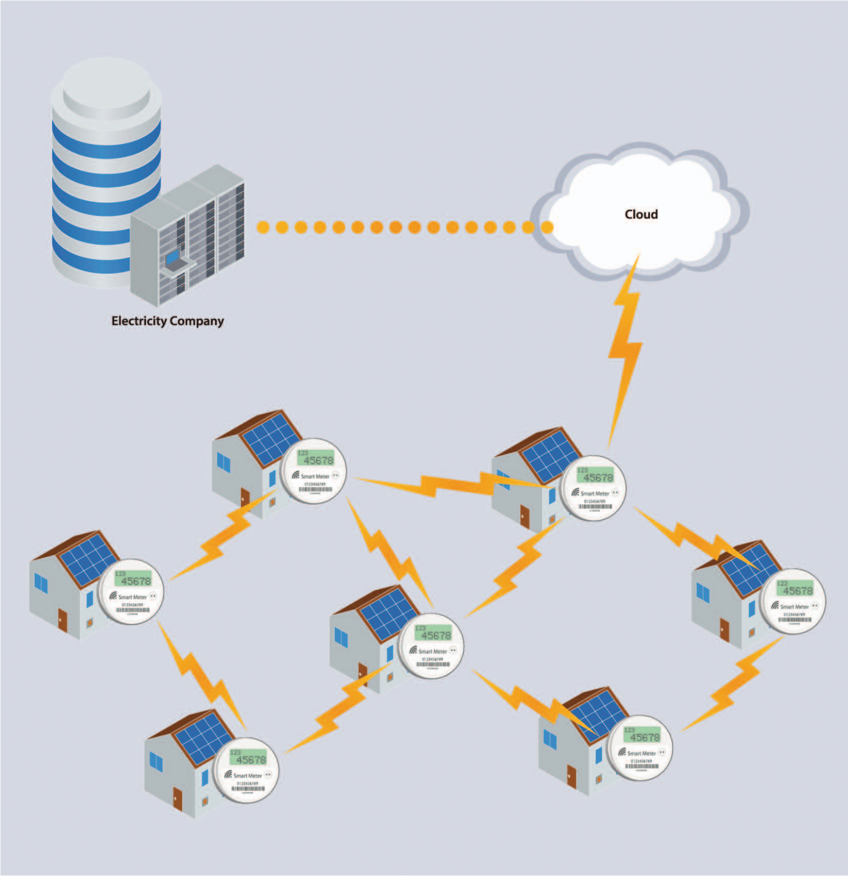
IMAGE LICENSED BY INGRAM PUBLISHING

Predicting energy consumption through machine learning using a smart-metering architecture

Paraskevas Deligiannis, Stelios Koutroubinas, and George Koronias

Extensive Internet of Things (IoT) networks consisting of billions of smart interconnected devices can serve a plethora of functions. The scale of these networks poses several architectural challenges, especially when combined with the essential requirements of reliable device telemetry, automated remote management, and multilayer security. In this article, we outline a flexible smart-metering architecture that can provide device monitoring and management in a unified manner over disparate underlying network technologies, such as narrow-band IoT (NB-IoT), LTE-Cat-M1, Zigbee, Wi-Fi, Wireless Smart Ubiquitous Network (Wi-SUN), long-range wide area network (LoRaWAN), and Sigfox.



©ISTOCKPHOTO.COM/CHOMBOSAN

The specifics of the underlying physical and link layers are abstracted away by using uniform, lightweight application layer protocols, such as Constrained Application Protocol (CoAP) and Message Queuing Telemetry Transport (MQTT). Then, we describe how this IoT architecture can be used to predict future energy consumption using past measurements and publicly available weather data. More specifically, we use energy measurements taken at 15-min intervals and build a random forest model that slightly outperforms the best autoregressive baseline. We use the test mean absolute error (MAE) and mean absolute percentage error (MAPE) as performance metrics. The significance of forecasting the future power demand lies in that it enables both consumers and grid operators to meet such demand in a more economical, efficient, and environmentally friendly way.

Digital Object Identifier 10.1109/MPOT.2018.2852564
Date of publication: 6 March 2019

0278-6648/19©2019IEEE

The scale of these networks poses several architectural challenges, especially when combined with the essential requirements of reliable device telemetry, automated remote management, and multilayer security.

Motivation and outline

IoT networks have been expanding rapidly during the past few years, and the strong upward trend is expected to continue well into the future. For example, Middleton et al. estimate that there will be more than 25 billion deployed IoT end devices by 2021, with almost 8 billion new devices being shipped during that year. Reliably connecting, managing, and securing that many endpoints gives rise to multiple challenges related to network architecture, protocols, and infrastructure.

A variety of different solutions have been proposed, each boasting unique advantages and suffering from particular weaknesses. In the first part of this article, we present a flexible architectural paradigm that is easily extensible and can be combined with a variety of underlying network technologies. To this end, we briefly discuss some key aspects of these protocols and move on to the unifying application layer mechanisms for device monitoring, management, and security.

This kind of architecture can power multiple useful applications, ranging from smart homes and monitoring electrical grids to tracking and remotely managing assets as well as performing predictive maintenance. In the second part of this article, we describe one such application, i.e., predicting future energy consumption. Successfully forecasting power demand can have tremendous economic and environmental benefits and, therefore, this research area is being extensively studied. Ahmad et al., Fumo, and Zhao and Magoules all explore different statistical and machine-learning prediction techniques and highlight the associated challenges due to the multitude of factors affecting energy consumption. In this article, we develop a random forest machine-learning model and compare its performance against a simple autoregressive baseline. The model uses past consumption measurements and weather data as features to predict the energy consumption of the following day at 15-min intervals.

Network and cloud architecture

Figure 1 shows a schematic diagram of our architecture encompassing everything from smart-metering devices to the data platform and the end user. On the bottom left of the figure are various field-deployment options using different network protocols that are mostly related to the physical and link layer of the open systems interconnection communications model. They are grouped together in four categories based on the network topology, the utilized spectrum, and the achievable communication data rates. The middle of the figure illustrates the high-level components of the cloud architecture and their respective roles. The top features the web interface for both administrators and regular users as well as an interface to separate, possibly third-party, data platforms and servers over a protocol bridge.

Brief overview of underlying network technologies

The different competing technologies differ in various areas, such as the network topology (star versus mesh), bandwidth and specific frequency bands used for transmission, required power, maximum packet payload size, achievable data rates, maximum communication range, maximum number of communicating devices, and security mechanisms. Due to these differences, the technologies all offer distinct advantages making them more suitable for particular applications than others. There is no universally superior technology. Since a full comparison of the available technologies is out of the scope of this article, we will focus on a few key elements, mainly regarding the network topology.

Cellular networks are composed of cellular base stations, each serving multiple end devices. An endpoint device can switch to a different base station, but endpoints cannot directly communicate with each other. Therefore, a cellular network has a star topology. There is currently a trend away from traditional protocols, such as general packet radio service, and toward the modern alternatives of NB-IoT and LTE Cat-M1, which were designed specifically for IoT devices and can handle massive deployments of low-power IoT devices in remote areas with intermittent connectivity. In contrast to the rest of the technologies, which use unlicensed spectrum for their operation, possibly resulting in unwanted interference and performance degradation, cellular networks operate on legally protected, licensed spectrum.

In mesh networks, each device can communicate with any other nearby device, and messages can be routed dynamically, thus mitigating connectivity loss even if some of the devices are offline. A protocol-specific gateway, coordinator, or base station acts as the bridge between the mesh network and the Internet. Zigbee is a very popular open standard that finds many applications in smart homes and home automation. The Wi-SUN Alliance promotes a competing mesh communication protocol.

Traditional Wi-Fi can be used for small-scale deployments, usually within a small building. Its main advantage is the extensive, already-existing infrastructure in urban areas, which makes device deployment easy and accessible to practically everyone. All one needs is a Wi-Fi access point (acting as a field gateway) with Internet access. Bluetooth Low Energy offers three communication profiles: point-to-point (i.e., one-to-one device communication), broadcast (i.e., one-to-many), and mesh. Therefore, both star and mesh topologies are achievable. Finally, popular ultranarrow-band IoT networks, such as LoRaWAN and Sigfox, exhibit star topologies using protocol-specific field gateways and base stations. They operate on unlicensed spectrum from 169 to 925 MHz depending on the area and the specific protocol.

IoT networks have been expanding rapidly during the past few years, and the strong upward trend is expected to continue well into the future.

Cloud services and infrastructure

The cloud architecture consists of several backend services that communicate with the outside world through three interfaces. The cloud gateway is responsible for the communication with the devices, the application web server is responsible for the user and administrator interface, and the protocol bridge connects this platform to an external data server or system.

The cloud gateway ingests device telemetry and ensures that control messages are reliably received by the target devices. It allows the abstraction of the underlying network technology by presenting a uniform application layer to the backend services. The protocol bridge translates between the common application protocol used by the current platform and the application protocols used by external systems.

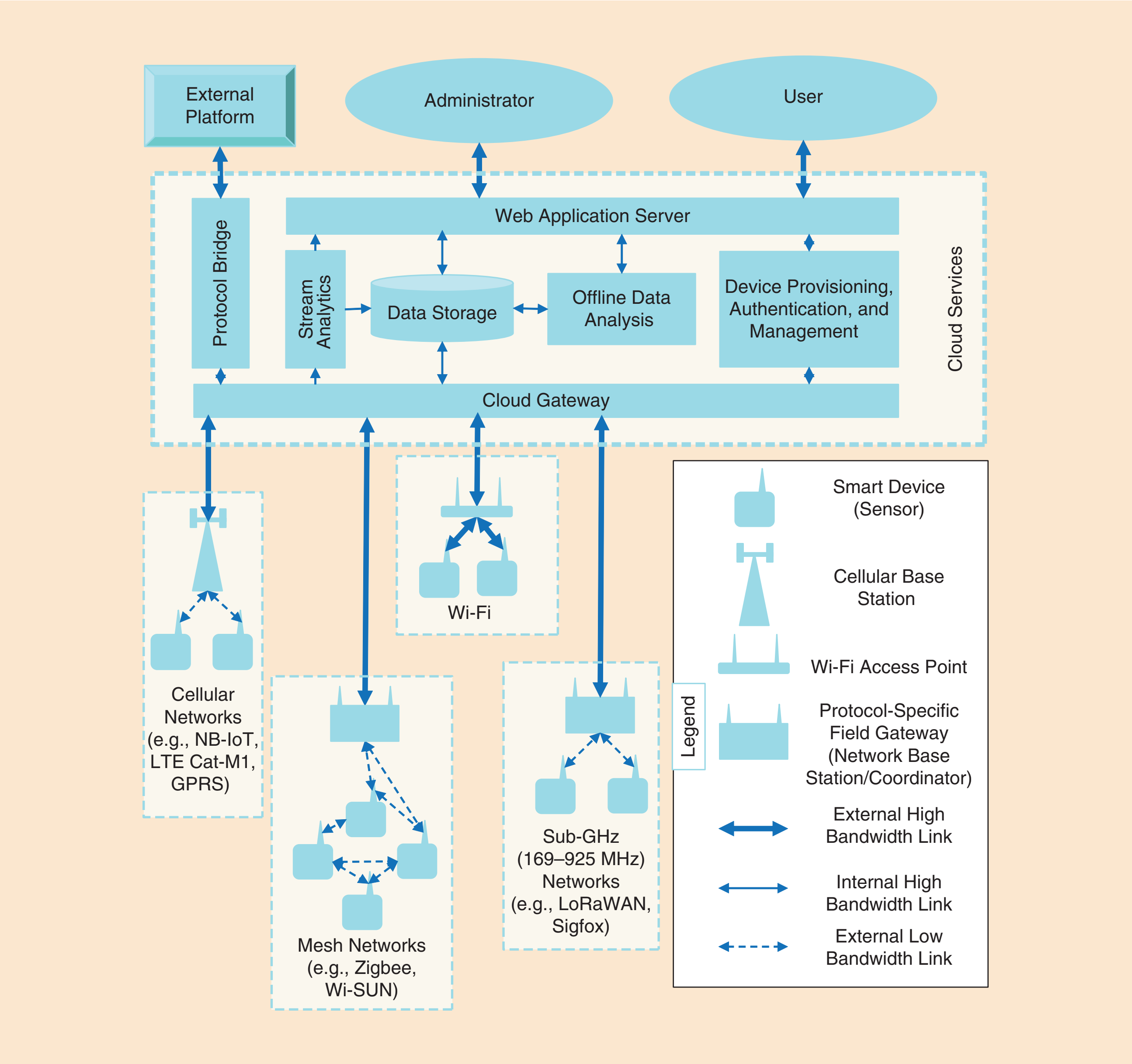

External Platform
Administrator
User
Web Application Server
Protocol Bridge
Stream Analytics
Data Storage
Offline Data Analysis
Device Provisioning, Authentication, and Management
Cloud Services
Cloud Gateway
Wi-Fi
Cellular Networks (e.g., NB-IoT, LTE Cat-M1, GPRS)
Mesh Networks (e.g., Zigbee, Wi-SUN)
Sub-GHz (169–925 MHz) Networks (e.g., LoRaWAN, Sigfox)
Legend
Smart Device (Sensor)
Cellular Base Station
Wi-Fi Access Point
Protocol-Specific Field Gateway (Network Base Station/Coordinator)
External High Bandwidth Link
Internal High Bandwidth Link
External Low Bandwidth Link

FIG1 An example of a flexible network architecture combining different physical and link layer protocols. GPRS: general packet radio service.

There is currently a trend away from traditional protocols, such as general packet radio service, and toward the modern alternatives of NB-IoT and LTE Cat-M1.

The web application server provides the user interface and application programming interface necessary for end-user applications and device management and monitoring as well as platform administration. It is also responsible for securing these interfaces. The data storage service is a hierarchical database management system mainly storing device telemetry and metadata.

The device provisioning, authentication, and management service is essential for securely controlling devices. It handles the device credentials used for authentication and encryption as well as the device configuration and lifecycle management. Finally, data analytics are performed both online, as data arrive (using stream analysis algorithms), and offline in batch mode. They can be used for energy consumption prediction and predictive maintenance, among other purposes.

Transport and application layer protocols—Device telemetry and management

In this section, we briefly discuss two lightweight machine-to-machine IoT application protocols. When the underlying transport protocol is User Datagram Protocol (UDP) [e.g., when Transmission Control Protocol (TCP) is too expensive in terms of device and network resources], then the extremely lightweight CoAP can be used. It offers a Representational-State-Transfer-like interface, similar to the traditional HyperText Transfer Protocol but using a fraction of the packet size.

When TCP is used at the transport layer, then MQTT carrying a JavaScript Object Notation payload is the preferred way. MQTT is a publish-subscribe protocol that was designed with massive IoT deployments in mind. Multiple clients can publish messages to a channel of the MQTT broker. Then, the broker sends these messages to all of the clients that have subscribed to that channel. Thus, many-to-many device communication is possible. If necessary, CoAP can be later translated to MQTT, e.g., at the cloud gateway.

Security considerations

No IoT solution is complete without some robust security mechanisms. The fact that IoT devices often use cheap, low-end hardware to communicate over the air and can be deployed in remote areas with minimum supervision makes them prime targets for cyberattacks. Every aforementioned network technology provides its own security mechanisms for over-the-air transmissions, even if they differ based on the level of protection they offer. However, an IoT solution greatly benefits from a uniform, extra layer of end-to-end security irrespective of the underlying network protocol.

The battle-tested industry standard for securing TCP connections is the Transport Layer Security (TLS) protocol and, more specifically, its latest versions with the appropriate cryptographic ciphers. The equivalent for UDP-based communication is the related Datagram TLS protocol. These protocols offer end-to-end encryption and authentication and provide options for robust public key cryptography.

The sensitive cryptographic device keys should be stored in secure hardware modules, e.g., in the embedded universal integrated circuit card (subscriber identity module) chip, and the device should provide antitampering protection. In addition, extra care should be taken to secure over-the-air firmware updates, due to their sensitive nature. A solution is to have them digitally signed using the application server's X.509 certificates.

Finally, the user and administrator web interface can be secured by using HyperText Transfer Protocol Secure with the appropriate TLS ciphers and Secure Sockets Layer certificates, in conjunction with two-factor authentication for the end users.

Energy consumption prediction

We will now show how the data gathered through an IoT smart-metering architecture can be used to forecast power demand. We start by visually presenting the available consumption measurement data and proceed with describing the models used. Then, we formally define the relevant performance metrics before providing the test results.

Our goal is to predict the energy consumption of each day at 15-min intervals (four intervals/h × 24 h/day = 96 prediction values/day) given the past consumption measurement data up to and including the previous day as well as the temperature, humidity, precipitation, wind speed, and direction data (both the historical data as well as the forecast for the next day). All measurement data have a resolution of 15 min. Historical weather data come from Modern-Era Retrospective Analysis for Research and Applications, Version 2 and forecast data are available from the Global Forecast System. Our measurement data were obtained from a large office building during a period of almost two years.

Data visualization

Figure 2 presents the energy consumption of a typical week, according to our measurement data. Specifically, it presents the consumption for each 15-min period of the week, averaged across all of the weeks in our sample. We can clearly observe a significant difference in consumption between working hours (approximately 8 a.m.–4 p.m. on weekdays) and nonworking hours (weekdays outside that hour range and weekends).

Models

We have chosen random forests as the basis of our machine-learning model because of their simplicity and reputation for achieving good results in a variety of machine-learning problems with minimal parameter fine-tuning. Other authors have chosen artificial neural networks and support vector machines (SVMs) for power demand forecasting. To limit the training data set to only the most relevant measurement data, we have decided to ignore data more than 30 days old, thus training the random forest using a sliding 30-day window. This approach has the added benefit of reducing the computation costs for training and testing the model. We use the scikit-learn Python package to train and test this model with ten trees. We call this model *RF-30*.

The battle-tested industry standard for securing TCP connections is the Transport Layer Security protocol.

A simple autoregressive model is used as a baseline, against which we will evaluate the performance of our model. After a brief parameter search, we chose the autoregressive model whose prediction equals the consumption of the same 15-min period during the previous day, e.g., if we need to predict the consumption on 15 January 2017, between 10:00 a.m. and 10:15 a.m., then the baseline model's prediction will equal the known consumption on 14 January 2017 during the same period. We call this model *AR-1*.

Performance metrics

Two performance metrics stand out as the most appropriate for energy consumption prediction: the test MAE and the test MAPE. We also present the MAE as a percentage of the mean consumption across all our samples, so as to have a better understanding of the relative prediction error.

MAE refers to the average absolute difference between the actual and predicted consumptions across all time intervals. MAPE is similar, but differs from MAE in that each absolute difference is normalized by dividing it with the actual consumption during the same time interval.

Results

Table 1 presents the results (MAE and MAPE) for the two models. We observe that RF-30 slightly outperforms AR-1 in both metrics. Figure 3 shows the histogram of the absolute error distribution for the RF-30 model.

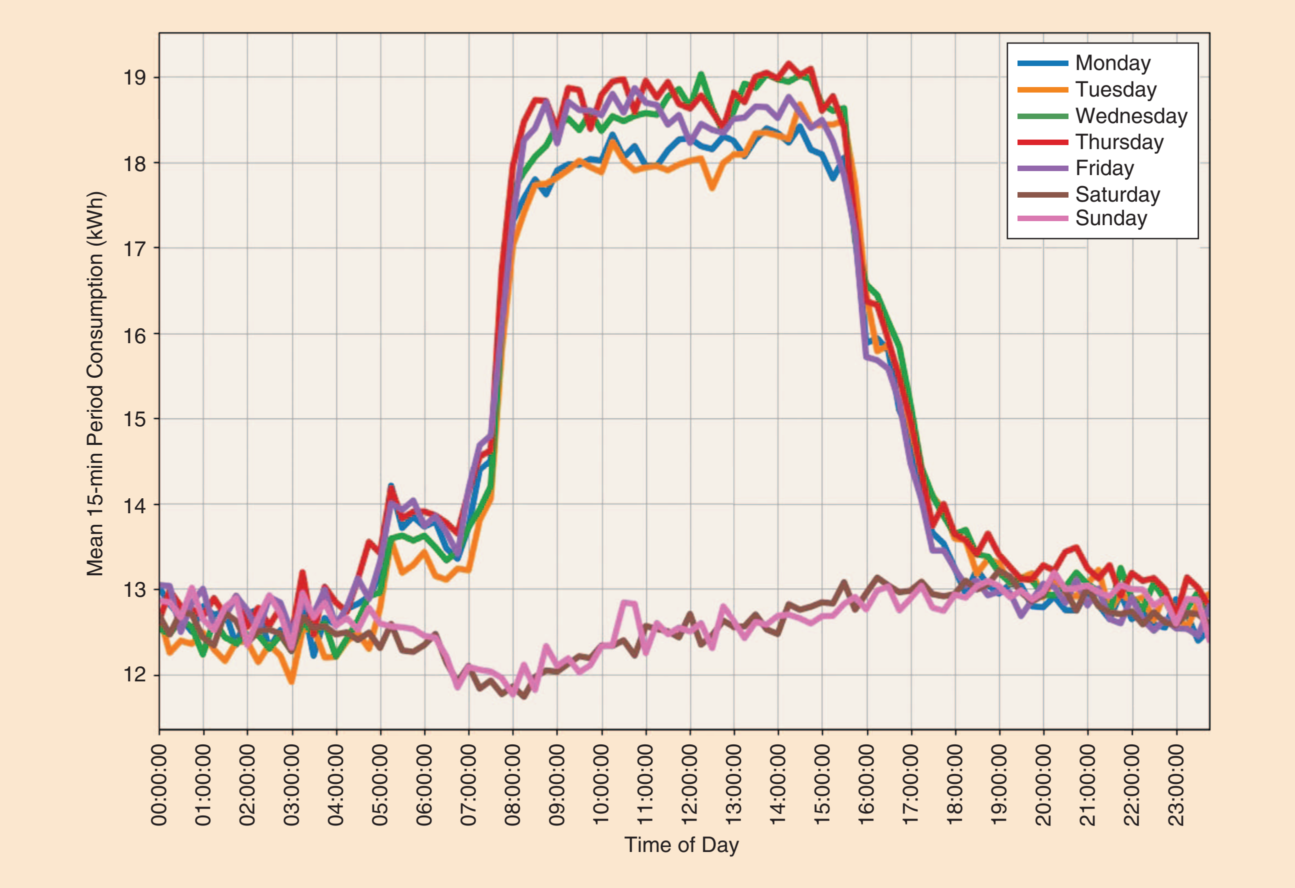


FIG2 The average consumption for each 15-min period of the week. This can be interpreted as a typical week pattern when seasonal trends are averaged away.

Two performance metrics stand out as the most appropriate for energy consumption prediction: the test MAE and the test MAPE.

TABLE 1. The performance results.

METRIC	AR-1	RF-30
MAE (kWh)	1.98	1.71
MAE (% of mean)	12.96	11.23
MAPE (%)	17.98	16.16

Conclusion

We have discussed some requirements and challenges of an IoT architecture and presented a complete and flexible solution incorporating disparate network technologies. We then explored an IoT and machine-learning application in a power-demand forecast.

Additional work could assess the suitability of the architectural solution proposed for different use cases and compare it with alternative architectures. Moreover, the energy consumption prediction application is meant only as a starting point. Other directions include fine-tuning the parameters of the random forest as well as exploring different approaches, such as time series analysis, deep neural networks, and SVMs. An interesting idea would be to measure the performance of algorithms for stream analytics, as this would provide real-time predictions.

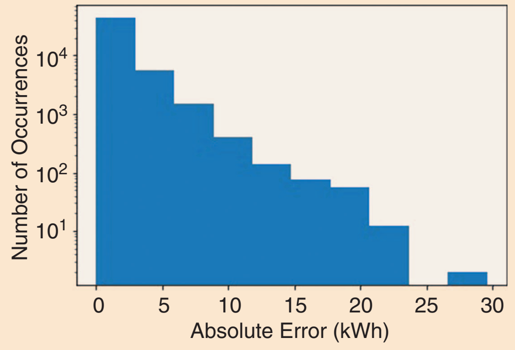


FIG3 A histogram of absolute error distribution. The number of occurrences is shown in log scale.

Read more about it

- A. S. Ahmadet et al., “A review on applications of ANN and SVM for building electrical energy consumption forecasting,” *Renewable Sustainable Energy Rev.*, vol. 33, pp. 102–109, 2014. doi: 10.1016/j.rser.2014.01.069.
- N. Fumo, “A review on the basics of building energy estimation,” *Renewable Sustainable Energy Rev.*, vol. 31, pp. 53–60, Mar. 2014. doi: 10.1016/j.rser.2013.11.040.
- H.-X. Zhao and F. Magoules, “A review on the prediction of building energy consumption,” *Renewable Sustainable Energy Rev.*, vol. 16, pp. 3586–3592, Aug. 2014. doi: 10.1016/j.rser.2012.02.049.
- F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *J. Mach. Learning Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
- P. Middleton, T. Tsai, M. Yamaji, A. Gupta, and D. Rueb. (2017). Forecast: Internet of Things—Endpoints and associated services, worldwide, 2017. Gartner, Stamford, CT. [Online]. Available: https://www.gartner.com/doc/3840665/forecast-internet-things–endpoints
- Global Modeling and Assimilation Office. (2015). MERRA-2 tavg1_2d_slv_Nx: 2d,1-Hourly, Time-Averaged, Single-Level, Assimilation, Single-Level Diagnostics V5.12.4. Goddard Earth Sciences Data and Information Services Center, Greenbelt, MD. [Online]. Available: https://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=NASA&KeywordPath=%5BProject%3A+Short_Name%3D%26%23039%3BREASON%26%23039%3B%5D&OrigMetadataNode=GCMD&EntryId=GES_DISC_M2T1NXSLV_V5.12.4&MetadataView=Full&MetadataType=0&lbnode=mdlb4 Accessed March 2, 2018, DOI:10.5067/VJAFPLI1CSIV

About the authors

Paraskevas Deligiannis (paras.delig@gmail.com) is a solution architect with Meazon SA, Patras, Greece.

Stelios Koutroubinas (s.koutroubinas@meazon.com) is a cofounder and the chief executive officer of Meazon SA, Patras, Greece.

George Koronias (g.koronias@meazon.com) is a cofounder and the executive chairman of Meazon SA, Patras, Greece.