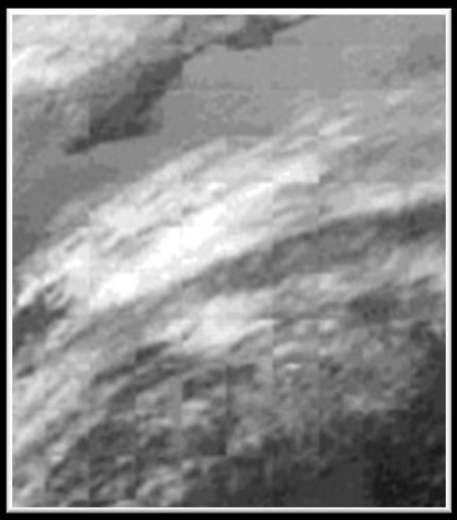
Forecasting of Solar Power Generation using Real-time Meteorological Information

Siraj Sabihuddin, Alexander Grishin, Ho Ko (Ubiik Inc.)

April 14, 2020



An overview of Ubiik's solution for prediction of real solar PV generation capacity in Taiwan is provided. Such a solution consists of a back-end data management and communication infrastructure; a meteorological data pre-processing and conditioning system; an analytical engine and machine learning system to learn from datasets and produce a highly functional learning model; a post-processing system to convert real-time output into estimates of real solar power generation capacity with known predictive accuracy; and a visualization system and service platform that provides actionable data analytics.

Traditionally, power plants have utilized generator systems with fairly predictable and deterministic power outputs. The increasing penetration of photovoltaic (PV) plants has changed this situation drastically. The output of such plants, after taking into consideration solar irradiance in the upper atmosphere, is highly dependent on weather conditions. These weather conditions can rapidly alter the output of Solar Photovoltaic (PV) plants. The result of these alterations is instabilities in the electricity grid and with potential resulting problems with downstream equipment and devices. This article attempts to provide an overview of Ubiik's complete and comprehensive solution for the prediction of real solar generation capacity in Taiwan and around the world.

www.ubiik.com

19F, No. 17, Sec. 1, Chengde Rd., Datong Dist., Taipei City 103, Taiwan

Taipei: +886-2-7751-5855
Hsinchu: +886-3-668-4886

UBIIK

Figure 1 below provides a general overview of this solution. Ubiik's solar PV forecasting system takes in potential locations for future or current solar installations and provides a series of cloud services that can be utilized by customers for instant querying and of expected output of an installation. The solution consists of:

A. A back-end data management and communication infrastructure to provide secure and reliable cloud services.

B. A meteorological data pre-processing and conditioning system used to produce consistent reliable model datasets for learning and analysis

C. An analytical engine and machine learning system to learn from datasets and produce a highly functional learning model

D. A post-processing system to convert the real-time output (from the model) into estimates of second-by-second real solar power generation capacity estimates with known predictive accuracy; and

E. A visualization system and service platform that provides actionable data for use by the Customers for planning for and operation of solar PV systems.

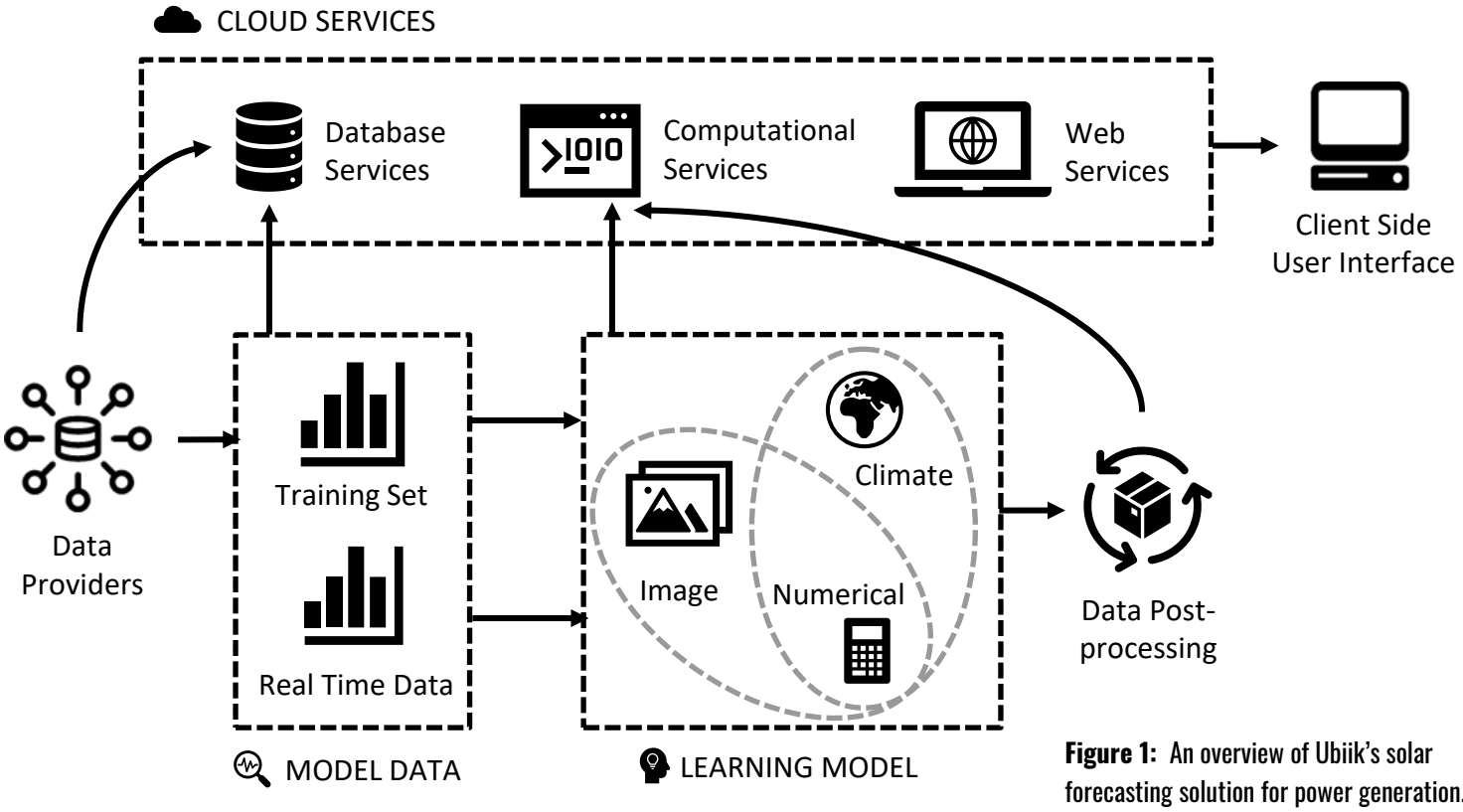


Figure 1: An overview of Ubiik's solar forecasting solution for power generation.

A. Backend Data Management & Communication

The backend system is based around Ubiik's existing database services, computational services, web services and a communication backbone infrastructure. Figure 2 below demonstrates the overall infrastructure in greater detail. The backend database, computation and web services are duplicated at two different locations. The first location provides a stable primary backend system via a redundant and firewalled switching network that interconnects via VPN to the customer's remote prediction service portal or user interface. This service portal accesses the backend cloud prediction services provided by Ubiik.

Ubiik development team periodically does research and development to test improvements in performance and new algorithms and to ensure that, in the event of failure in the prediction accuracy of the system, improvements can be made quickly. For this purpose, a secondary backend (cloud) system is available at a second location. This secondary system provides a mechanism for testing algorithm performance and correspondingly improving predictions.

To ensure that customers always gets the best possible service, updates are rolled out by Ubiik into the secondary R&D backend system first. When a stable release of the new algorithm and prediction system has been released, the secondary system is synchronized with the primary system to provide seamless transition to a new prediction system that can be accessed through the remote predictive services portal or user interface. In the background, data collection and pre-processing services provide the backend with continuously improving datasets from satellite and other meteorological weather data providers.

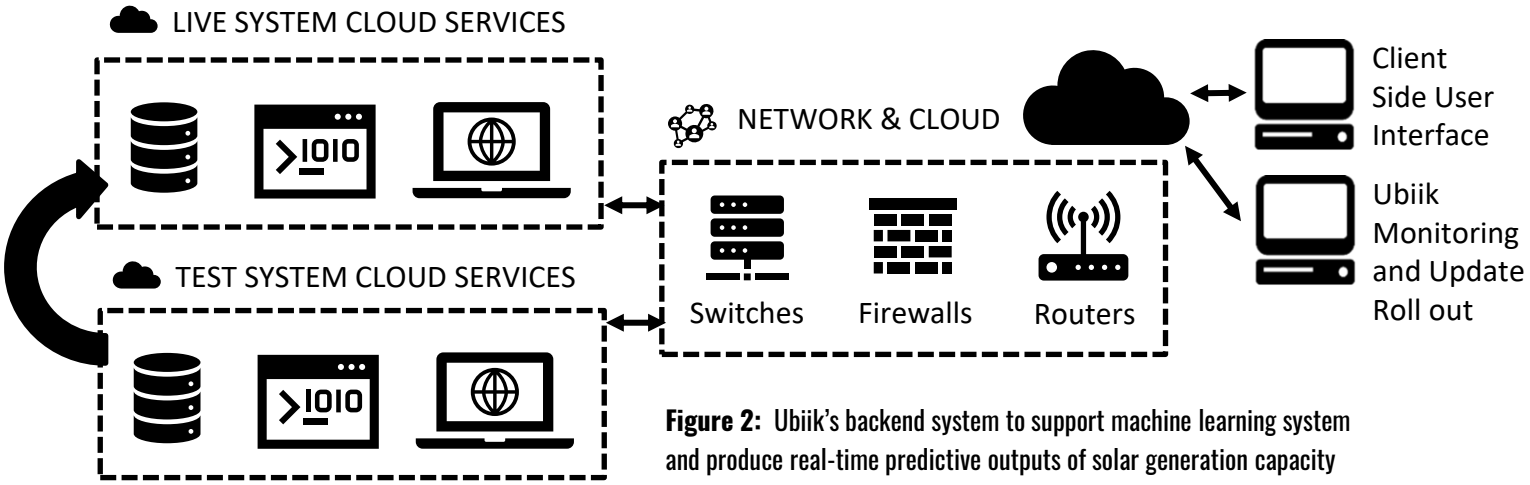


Figure 2: Ubiik's backend system to support machine learning system and produce real-time predictive outputs of solar generation capacity

B. Pre-Processing Infrastructure & Data Providers

The fundamental part of any complex machine learning system is the utilization of large amounts of data from diverse sources. These data providers are used to extract model relevant data for prediction and output of PV generation capacity. Unlike other modeling solutions and approaches that rely mostly on a single data provider, Ubiik's solution takes advantage of wide ranges of diverse and independent data sources. To do this, a significant data pre-processing infrastructure is also needed. Input data from these diverse sources cannot be easily integrated without a framework similar to one provided by Ubiik.

1. Data Sources & Providers

Input data sources normally incorporate a range of different kinds of information. The single biggest impact on power generation capacity is from solar irradiance (Figure 3) and variations away from this correlation are a result of weather conditions specifically.

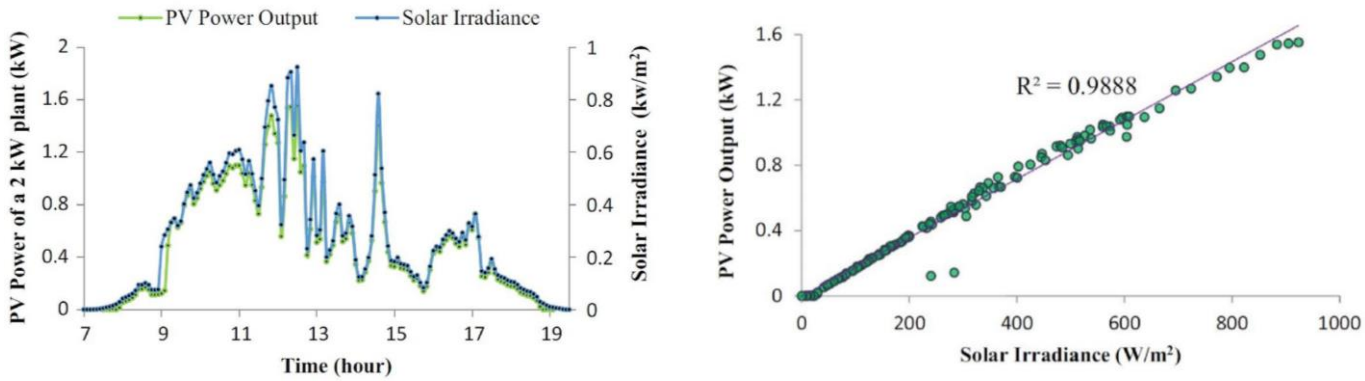


Figure 3: impact of solar irradiance on solar PV output power. Data from [1]

After considering variations in irradiance in the upper atmosphere, any surface variations in Global Horizontal Irradiance (GHI) will be heavily influenced by the additional factors shown in Table 1. Predictive modeling needs to take these factors into consideration. Historical data fed in as input into this modeling can lead to poor accuracy for even highly functional models.

Ubiik's solution has a unique selling point. Most models only make use of globally available satellite data. As such these existing approaches are unable to cater for the unique on-the-ground knowledge of the customer's local power generation and power systems data and obtain accurate location data for later prediction. By incorporating this local information, a much higher degree of accuracy can be achieved in predictions. Ubiik's solution does just this. It includes a mechanism for fast data collection and processing (from data providers) that also ensures minimization of latency artificially introduced by delays between data availability and utilization.

Meteorological factor	Correlation coefficient
Solar irradiance	0.9840
Air-temperature	0.7615
Cloud type	-0.4847
Dew point	0.6386
Relative humidity	-0.4918
Precipitable water	0.3409
Wind direction	0.1263
Wind speed	0.1970
Air pressure	0.0815

Table 1: The impact of various variables on solar PV plant output power. Data from [1].

2. Data Pre-processing

Each type of data source has benefits, which, when combined together, enhance the AI-based performance and increase effective temporal and spatial resolution as well as robustness. To accommodate the various local sources of information with the global data and provide useful datasets for training and real-time prediction of Solar PV output, Ubiik has implemented additional pre-processing algorithms.

These algorithms take the multiple local data sources, that are of low quality, and correlates them to generate high quality training sets. The training sets are then used as a tool for training the predictive model while real-time datasets. This provides the necessary input data to produce the real-time predictions of PV output and deliver them to the customer on an intra-hour schedule. Ubiik's pre-processing approach addresses four main problems that classical prediction tools are unable to handle robustly.

1. Often data providers do not provide uniformly structured data. Compounding this problem is the possible inconsistency of data format used by a single provider or the change in format introduced by the provider without notice. Data from various providers is, semi-automatically reformatted and restructured before input into the predictive model.
2. Datasets may also be missing data for certain time frames. Any missing elements are identified and appropriately removed from the datasets.
3. Data points may be incorrectly labeled. This is particularly problematic for training data as it reduces the predicative capability. Ubiik's algorithm provides a method for identifying these points.
4. Datasets may be aggregated across wide geographic regions and may not be labeled as such. PV predictive accuracy depends highly on training data that is specific to one location. These datasets are eliminated before prediction.

A key part of the algorithm for solving these steps relies on a multi-model approach for identifying good quality versus bad quality data. This is demonstrated in Figure 4. Changes in the accuracy of these models and related correlation of model behavior, after addition of new training data, provides an automated method for quickly eliminating poorly formed datasets.

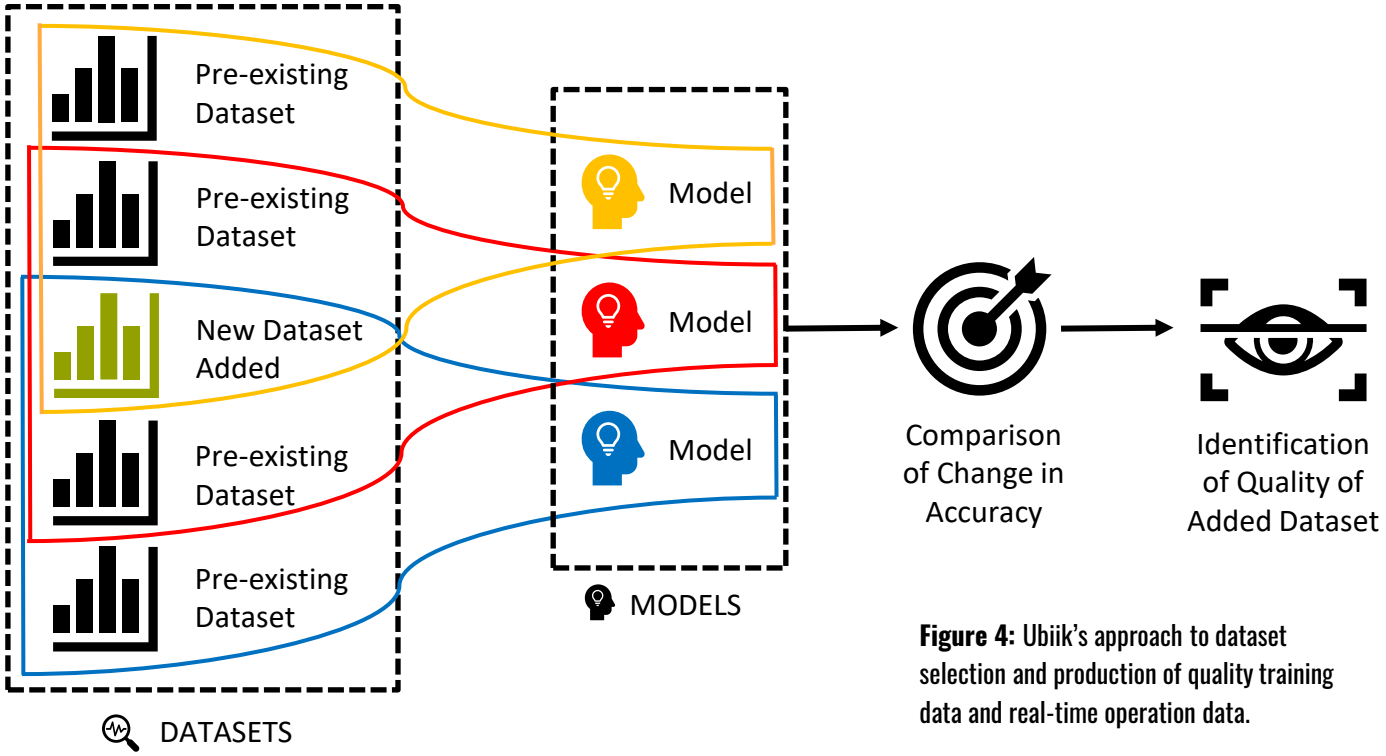


Figure 4: Ubiik's approach to dataset selection and production of quality training data and real-time operation data.

3. Training Data Volume Enlargement

Ubiik's data cleaning and pre-processing algorithms ensure that data that is incorporated into the training dataset is of high quality. However, since data quantity is a very limited resource, Ubiik uses a novel technique of artificially enlarging the size of the training dataset by localizing and cropping data for different locations from satellite images and feeding them into one model instead of using a single model for each location. The distinction between the one model for different locations approach versus Ubiik's multiple models for different locations scheme is shown in Figure 5.

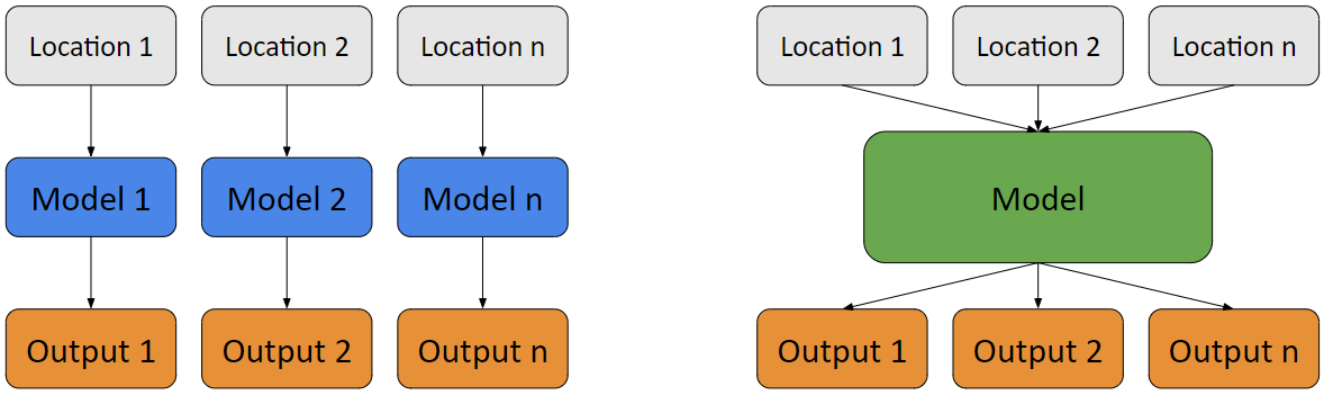


Figure 5: Multiple models for multiple locations vs one model for multiple location

The benefits of this approach are as follows:

1. Number of training samples can be effectively increased by adding new locations in the training dataset.
2. Does not require historical records for every location.
3. Can be applied to a new locations without going through the training process (transferred learning).
4. More robust to variations while producing stable results.
5. Better ability to generalize to a wider variety of geographic locations.
6. Simpler process for updating elements of the model for a wider range of applications.

C. Predictive Model

Typically, a 2 to 3-day forecast has been used for scheduling, dispatch and regulation of the grid by operators – this forecast takes the form of intra-hour (seconds to hours), intra-day (hours) and day-ahead (hours to days) forecasting. Ubiik's predictive solution can achieve prediction into the 72 hour window required by some electrical utility companies - Ubiik can effectively cover all three forecast time intervals. This is done through the use of a modular multistage adaptive AI prediction system. Figure 6 elaborates, briefly, on the general architecture of this multistage modeling approach.

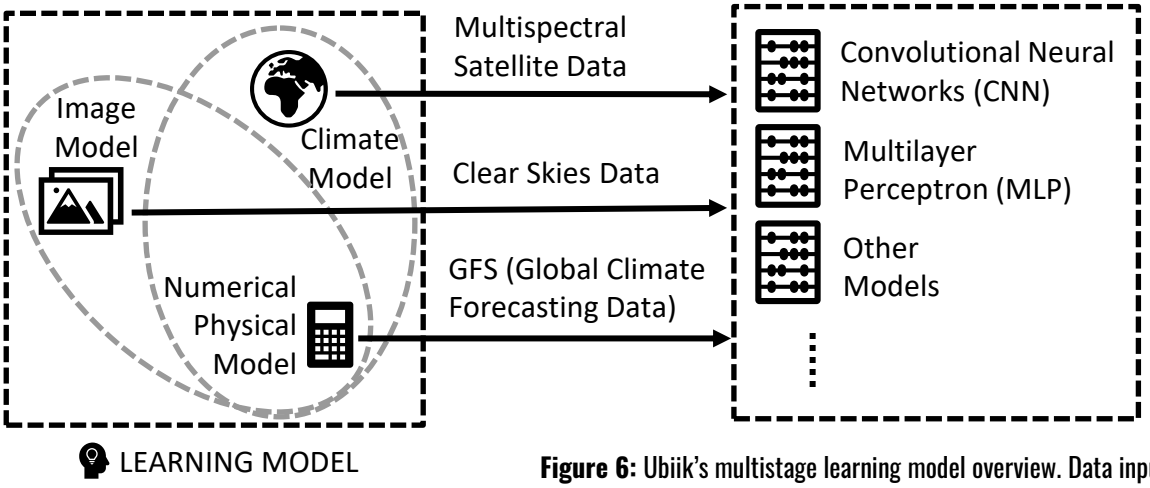


Figure 6: Ubiik's multistage learning model overview. Data input is fed into three distinct models. These act to reduce the dimensionality of the data for a second layer of processing with convolutional neural networks supported by a variety of other stages of modeling techniques.

Largely, existing approaches can be classified into the following main categories: (a) Persistence models (that are computationally simple and mimic past behavior as a proxy for future behavior); (b) statistical models (that utilize historical and real time time-series data for predictive analysis); (c) physical models (that take into account the wider arena of real-world physical phenomenon for weather prediction). In practice a hybridized approach is more effective – one that combines all three categories of modeling - to support high accuracy PV power output prediction [4][1].

It is this hybridized approach that Ubiik utilizes through multistage modeling to achieve its unprecedented accuracy. In fact, a recent review from 2020 [1] of existing solar PV prediction techniques clearly demonstrates this.

A review of about 19 distinct novel models demonstrated a range of Normalized Root Mean Square Error (NRMSE) prediction accuracies ranging from 4% to 54%. Of these, only three demonstrated accuracies at or below 5% error. All such approaches used some combination of deep learning, convolutional neural networks, and other hybridizations of physical and statistical models. It is a combination of these approaches that Ubiik also makes use of for its predictive algorithms.

The Ubiik solution, thus, stands out among the top-most effective techniques in terms of modeling. It uses four main classes of models for its multi-stage approach (see Figure 6). These provide tools for dimensionality reduction and higher accuracy via the following approaches:

1. Multispectral satellite image modeling for exact plant location prediction.
2. Global climate forecast model for weather related humidity, dew point, wind, pressure and cloud movement predictions.
3. Numerical physical based models for deterministic clear skies, global horizontal irradiance predictions.
4. Deep Learning (Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Ensemble learning, Active learning)
5. Flexible Model Selection

Collectively these stages form a highly effective multistage modular adaptive predictive AI system. At the core of this system is independent modules that extract features from independent non-uniform diverse data sources. The modularity of the system allows the use of machine learning models of different complexity to be used. Further it allows matching of the existing volume of training data from each data source independently while maximizing the performance of each module individually. The result is an overall system that eliminates the effect of a bottleneck. A modular approach allows the reconfigurable selection of the most suitable machine learning algorithms for data providers and sources. The sections following elaborate.

1. Satellite Images

Himawari8 is a geosynchronous multi-spectral satellite that provides data on typhoons, rainstorm, weather forecasts and other related reports for Japan, East Asia and the Western Pacific region. It has 16 different bands in total (see Figure 7).

Each band is analyzed by Ubiik to account for the impact of water vapor, ozone, carbon dioxide and more. The combination of these variables allow the extraction of points of interest for prediction of exact plant location at a rate of every 10 minutes.

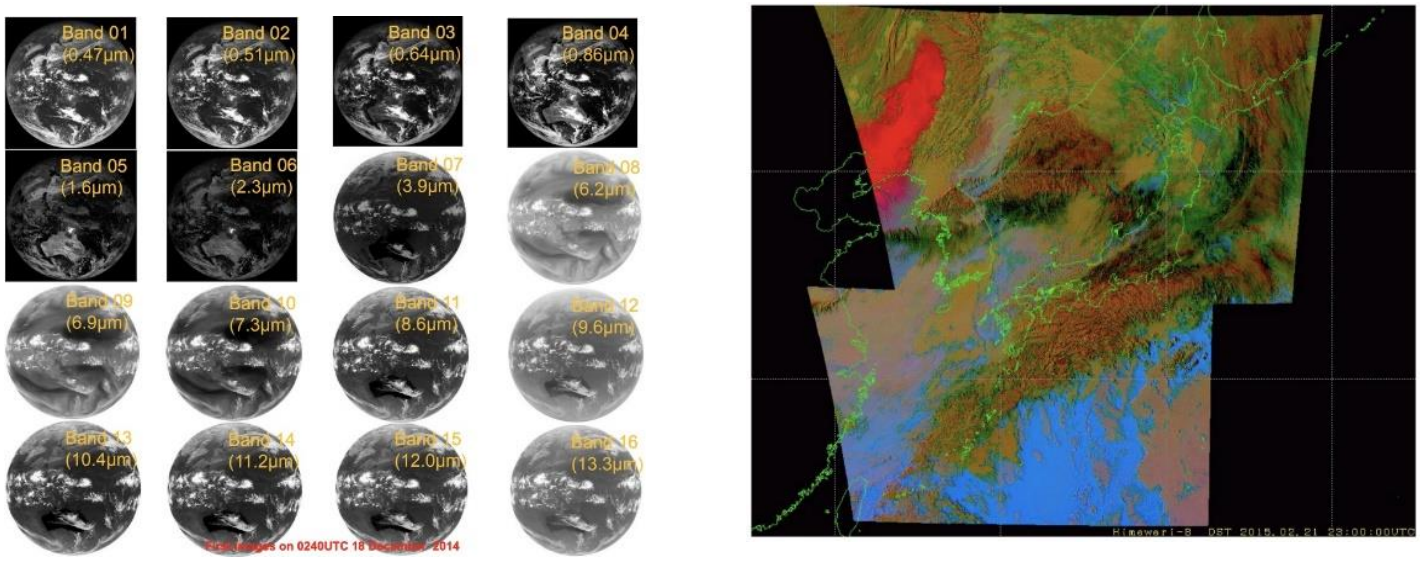


Figure 7: Ubiik's satellite models: On the left, 16 spectral bands utilized from satellite data; on the right, an example of desert dust visualization using multi-spectral analysis.

2. Global Climate Forecasts

Ubiik's selection of climate forecast model achieves very high performance over a wide range of forecasting intervals between intra-hour and day-ahead. To do this a clever selection of climate model is used. This climate mode is the Global Climate Forecast System (GFS). This forecast approach is highly suited to day-ahead and longer-term predictions (see Figure 8). Other similar long-time scale models have significant problems in terms of optimality for solar prediction while also not being fully open – GFS provides a fully open platform for forecasting. Ubiik's utilization of GFS keeps in mind some of the most important parameterizations necessary for solar PV prediction as seen in the figure. This is reflected in the higher accuracy of prediction at lower latitude areas (e.g. in Taiwan). The high hourly resolution is another important tool that Ubiik uses for prediction accuracy. The localization and spatial resolution of the GFS model is low – Ubiik's solution leverages other data providers and modeling techniques to take this low spatial resolution and increase the accuracy of localized weather predictions.

This is something unique to Ubiik's approach for prediction. To address the shorter-term prediction, the GFS is complimented by Ubiik's statistical predictions in the other stages. By modular selection of the specific models, it is possible to produce high prediction accuracy of PV generation for all time scales (see Figure 8).

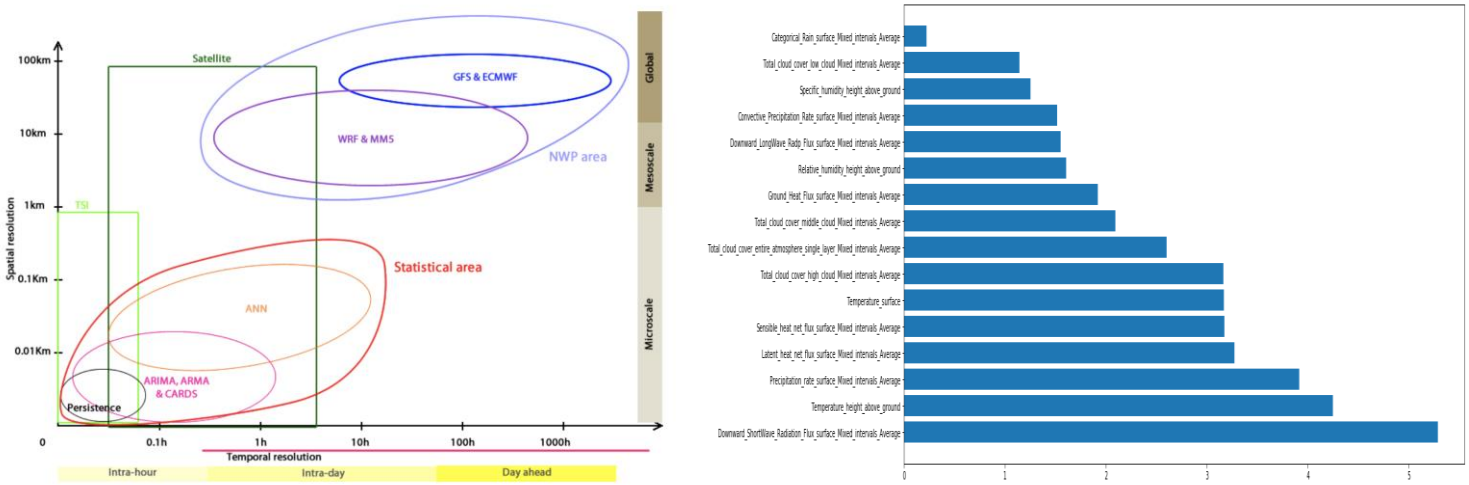


Figure 8: Ubiik's choice of multi-stage modeling provides very wide range forecasting from short term intra-hour to long term day-ahead. On the left: The statistical models ANN provide short-term intra-hour forecasting while the global climate model (GFS) provides long-term day-ahead forecasting. On the right: Parameters in GFS modeling, used by Ubiik, ordered by relative importance.

3. Numerical Physical Based Models

Ubiik leverages physical deterministic modeling during clear sky events. This clear sky solar radiation model uses ground truth data such as the position of the sun, diffusion rate of light, and historical turbidity data in corresponding time of the year to calculate the solar radiation in an ideal clear sky situation. This approach helps Ubiik increase the accuracy compared to purely albedo dependent satellite data and estimation of radiative flux in climate models. The result of the numerical and physical modeling thus largely reduces the error rate from other approaches.

4. Deep Learning

Deep Learning architectures demonstrate better efficiency when compared with other approaches and provides a high degree of nonlinear transformations to the input data, with fewer learning parameters (i.e. lower computational complexity) when compared to competitor algorithms

when used for PV prediction. Deep Learning networks are used in such fields as computer vision, speech recognition, natural language processing, audio recognition, machine translation, bioinformatics, with results compared or overcoming human abilities [6][8]. In Deep Learning architectures each subsequent layer of neurons (computational units of the network) learns to transform its input data into a more abstract and composite information representation. The term "Deep" refers to the number of subsequent layers in the network architecture. Ubiik uses its internally developed Deep learning network topology to achieve the higher accuracy prediction rates as shown in the following subsections.

1. Convolutional Neural Networks (CNN)

The main algorithmic unit in the Ubiik AI system (related to the satellite image processing) is a feature extraction unit, which provides some intelligent information, obtained from a patch of a single satellite image as an input to narrow down the area over which the actual solar power output has to be predicted. This problem can be reduced to a known task of image classification. The state-of-the-art approaches in modern machine learning for this kind of problem almost exclusively utilizes convolutional neural networks (CNN) [5][6][7][1]. Ubiik utilizes a similar approach as highlighted in Figure 9 below.

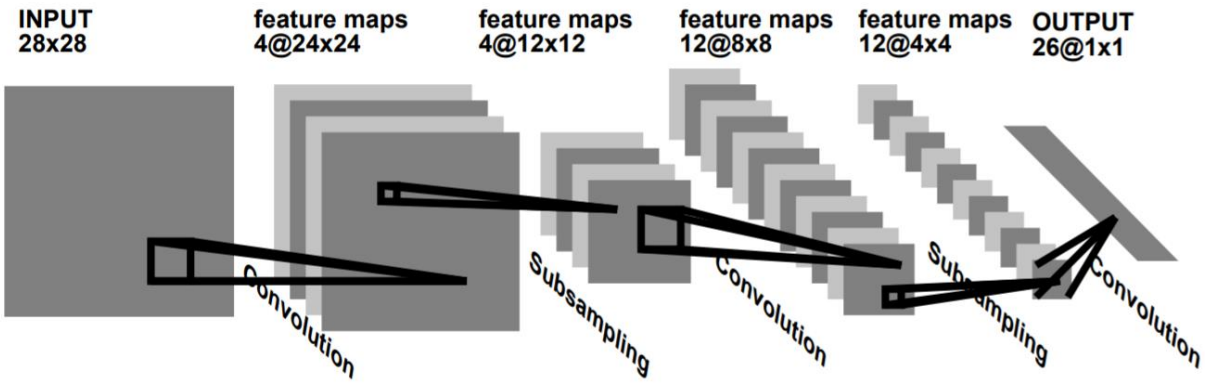


Figure 9: Ubiik's approach to performing additional image processing via convolutional neural networks to assist in evaluating weather conditions at a specific plant location [7].

A CNN is a subclass of Deep Learning algorithms. The main distinctive features of CNN architectures is the introduction of input distortion invariance thus allowing some degree of variance to the input image at different zoom levels, camera angles, lighting, spatial resolution, object position, etc. The base concept of CNN architecture implies using kernels of different sizes (receptive field), which slide over the input image and respond to particular visual features.

The kernels then take responsibility of these features during training. CNN networks usually contain a large amount of subsequent convolutional layers. These additional layers involve repeated convolution in such a way that spatial information becomes less and less relevant and is replaced by some abstract semantic representation for final classification. CNNs tend to require fewer weights (training parameters) than traditional ANNs.

The convolution blocks in the Ubiik CNN contain several hidden layers (convolutional layers, pooling layers, fully connected layers, and normalization layers). The blocks are used in the prediction algorithm as a single computational unit. Blocks have very limited reception fields of only a few pixels but with Ubiik's approach to multiple stacking, the reception field becomes wider and wider as convolutional layers increase.

Altogether, the properties of CNN architecture make it well suited for localizing the solar PV plant of interest and utilizing highly targeted tracking of weather data to improve prediction accuracy. At the same time the architecture reduces computational resources demands so that real-time prediction can be done for very high-resolution generation capacity predictions.

2. Ensemble Learning

Artificial Neural Networks with a sufficiently large number of parameters are prone to overfitting during training and can 'memorize' some features in training data to optimize learning metrics, such as maximizing accuracy and/or minimizing a loss function. Dependence on random initial parameters introduces random error in results. The performance of those models become difficult to evaluate since they now encode random factors that are difficult to control.

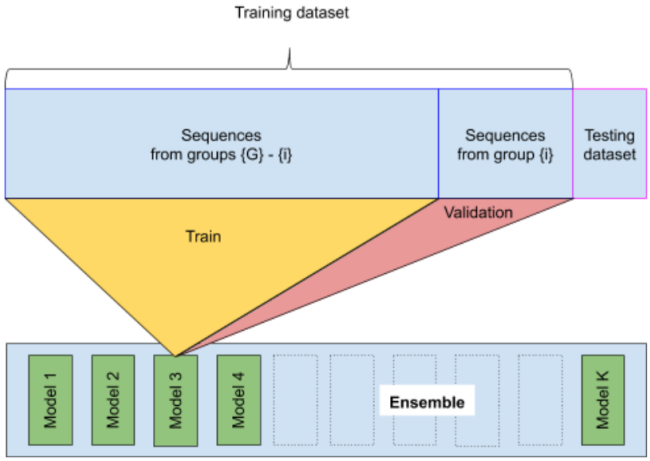


Figure 10: Ubiik's approach to data spliting for ensemble learning.

Even if initial weights are fixed, the variance of a model's performance after training is still quite significant and thus makes interpreting results with high confidence a challenge. Ubiik's approach of using ensemble learning aims to solve this issue by making full utilization of training data. When using K-fold cross-validation and ensemble learning, it is possible to do this by splitting into non-overlapping training and validation sets for a different model in an ensemble.

Sequences in training datasets are sampled for non-overlapping K groups in set {G}. For each model in an ensemble different groups are selected as a validation subset {i}, whereas the rest of the sequence groups are used as the training subset {G} - {i}. This is shown in Figure 10.

3. Active Learning

Ubiik's PV generation prediction system is a dynamically evolving adaptive AI solution, which constantly improves performance by utilizing new data obtained during system operation. The deep learning models' performance is highly limited by the training data – in real world situations this data is a highly limited resource. New data, collected during the active operation phase, is reused for further training and the ANN weights are readjusting accordingly. This allows deployment of the model at any stage of initial data accumulation while keeping its performance constantly improving over time. Active learning techniques are shown to be effective in [9]. Ubiik AI model uses two types of active learning: (a) Global active learning; (b) Momentary PV performance evaluation adaptive mechanism.

- **Global active learning:**

  The global active learning algorithm incorporates newly obtained data (satellite imagery, climate forecast data, AMI data) into existing training datasets. The increasing number of training data improve prediction accuracy and robustness. This mechanism requires a periodical execution of the training process on a monthly basis, comparing the achieved results with the existing model and updating if performance improvement is achieved.

- **Momentary PV Performance evaluation mechanism:**

  The momentary PV performance evaluation adaptive mechanism monitors the performance of every PV panel and identifies trends and the changes in it. This allows the automatic detection of any abnormalities such as panel degradation or malfunction. This mechanism also enables estimation of a real PV panel capacity, which can be different from a claimed capacity. Using actual PV panel capacity, estimated from actual data, in the training process increases the overall performance of the AI model since inaccurate PV panel capacity in the training dataset is an example of data mis-labelling and hence leads to decreased data quality and decreased accuracy of prediction.

5. Flexible Sub-model Selection

Ubiik uses a modular AI system, where each of its components is applied to the most optimal scope of input and output space. In our prediction system, the error rate of different components is non-uniform. Predictions, based on satellite imagery, are of higher precision for a short-term range (from 10 minutes to roughly 3 hours ahead) in comparison to the predictions based on the climate data. Climate data-based predictions, on the other hand, provide better results in the scope of mid-range forecasting (from 3 hours to 72 hours ahead). In practice, the boundary between the two models is smooth and we use a flexible model selection to make the transition seamless and avoid the introduction of artifacts caused by a disagreement between models. This approach increases robustness and decreases overall error in the transitional range of predictions. Figure 11 illustrates non-uniform precision decay for different models and the proportion of each model output's contribution to the combined output.

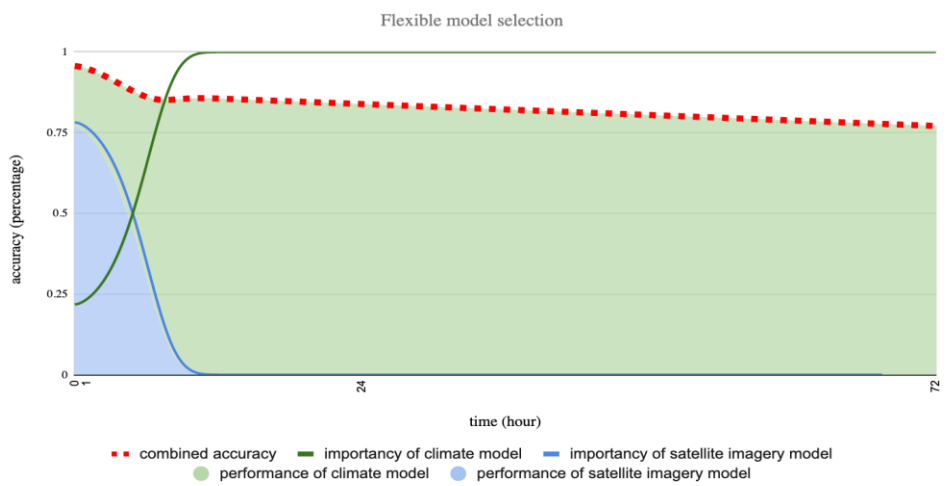


Figure 11: Flexible model selection approach to ensure seamless transition between different time-scale prediction of PV output. .

D. Post Processing System: Performance & Accuracy

Ubiik's post processing system aims to convert model outputs and aggregate them to produce a forecast of the solar PV generation power output at a given time and at a given GPS coordinate. In addition, the post-processor provides a measure of the accuracy of the prediction produced. Accuracy metrics may be used for the purposes of testing and verifying that the accuracy falls within the range required by customers so that real actionable decisions can be made. The next sub-sections demonstrate these metrics and Ubiik's performance as well.

1. Accuracy

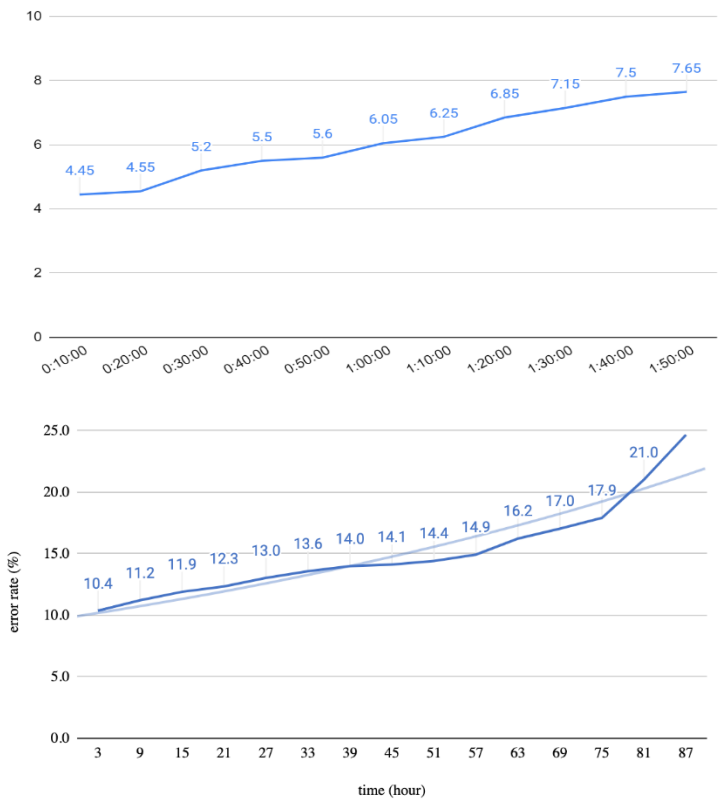
A list of the most common metrics for accuracy prediction is shown in the table below [2][3]. These metrics are built based on the actual output power (p_i) and an existing time series of data points of some size (i=1..N). From this data an estimate of the output power ($\hat{p}_i$) is produced. Algorithms largely aim to minimize the error between prediction ($\hat{p}_i$) and actual output power (p_i) or increase correlation. To evaluate the end accuracy of the prediction over some timescale thus requires an aggregation of the individual predictions. The error between predication and actual power acts as proxy for the accuracy upon occasion. When comparing various algorithms and approaches for power output prediction accuracy (and error) must then be evaluated using multiple metrics each with their own particular use cases and relevance. Table 2 provides a summary of the different approaches to prediction error calculation. Ubiik will make use primarily of the RMSE and NRMSE metrics as required by some utility companies for its measurement of accuracy.

Accuracy Metric	Description	Accuracy Formulation	Reference Values for Existing Model Accuracy 100 MW PV Plant in %	
			Intra-hour	Intra-day
Pearson's correlation coefficient	Measures the similarity between the overall trend of the forecasts and actual values, though it does not account for relative magnitudes (i.e., biases).	$\rho = \frac{\mathrm{cov}(p, \hat{p})}{\sigma_p \sigma_{\hat{p}}}$	76%	65%
Root Mean Squared Error or Normalized Root Mean Squared Error	Measures overall accuracy of the forecasting models. However, square order increases the prediction error rate.	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p_i)^2}$ $NRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{p}_i - p_i}{p_r}\right)^2}$	17%	22%
Maximum absolute error	Useful to evaluate forecasts of short-term extreme events in the power system. May give too much weight to extreme events	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert\hat{p}_i - p_i\rvert$	74%	84%
Mean absolute error or Mean absolute percentage error	Global error metric which measures the overall accuracy of the forecast, but it does not punish larger forecast errors compared to the RMSE metric. Large number of small errors can overwhelm small number of large errors	$MaxAE = \max_{i=1,2,\cdots,N}\lvert\hat{p}_i - p_i\rvert$ $MAPE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{\hat{p}_i - p_i}{capacity}\right\rvert$	11%	15%

Table 2: Various statistical error metrics to evaluate accuracy. These metrics directly rely on the actual power output (p_i) and the estimator for power output ($\hat{p}_i$) over some time-series to the present moment (i). Correlation is computed with support of the standard deviations of the estimator ($\sigma_{\hat{p}}$) and actual power (σ_p) dataset [1] [2][3]. Highlighted technique will be the approach by which models will be evaluated in the project proposed by Ubiik.

2. Performance of the Ubiik System

Although the Ubiik model is designed to output forecasts in the range of 0 to 72 hours with 15 minutes resolution, the focus is on two major reference points: (a) RMSE Error rates at 1 hour and (b) at 24 hours. The 1-hour forecasts mostly rely on the output of the AI model, which is based on satellite imagery data and numerical clear sky models. This performance is evaluated at the locations indicated in Figure 12 (right) and over samples taken from a randomized selection of dates. The input data for the locations was obtained from Taiwanese government public data repositories located at [10].



Location
七股
中科ES
北儲
台中生水池
台中龍井
大潭
尖山
彰林ES
核三生水池
竹工ES
興達生水池
路北
金門金沙
龍潭ES

Randomly Selected Dates	
2019-04-24	2019-11-06
2019-04-30	2019-11-07
2019-05-10	2019-12-08
2019-05-25	2019-12-10
2019-06-15	2020-01-18
2019-06-20	2020-01-23
2019-07-13	2020-02-08
2019-07-15	2020-02-15
2019-08-05	2020-03-04
2019-08-06	2020-03-09
2019-09-02	
2019-09-07	
2019-10-04	
2019-10-24	

Figure 12: On the right, locations and randomized dates over which data was sampled and aggregated for estimating accuracy. On the left, RMSE error over various forecasting scales (in % of PV panel capacity). The top image shows intra-hour time scales and the bottom image shows intra-day and day-ahead predictions. Note that the x-axis is time and the y-axis is the Normalized RMSE error.

Predictions based on satellite imagery data have an error rate of 6.05% for 1-hour forecasts, however, in a real-life situation, satellite images become available not immediately at the reference time, but rather with some delay, which is introduced by the satellite image provider. In cases where publicly available open-source data is used, this delay is guaranteed to be within 40 minutes (usually around 30 minutes). In the case where this data is obtained from a commercial data provider, this delay can be reduced to 20 minutes. Therefore, the actual error rate for 1-hour forecasts will be 7.50% and 6.85% respectively and significantly less for the shorter 15min period.

E. Visualization System & Prediction System

The GUI in Ubiik's prediction system will be divided into a test station and formal station, which will be set up on different virtual machines respectively, and a continuous payment mode will be used for module modification and layout. The front-end and backend is highly customizable to needs of specific customers.

The prediction system can be browsed by a web browser (at minimum Google Chrome) and utilize HTTPS secure connections for users to operate – Ubiik provides support for the latest versions of browsers normally under use by customers).

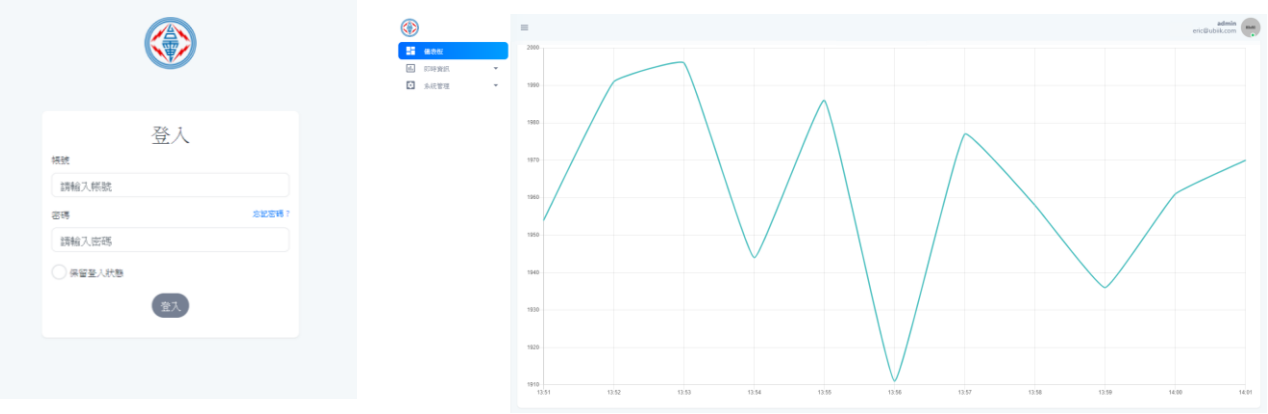


Figure 13: The visual interface for the service portal. On the left, the login screen for the remote user; On the right, the forecasting screen for visualization and prediction analytics.

1. Login System

The system and database control mechanism can be achieved by setting user permissions (for example, users need to log in to obtain the data generated by the system, and system managers have the right to upload data), so as to protect the system from accidental or malicious access, use, modification or damage. A sample login screen is shown in Figure 13 (left).

The login system uses a password longer than 8-digit composed of English alphanumeric characters as well as special characters and has the following functions

1. The system database does not store the user's original password in plain text format. A hash function is used for secure password storage.
2. If the password is entered incorrectly for 3 consecutive attempts, the system restricts the account login for a period of 15 minutes.
3. The system has a password reset mechanism and the account user can change the password.
4. The system administrator can observe all users and their usage of the web interface.

2. Forecast Data Display

Forecast data will be screened and displayed according to region and time interval and will mainly be displayed by a line chart. This however, is highly customizable to the needs of the end user. Figure 13 (right) demonstrates a sample of the forecasting screen.

3. API for Integrating into Other Solutions

The Ubiik prediction system provides a data format based on JSON. Therefore, all APIs provide encrypted transmission via HTTPS. Users are verified using JSON web tokens (JWT). API resource access URLs are presented in restful style with the API providing data download functionality. This download functionality includes items such as forecast value, actual value, forecast accuracy evaluation and more.

References

[1] R. Ahmed, V. Sreeram, Y. Mishra, M.D. Arif. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. Renewable and Sustainable Energy Reviews. Volume 124, 2020.

[2] Jensen, & Fowler, T. & Brown, Barbara & Lazo, Jeffrey & Haupt, Sue. (2016). Metrics for Evaluation of Solar Energy Forecasts.

[3] J. Zhang, B.-M. Hodge, A. Florita, S. Lu, H. F. Hamann, V. Banunarayanan. Metrics for Evaluating the Accuracy of Solar Power Forecasting. 3rd International Workshop on Integration of Solar Power into Power Systems. 2013.

[4] Hadja Maïmouna Diagne, Mathieu David, Philippe Lauret, John Boland, Nicolas Schmutz. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews, Elsevier, 2013, 27, pp.65 - 76.

[5] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. 2014.

[6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 2012, pp. 1097–1105

[7] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. 1995.

[8] Dan Ciresan, U. Meier, J. Schmidhuber. Multi-column deep neural networks for image classification. 2012 IEEE Conference on Computer Vision and Pattern Recognition. 2012.

[9] Ting Yang, Liyuan Zhao, Wei Li, Albert Y. Zomaya. Reinforcement learning in sustainable energy and electric systems: a survey. Annual Reviews in Control, 2020.

[10] 台灣電力公司. 台電系統各機組發電量（單位 MW）. https://www.taipower.com.tw/d006/loadGraph/loadGraph/genshx_.html. [Accessed: April 2020]