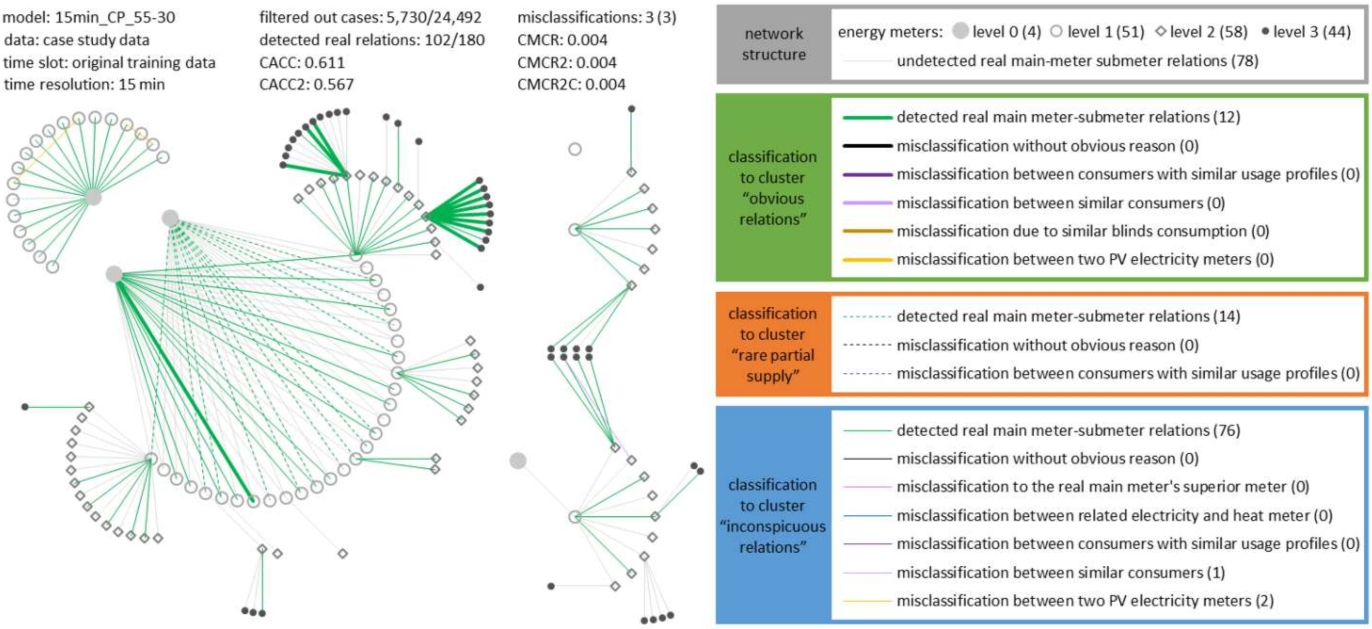
data: **synthetically generated data** time resolution: **60 min**

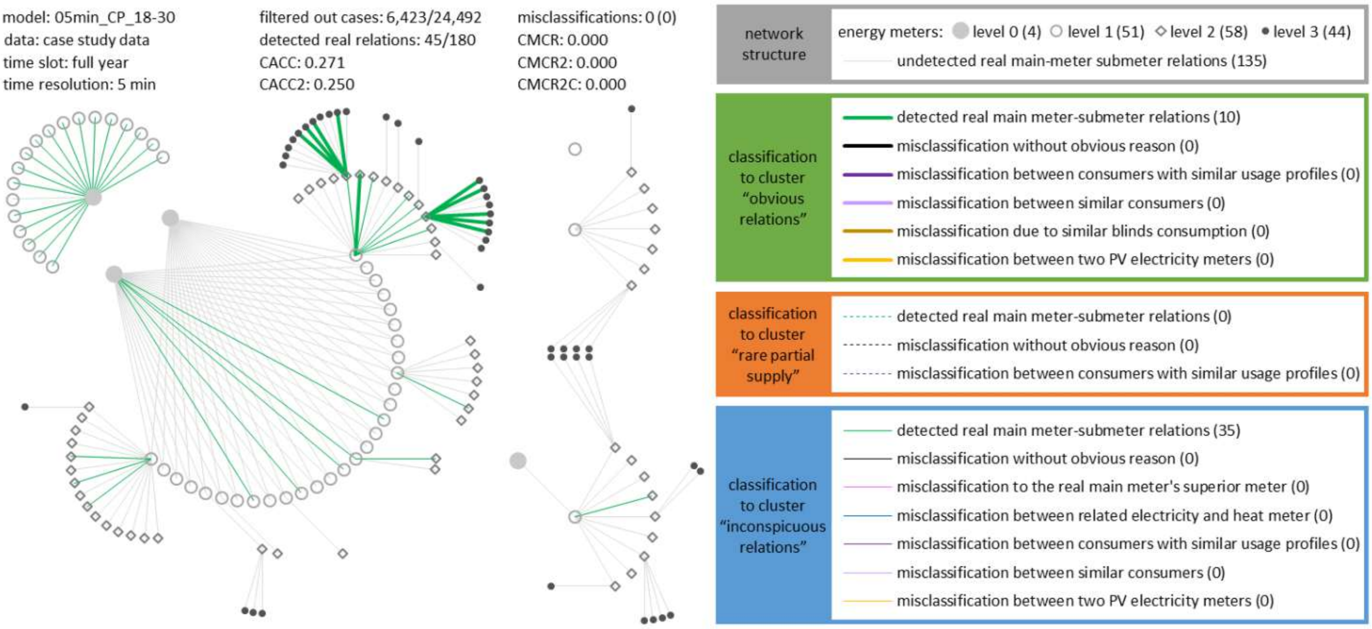
		mean CMCR2C						mean CACC2					
	time slot	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**	**full_year**	**6_months**	**3_months**	**5_weeks**	**2_weeks**	**1_week**
best classifier candidates	**15min_CP_55-30**	0.249	0.250	0.262	0.280	0.319	0.475	0.247	0.269	0.271	0.275	0.275	0.262
	15min_CP_53-5	0.234	0.237	0.247	0.269	0.307	0.405	0.212	0.236	0.237	0.245	0.238	0.236
	15min_CP_45-5	0.239	0.242	0.255	0.273	0.318	0.412	0.206	0.234	0.236	0.246	0.238	0.230
	15min_CP_25-13	0.246	0.240	0.251	0.275	0.356	0.477	0.285	0.302	0.307	0.303	0.296	0.284
	15min_CP_25-4	0.210	0.210	0.214	0.224	0.239	0.260	0.196	0.207	0.204	0.216	0.207	0.212
	15min_CP_13-21	0.339	0.374	0.383	0.432	0.516	0.557	0.304	0.310	0.294	0.282	0.286	0.276
	15min_CP_8-8	0.205	0.265	0.306	0.326	0.485	0.536	0.288	0.285	0.258	0.241	0.227	0.214
	15min_CP_b_55-30	0.547	0.579	0.599	0.646	0.716	0.813	0.307	0.304	0.307	0.301	0.307	0.295
	15min_CP_b_18-4	0.608	0.617	0.626	0.644	0.712	0.789	0.282	0.282	0.274	0.273	0.268	0.269
	15min_CP_b_7-10	0.589	0.585	0.621	0.682	0.760	0.859	0.278	0.267	0.254	0.241	0.229	0.219
	15min_CP_b_2-30	0.565	0.593	0.614	0.664	0.751	0.863	0.266	0.255	0.248	0.232	0.217	0.208
	15min_C_28-30	0.228	0.234	0.240	0.268	0.304	0.372	0.190	0.188	0.196	0.214	0.217	0.214
	15min_C_3-22	0.622	0.723	0.764	0.767	0.784	0.801	0.127	0.123	0.127	0.136	0.146	0.152
	15min_C_b_28-30	0.623	0.623	0.617	0.695	0.786	0.849	0.190	0.203	0.222	0.237	0.234	0.235
	15min_C_b_24-2					0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001
	15min_C_b_8-19	0.634	0.653	0.652	0.702	0.793	0.831	0.225	0.215	0.223	0.232	0.227	0.235
	15min_C_b_3-3	0.769	0.761	0.749	0.741	0.760	0.820	0.038	0.040	0.042	0.043	0.044	0.046
	30min_CP_55-30	0.224	0.229	0.227	0.221	0.227	0.289	0.057	0.071	0.085	0.103	0.112	0.119
	30min_CP_41-30	0.224	0.230	0.222	0.215	0.219	0.305	0.076	0.082	0.095	0.105	0.104	0.101
	30min_CP_27-15	0.225	0.226	0.232	0.272	0.320	0.397	0.076	0.084	0.098	0.108	0.125	0.131
	30min_CP_20-20	0.206	0.199	0.206	0.287	0.365	0.471	0.101	0.109	0.111	0.121	0.137	0.150
	30min_CP_8-29	0.431	0.390	0.464	0.584	0.593	0.573	0.104	0.119	0.120	0.123	0.136	0.146
	30min_CP_5-7	0.913	0.802	0.633	0.502	0.646	0.706	0.063	0.065	0.063	0.070	0.075	0.086
	30min_CP_b_55-30	0.479	0.475	0.482	0.468	0.493	0.564	0.120	0.125	0.143	0.172	0.180	0.178
	30min_CP_b_45-30	0.568	0.555	0.541	0.551	0.556	0.623	0.101	0.111	0.127	0.152	0.172	0.177
	30min_CP_b_45-18	0.568	0.555	0.541	0.550	0.556	0.622	0.101	0.111	0.127	0.152	0.172	0.177
	30min_CP_b_34-30	0.590	0.577	0.572	0.573	0.587	0.657	0.101	0.111	0.127	0.152	0.173	0.175
	30min_C_28-28	0.202	0.179	0.159	0.157	0.211	0.363	0.038	0.036	0.046	0.044	0.043	0.037
	30min_C_28-17	0.198	0.176	0.160	0.158	0.213	0.362	0.038	0.040	0.047	0.045	0.043	0.039
	30min_C_b_28-30	0.355	0.368	0.352	0.322	0.407	0.523	0.063	0.066	0.071	0.076	0.077	0.069
	30min_C_b_15-30	0.517	0.429	0.419	0.393	0.404	0.477	0.044	0.047	0.055	0.055	0.057	0.060
	30min_C_b_15-12	0.519	0.455	0.409	0.399	0.385	0.478	0.041	0.047	0.052	0.051	0.056	0.056
	30min_C_b_13-12	0.417	0.405	0.433	0.423	0.434	0.532	0.066	0.060	0.063	0.059	0.063	0.065
	60min_CP_3-30	0.667	0.424	0.605	0.637	0.744	0.736	0.025	0.027	0.028	0.029	0.028	0.031
	60min_CP_3-15	0.667	0.091	0.605	0.557	0.740	0.767	0.025	0.027	0.028	0.030	0.028	0.029
	60min_CP_b_55-30	0.000	0.000	0.000	0.021	0.094	0.175	0.003	0.009	0.011	0.025	0.030	0.035
	60min_CP_b_37-30	0.000	0.000	0.000	0.019	0.106	0.242	0.006	0.014	0.013	0.028	0.031	0.035
	60min_CP_b_37-20	0.000	0.000	0.000	0.019	0.106	0.242	0.006	0.014	0.013	0.028	0.031	0.035
	60min_CP_b_16-30	0.000	0.091	0.156	0.195	0.368	0.494	0.032	0.030	0.037	0.036	0.041	0.040
	60min_CP_b_16-14	0.000	0.091	0.143	0.198	0.378	0.505	0.032	0.028	0.037	0.035	0.038	0.039
	60min_C_4-27	0.600	0.536	0.613	0.492	0.563	0.514	0.006	0.009	0.006	0.006	0.005	0.006
	60min_C_4-14	0.600	0.536	0.564	0.467	0.604	0.510	0.006	0.009	0.006	0.006	0.005	0.007
	60min_C_b_28-30		0.000	0.000	0.000	0.000	0.125	0.000	0.003	0.002	0.003	0.006	0.005
	60min_C_b_22-14		0.000	0.000	0.000	0.000	0.125	0.000	0.003	0.004	0.004	0.005	0.004
	60min_C_b_6-30	0.000	0.000	0.000	0.000	0.175	0.558	0.016	0.011	0.012	0.013	0.008	0.009

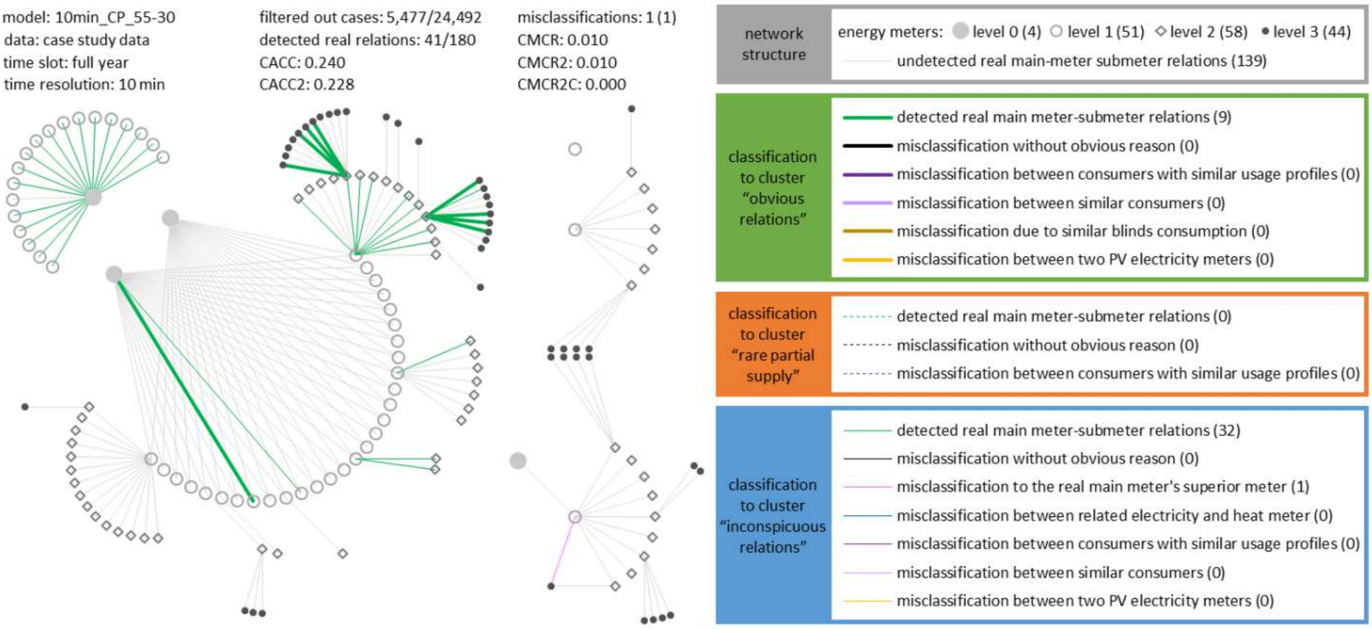
13 Appendix E: Examples of the final classifiers' classification performances

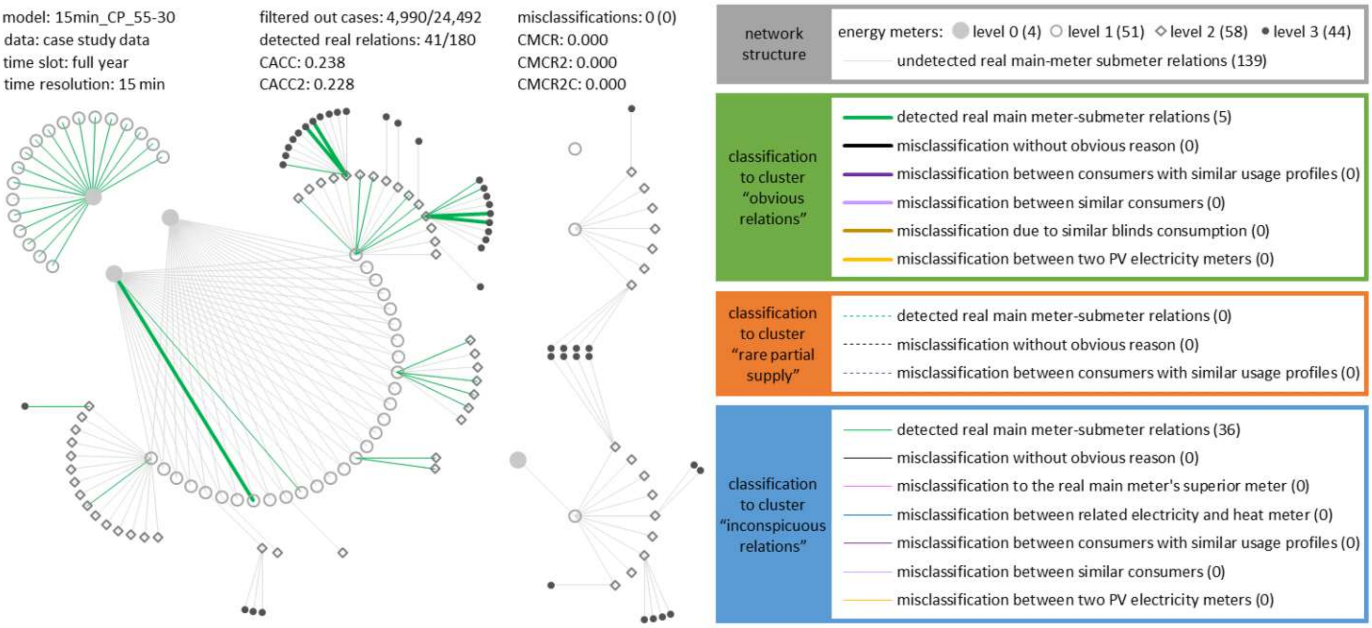


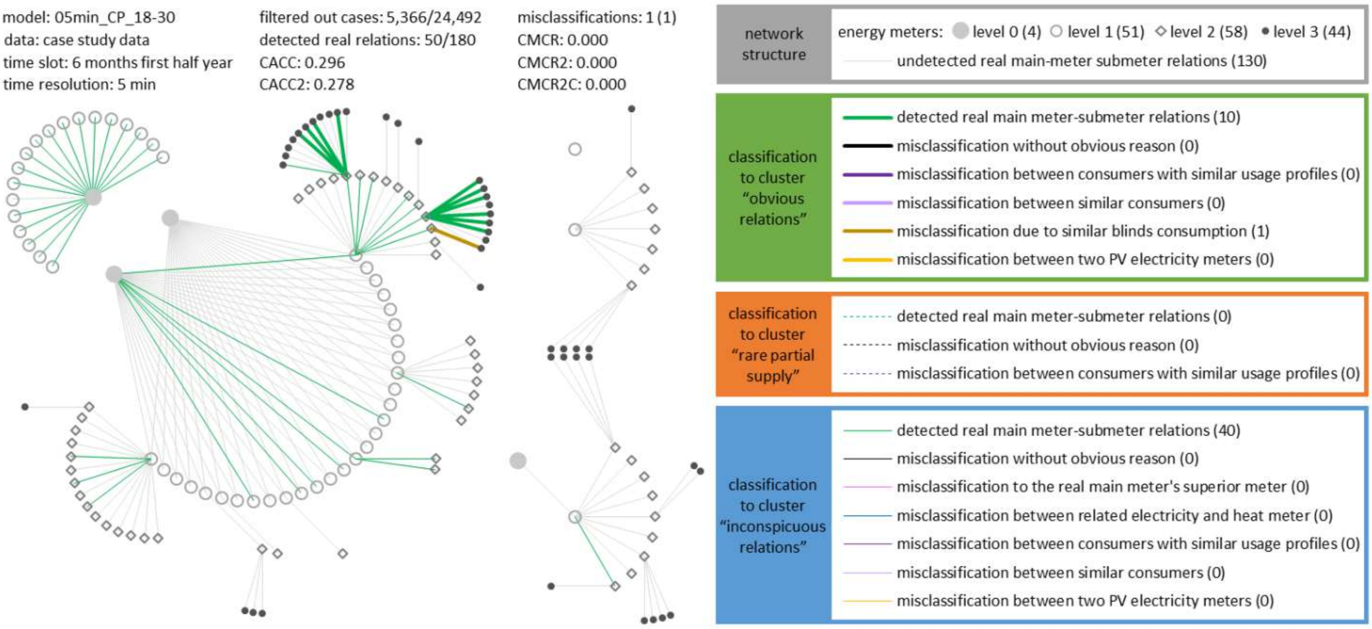


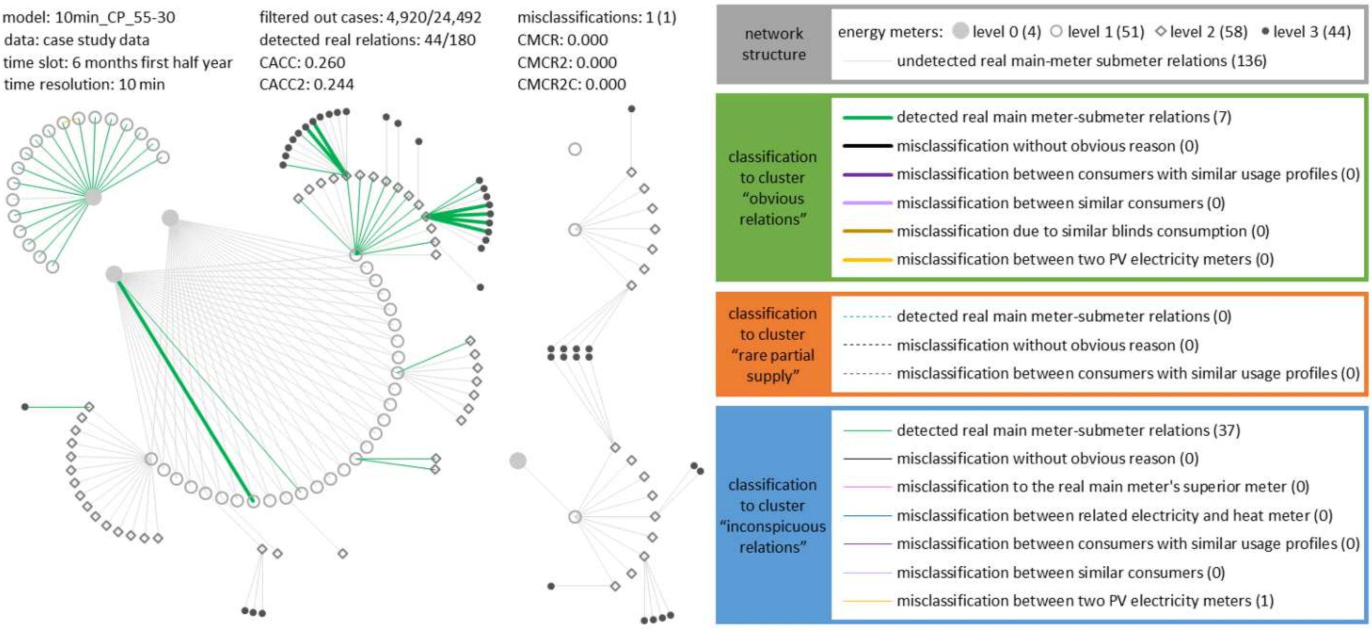


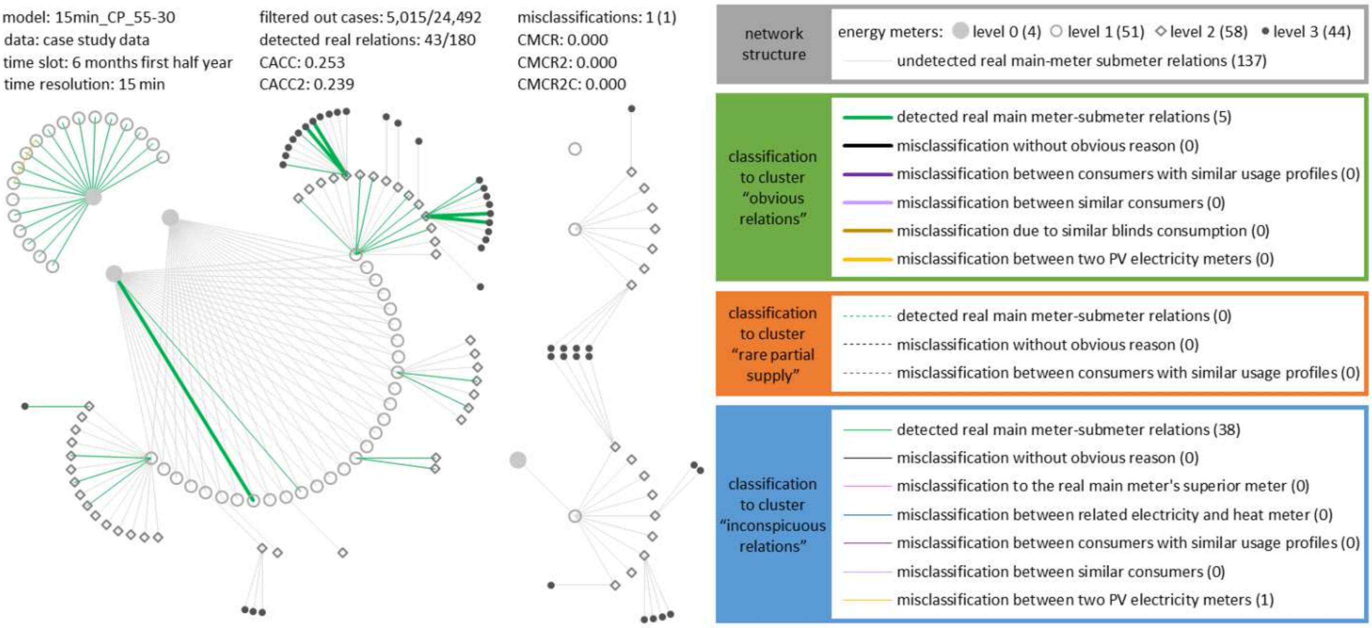

model: 05min_CP_18-30
data: case study data
time slot: full year
time resolution: 5 min
filtered out cases: 6,423/24,492
detected real relations: 45/180
CACC: 0.271
CACC2: 0.250
misclassifications: 0 (0)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (135)
classification to cluster "obvious relations"
detected real main meter-submeter relations (10)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (35)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (0)

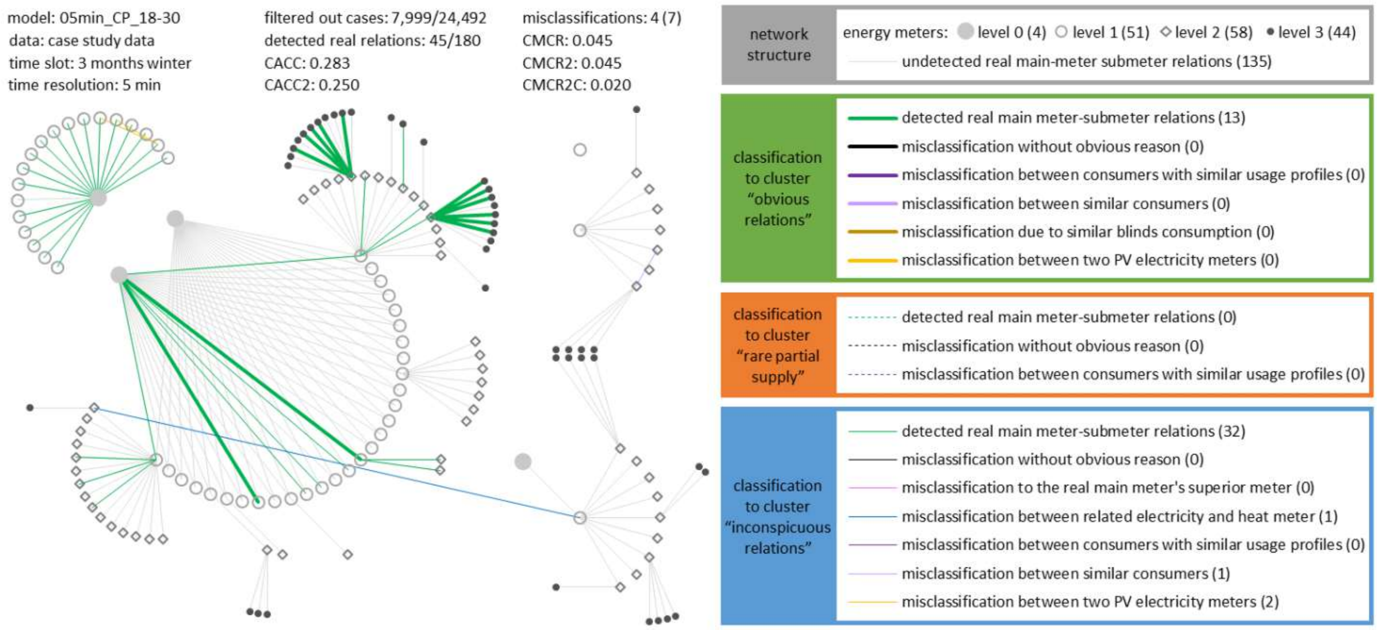

model: 10min_CP_55-30
data: case study data
time slot: full year
time resolution: 10 min
filtered out cases: 5,477/24,492
detected real relations: 41/180
CACC: 0.240
CACC2: 0.228
misclassifications: 1 (1)
CMCR: 0.010
CMCR2: 0.010
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (139)
classification to cluster "obvious relations"
detected real main meter-submeter relations (9)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (32)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (1)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (0)

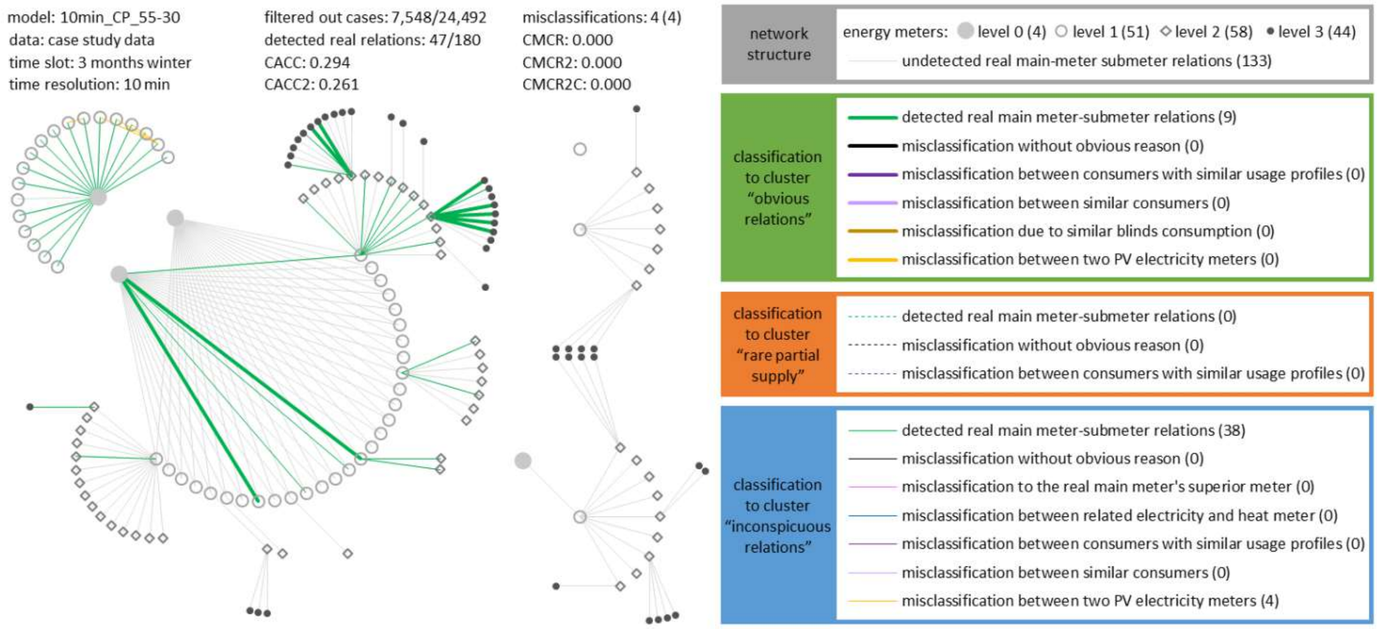

model: 15min_CP_55-30
data: case study data
time slot: full year
time resolution: 15 min
filtered out cases: 4,990/24,492
detected real relations: 41/180
CACC: 0.238
CACC2: 0.228
misclassifications: 0 (0)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (139)
classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (36)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (0)

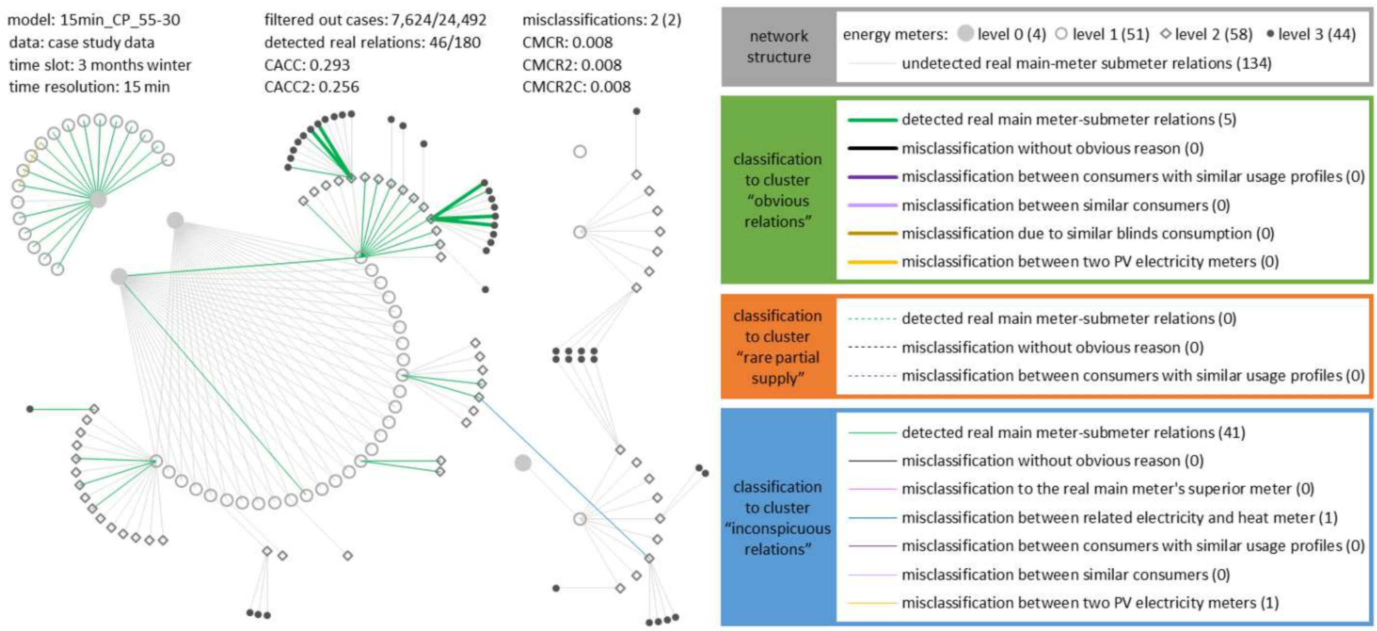

model: 05min_CP_18-30
data: case study data
time slot: 6 months first half year
time resolution: 5 min
filtered out cases: 5,366/24,492
detected real relations: 50/180
CACC: 0.296
CACC2: 0.278
misclassifications: 1 (1)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (130)
classification to cluster "obvious relations"
detected real main meter-submeter relations (10)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (1)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (40)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (0)

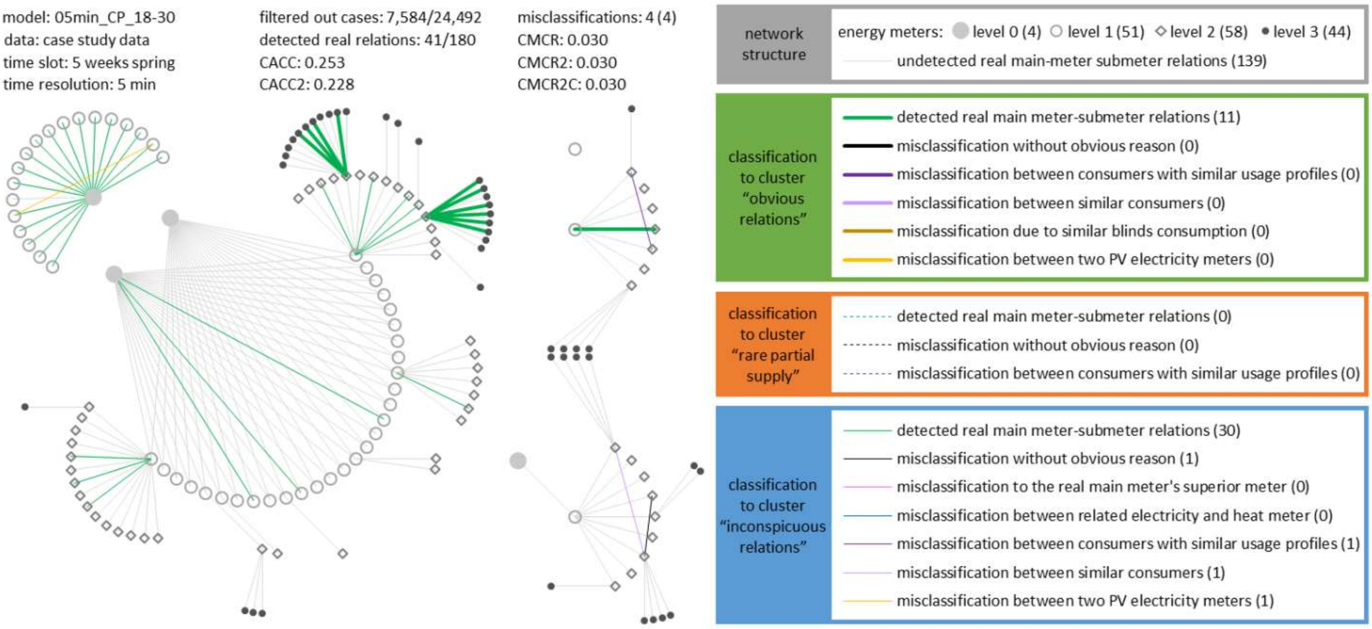

model: 10min_CP_55-30
data: case study data
time slot: 6 months first half year
time resolution: 10 min
filtered out cases: 4,920/24,492
detected real relations: 44/180
CACC: 0.260
CACC2: 0.244
misclassifications: 1 (1)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (136)
classification to cluster "obvious relations"
detected real main meter-submeter relations (7)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (37)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (1)

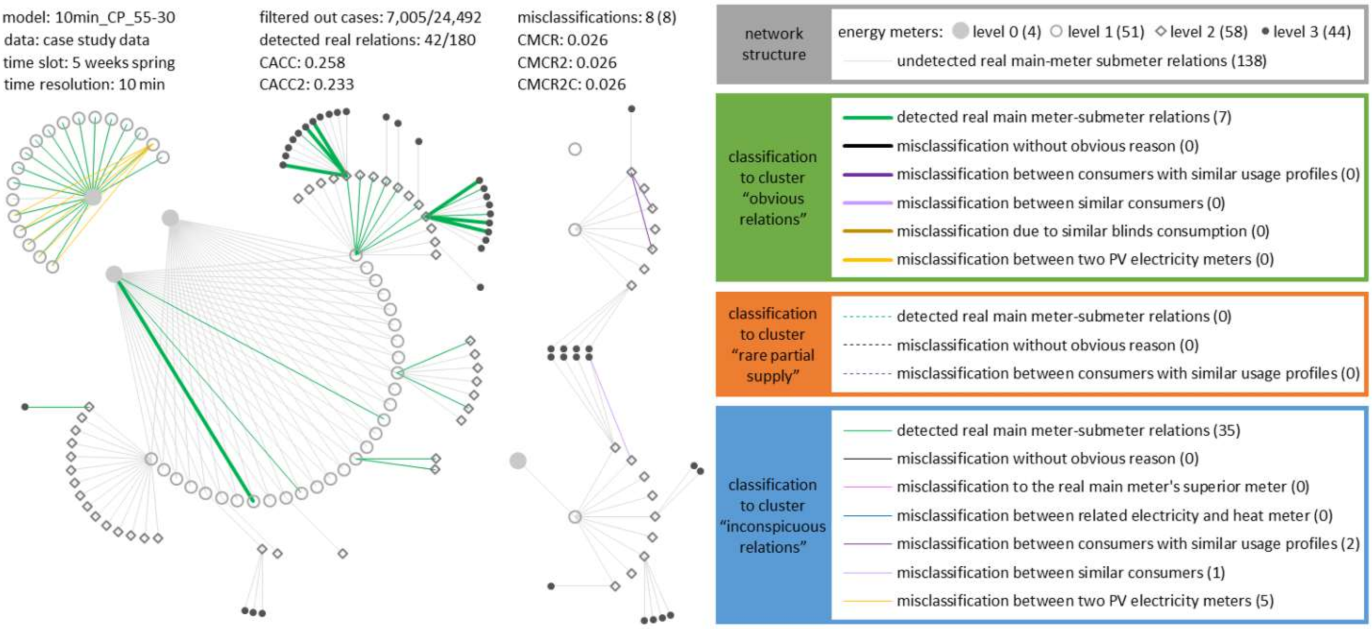

model: 15min_CP_55-30
data: case study data
time slot: 6 months first half year
time resolution: 15 min
filtered out cases: 5,015/24,492
detected real relations: 43/180
CACC: 0.253
CACC2: 0.239
misclassifications: 1 (1)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (137)
classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (38)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (1)

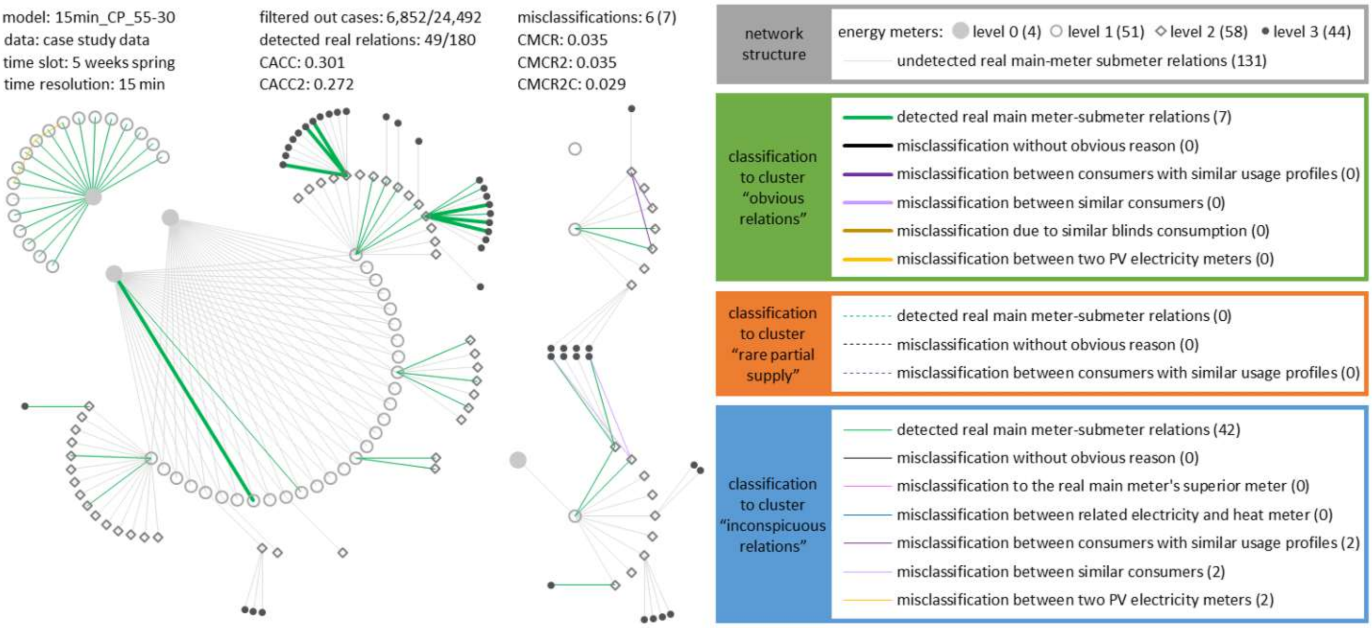

model: 05min_CP_18-30
data: case study data
time slot: 3 months winter
time resolution: 5 min
filtered out cases: 7,999/24,492
detected real relations: 45/180
CACC: 0.283
CACC2: 0.250
misclassifications: 4 (7)
CMCR: 0.045
CMCR2: 0.045
CMCR2C: 0.020
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (135)
classification to cluster "obvious relations"
detected real main meter-submeter relations (13)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (32)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (1)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (2)

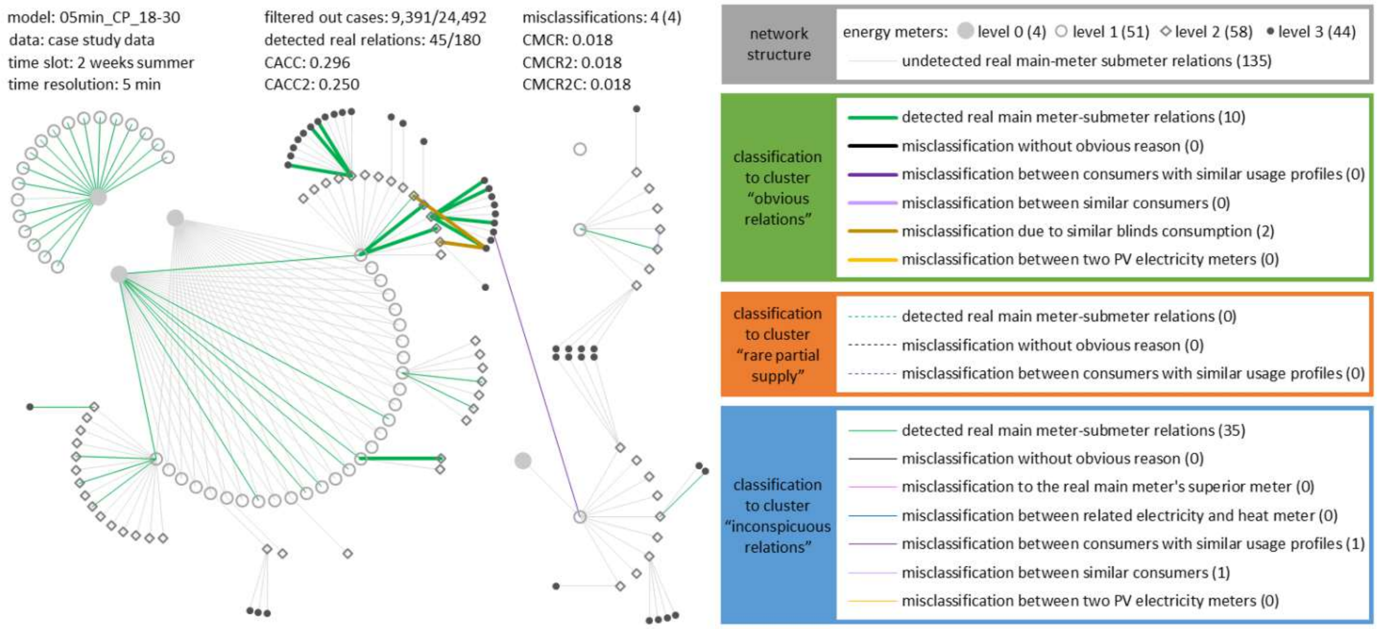

model: 10min_CP_55-30
data: case study data
time slot: 3 months winter
time resolution: 10 min
filtered out cases: 7,548/24,492
detected real relations: 47/180
CACC: 0.294
CACC2: 0.261
misclassifications: 4 (4)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (133)
classification to cluster "obvious relations"
detected real main meter-submeter relations (9)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (38)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (4)

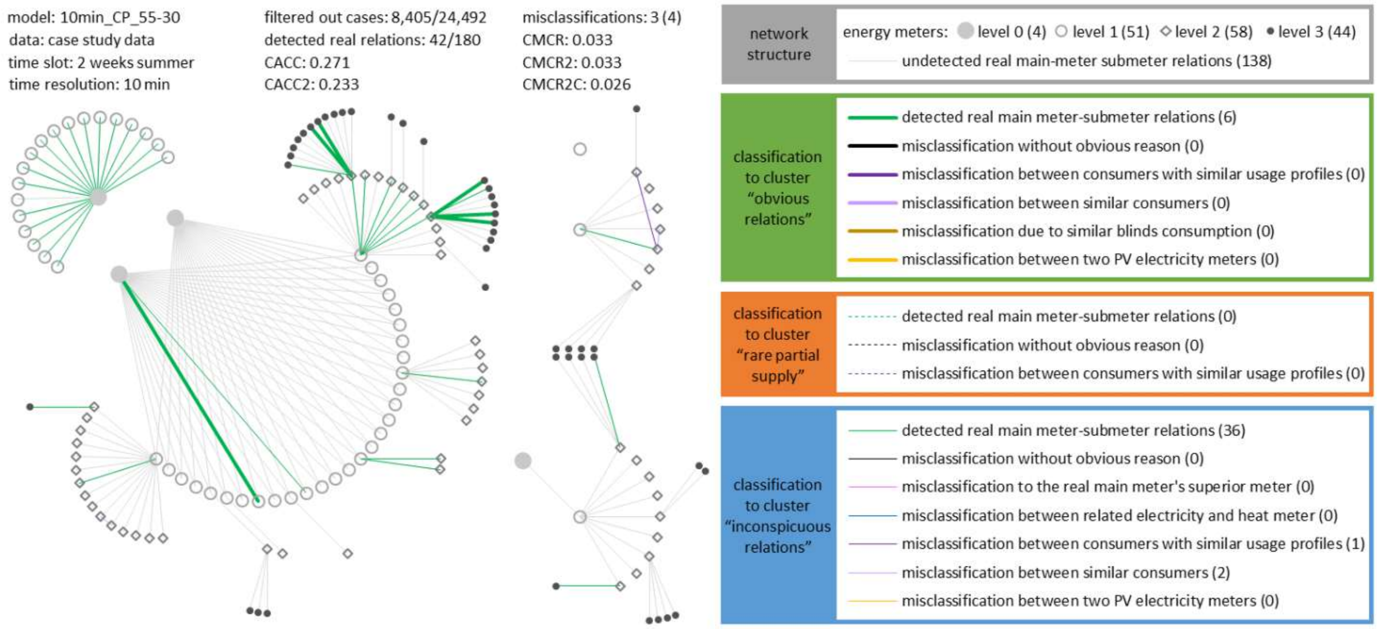

model: 15min_CP_55-30
data: case study data
time slot: 3 months winter
time resolution: 15 min
filtered out cases: 7,624/24,492
detected real relations: 46/180
CACC: 0.293
CACC2: 0.256
misclassifications: 2 (2)
CMCR: 0.008
CMCR2: 0.008
CMCR2C: 0.008
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (134)
classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (41)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (1)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (1)

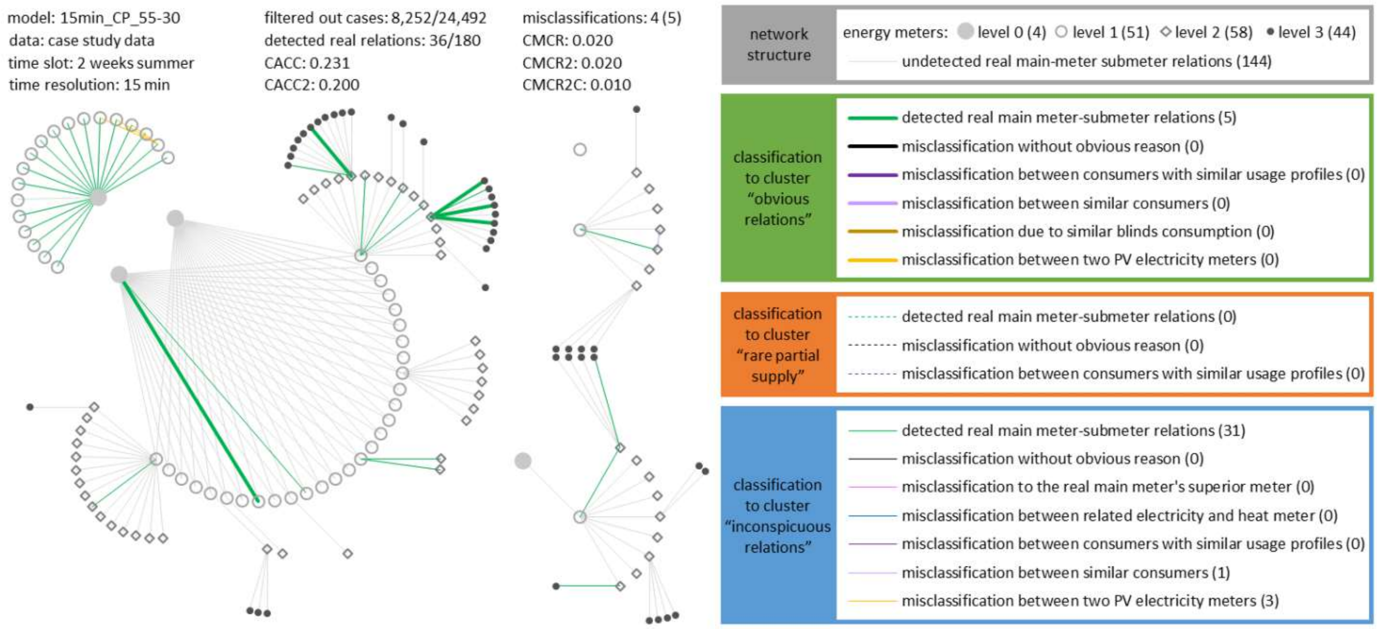

model: 05min_CP_18-30
data: case study data
time slot: 5 weeks spring
time resolution: 5 min
filtered out cases: 7,584/24,492
detected real relations: 41/180
CACC: 0.253
CACC2: 0.228
misclassifications: 4 (4)
CMCR: 0.030
CMCR2: 0.030
CMCR2C: 0.030
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (139)
classification to cluster "obvious relations"
detected real main meter-submeter relations (11)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (30)
misclassification without obvious reason (1)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (1)

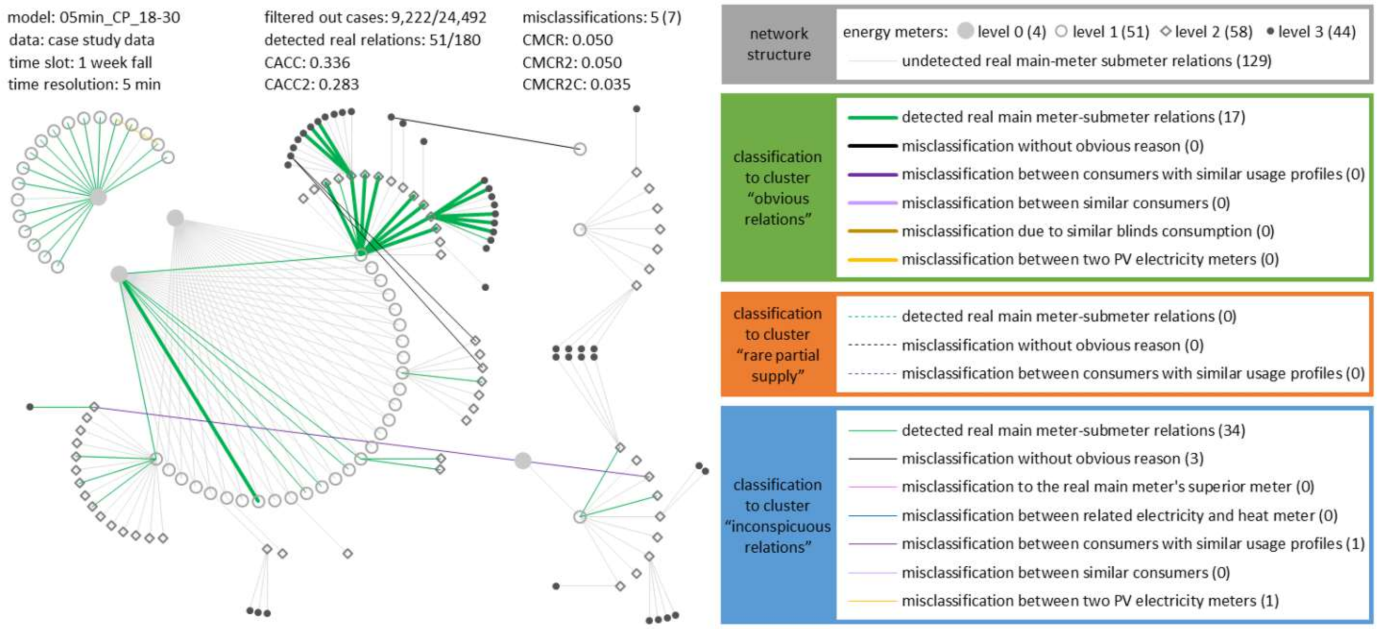

model: 10min_CP_55-30
data: case study data
time slot: 5 weeks spring
time resolution: 10 min
filtered out cases: 7,005/24,492
detected real relations: 42/180
CACC: 0.258
CACC2: 0.233
misclassifications: 8 (8)
CMCR: 0.026
CMCR2: 0.026
CMCR2C: 0.026
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (138)
classification to cluster "obvious relations"
detected real main meter-submeter relations (7)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (35)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (2)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (5)

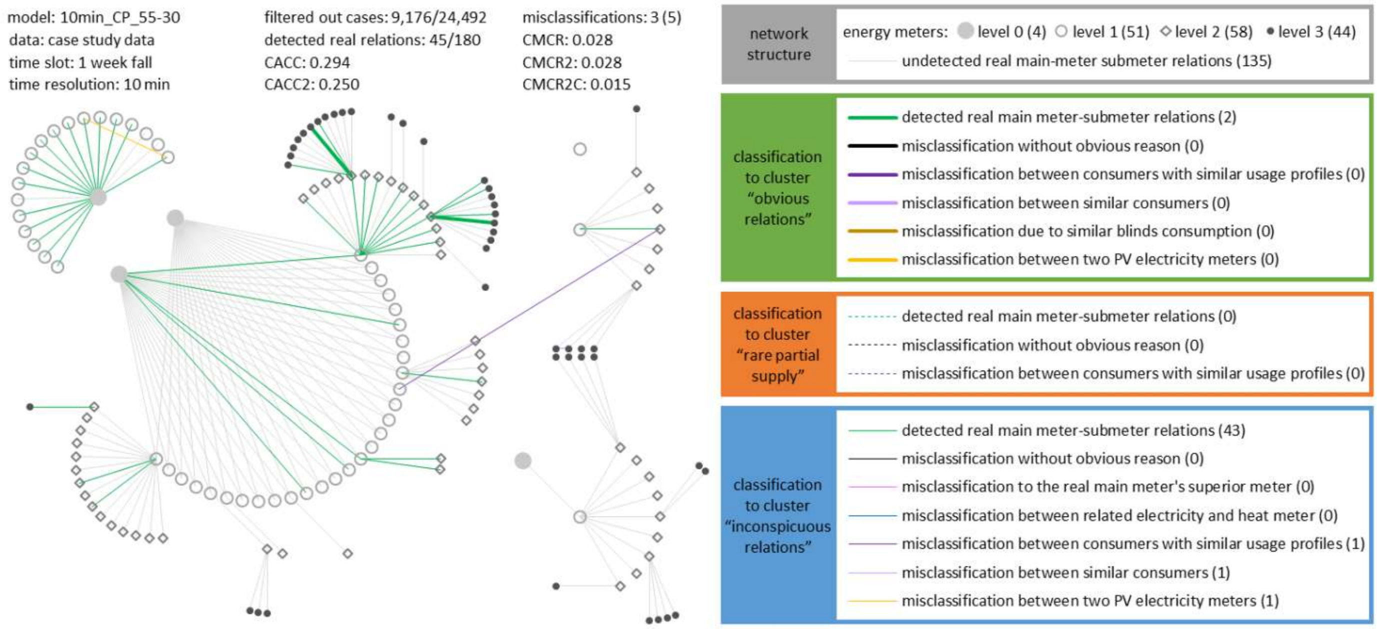

model: 15min_CP_55-30
data: case study data
time slot: 5 weeks spring
time resolution: 15 min
filtered out cases: 6,852/24,492
detected real relations: 49/180
CACC: 0.301
CACC2: 0.272
misclassifications: 6 (7)
CMCR: 0.035
CMCR2: 0.035
CMCR2C: 0.029
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (131)
classification to cluster "obvious relations"
detected real main meter-submeter relations (7)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (42)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (2)
misclassification between similar consumers (2)
misclassification between two PV electricity meters (2)

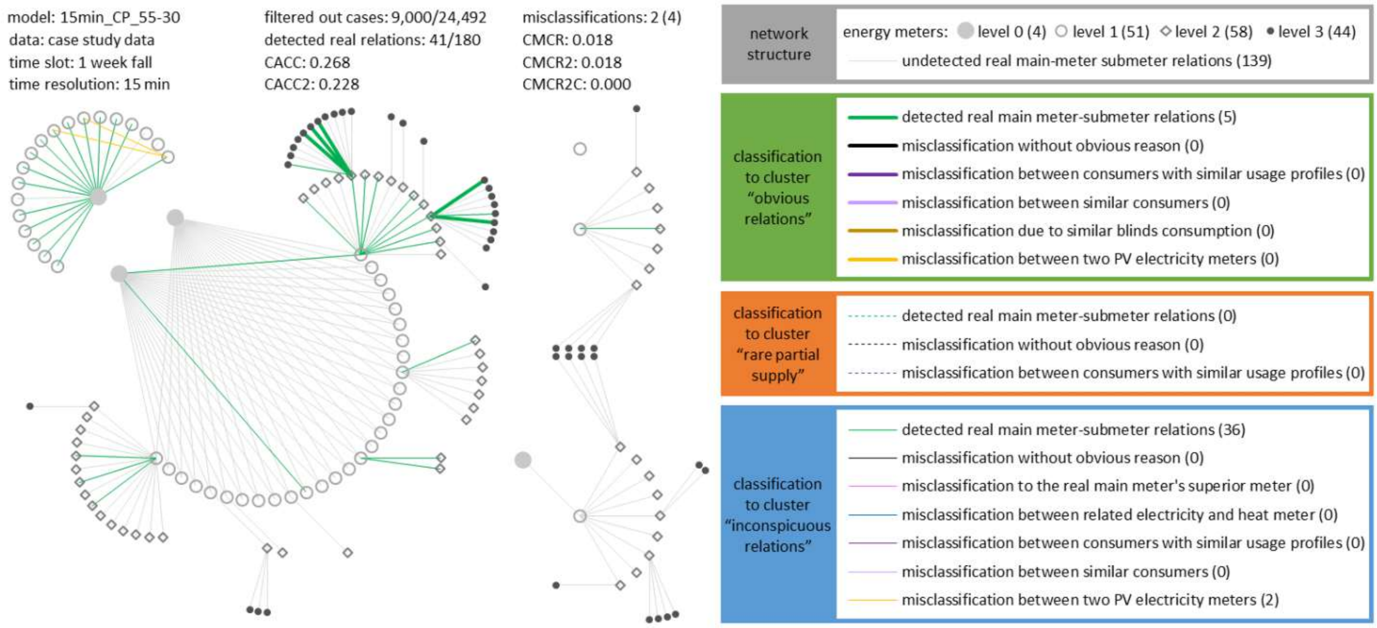

model: 05min_CP_18-30
data: case study data
time slot: 2 weeks summer
time resolution: 5 min
filtered out cases: 9,391/24,492
detected real relations: 45/180
CACC: 0.296
CACC2: 0.250
misclassifications: 4 (4)
CMCR: 0.018
CMCR2: 0.018
CMCR2C: 0.018
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (135)
classification to cluster "obvious relations"
detected real main meter-submeter relations (10)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (2)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (35)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (0)


model: 10min_CP_55-30
data: case study data
time slot: 2 weeks summer
time resolution: 10 min
filtered out cases: 8,405/24,492
detected real relations: 42/180
CACC: 0.271
CACC2: 0.233
misclassifications: 3 (4)
CMCR: 0.033
CMCR2: 0.033
CMCR2C: 0.026
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (138)
classification to cluster "obvious relations"
detected real main meter-submeter relations (6)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (36)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (2)
misclassification between two PV electricity meters (0)


model: 15min_CP_55-30
data: case study data
time slot: 2 weeks summer
time resolution: 15 min
filtered out cases: 8,252/24,492
detected real relations: 36/180
CACC: 0.231
CACC2: 0.200
misclassifications: 4 (5)
CMCR: 0.020
CMCR2: 0.020
CMCR2C: 0.010
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (144)
classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (31)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (3)


model: 05min_CP_18-30
data: case study data
time slot: 1 week fall
time resolution: 5 min
filtered out cases: 9,222/24,492
detected real relations: 51/180
CACC: 0.336
CACC2: 0.283
misclassifications: 5 (7)
CMCR: 0.050
CMCR2: 0.050
CMCR2C: 0.035
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (129)
classification to cluster "obvious relations"
detected real main meter-submeter relations (17)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (34)
misclassification without obvious reason (3)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (1)

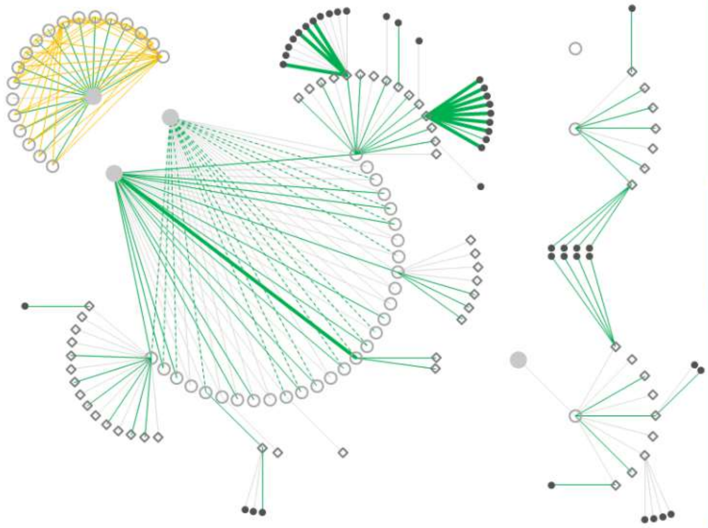

model: 10min_CP_55-30
data: case study data
time slot: 1 week fall
time resolution: 10 min
filtered out cases: 9,176/24,492
detected real relations: 45/180
CACC: 0.294
CACC2: 0.250
misclassifications: 3 (5)
CMCR: 0.028
CMCR2: 0.028
CMCR2C: 0.015
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (135)
classification to cluster "obvious relations"
detected real main meter-submeter relations (2)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (43)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (1)


model: 15min_CP_55-30
data: case study data
time slot: 1 week fall
time resolution: 15 min
filtered out cases: 9,000/24,492
detected real relations: 41/180
CACC: 0.268
CACC2: 0.228
misclassifications: 2 (4)
CMCR: 0.018
CMCR2: 0.018
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (139)
classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (36)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (2)

model: 05min_C_7-30
data: case study data
time slot: original training data
time resolution: 5 min

filtered out cases: 6,524/24,492
detected real relations: 106/180
CACC: 0.631
CACC2: 0.589

misclassifications: 41 (41)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000



network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (74)

classification to cluster "obvious relations"
detected real main meter-submeter relations (13)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)

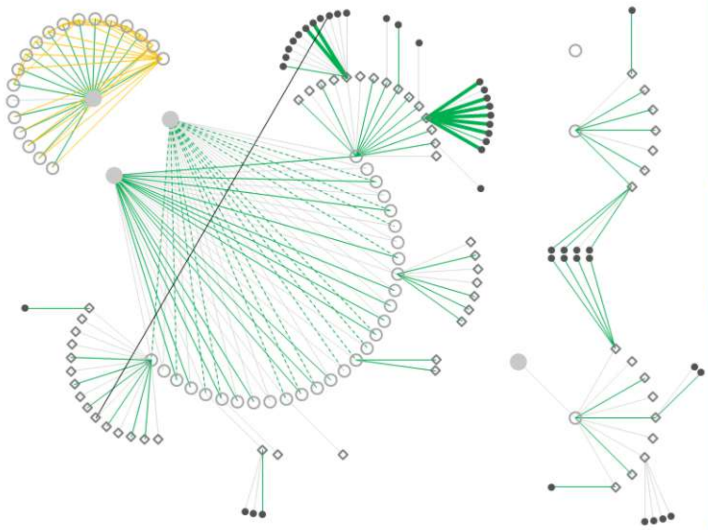
classification to cluster "rare partial supply"
detected real main meter-submeter relations (15)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (78)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (41)

model: 10min_C_28-30
data: case study data
time slot: original training data
time resolution: 10 min

filtered out cases: 5,978/24,492
detected real relations: 103/180
CACC: 0.617
CACC2: 0.572

misclassifications: 31 (31)
CMCR: 0.004
CMCR2: 0.004
CMCR2C: 0.004



network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (77)

classification to cluster "obvious relations"
detected real main meter-submeter relations (9)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)

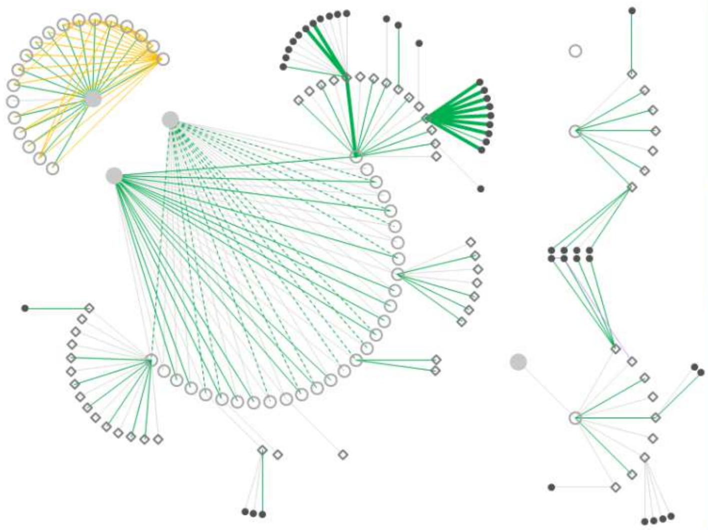
classification to cluster "rare partial supply"
detected real main meter-submeter relations (15)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (79)
misclassification without obvious reason (1)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (30)

model: 15min_C_28-30
data: case study data
time slot: original training data
time resolution: 15 min

filtered out cases: 5,730/24,492
detected real relations: 101/180
CACC: 0.605
CACC2: 0.561

misclassifications: 31 (31)
CMCR: 0.007
CMCR2: 0.007
CMCR2C: 0.007

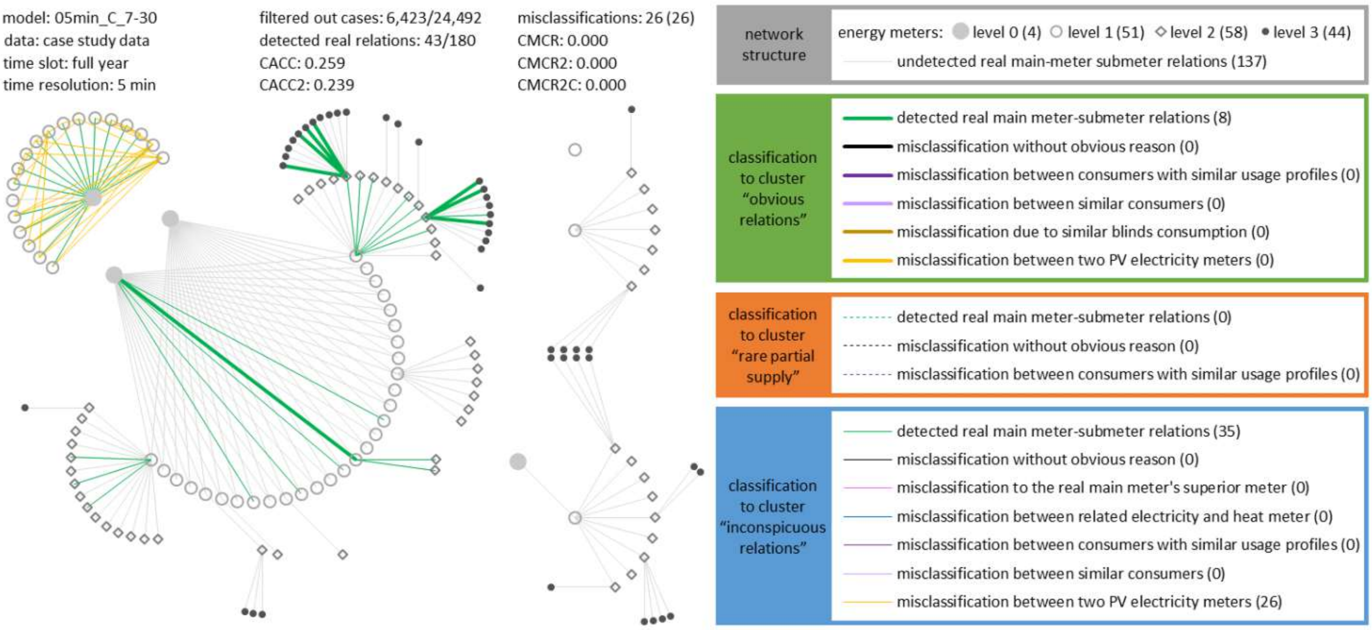


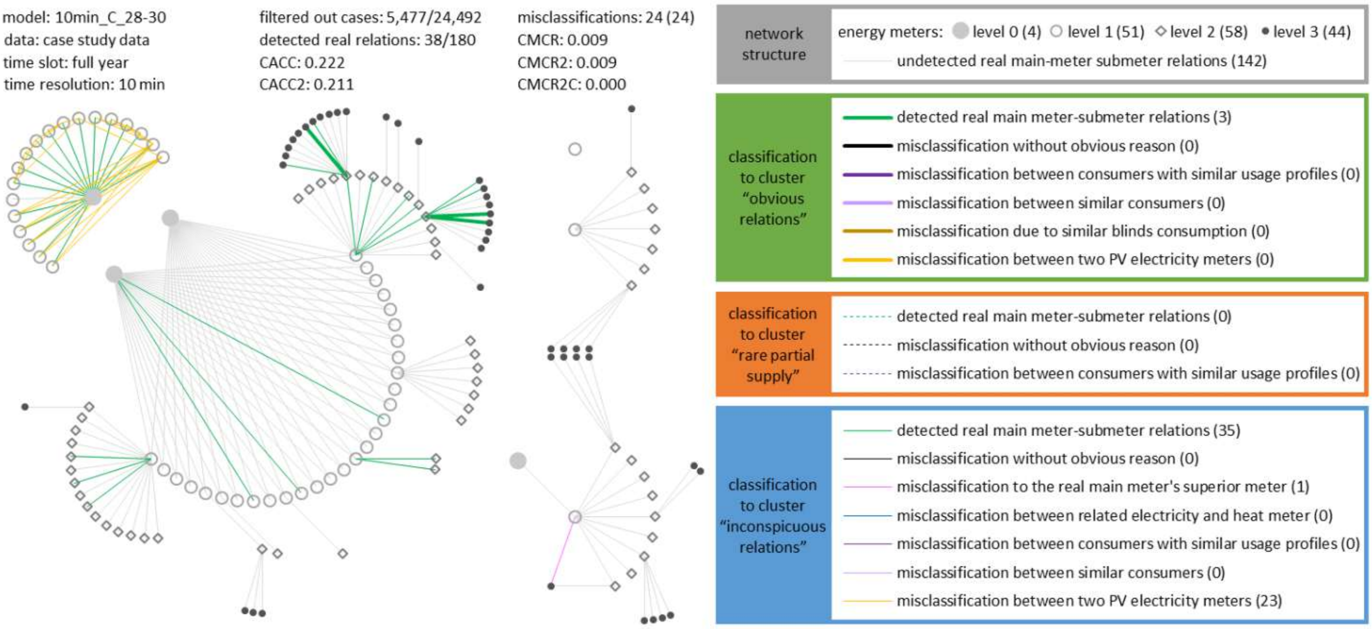
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (79)

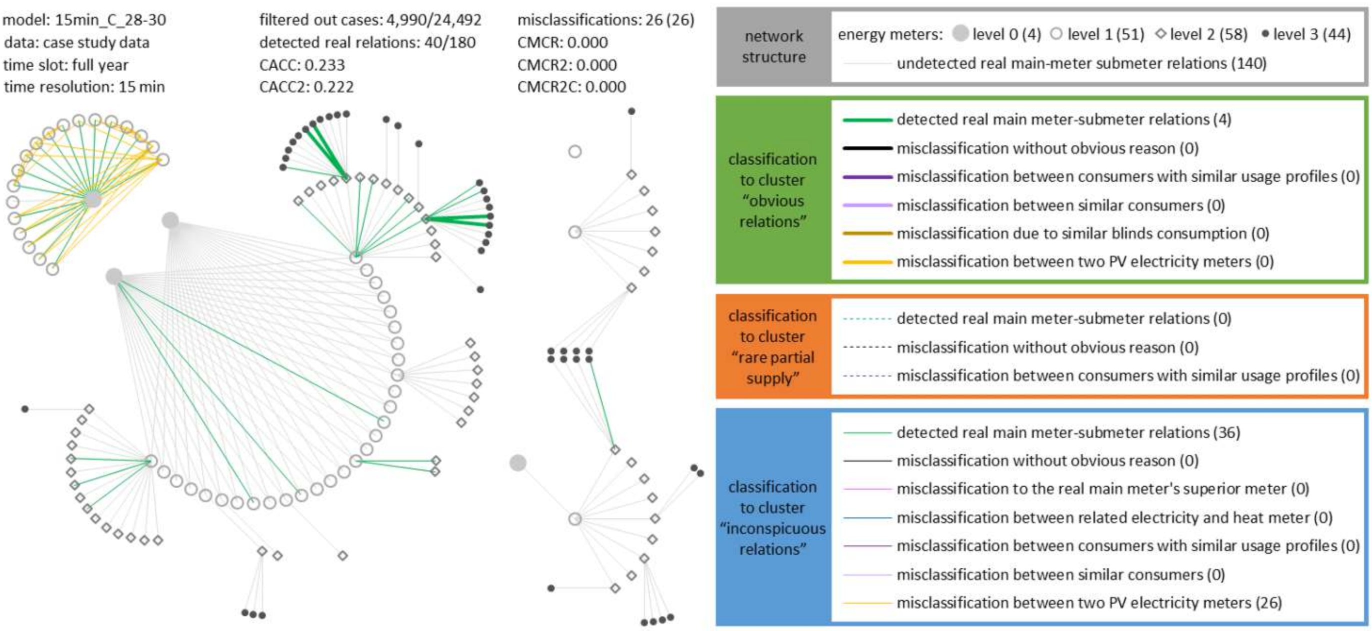
classification to cluster "obvious relations"
detected real main meter-submeter relations (11)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)

classification to cluster "rare partial supply"
detected real main meter-submeter relations (14)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (76)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (2)
misclassification between two PV electricity meters (29)


model: 05min_C_7-30
data: case study data
time slot: full year
time resolution: 5 min
filtered out cases: 6,423/24,492
detected real relations: 43/180
CACC: 0.259
CACC2: 0.239
misclassifications: 26 (26)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (137)
classification to cluster "obvious relations"
detected real main meter-submeter relations (8)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (35)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (26)

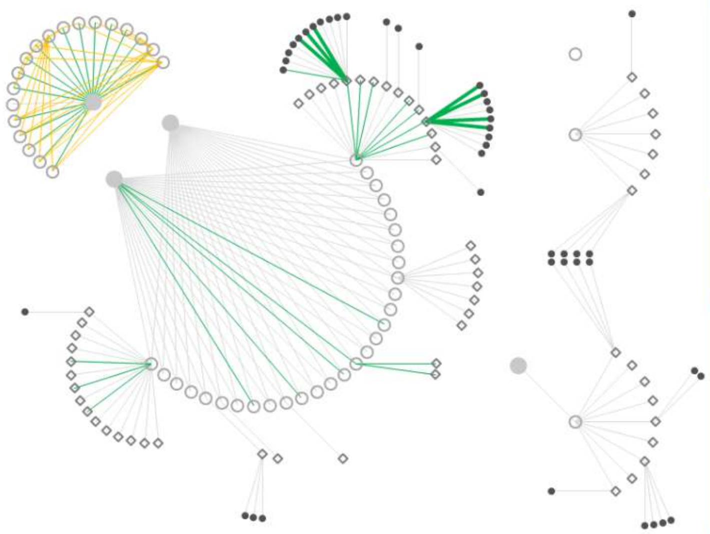

model: 10min_C_28-30
data: case study data
time slot: full year
time resolution: 10 min
filtered out cases: 5,477/24,492
detected real relations: 38/180
CACC: 0.222
CACC2: 0.211
misclassifications: 24 (24)
CMCR: 0.009
CMCR2: 0.009
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (142)
classification to cluster "obvious relations"
detected real main meter-submeter relations (3)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (35)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (1)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (23)


model: 15min_C_28-30
data: case study data
time slot: full year
time resolution: 15 min
filtered out cases: 4,990/24,492
detected real relations: 40/180
CACC: 0.233
CACC2: 0.222
misclassifications: 26 (26)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (140)
classification to cluster "obvious relations"
detected real main meter-submeter relations (4)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (36)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (26)

model: 05min_C_7-30
data: case study data
time slot: 6 months first half year
time resolution: 5 min

filtered out cases: 5,366/24,492
detected real relations: 43/180
CACC: 0.254
CACC2: 0.239

misclassifications: 26 (26)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000



network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (137)

classification to cluster "obvious relations"
detected real main meter-submeter relations (7)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)

classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (36)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (26)

model: 10min_C_28-30
data: case study data
time slot: 6 months first half year
time resolution: 10 min

filtered out cases: 4,920/24,492
detected real relations: 44/180
CACC: 0.260
CACC2: 0.244

misclassifications: 22 (22)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000

network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (136)

classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)

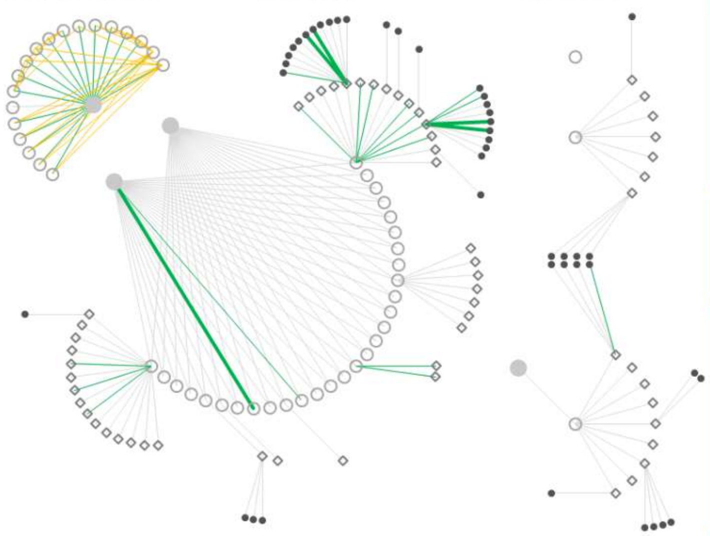
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (39)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (22)

model: 15min_C_28-30
data: case study data
time slot: 6 months first half year
time resolution: 15 min

filtered out cases: 5,015/24,492
detected real relations: 40/180
CACC: 0.235
CACC2: 0.222

misclassifications: 27 (27)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000



network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (140)

classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)

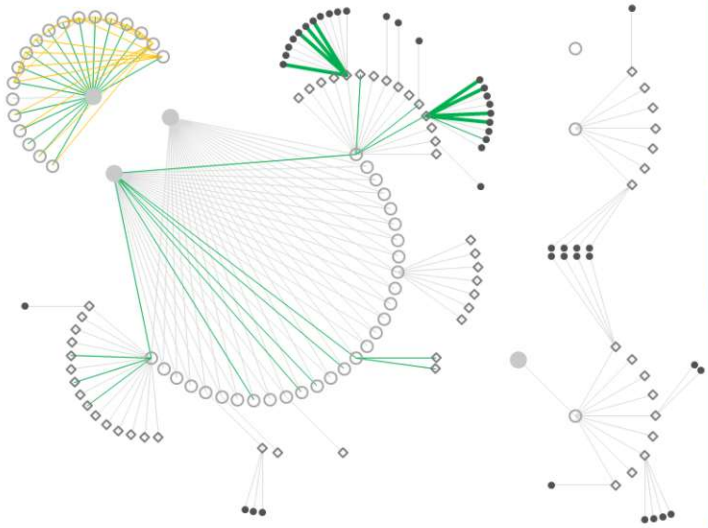
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (35)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (27)

model: 05min_C_7-30
data: case study data
time slot: 3 months winter
time resolution: 5 min

filtered out cases: 7,999/24,492
detected real relations: 42/180
CACC: 0.264
CACC2: 0.233

misclassifications: 23 (26)
CMCR: 0.093
CMCR2: 0.093
CMCR2C: 0.000



network structure
- energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
- undetected real main-meter submeter relations (138)

classification to cluster "obvious relations"
- detected real main meter-submeter relations (8)
- misclassification without obvious reason (0)
- misclassification between consumers with similar usage profiles (0)
- misclassification between similar consumers (0)
- misclassification due to similar blinds consumption (0)
- misclassification between two PV electricity meters (0)

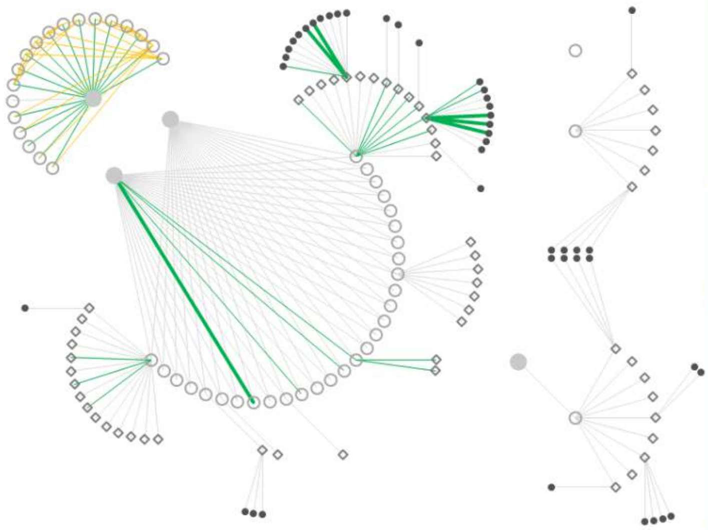
classification to cluster "rare partial supply"
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)
- misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
- detected real main meter-submeter relations (34)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)
- misclassification between related electricity and heat meter (0)
- misclassification between consumers with similar usage profiles (0)
- misclassification between similar consumers (0)
- misclassification between two PV electricity meters (23)

model: 10min_C_28-30
data: case study data
time slot: 3 months winter
time resolution: 10 min

filtered out cases: 7,548/24,492
detected real relations: 42/180
CACC: 0.263
CACC2: 0.233

misclassifications: 24 (24)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000



network structure
- energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
- undetected real main-meter submeter relations (138)

classification to cluster "obvious relations"
- detected real main meter-submeter relations (6)
- misclassification without obvious reason (0)
- misclassification between consumers with similar usage profiles (0)
- misclassification between similar consumers (0)
- misclassification due to similar blinds consumption (0)
- misclassification between two PV electricity meters (0)

classification to cluster "rare partial supply"
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)
- misclassification between consumers with similar usage profiles (0)

classification to cluster "inconspicuous relations"
- detected real main meter-submeter relations (36)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)
- misclassification between related electricity and heat meter (0)
- misclassification between consumers with similar usage profiles (0)
- misclassification between similar consumers (0)
- misclassification between two PV electricity meters (24)

model: 15min_C_28-30
data: case study data
time slot: 3 months winter
time resolution: 15 min

filtered out cases: 7,624/24,492
detected real relations: 42/180
CACC: 0.268
CACC2: 0.233

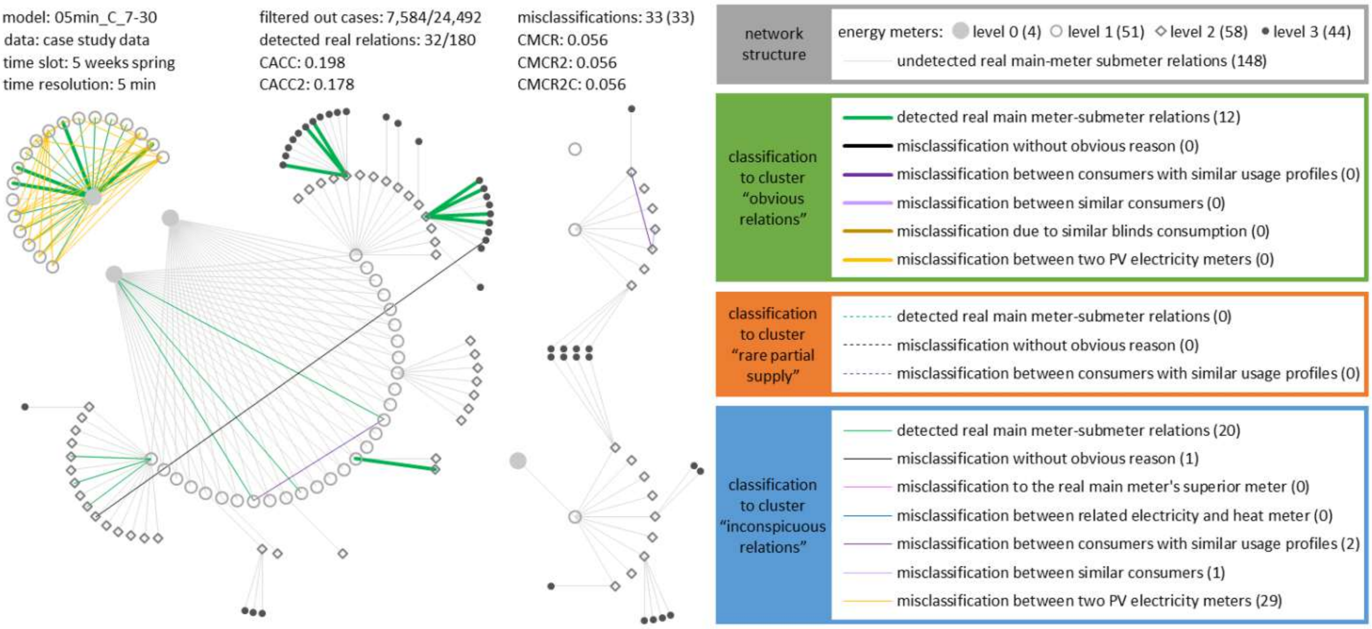
misclassifications: 20 (20)
CMCR: 0.000
CMCR2: 0.000
CMCR2C: 0.000

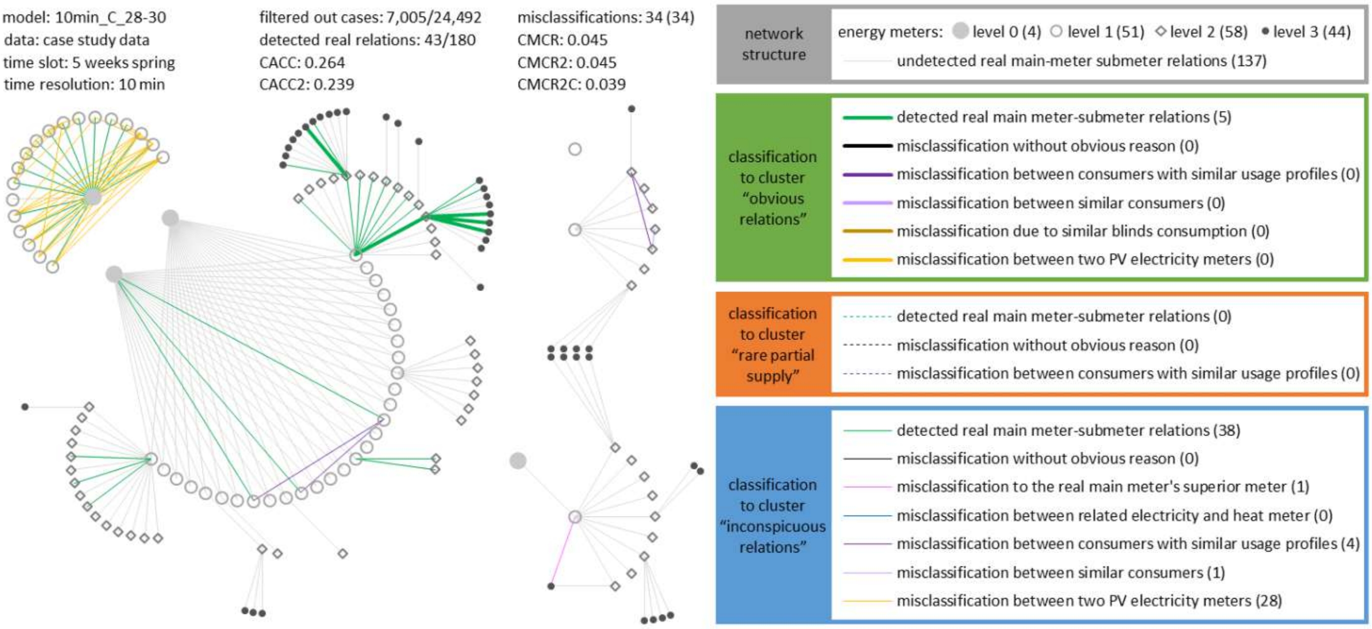
network structure
- energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
- undetected real main-meter submeter relations (138)

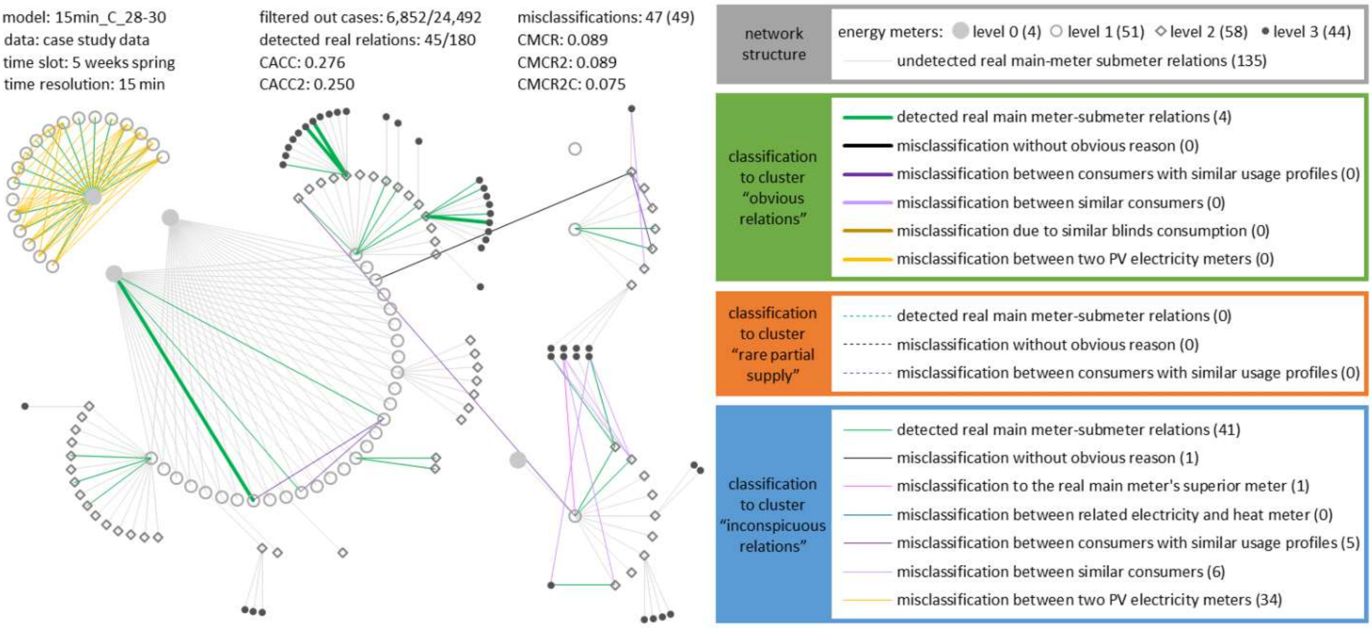
classification to cluster "obvious relations"
- detected real main meter-submeter relations (5)
- misclassification without obvious reason (0)
- misclassification between consumers with similar usage profiles (0)
- misclassification between similar consumers (0)
- misclassification due to similar blinds consumption (0)
- misclassification between two PV electricity meters (0)

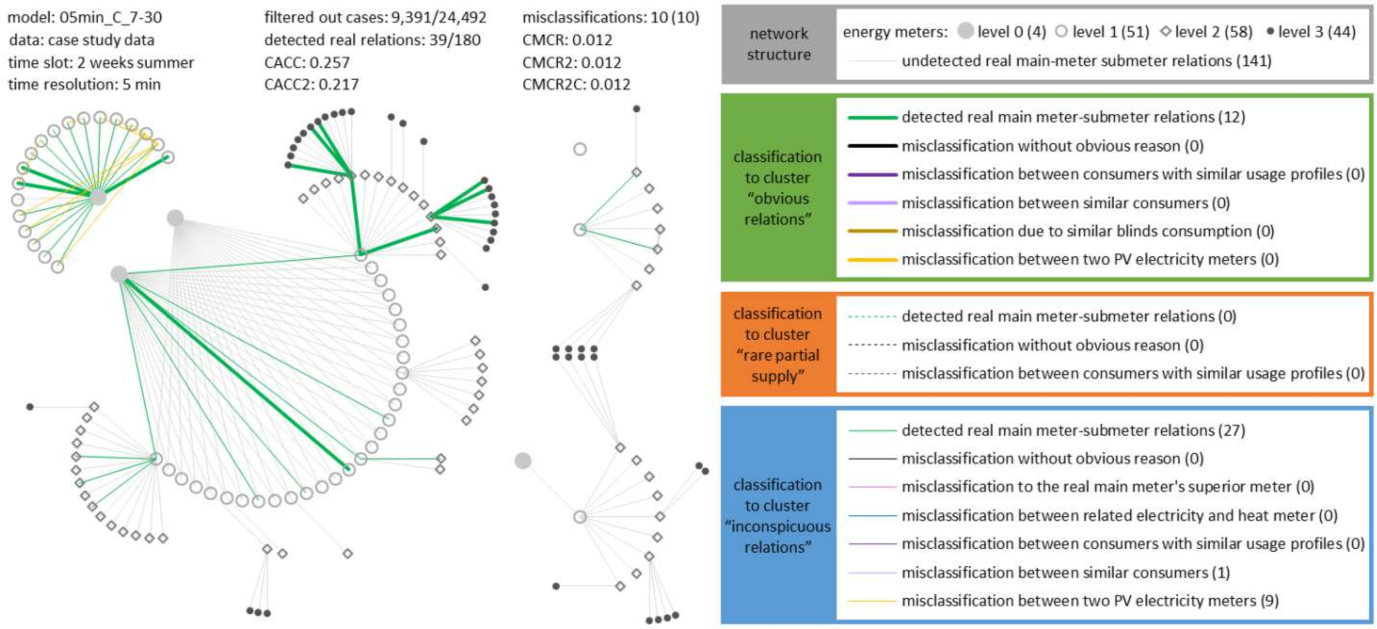
classification to cluster "rare partial supply"
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)
- misclassification between consumers with similar usage profiles (0)

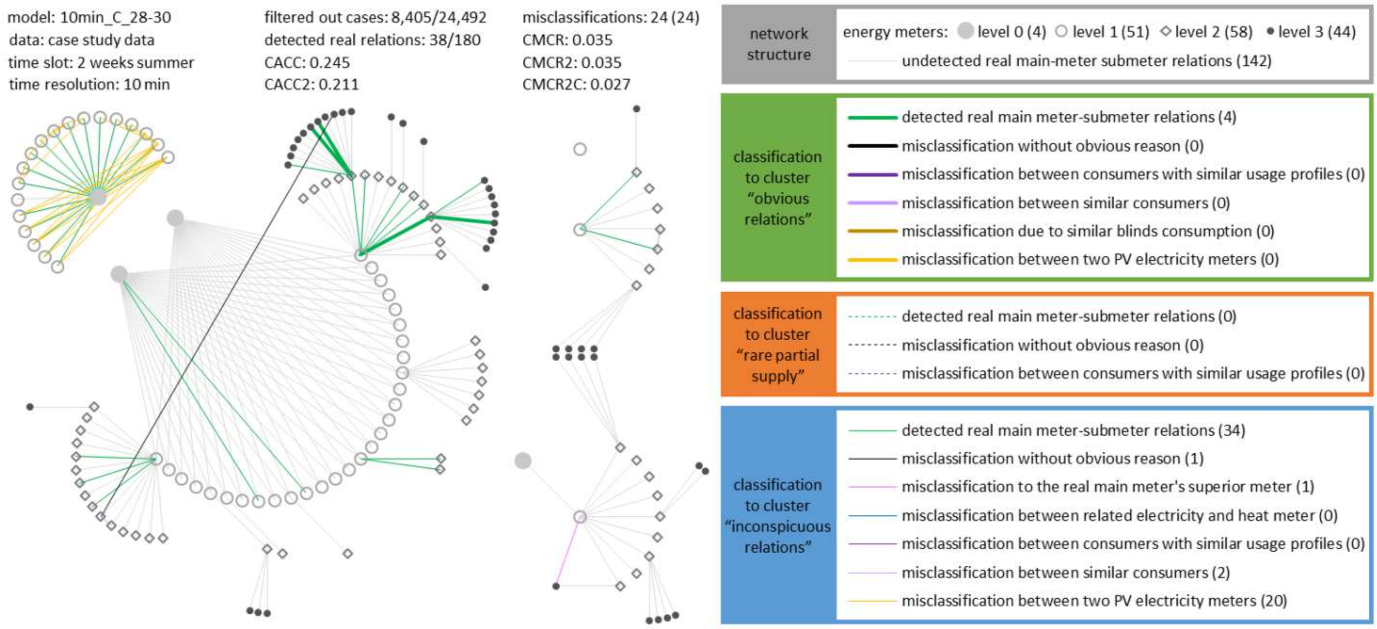
classification to cluster "inconspicuous relations"
- detected real main meter-submeter relations (37)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)
- misclassification between related electricity and heat meter (0)
- misclassification between consumers with similar usage profiles (0)
- misclassification between similar consumers (0)
- misclassification between two PV electricity meters (20)

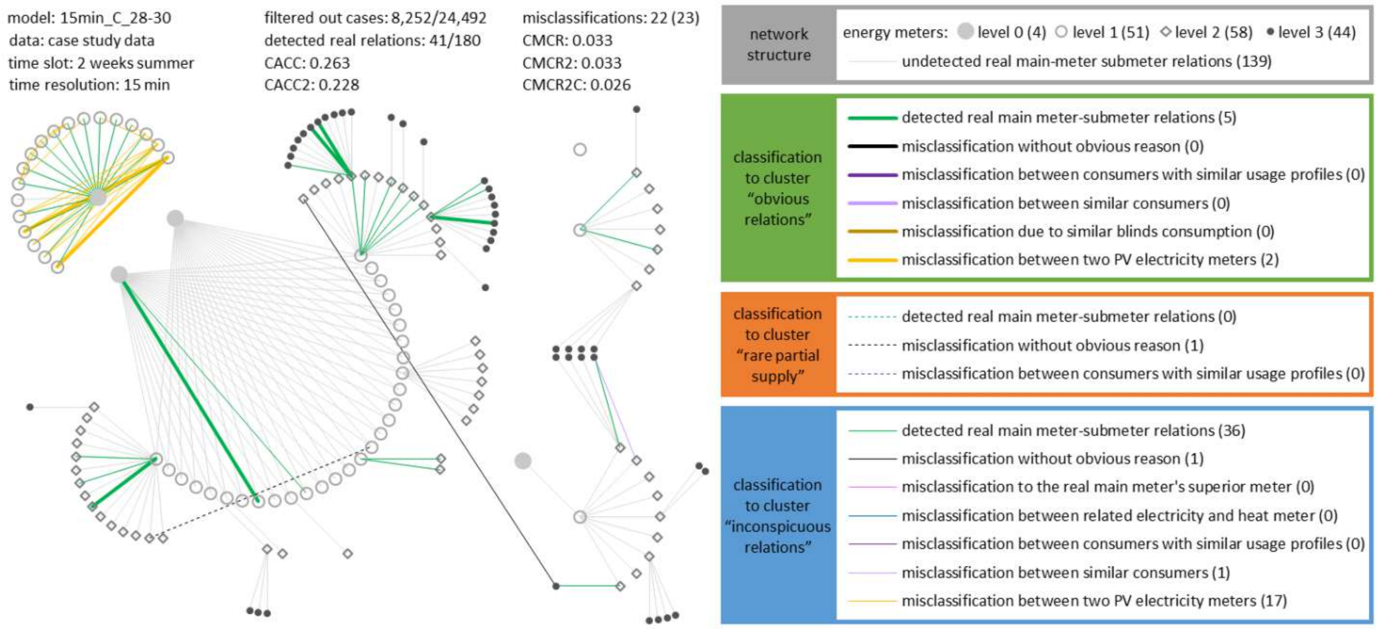

model: 05min_C_7-30
data: case study data
time slot: 5 weeks spring
time resolution: 5 min
filtered out cases: 7,584/24,492
detected real relations: 32/180
CACC: 0.198
CACC2: 0.178
misclassifications: 33 (33)
CMCR: 0.056
CMCR2: 0.056
CMCR2C: 0.056
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (148)
classification to cluster "obvious relations"
detected real main meter-submeter relations (12)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (20)
misclassification without obvious reason (1)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (2)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (29)

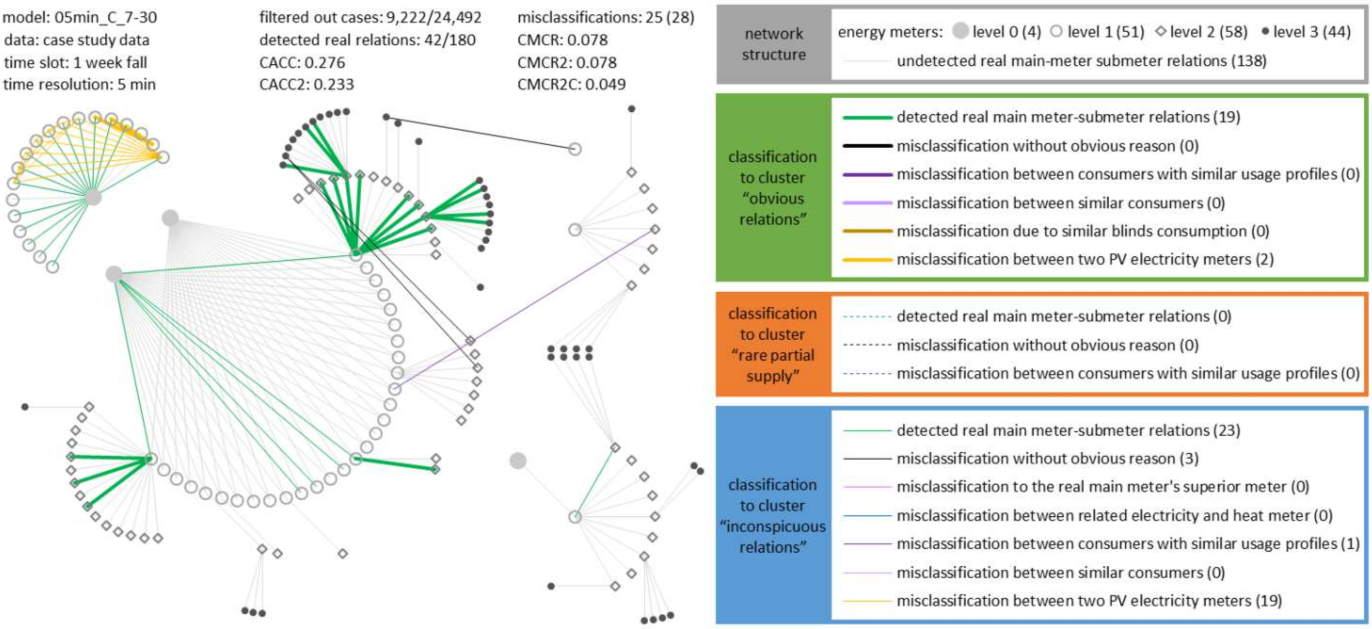

model: 10min_C_28-30
data: case study data
time slot: 5 weeks spring
time resolution: 10 min
filtered out cases: 7,005/24,492
detected real relations: 43/180
CACC: 0.264
CACC2: 0.239
misclassifications: 34 (34)
CMCR: 0.045
CMCR2: 0.045
CMCR2C: 0.039
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (137)
classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (38)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (1)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (4)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (28)

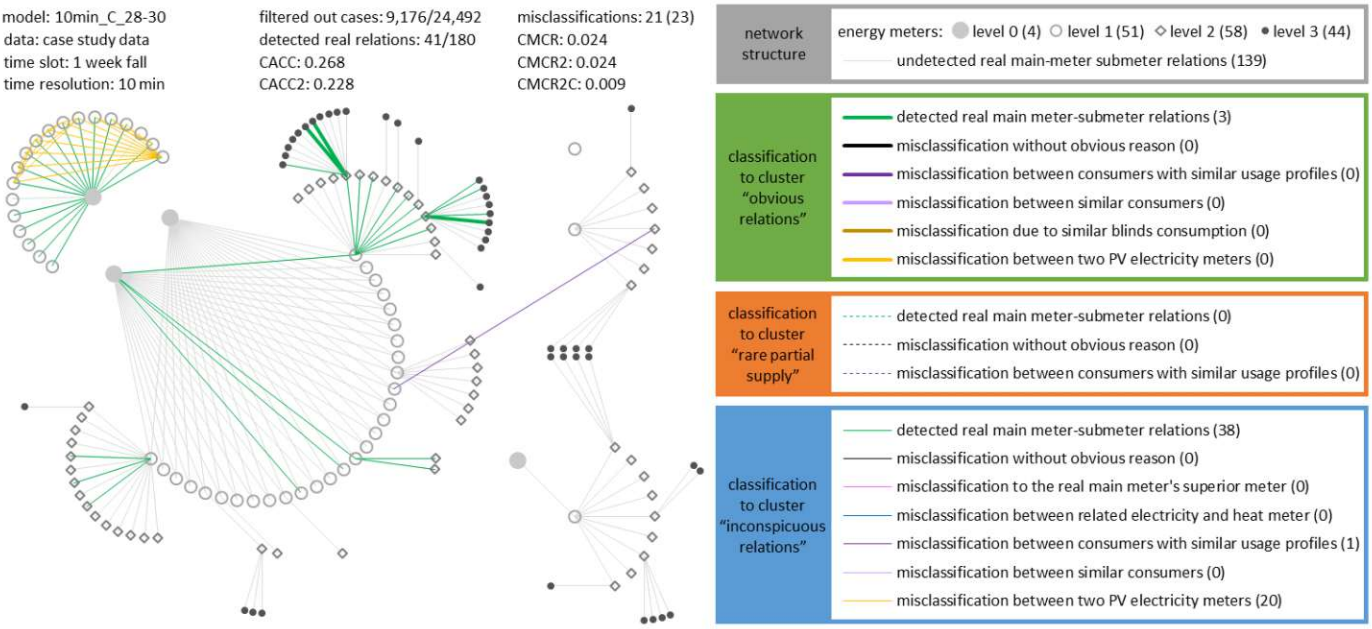

model: 15min_C_28-30
data: case study data
time slot: 5 weeks spring
time resolution: 15 min
filtered out cases: 6,852/24,492
detected real relations: 45/180
CACC: 0.276
CACC2: 0.250
misclassifications: 47 (49)
CMCR: 0.089
CMCR2: 0.089
CMCR2C: 0.075
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (135)
classification to cluster "obvious relations"
detected real main meter-submeter relations (4)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (41)
misclassification without obvious reason (1)
misclassification to the real main meter's superior meter (1)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (5)
misclassification between similar consumers (6)
misclassification between two PV electricity meters (34)

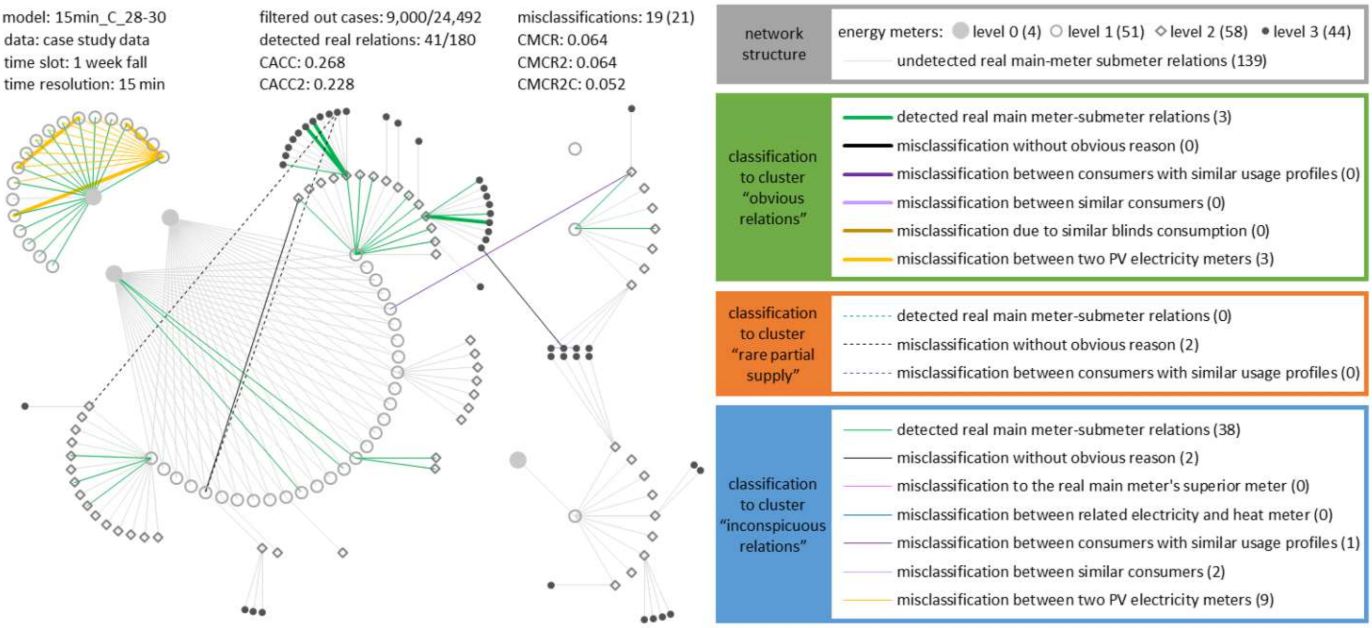

model: 05min_C_7-30
data: case study data
time slot: 2 weeks summer
time resolution: 5 min
filtered out cases: 9,391/24,492
detected real relations: 39/180
CACC: 0.257
CACC2: 0.217
misclassifications: 10 (10)
CMCR: 0.012
CMCR2: 0.012
CMCR2C: 0.012
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (141)
classification to cluster "obvious relations"
detected real main meter-submeter relations (12)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (27)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (9)


model: 10min_C_28-30
data: case study data
time slot: 2 weeks summer
time resolution: 10 min
filtered out cases: 8,405/24,492
detected real relations: 38/180
CACC: 0.245
CACC2: 0.211
misclassifications: 24 (24)
CMCR: 0.035
CMCR2: 0.035
CMCR2C: 0.027
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (142)
classification to cluster "obvious relations"
detected real main meter-submeter relations (4)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (34)
misclassification without obvious reason (1)
misclassification to the real main meter's superior meter (1)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (2)
misclassification between two PV electricity meters (20)


model: 15min_C_28-30
data: case study data
time slot: 2 weeks summer
time resolution: 15 min
filtered out cases: 8,252/24,492
detected real relations: 41/180
CACC: 0.263
CACC2: 0.228
misclassifications: 22 (23)
CMCR: 0.033
CMCR2: 0.033
CMCR2C: 0.026
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (139)
classification to cluster "obvious relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (2)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (1)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (36)
misclassification without obvious reason (1)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (1)
misclassification between two PV electricity meters (17)


model: 05min_C_7-30
data: case study data
time slot: 1 week fall
time resolution: 5 min
filtered out cases: 9,222/24,492
detected real relations: 42/180
CACC: 0.276
CACC2: 0.233
misclassifications: 25 (28)
CMCR: 0.078
CMCR2: 0.078
CMCR2C: 0.049
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (138)
classification to cluster "obvious relations"
detected real main meter-submeter relations (19)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (2)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (23)
misclassification without obvious reason (3)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (19)

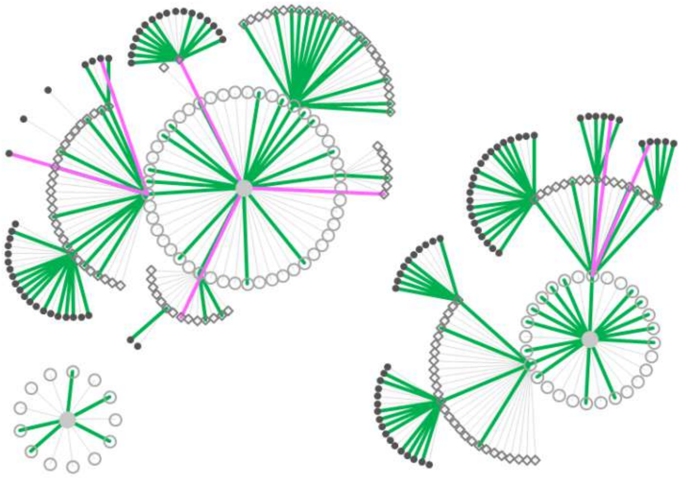

model: 10min_C_28-30
data: case study data
time slot: 1 week fall
time resolution: 10 min
filtered out cases: 9,176/24,492
detected real relations: 41/180
CACC: 0.268
CACC2: 0.228
misclassifications: 21 (23)
CMCR: 0.024
CMCR2: 0.024
CMCR2C: 0.009
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (139)
classification to cluster "obvious relations"
detected real main meter-submeter relations (3)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (38)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (0)
misclassification between two PV electricity meters (20)


model: 15min_C_28-30
data: case study data
time slot: 1 week fall
time resolution: 15 min
filtered out cases: 9,000/24,492
detected real relations: 41/180
CACC: 0.268
CACC2: 0.228
misclassifications: 19 (21)
CMCR: 0.064
CMCR2: 0.064
CMCR2C: 0.052
network structure
energy meters: level 0 (4) level 1 (51) level 2 (58) level 3 (44)
undetected real main-meter submeter relations (139)
classification to cluster "obvious relations"
detected real main meter-submeter relations (3)
misclassification without obvious reason (0)
misclassification between consumers with similar usage profiles (0)
misclassification between similar consumers (0)
misclassification due to similar blinds consumption (0)
misclassification between two PV electricity meters (3)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (2)
misclassification between consumers with similar usage profiles (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (38)
misclassification without obvious reason (2)
misclassification to the real main meter's superior meter (0)
misclassification between related electricity and heat meter (0)
misclassification between consumers with similar usage profiles (1)
misclassification between similar consumers (2)
misclassification between two PV electricity meters (9)

model: 05min_CP_18-30
data: synthetically generated data
time slot: full year
time resolution: 5 min

filtered out cases: 2,741/103,362
detected real relations: 137/319
CACC: 0.434
CACC2: 0.430

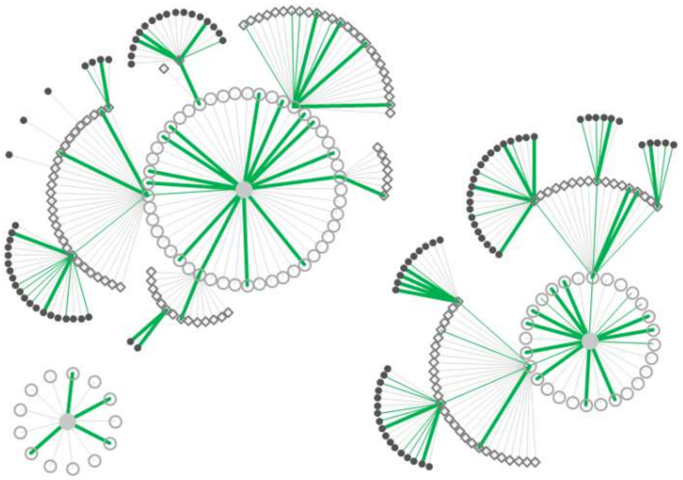
misclassifications: 11 (38)
CMCR: NaN
CMCR2: 0.217
CMCR2C: 0.000



model: 10min_CP_55-30
data: synthetically generated data
time slot: full year
time resolution: 10 min

filtered out cases: 2,035/103,362
detected real relations: 104/319
CACC: 0.329
CACC2: 0.326

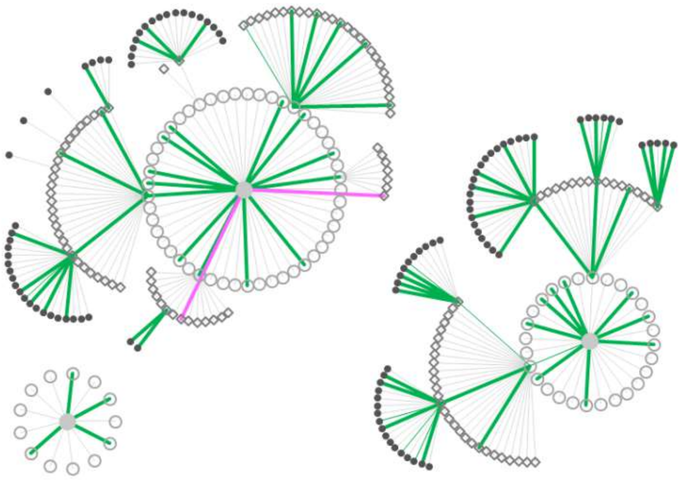
misclassifications: 0 (14)
CMCR: 0.118
CMCR2: 0.118
CMCR2C: 0.000



model: 15min_CP_55-30
data: synthetically generated data
time slot: full year
time resolution: 15 min

filtered out cases: 2,031/103,362
detected real relations: 80/319
CACC: 0.253
CACC2: 0.251

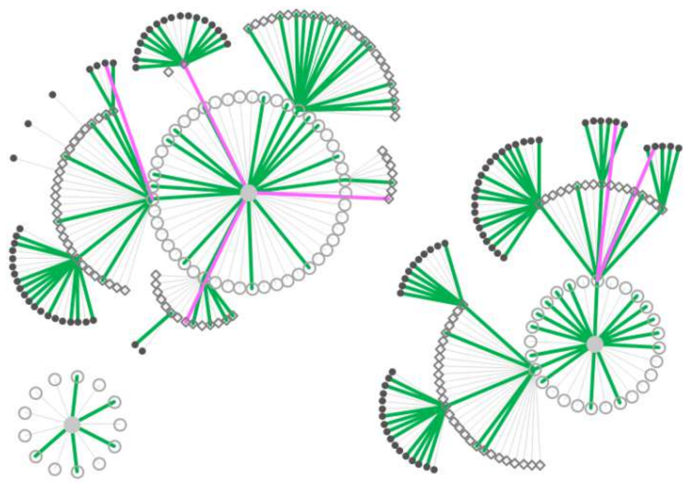
misclassifications: 4 (6)
CMCR: 0.052
CMCR2: 0.052
CMCR2C: 0.000



model: 05min_CP_18-30
data: synthetically generated data
time slot: 6 months first half year
time resolution: 5 min

filtered out cases: 3,386/103,362
detected real relations: 137/319
CACC: 0.434
CACC2: 0.430

misclassifications: 9 (37)
CMCR: 0.143
CMCR2: 0.143
CMCR2C: 0.000

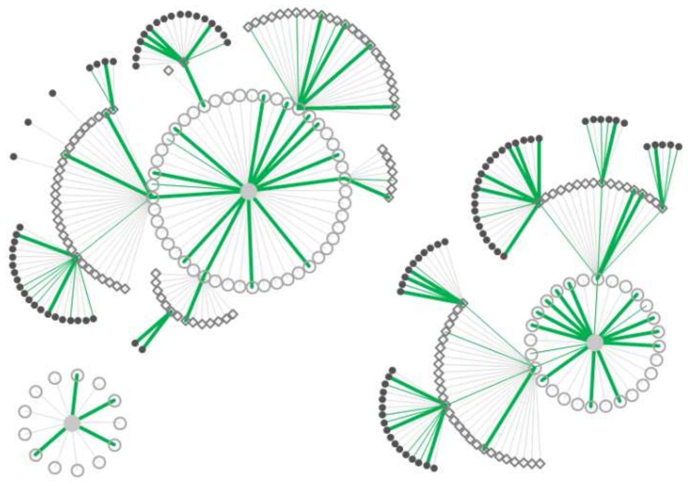


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (182)
classification to cluster "obvious relations"	detected real main meter-submeter relations (136) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (9) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (1) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 10min_CP_55-30
data: synthetically generated data
time slot: 6 months first half year
time resolution: 10 min

filtered out cases: 2,714/103,362
detected real relations: 104/319
CACC: 0.329
CACC2: 0.326

misclassifications: 0 (14)
CMCR: 0.113
CMCR2: 0.113
CMCR2C: 0.000

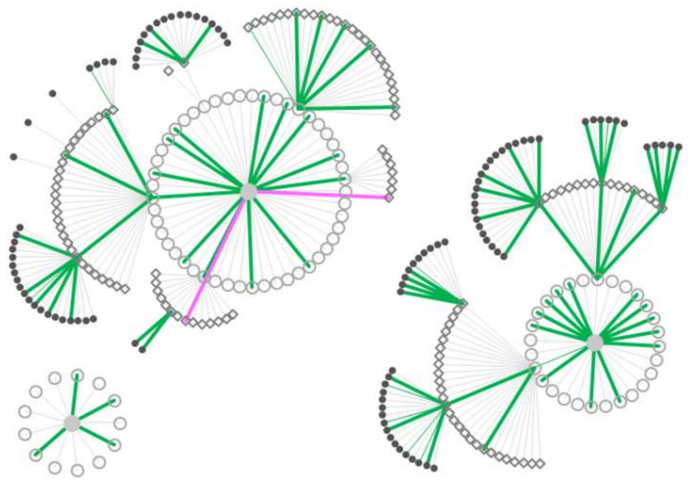


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (215)
classification to cluster "obvious relations"	detected real main meter-submeter relations (64) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (40) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 15min_CP_55-30
data: synthetically generated data
time slot: 6 months first half year
time resolution: 15 min

filtered out cases: 2,421/103,362
detected real relations: 84/319
CACC: 0.266
CACC2: 0.263

misclassifications: 4 (6)
CMCR: 0.050
CMCR2: 0.050
CMCR2C: 0.000

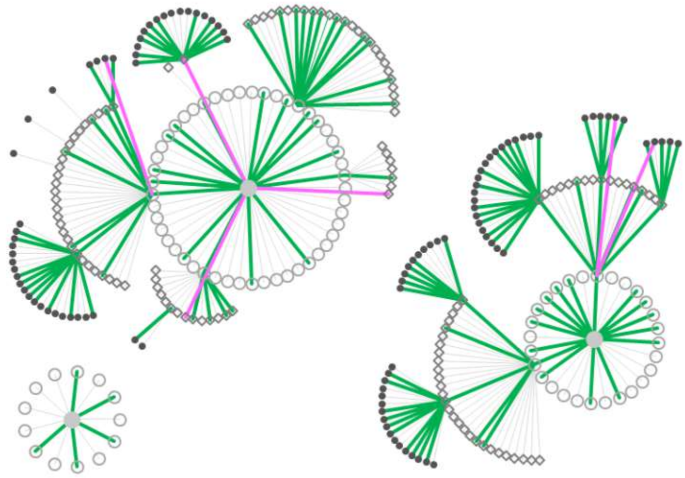


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (235)
classification to cluster "obvious relations"	detected real main meter-submeter relations (74) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (4) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (10) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 05min_CP_18-30
data: synthetically generated data
time slot: 3 months first qtr. year
time resolution: 5 min

filtered out cases: 5,418/103,362
detected real relations: 134/319
CACC: 0.424
CACC2: 0.420

misclassifications: 9 (33)
CMCR: 0.133
CMCR2: 0.133
CMCR2C: 0.000



network structure
energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
undetected real main-meter submeter relations (185)

classification to cluster "obvious relations"
detected real main meter-submeter relations (133)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (9)
misclassification to the real main meter's superior meter's superior meter (0)

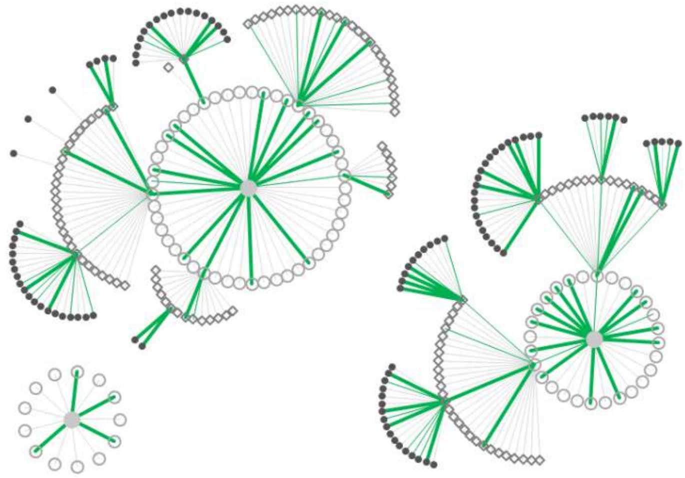
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (1)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)

model: 10min_CP_55-30
data: synthetically generated data
time slot: 3 months first qtr. year
time resolution: 10 min

filtered out cases: 4,723/103,362
detected real relations: 108/319
CACC: 0.342
CACC2: 0.339

misclassifications: 0 (18)
CMCR: 0.128
CMCR2: 0.128
CMCR2C: 0.000



network structure
energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
undetected real main-meter submeter relations (211)

classification to cluster "obvious relations"
detected real main meter-submeter relations (69)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification to the real main meter's superior meter's superior meter (0)

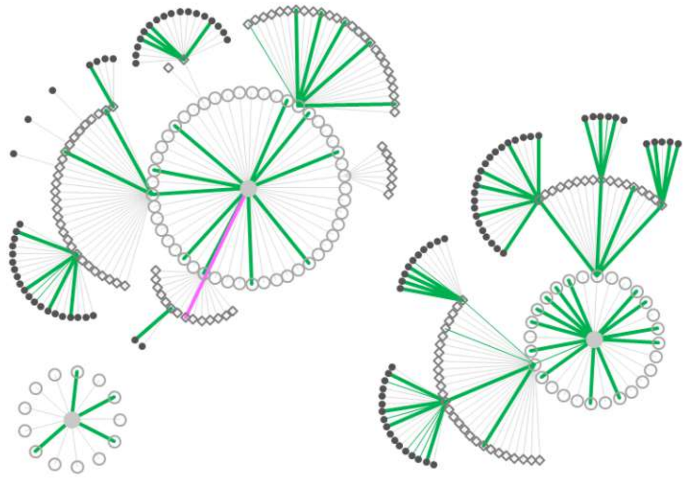
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (39)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)

model: 15min_CP_55-30
data: synthetically generated data
time slot: 3 months first qtr. year
time resolution: 15 min

filtered out cases: 4,095/103,362
detected real relations: 84/319
CACC: 0.266
CACC2: 0.263

misclassifications: 2 (8)
CMCR: 0.068
CMCR2: 0.068
CMCR2C: 0.000



network structure
energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
undetected real main-meter submeter relations (235)

classification to cluster "obvious relations"
detected real main meter-submeter relations (71)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (2)
misclassification to the real main meter's superior meter's superior meter (0)

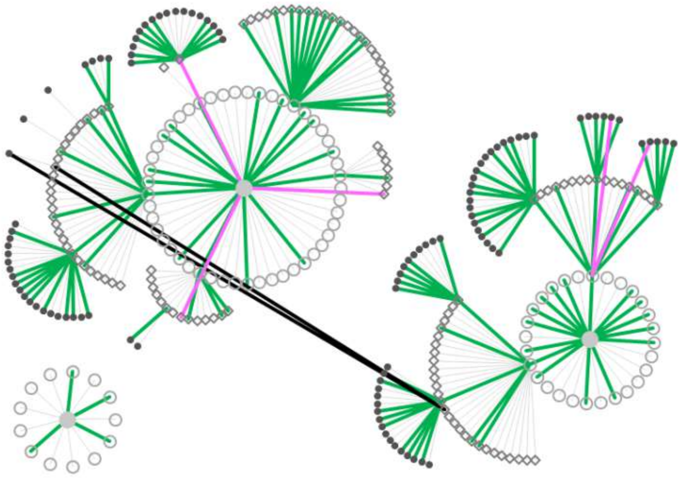
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)

classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (13)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)

model: 05min_CP_18-30
data: synthetically generated data
time slot: 5 weeks second qtr. year
time resolution: 5 min

filtered out cases: 7,628/103,362
detected real relations: 129/319
CACC: 0.408
CACC2: 0.404

misclassifications: 9 (30)
CMCR: NaN
CMCR2: 0.189
CMCR2C: 0.015

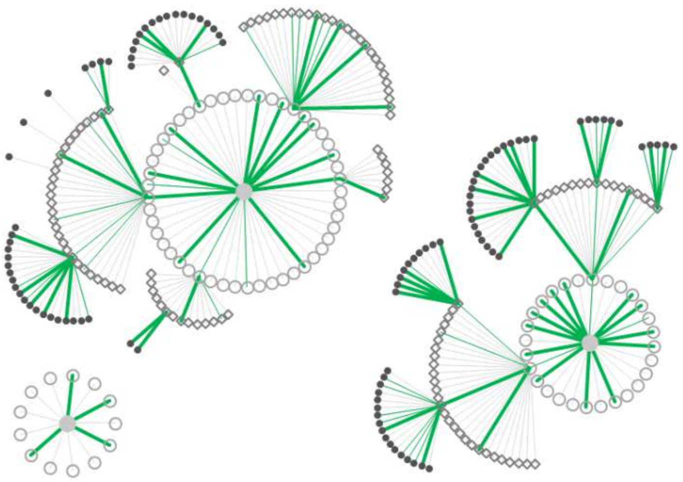


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (190)
classification to cluster "obvious relations"	detected real main meter-submeter relations (129) misclassification without obvious reason (2) misclassification to the real main meter's superior meter (7) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (0) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 10min_CP_55-30
data: synthetically generated data
time slot: 5 weeks second qtr. year
time resolution: 10 min

filtered out cases: 7,350/103,362
detected real relations: 111/319
CACC: 0.351
CACC2: 0.348

misclassifications: 0 (18)
CMCR: 0.134
CMCR2: 0.134
CMCR2C: 0.000

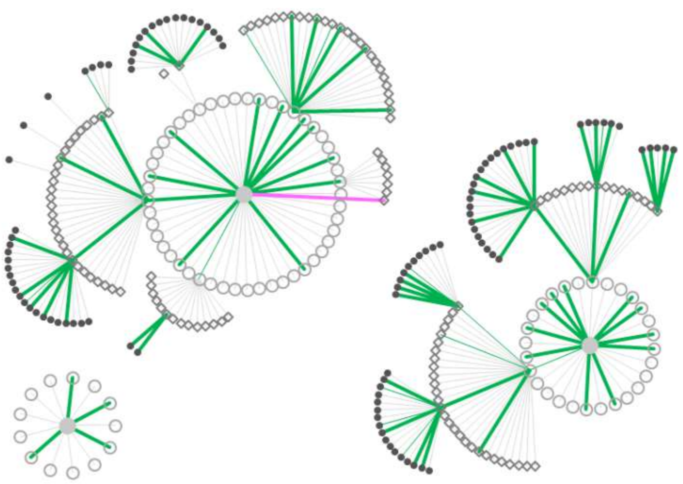


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (208)
classification to cluster "obvious relations"	detected real main meter-submeter relations (70) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (41) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 15min_CP_55-30
data: synthetically generated data
time slot: 5 weeks second qtr. year
time resolution: 15 min

filtered out cases: 6,720/103,362
detected real relations: 83/319
CACC: 0.263
CACC2: 0.260

misclassifications: 2 (6)
CMCR: 0.053
CMCR2: 0.053
CMCR2C: 0.000

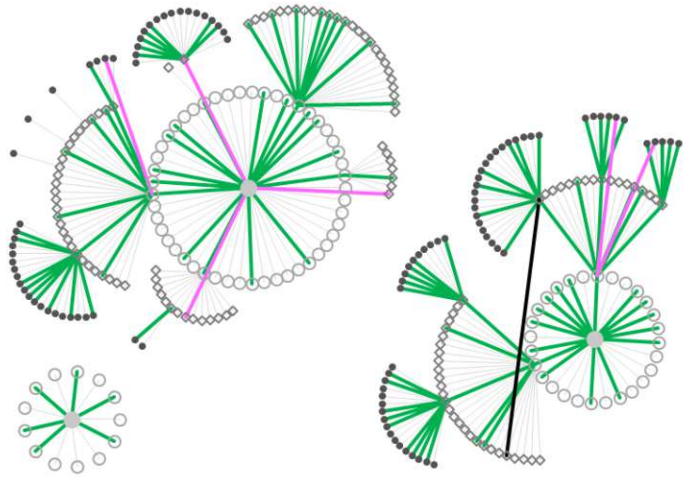


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (236)
classification to cluster "obvious relations"	detected real main meter-submeter relations (70) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (2) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (13) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 05min_CP_18-30
data: synthetically generated data
time slot: 2 weeks third qtr. year
time resolution: 5 min

filtered out cases: 12,787/103,362
detected real relations: 116/319
CACC: 0.367
CACC2: 0.364

misclassifications: 10 (29)
CMCR: 0.134
CMCR2: 0.134
CMCR2C: 0.006



network structure
- energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
- undetected real main-meter submeter relations (203)

classification to cluster "obvious relations"
- detected real main meter-submeter relations (115)
- misclassification without obvious reason (1)
- misclassification to the real main meter's superior meter (9)
- misclassification to the real main meter's superior meter's superior meter (0)

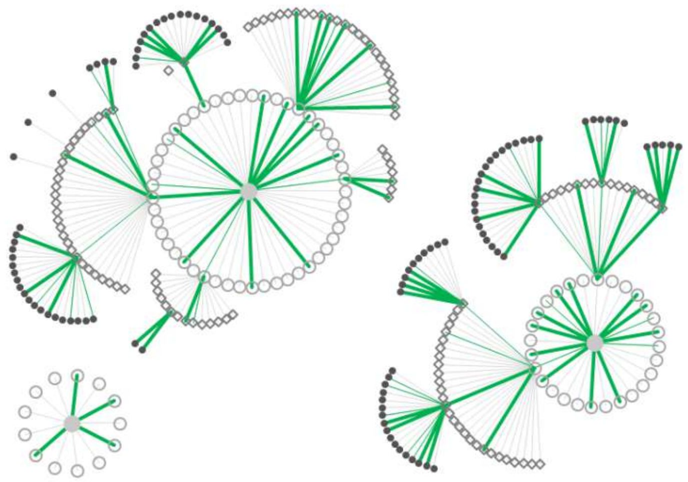
classification to cluster "rare partial supply"
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)

classification to cluster "inconspicuous relations"
- detected real main meter-submeter relations (1)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)

model: 10min_CP_55-30
data: synthetically generated data
time slot: 2 weeks third qtr. year
time resolution: 10 min

filtered out cases: 12,130/103,362
detected real relations: 104/319
CACC: 0.329
CACC2: 0.326

misclassifications: 0 (20)
CMCR: 0.159
CMCR2: 0.159
CMCR2C: 0.000



network structure
- energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
- undetected real main-meter submeter relations (215)

classification to cluster "obvious relations"
- detected real main meter-submeter relations (72)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)
- misclassification to the real main meter's superior meter's superior meter (0)

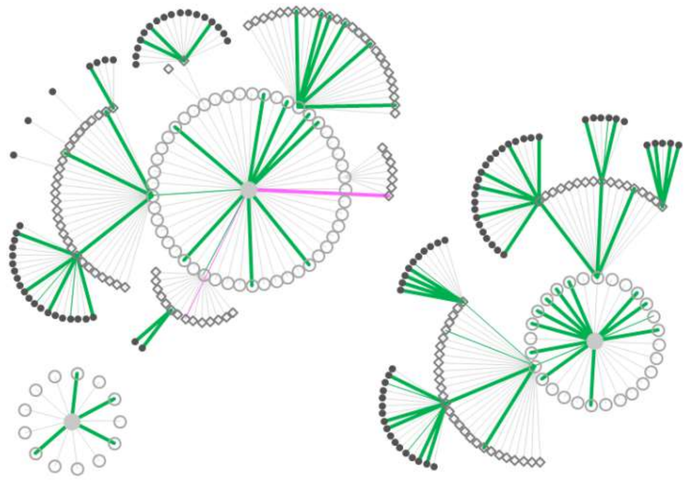
classification to cluster "rare partial supply"
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)

classification to cluster "inconspicuous relations"
- detected real main meter-submeter relations (32)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)

model: 15min_CP_55-30
data: synthetically generated data
time slot: 2 weeks third qtr. year
time resolution: 15 min

filtered out cases: 11,959/103,362
detected real relations: 80/319
CACC: 0.253
CACC2: 0.251

misclassifications: 4 (6)
CMCR: 0.071
CMCR2: 0.071
CMCR2C: 0.000



network structure
- energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
- undetected real main-meter submeter relations (239)

classification to cluster "obvious relations"
- detected real main meter-submeter relations (68)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (2)
- misclassification to the real main meter's superior meter's superior meter (0)

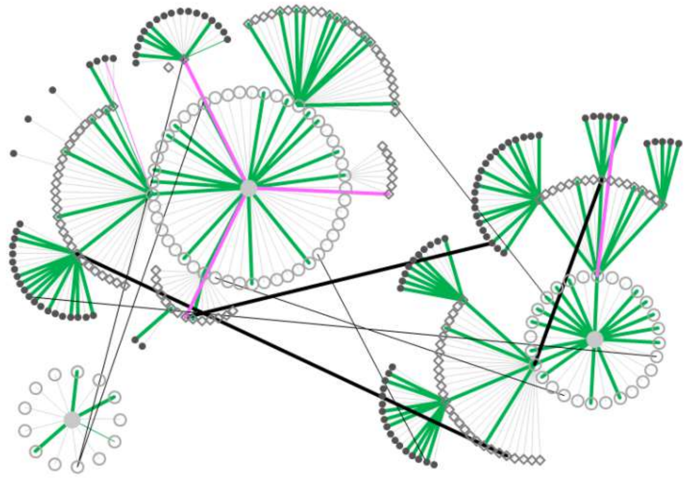
classification to cluster "rare partial supply"
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)

classification to cluster "inconspicuous relations"
- detected real main meter-submeter relations (12)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (2)

model: 05min_CP_18-30
data: synthetically generated data
time slot: 1 week fourth qtr. year
time resolution: 5 min

filtered out cases: 15,961/103,362
detected real relations: 116/319
CACC: 0.368
CACC2: 0.364

misclassifications: 17 (33)
CMCR: 0.367
CMCR2: 0.367
CMCR2C: 0.251



network structure
- energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
- undetected real main-meter submeter relations (203)

classification to cluster “obvious relations”
- detected real main meter-submeter relations (113)
- misclassification without obvious reason (3)
- misclassification to the real main meter's superior meter (6)
- misclassification to the real main meter's superior meter's superior meter (0)

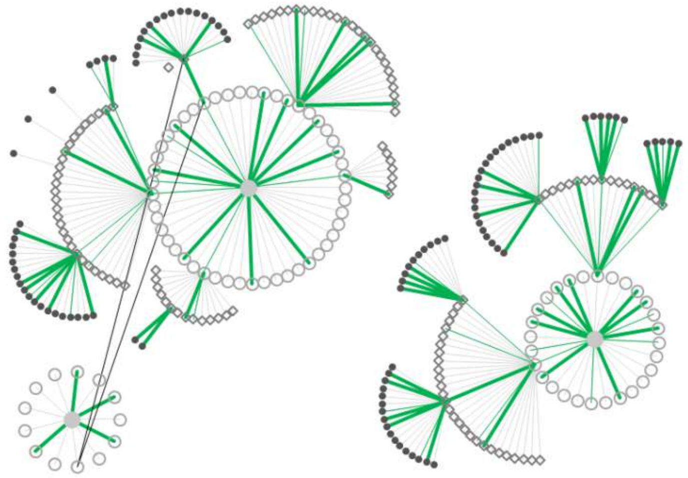
classification to cluster “rare partial supply”
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)

classification to cluster “inconspicuous relations”
- detected real main meter-submeter relations (3)
- misclassification without obvious reason (6)
- misclassification to the real main meter's superior meter (2)

model: 10min_CP_55-30
data: synthetically generated data
time slot: 1 week fourth qtr. year
time resolution: 10 min

filtered out cases: 15,153/103,362
detected real relations: 105/319
CACC: 0.333
CACC2: 0.329

misclassifications: 2 (21)
CMCR: 0.145
CMCR2: 0.145
CMCR2C: 0.018



network structure
- energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
- undetected real main-meter submeter relations (214)

classification to cluster “obvious relations”
- detected real main meter-submeter relations (70)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)
- misclassification to the real main meter's superior meter's superior meter (0)

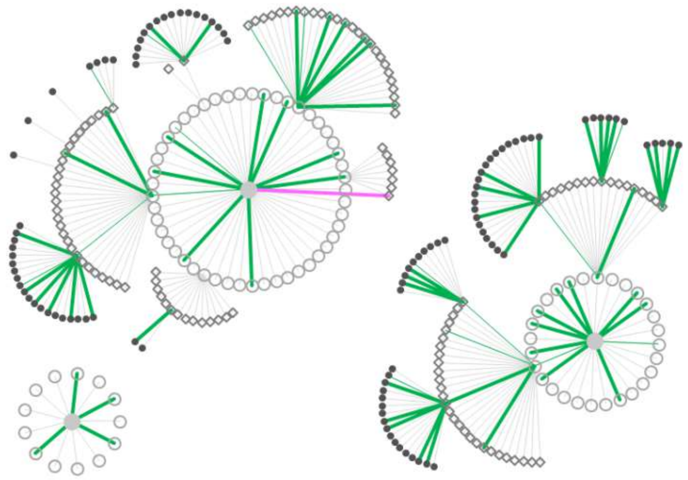
classification to cluster “rare partial supply”
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)

classification to cluster “inconspicuous relations”
- detected real main meter-submeter relations (35)
- misclassification without obvious reason (2)
- misclassification to the real main meter's superior meter (0)

model: 15min_CP_55-30
data: synthetically generated data
time slot: 1 week fourth qtr. year
time resolution: 15 min

filtered out cases: 15,448/103,362
detected real relations: 78/319
CACC: 0.247
CACC2: 0.245

misclassifications: 2 (6)
CMCR: 0.077
CMCR2: 0.077
CMCR2C: 0.000



network structure
- energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
- undetected real main-meter submeter relations (241)

classification to cluster “obvious relations”
- detected real main meter-submeter relations (62)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (2)
- misclassification to the real main meter's superior meter's superior meter (0)

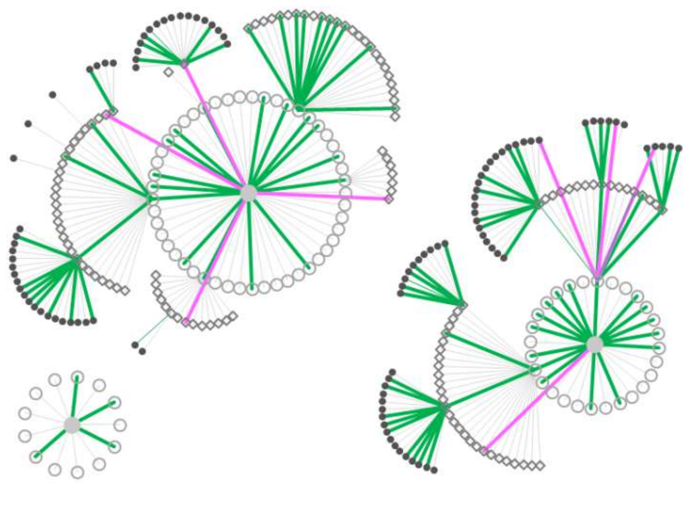
classification to cluster “rare partial supply”
- detected real main meter-submeter relations (0)
- misclassification without obvious reason (0)

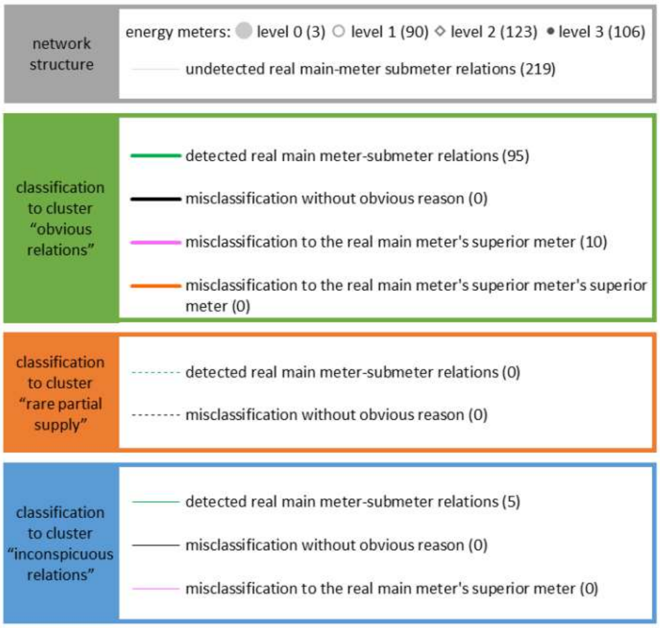
classification to cluster “inconspicuous relations”
- detected real main meter-submeter relations (16)
- misclassification without obvious reason (0)
- misclassification to the real main meter's superior meter (0)

model: 05min_C_7-30
data: synthetically generated data
time slot: full year
time resolution: 5 min

filtered out cases: 2,741/103,362
detected real relations: 100/319
CACC: 0.317
CACC2: 0.314

misclassifications: 10 (19)
CMCR: 0.162
CMCR2: 0.162
CMCR2C: 0.000

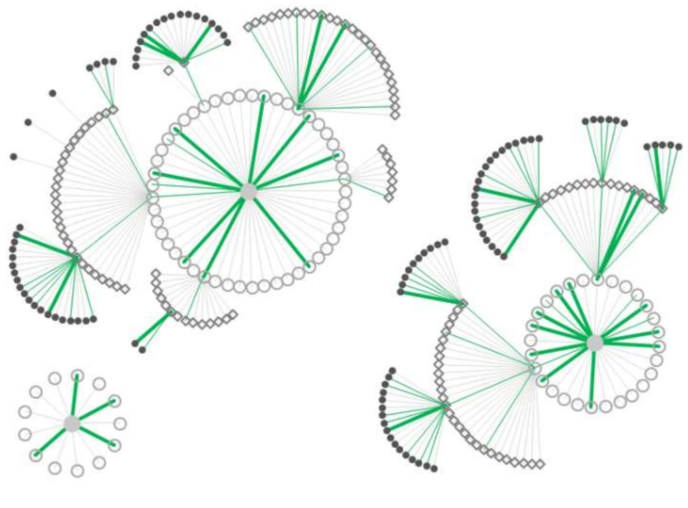


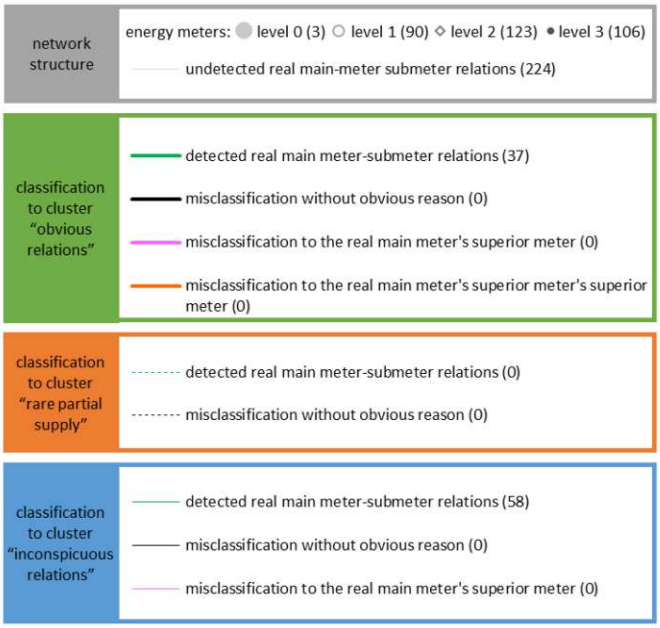


model: 10min_C_28-30
data: synthetically generated data
time slot: full year
time resolution: 10 min

filtered out cases: 2,035/103,362
detected real relations: 95/319
CACC: 0.301
CACC2: 0.298

misclassifications: 0 (13)
CMCR: 0.129
CMCR2: 0.129
CMCR2C: 0.000

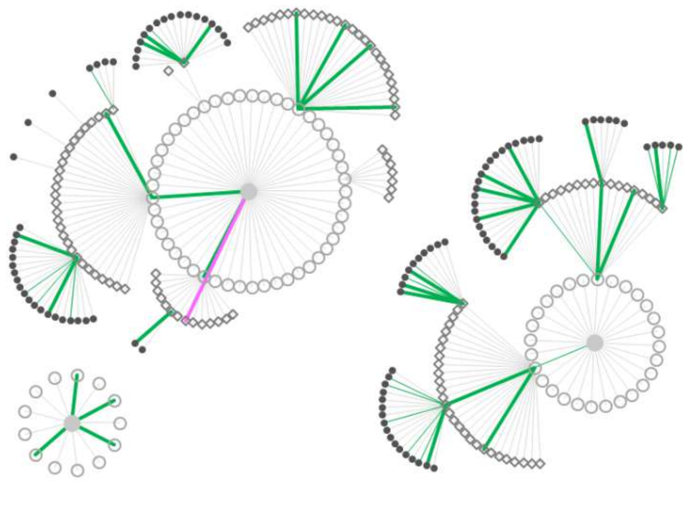


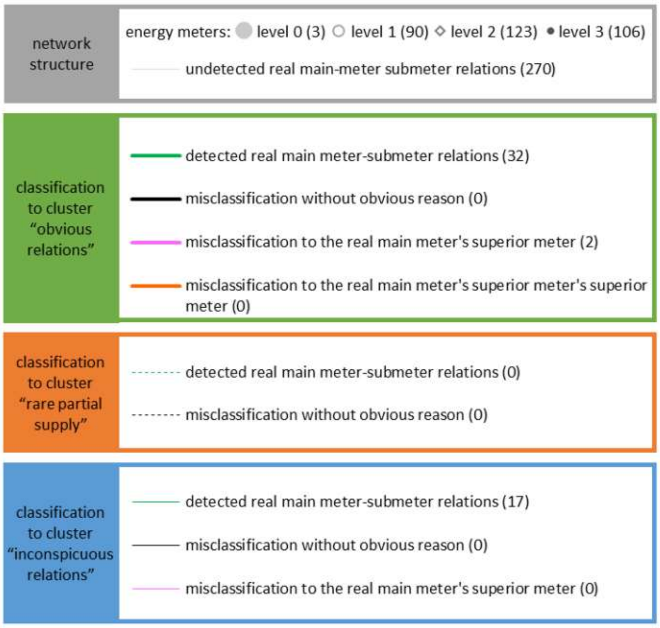


model: 15min_C_28-30
data: synthetically generated data
time slot: full year
time resolution: 15 min

filtered out cases: 2,031/103,362
detected real relations: 49/319
CACC: 0.155
CACC2: 0.154

misclassifications: 2 (5)
CMCR: 0.093
CMCR2: 0.093
CMCR2C: 0.000

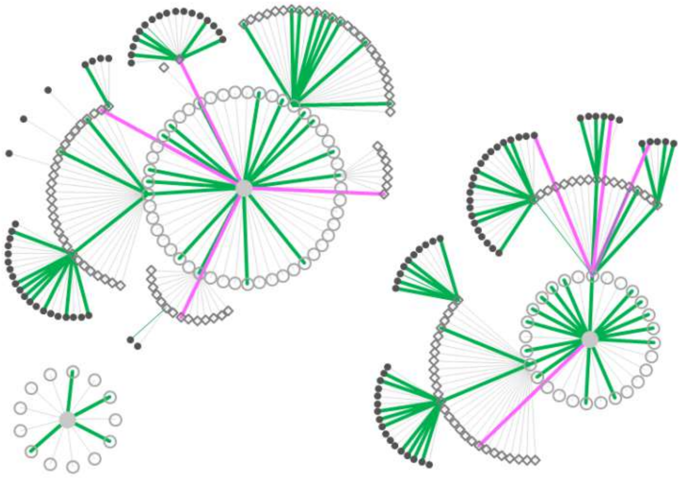




model: 05min_C_7-30
data: synthetically generated data
time slot: 6 months first half year
time resolution: 5 min

filtered out cases: 3,386/103,362
detected real relations: 99/319
CACC: 0.313
CACC2: 0.310

misclassifications: 10 (19)
CMCR: 0.163
CMCR2: 0.163
CMCR2C: 0.000

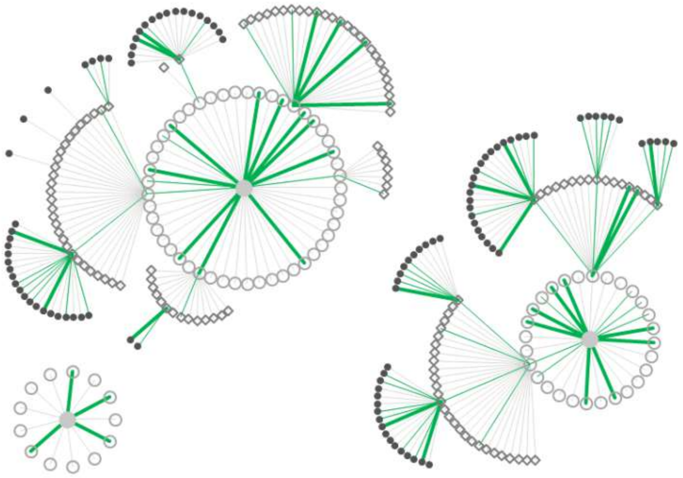


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (220)
classification to cluster "obvious relations"	detected real main meter-submeter relations (94) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (10) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (5) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 10min_C_28-30
data: synthetically generated data
time slot: 6 months first half year
time resolution: 10 min

filtered out cases: 2,714/103,362
detected real relations: 96/319
CACC: 0.304
CACC2: 0.301

misclassifications: 0 (12)
CMCR: 0.111
CMCR2: 0.111
CMCR2C: 0.000

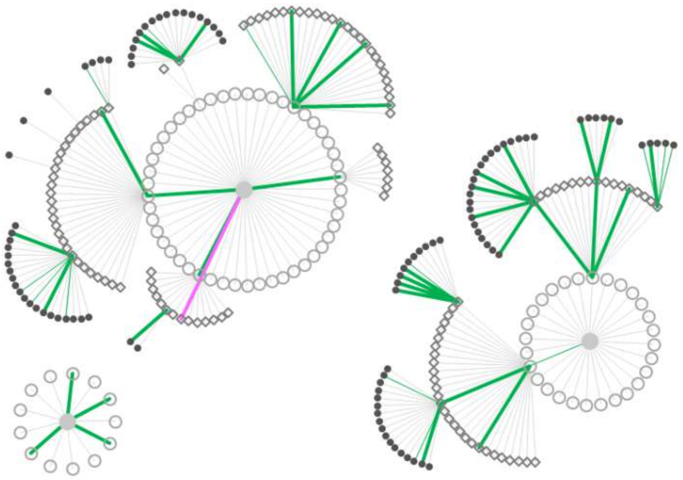


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (223)
classification to cluster "obvious relations"	detected real main meter-submeter relations (40) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (56) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 15min_C_28-30
data: synthetically generated data
time slot: 6 months first half year
time resolution: 15 min

filtered out cases: 2,421/103,362
detected real relations: 49/319
CACC: 0.155
CACC2: 0.154

misclassifications: 2 (5)
CMCR: 0.090
CMCR2: 0.091
CMCR2C: 0.000

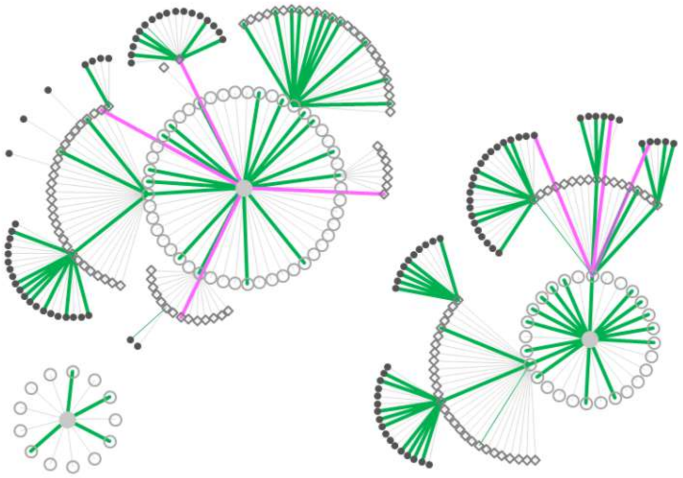


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (270)
classification to cluster "obvious relations"	detected real main meter-submeter relations (36) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (2) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (13) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 05min_C_7-30
data: synthetically generated data
time slot: 3 months first qtr. year
time resolution: 5 min

filtered out cases: 5,418/103,362
detected real relations: 102/319
CACC: 0.323
CACC2: 0.320

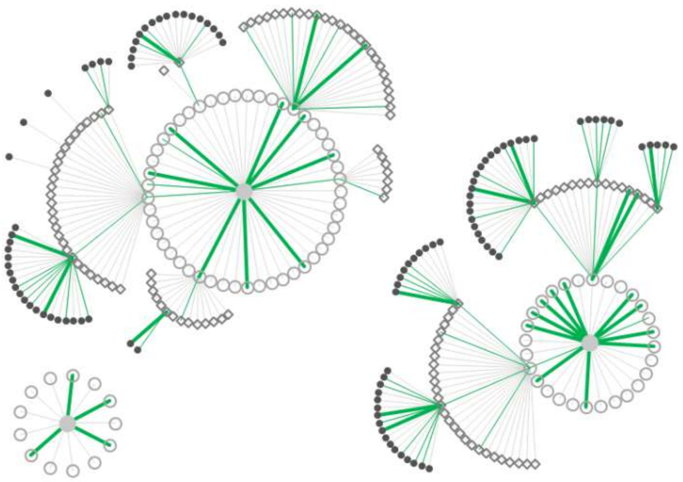
misclassifications: 9 (19)
CMCR: 0.160
CMCR2: 0.160
CMCR2C: 0.000



model: 10min_C_28-30
data: synthetically generated data
time slot: 3 months first qtr. year
time resolution: 10 min

filtered out cases: 4,723/103,362
detected real relations: 93/319
CACC: 0.294
CACC2: 0.292

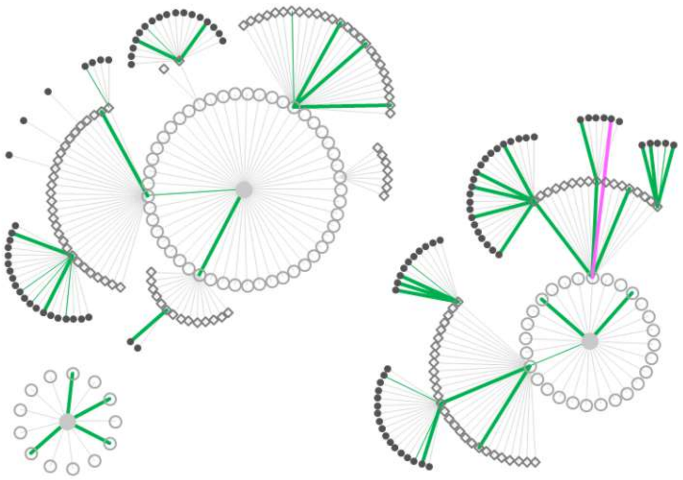
misclassifications: 0 (12)
CMCR: 0.096
CMCR2: 0.096
CMCR2C: 0.000

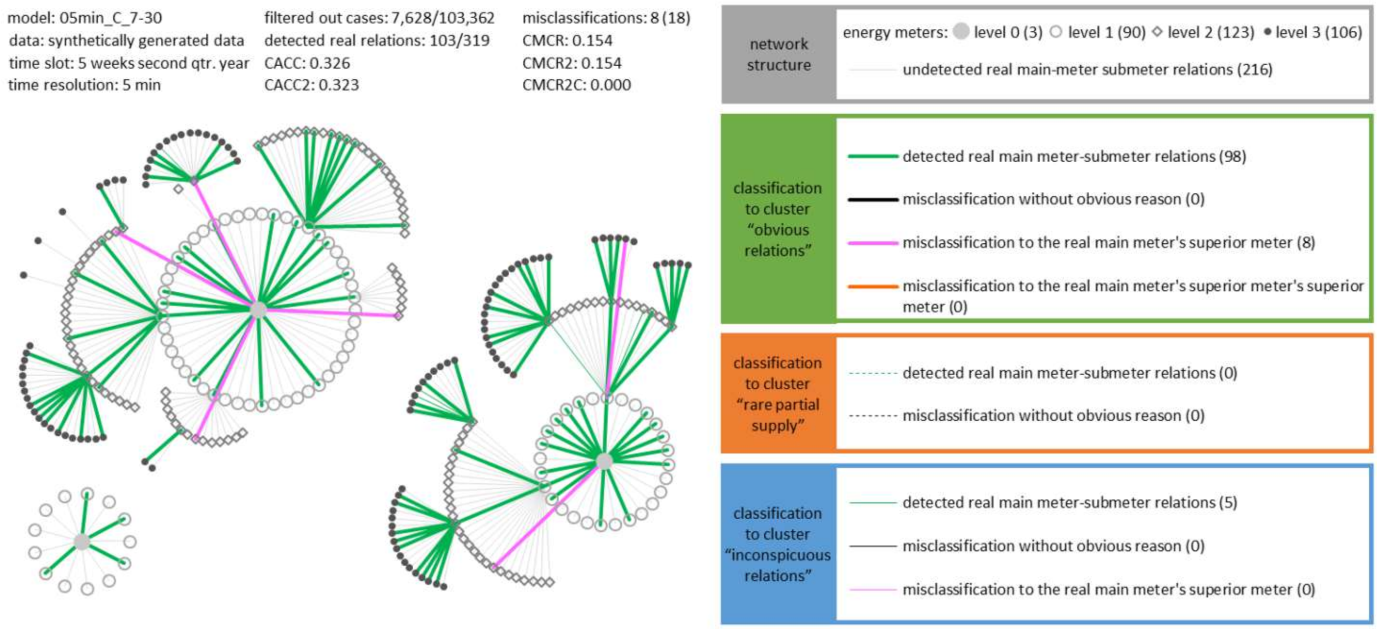


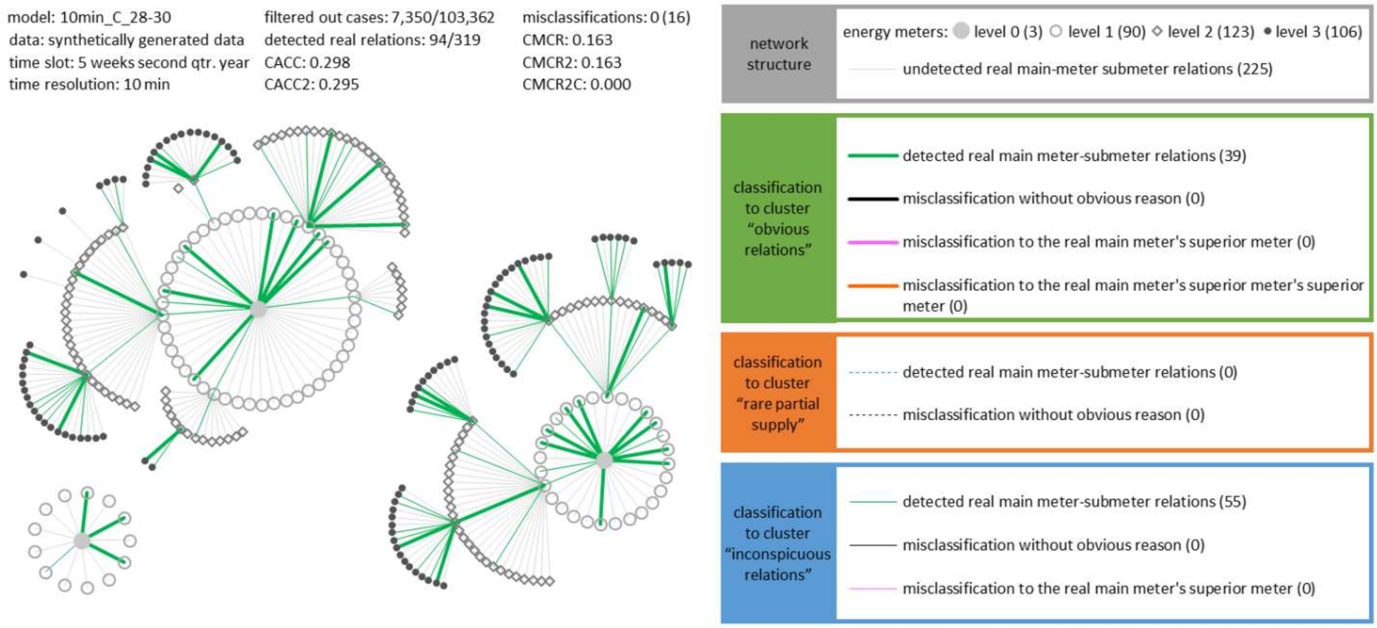
model: 15min_C_28-30
data: synthetically generated data
time slot: 3 months first qtr. year
time resolution: 15 min

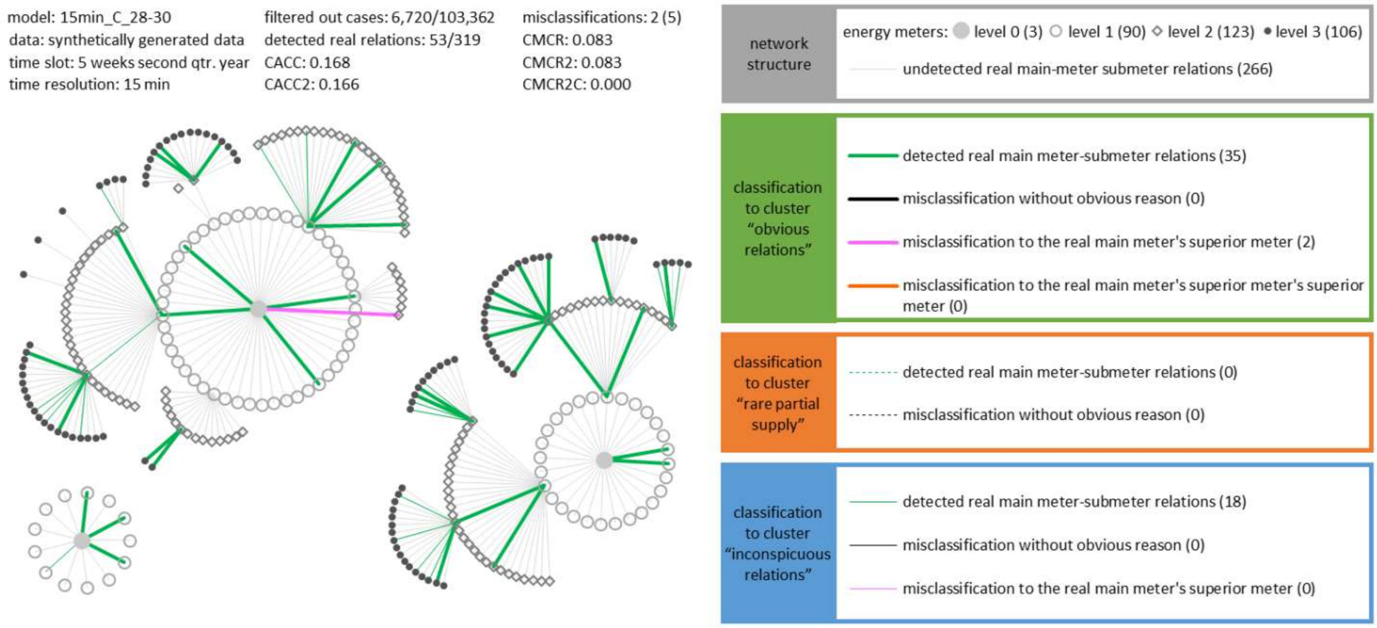
filtered out cases: 4,095/103,362
detected real relations: 46/319
CACC: 0.146
CACC2: 0.144

misclassifications: 1 (4)
CMCR: 0.080
CMCR2: 0.080
CMCR2C: 0.000




model: 05min_C_7-30
data: synthetically generated data
time slot: 5 weeks second qtr. year
time resolution: 5 min
filtered out cases: 7,628/103,362
detected real relations: 103/319
CACC: 0.326
CACC2: 0.323
misclassifications: 8 (18)
CMCR: 0.154
CMCR2: 0.154
CMCR2C: 0.000
network structure
energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
undetected real main-meter submeter relations (216)
classification to cluster "obvious relations"
detected real main meter-submeter relations (98)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (8)
misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (5)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)

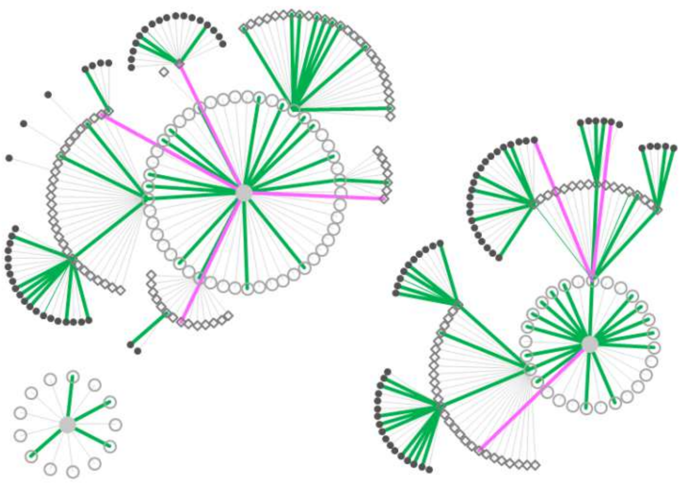

model: 10min_C_28-30
data: synthetically generated data
time slot: 5 weeks second qtr. year
time resolution: 10 min
filtered out cases: 7,350/103,362
detected real relations: 94/319
CACC: 0.298
CACC2: 0.295
misclassifications: 0 (16)
CMCR: 0.163
CMCR2: 0.163
CMCR2C: 0.000
network structure
energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
undetected real main-meter submeter relations (225)
classification to cluster "obvious relations"
detected real main meter-submeter relations (39)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)
misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (55)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)


model: 15min_C_28-30
data: synthetically generated data
time slot: 5 weeks second qtr. year
time resolution: 15 min
filtered out cases: 6,720/103,362
detected real relations: 53/319
CACC: 0.168
CACC2: 0.166
misclassifications: 2 (5)
CMCR: 0.083
CMCR2: 0.083
CMCR2C: 0.000
network structure
energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106)
undetected real main-meter submeter relations (266)
classification to cluster "obvious relations"
detected real main meter-submeter relations (35)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (2)
misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"
detected real main meter-submeter relations (0)
misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"
detected real main meter-submeter relations (18)
misclassification without obvious reason (0)
misclassification to the real main meter's superior meter (0)

model: 05min_C_7-30
data: synthetically generated data
time slot: 2 weeks third qtr. year
time resolution: 5 min

filtered out cases: 12,787/103,362
detected real relations: 98/319
CACC: 0.310
CACC2: 0.307

misclassifications: 9 (16)
CMCR: 0.148
CMCR2: 0.148
CMCR2C: 0.000

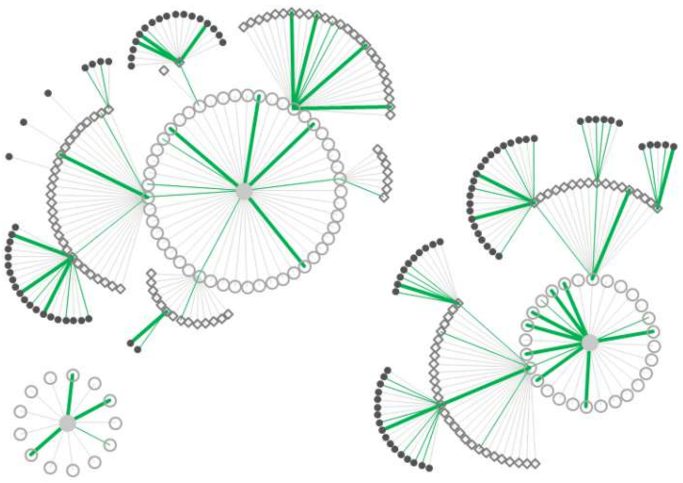


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (221)
classification to cluster "obvious relations"	detected real main meter-submeter relations (93) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (9) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (5) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 10min_C_28-30
data: synthetically generated data
time slot: 2 weeks third qtr. year
time resolution: 10 min

filtered out cases: 12,130/103,362
detected real relations: 83/319
CACC: 0.263
CACC2: 0.260

misclassifications: 0 (9)
CMCR: 0.101
CMCR2: 0.101
CMCR2C: 0.000

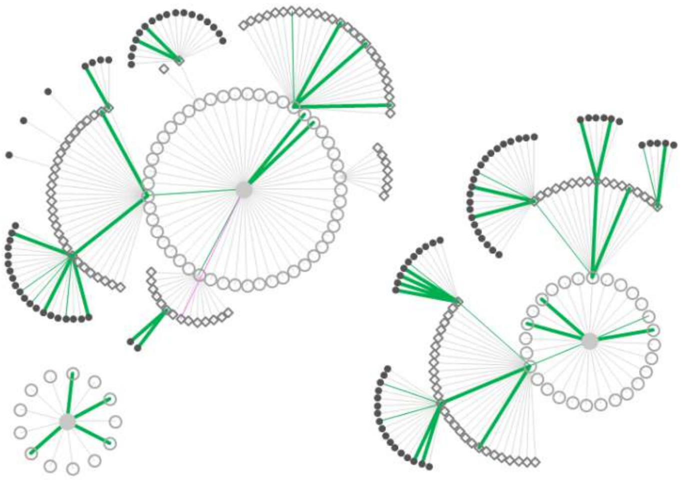


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (236)
classification to cluster "obvious relations"	detected real main meter-submeter relations (34) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (49) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0)

model: 15min_C_28-30
data: synthetically generated data
time slot: 2 weeks third qtr. year
time resolution: 15 min

filtered out cases: 11,959/103,362
detected real relations: 52/319
CACC: 0.165
CACC2: 0.163

misclassifications: 2 (4)
CMCR: 0.071
CMCR2: 0.071
CMCR2C: 0.000

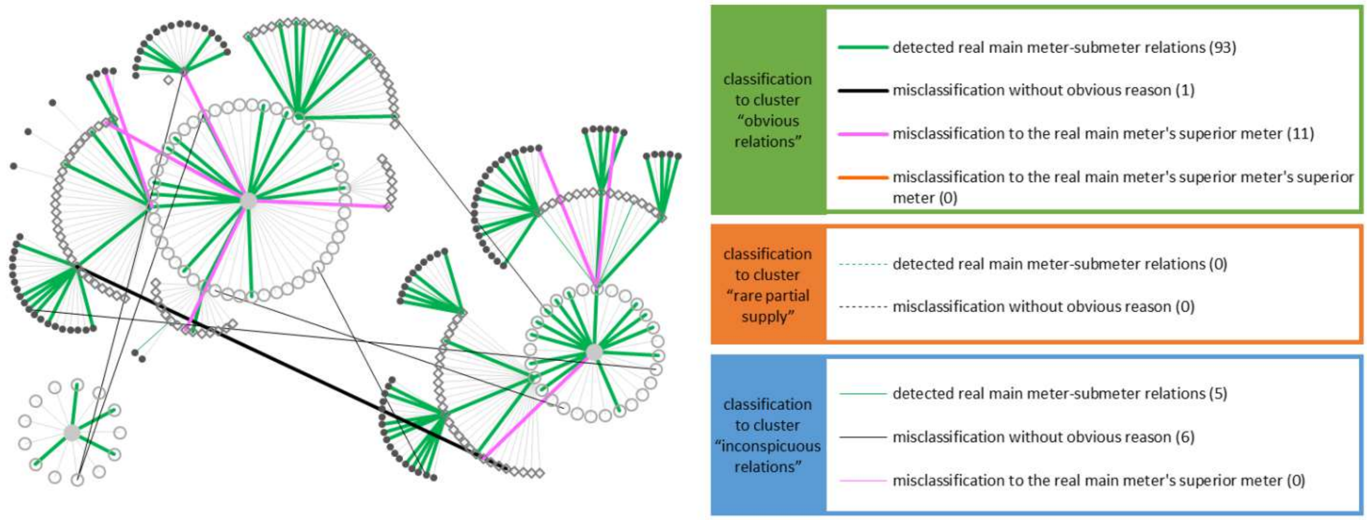


network structure	energy meters: level 0 (3) level 1 (90) level 2 (123) level 3 (106) undetected real main-meter submeter relations (267)
classification to cluster "obvious relations"	detected real main meter-submeter relations (37) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (0) misclassification to the real main meter's superior meter's superior meter (0)
classification to cluster "rare partial supply"	detected real main meter-submeter relations (0) misclassification without obvious reason (0)
classification to cluster "inconspicuous relations"	detected real main meter-submeter relations (15) misclassification without obvious reason (0) misclassification to the real main meter's superior meter (2)

model: 05min_C_7-30
data: synthetically generated data
time slot: 1 week fourth qtr. year
time resolution: 5 min

filtered out cases: 15,961/103,362
detected real relations: 98/319
CACC: 0.311
CACC2: 0.307

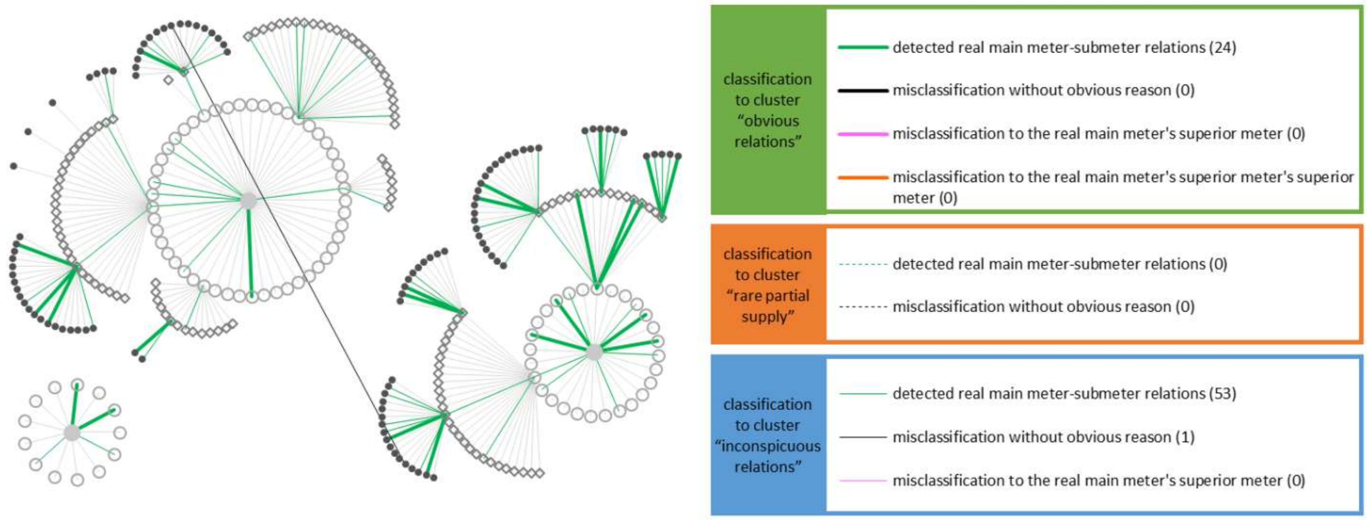
misclassifications: 18 (31)
CMCR: 0.349
CMCR2: 0.349
CMCR2C: 0.242



model: 10min_C_28-30
data: synthetically generated data
time slot: 1 week fourth qtr. year
time resolution: 10 min

filtered out cases: 15,153/103,362
detected real relations: 77/319
CACC: 0.244
CACC2: 0.241

misclassifications: 1 (12)
CMCR: 0.104
CMCR2: 0.104
CMCR2C: 0.006



model: 15min_C_28-30
data: synthetically generated data
time slot: 1 week fourth qtr. year
time resolution: 15 min

filtered out cases: 15,448/103,362
detected real relations: 49/319
CACC: 0.155
CACC2: 0.154

misclassifications: 0 (6)
CMCR: 0.120
CMCR2: 0.120
CMCR2C: 0.051