Professional Certificate in AI for Renewable Energy Forecasting (Thailand)

Fundamentals of Renewable Energy Forecasting

Renewable Energy Forecasting is the discipline that uses statistical, physical, and data-driven methods to predict the future output of energy sources such as solar, wind, hydro, and biomass. Accurate forecasts enable grid operators, market participants, and policy makers to balance supply and demand, plan maintenance, and price electricity appropriately. The following key terms and vocabulary form the foundation for any professional working in this field. Each entry includes a concise definition, illustrative example, practical application, and common challenges encountered in real-world settings.

Solar Irradiance refers to the power per unit area received from the Sun's rays at the Earth's surface, typically expressed in watts per square meter (W/m²). It varies with time of day, atmospheric conditions, and geographic location. For instance, a clear sky at noon in Bangkok may deliver 1,000 W/m², while a partly cloudy afternoon might drop to 400 W/m². In forecasting, solar irradiance is the primary driver of photovoltaic (PV) generation models. Accurate measurement or estimation of irradiance is essential for short-term (minutes to hours) PV output prediction. Challenges include cloud dynamics, aerosol scattering, and sensor calibration errors.

Capacity Factor is the ratio of actual energy produced by a plant over a period to the maximum possible energy it could have produced if it operated at full name-plate capacity continuously. Mathematically, Capacity Factor = (Actual Energy Output) / (Name-Plate Capacity × Time). A solar farm in Thailand might have a capacity factor of 18 % because it only generates electricity during daylight and is affected by seasonal cloud cover. Capacity factor is useful for comparing the efficiency of different technologies and for financial modeling. The main difficulty lies in accounting for variability in weather and equipment downtime.

Levelized Cost of Energy (LCOE) represents the average cost per megawatt-hour (MWh) of electricity generated over the lifetime of an asset, incorporating capital expenditures, operation and maintenance, fuel (if any), and discount rates. For a wind turbine, the LCOE may be around 0.07 USD/kWh, while a coal plant may be higher due to fuel costs and emissions penalties. LCOE is a benchmark for investment decisions and policy incentives. However, it does not directly capture the value of flexibility or the cost of forecast errors, which can be significant for intermittent renewables.

Forecast Horizon defines the temporal length over which a prediction is made. Common horizons include very-short term (seconds to minutes), short-term (minutes to hours), intra-day (hours to a day), day-ahead, week-ahead, and seasonal forecasts. A day-ahead solar PV forecast helps market operators schedule generation, while a week-ahead wind forecast assists in maintenance planning. Selecting an appropriate horizon is crucial because model accuracy typically declines as the horizon lengthens, and different horizons may require distinct modeling techniques.

Persistence Model is the simplest forecasting approach, assuming that future conditions will be identical to the most recent observation. For solar PV, a persistence forecast would predict that the irradiance for the next hour will equal the measured irradiance at the current hour. Although naïve, persistence often serves as a baseline to evaluate more sophisticated models. Its limitations become evident during rapid weather changes, where the model fails to capture the transition from sunny to cloudy conditions.

Physical Model utilizes the underlying physics of atmospheric processes to estimate renewable generation. In solar forecasting, a physical model may incorporate the Clear-Sky Index, atmospheric transmittance, and sun position algorithms. For wind, the model might use the power curve of the turbine combined with wind speed and air density calculations. Physical models are valuable because they can extrapolate beyond observed data, but they require extensive meteorological inputs and may be computationally intensive.

Statistical Model relies on historical data to infer relationships between input variables (e.g., weather observations) and output (e.g., power generation). Techniques include linear regression, autoregressive integrated moving average (ARIMA), and exponential smoothing. For example, a linear regression might predict PV output based on measured global horizontal irradiance and ambient temperature. Statistical models are relatively easy to implement, yet they can struggle with non-linear patterns and may not adapt quickly to changing climate trends.

Machine Learning encompasses a broad set of algorithms that learn complex, often non-linear relationships from data. In renewable forecasting, common methods include decision trees, random forests, gradient boosting machines, support vector machines, and deep neural networks. A deep learning model, such as a convolutional neural network (CNN), can ingest satellite imagery to predict cloud movement and thus solar output several hours ahead. Machine learning models often outperform traditional statistical methods when sufficient high-quality data are available, but they can be opaque (“black-box”) and prone to over-fitting if not properly regularized.

Ensemble Forecasting combines multiple individual forecasts to produce a single, often more accurate prediction. Ensembles may be formed by averaging, weighted averaging, or more sophisticated techniques like Bayesian model averaging. For instance, an ensemble might merge a physical model, a statistical model, and a neural network output to generate a day-ahead wind forecast. The diversity of methods helps mitigate the weaknesses of any single model, though ensemble construction introduces complexity in weighting and error propagation.

Numerical Weather Prediction (NWP) is the use of computer simulations of the atmosphere based on fluid dynamics equations to forecast weather variables such as temperature, wind speed, and cloud cover. NWP outputs are essential inputs for renewable forecasting because they provide spatially resolved forecasts of the meteorological drivers. Global models (e.g., ECMWF, GFS) provide forecasts up to 10 days, while regional models (e.g., WRF) can deliver higher resolution for shorter horizons. Limitations include model bias, coarse spatial resolution, and the need for post-processing to align with local conditions.

Post-Processing refers to the adjustment of raw NWP outputs to improve their relevance for renewable generation forecasts. Techniques include bias correction, statistical downscaling, and machine-learning-based calibration. For example, a bias-corrected wind forecast may subtract a systematic under-prediction observed over the past year. Post-processing can significantly reduce forecast error, but it requires historical verification data and careful handling of non-stationary biases.

Clear-Sky Model predicts the solar irradiance that would be received under cloud-free conditions. It serves as a reference to quantify the effect of clouds on actual irradiance. Common clear-sky models include the Ineichen-Perez and the REST2 models. By comparing measured irradiance to the clear-sky estimate, forecasters can compute the cloud attenuation factor, which is then used in forecasting algorithms. The accuracy of clear-sky models depends on atmospheric parameters such as aerosol optical depth and water vapor content.

Cloud Motion Vector (CMV) describes the speed and direction of cloud movement across the sky, typically derived from successive satellite images. CMVs are crucial for short-term solar forecasting because they allow the extrapolation of cloud shadows onto PV arrays. A typical CMV might indicate clouds moving westward at 30 km/h. Extracting reliable CMVs requires image processing techniques and may be hindered by low-resolution satellite data or rapid cloud evolution.

Power Curve is a manufacturer-provided relationship that maps wind speed to expected electrical power output for a specific turbine model. The curve is usually defined for a range of wind speeds from cut-in to cut-out. For a 2 MW turbine, the power curve might show 1 MW output at 8 m/s wind speed. In forecasting, the power curve translates wind speed forecasts into power forecasts. However, actual turbine performance can deviate due to turbulence, air density variations, and wear, necessitating curve adjustments.

Air Density influences the kinetic energy of wind and therefore the power that can be extracted. It is a function of temperature, pressure, and humidity. In hot, humid conditions typical of tropical Thailand, air density may be lower than in cooler climates, reducing wind turbine output for a given wind speed. Forecasting models that incorporate real-time air density calculations can improve accuracy, especially for high-altitude sites.

Ground-Based Measurement includes instruments such as pyranometers, anemometers, and sky cameras that collect local meteorological data. These observations are vital for calibrating and validating forecasts. A pyranometer might record a solar irradiance of 850 W/m^2 at 10 am, while an anemometer could measure a wind speed of 5 m/s at hub height. Ground-based data can also be used for real-time model correction (nowcasting). The main challenges involve sensor maintenance, data gaps, and ensuring representativeness of the measurement location.

Nowcasting is the process of generating ultra-short-term forecasts (seconds to a few hours) using real-time observations, often for immediate operational decisions. For solar PV, nowcasting may predict the next 15-minute output to manage inverter curtailment or battery dispatch. Techniques frequently involve cloud detection from sky cameras, rapid NWP updates, and machine-learning models trained on high-frequency data. The principal difficulty is handling the rapid evolution of weather phenomena and the limited lead time for corrective actions.

Deterministic Forecast provides a single predicted value for each variable at each time step, as opposed to a range or probability distribution. An example is a day-ahead wind speed forecast of 7 m/s at 12 UTC. Deterministic forecasts are straightforward to interpret and are often used in market settlements. However, they do not convey uncertainty, which can be critical for risk-averse stakeholders.

Probabilistic Forecast expresses the likelihood of different outcomes, often through quantiles, prediction intervals, or full probability density functions. A 90 % prediction interval for solar PV output might be 300 kW to 500 kW for a given hour. Probabilistic forecasts enable operators to assess risk, perform stochastic unit commitment, and optimize reserves. Generating reliable probabilistic forecasts requires ensemble methods, Bayesian techniques, or quantile regression, and they may be more difficult for end-users to interpret.

Quantile Regression is a statistical technique that predicts specific quantiles (e.g., 10th, 50th, 90th percentiles) of the target variable distribution, rather than the mean. In renewable forecasting, quantile regression can directly produce prediction intervals for PV output. For example, the 5th-percentile forecast may be used to set a conservative generation estimate. This method avoids assumptions of normality but can be computationally demanding for large datasets.

Bias denotes the systematic error where forecasts consistently over- or under-predict the observed values. Bias can be expressed as the mean error over a period. A wind forecast that is on average 0.5 m/s too low exhibits a negative bias. Identifying and correcting bias is essential for improving forecast reliability. Bias may arise from model mis-specification, sensor drift, or climate trends.

Root Mean Square Error (RMSE) quantifies the average magnitude of forecast errors, giving higher weight to larger errors. It is calculated as the square root of the average of squared differences between forecast and observation. An RMSE of 50 kW for a 500 kW PV system indicates moderate accuracy, whereas a lower RMSE would be desirable for high-value markets. RMSE is widely used for model evaluation but can be sensitive to outliers.

Mean Absolute Error (MAE) measures the average absolute difference between forecasts and observations, providing a more intuitive error metric than RMSE. An MAE of 30 kW may be easier to interpret for operators than an RMSE of 45 kW. MAE is less affected by extreme errors and is useful for comparing models across different scales.

Mean Absolute Percentage Error (MAPE) expresses forecast error as a percentage of the observed value, facilitating comparison across sites with different capacities. For a wind farm, a MAPE of 5 % indicates that forecasts are, on average, within 5 % of the actual generation. However, MAPE can be inflated when observed values are near zero, making it less reliable for low-output periods.

Skill Score measures the relative performance of a forecast against a reference model, such as persistence or climatology. It is often expressed as a percentage improvement; a skill score of 20 % means the forecast reduces error by 20 % compared to the reference. Skill scores are valuable for demonstrating the added value of advanced methods. The challenge lies in selecting an appropriate benchmark and ensuring statistical significance.

Climatology refers to the long-term average of a meteorological variable for a specific location and time of year. In forecasting, climatology can serve as a naïve baseline: for solar PV, the climatological irradiance for a given hour in March might be 800 W/m^2. While simplistic, climatology provides a useful reference for assessing model skill, especially during periods with limited data.

Temporal Resolution describes the frequency at which data points are recorded or forecasts are issued, such as 5 minutes, 15 minutes, hourly, or daily. Higher temporal resolution captures rapid fluctuations but requires more storage and computational resources. For battery management, 5-minute forecasts may be necessary, whereas long-term planning may rely on hourly or daily data.

Spatial Resolution indicates the size of the grid cell or area represented by a dataset, commonly expressed in kilometers. A 3 km NWP grid provides finer detail than a 25 km grid, allowing better representation of local terrain effects on wind. However, higher spatial resolution increases computational cost and may introduce noise if the underlying observations are sparse.

Data Assimilation is the process of integrating observations into a numerical model to improve its state estimate. In weather forecasting, data assimilation incorporates satellite radiances, surface stations, and radar measurements into NWP models. Effective assimilation can reduce forecast error, especially for short-term horizons critical to renewable forecasting. The complexity of assimilation algorithms and the need for real-time data pipelines pose operational challenges.

Training Set is the portion of historical data used to fit a forecasting model. For machine learning, the training set may consist of several years of solar irradiance, temperature, and power output records. Careful selection of the training period ensures that the model captures seasonal patterns while avoiding over-fitting to anomalous events.

Validation Set provides an independent dataset to tune model hyper-parameters and prevent over-fitting. For example, a wind forecast model might be trained on 2018-2020 data and validated on 2021 data. The validation set helps assess how well the model generalizes to unseen conditions. Improper validation can lead to overly optimistic performance estimates.

Test Set is a final dataset reserved for evaluating the model's performance after training and validation are complete. It should reflect the operational environment where the model will be deployed. Reporting test-set metrics such as RMSE or skill scores gives stakeholders confidence in the model's reliability. Leakage of test data into training compromises the credibility of the evaluation.

Cross-Validation involves partitioning the data into multiple folds and rotating training and validation sets to obtain a more robust estimate of model performance. K-fold cross-validation, where the data is split into K equal parts, is common. In renewable forecasting, cross-validation helps assess sensitivity to different seasonal patterns. The main drawback is increased computational load.

Hyper-Parameter denotes a configuration setting for a model that is not learned from the data during training, such as the depth of a decision tree, learning rate of a gradient boosting model, or number of hidden layers in a neural network. Hyper-parameters are typically tuned using grid search, random search, or Bayesian optimization. Selecting appropriate hyper-parameters is critical for balancing bias and variance.

Over-fitting occurs when a model learns noise and specific patterns in the training data that do not generalize to new data, resulting in poor performance on the test set. Indicators of over-fitting include a large gap between training error and validation error. Techniques such as regularization, early stopping, and dropout (for neural networks) mitigate over-fitting.

Under-fitting describes a model that is too simple to capture the underlying relationships in the data, leading to high error on both training and validation sets. Under-fitting may be addressed by increasing model complexity, adding relevant features, or reducing regularization strength.

Feature Engineering is the process of creating informative input variables (features) from raw data to improve model performance. In solar forecasting, features may include the sine and cosine of the solar zenith angle, cloud index, or lagged irradiance values. Effective feature engineering can significantly boost predictive accuracy, but it requires domain expertise and careful testing.

Lagged Variable is a feature that represents a past observation of a variable at a specified time offset. For wind speed, a 1-hour lagged variable would be the wind speed measured one hour earlier. Lagged variables capture temporal dependencies and are commonly used in time-series models such as ARIMA or recurrent neural networks.

Exogenous Variable (often abbreviated as exog) is an input variable that originates outside the target system but influences it. Weather variables (temperature, humidity, cloud cover) are exogenous to PV generation. Incorporating exogenous variables can enhance forecast skill, especially when the target variable is strongly driven by external factors.

Recurrent Neural Network (RNN) is a type of deep learning architecture designed to handle sequential data by maintaining a hidden state that captures information from previous time steps. Long Short-Term Memory (LSTM) networks are a widely used RNN variant for renewable forecasting because they can learn long-range dependencies in wind speed or solar irradiance time series. Training RNNs requires careful handling of vanishing gradients and may be computationally demanding.

Convolutional Neural Network (CNN) applies convolutional filters to extract spatial features from data such as satellite images or radar maps. In solar forecasting, a CNN can learn patterns of cloud movement across successive images to predict irradiance several hours ahead. CNNs excel at capturing local spatial correlations but may need to be combined with RNN layers for temporal modeling.

Hybrid Model integrates multiple modeling approaches, such as combining a physical NWP model with a machine-learning correction term. Hybrid models aim to leverage the strengths of each component: the physical realism of NWP and the data-driven adaptability of machine learning. Implementation can be complex, requiring synchronization of inputs and careful error analysis.

Transfer Learning involves reusing a model trained on one dataset or region for a different but related task, often with fine-tuning. For example, a wind forecasting neural network trained on European data could be adapted to Thai sites by retraining the final layers with local observations. Transfer learning reduces the need for large local datasets but may still suffer from domain mismatch.

Domain Adaptation is a broader concept that addresses differences between source and target data distributions. Techniques such as adversarial training can align feature representations from different regions, improving forecast portability. In renewable forecasting, domain adaptation helps deploy models across diverse climatic zones.

Data Quality encompasses the accuracy, completeness, consistency, and timeliness of the datasets used for forecasting. Issues such as missing values, sensor drift, time-stamp mismatches, and outliers can degrade model performance. Data cleaning procedures, including interpolation, outlier detection, and sensor validation, are essential steps before model development.

Missing Data Imputation fills gaps in time-series datasets using statistical or machine-learning methods. Simple approaches include linear interpolation or forward-fill; more advanced techniques employ k-nearest neighbors or Gaussian process regression. Accurate imputation preserves the integrity of training data, but inappropriate methods can introduce bias.

Outlier Detection identifies observations that deviate markedly from typical patterns, often caused by sensor malfunction or extreme weather events. Methods range from statistical thresholds (e.g., values beyond three standard deviations) to robust algorithms like isolation forests. Removing or correcting outliers prevents them from skewing model training.

Normalization scales features to a common range, such as 0-1 or –1 to 1, facilitating efficient training of machine-learning models. For wind speed, normalization may involve dividing by the maximum observed speed. Normalization reduces numerical instability and speeds up convergence, but the scaling parameters must be stored and applied consistently during inference.

Standardization transforms data to have zero mean and unit variance, which is particularly useful for algorithms sensitive to feature magnitude, such as support vector machines. Standardization requires computing the mean and standard deviation on the training set and applying those statistics to future data. Failure to standardize appropriately can degrade model performance.

Feature Selection reduces the dimensionality of the input space by retaining only the most informative variables. Techniques include correlation analysis, mutual information, recursive feature elimination, and regularization-based methods (e.g., Lasso). Feature selection improves model interpretability, reduces over-fitting risk, and lowers computational cost.

Dimensionality Reduction transforms high-dimensional data into a lower-dimensional representation while preserving essential structure. Principal Component Analysis (PCA) is a common linear method; autoencoders provide non-linear alternatives. In renewable forecasting, dimensionality reduction can compress large sets of meteorological variables into a few principal components for efficient modeling.

Time-Series Decomposition separates a series into trend, seasonal, and residual components. Decomposition helps isolate systematic patterns (e.g., daily solar cycle) from irregular fluctuations. Models can be built on the residual component, improving forecast accuracy. Seasonal decomposition of solar irradiance typically reveals a strong diurnal pattern.

Autoregressive Integrated Moving Average (ARIMA) is a classical time-series model that captures autocorrelation (AR), differencing to achieve stationarity (I), and moving-average components (MA). ARIMA models have been applied to wind speed forecasting, where they can model short-term persistence. However, ARIMA assumes linear relationships and may not handle non-stationary weather regimes well.

Seasonal ARIMA (SARIMA) extends ARIMA by incorporating seasonal terms to model periodic patterns. For solar PV, a SARIMA model might include a 24-hour seasonal lag to capture the daily irradiance cycle. SARIMA can improve accuracy over plain ARIMA when strong seasonality is present, but model selection becomes more intricate.

Exponential Smoothing methods, such as Holt-Winters, assign exponentially decreasing weights to older observations, allowing rapid adaptation to recent changes. These methods are computationally light and suitable for real-time forecasting. Their simplicity, however, limits their ability to capture complex non-linear dynamics.

Kalman Filter provides recursive estimation of a system’s state by combining a predictive model with noisy observations. In wind forecasting, a Kalman filter can merge NWP wind speed predictions with local anemometer readings to produce a refined estimate. Kalman filtering is optimal for linear Gaussian systems but extensions (e.g., extended Kalman filter) are needed for non-linear cases.

Particle Filter is a Monte-Carlo based technique for estimating the posterior distribution of a system’s state in non-linear, non-Gaussian contexts. Particle filters can be applied to solar forecasting where cloud dynamics introduce non-linearities. They are computationally intensive, requiring many particles to approximate the distribution accurately.

Digital Twin is a virtual replica of a physical asset that simulates its behavior in real time using sensor data and predictive models. For a wind farm, a digital twin may integrate turbine performance models, environmental inputs, and maintenance schedules to forecast output and detect anomalies. Digital twins enable scenario analysis but demand high-fidelity models and continuous data streams.

Smart Grid denotes an electricity network that uses information and communication technologies to monitor, control, and optimize the flow of electricity. Renewable forecasting is a core component of smart grids, providing the necessary data for demand-response programs, distributed energy resource (DER) integration, and automated dispatch. Implementing forecasting within a smart grid requires interoperability standards and low-latency data pipelines.

Distributed Energy Resource (DER) includes small-scale generation units such as rooftop PV, micro-wind turbines, and battery storage. Accurate forecasts for DERs support local grid stability, peak shaving, and participation in ancillary services markets. DER forecasting faces challenges due to the heterogeneity of installations, limited metering, and rapid changes in ownership.

Ancillary Services are support functions that maintain grid reliability, including frequency regulation, voltage control, and spinning reserve. Renewable generators can provide ancillary services when their output is predictable or when coupled with storage. Forecast uncertainty directly impacts the ability to commit renewables to these services, making probabilistic forecasting essential.

Reserve Margin is the excess generation capacity maintained to ensure reliability under unexpected demand spikes or generation outages. In systems with high renewable penetration, the reserve margin must account for forecast error. Operators may increase reserves during periods of high forecast uncertainty, affecting market prices and system costs.

Capacity Trading involves the buying and selling of generation capacity rights, often on day-ahead or month-ahead markets. Accurate forecasts enable generators to secure capacity contracts and reduce the risk of penalties for under-delivery. Forecast errors can lead to imbalance charges, emphasizing the economic importance of precise predictions.

Imbalance Settlement is the process by which deviations between scheduled and actual generation are financially reconciled. In many markets, participants with surplus generation receive payments, while those with deficits incur charges. Renewable forecasters aim to minimize imbalance by providing reliable forecasts and by incorporating uncertainty into dispatch decisions.

Curtailment occurs when renewable generators are instructed to reduce output due to grid constraints, oversupply, or market conditions. Forecasting helps anticipate curtailment events, allowing operators to schedule storage or demand-response actions. However, curtailment reduces revenue for renewable owners and can distort market signals for further investment.

Load Forecasting predicts electricity demand over various horizons and is closely linked to renewable generation forecasting. Accurate load forecasts enable better matching of supply and demand, especially when high-penetration renewables introduce variability. Techniques overlap with renewable forecasting, including statistical models, machine learning, and ensemble methods.

Net Load is the residual demand after subtracting renewable generation from the total load. Net load forecasting is critical for dispatching conventional generators and for planning reserve requirements. For example, a sunny day may result in a low net load, reducing the need for thermal generation. Net load uncertainty is driven by both load and renewable forecast errors.

Hybrid Energy System combines multiple generation technologies, such as solar PV, wind turbines, and diesel generators, often with storage. Forecasting in hybrid systems must consider the interaction among components, such as when solar output drops and a diesel generator ramps up. Optimization algorithms use forecasts to schedule the most cost-effective mix of resources.

Battery Energy Storage System (BESS) stores electrical energy for later use, providing flexibility to smooth renewable variability. Forecasting informs BESS operation strategies like charging during expected excess generation and discharging during anticipated deficits. Accurate forecasts reduce the need for conservative reserve sizing and improve economic returns.

Levelized Avoided Cost of Energy (LACE) measures the cost avoided by generating electricity from a particular renewable source instead of from conventional generation. LACE calculations incorporate forecast accuracy, as higher forecast errors increase the cost of balancing services. LACE helps policymakers evaluate the competitiveness of renewables relative to other options.

Grid-Forming Inverter is a power electronic device capable of establishing voltage and frequency reference in a weak or islanded grid, enabling renewable sources to support grid stability. Forecasting assists grid-forming inverters by predicting periods when they must operate in “grid-forming” mode versus “grid-following” mode. The challenge lies in integrating forecasts with inverter control algorithms in real time.

Grid-Following Inverter synchronizes with an existing grid voltage and frequency, injecting power when available. Accurate forecasts enable grid-following inverters to anticipate output and avoid over-generation that could cause voltage rise. In high-penetration scenarios, coordination between grid-forming and grid-following inverters becomes essential.

Power Purchase Agreement (PPA) is a contract between a renewable generator and a buyer, specifying the price and quantity of electricity to be delivered over a period. Forecast accuracy influences PPA negotiations, as parties may include clauses for performance guarantees or penalties based on forecast error. Robust forecasting reduces the risk of contractual disputes.

Renewable Portfolio Standard (RPS) is a policy that mandates a minimum share of electricity to come from renewable sources. Forecasting supports compliance with RPS targets by providing utilities with reliable generation estimates. Uncertainty in forecasts can affect the ability to meet mandated percentages, prompting the need for contingency planning.

Capacity Credit quantifies the contribution of a renewable resource to system reliability, expressed as a percentage of its name-plate capacity. Capacity credit is derived from statistical analysis of historical generation and demand data, often using loss-of-load probability (LOLP) calculations. Accurate forecasts improve the estimation of capacity credit, influencing planning decisions.

Loss-of-Load Probability (LOLP) measures the likelihood that the system will be unable to meet demand at a given time. Renewable forecasts feed into LOLP calculations by providing probabilistic generation estimates. Higher forecast uncertainty increases LOLP, potentially leading to higher reserve requirements.

Monte Carlo Simulation employs random sampling to assess the impact of uncertainties on system performance. In renewable forecasting, Monte Carlo techniques generate many possible generation scenarios based on probabilistic forecasts, allowing risk analysis for investment and operation. The method is computationally intensive and requires well-characterized probability distributions.

Scenario Analysis evaluates the performance of a system under distinct, predefined conditions such as “high wind,” “low solar,” or “extreme temperature.” Scenario analysis helps planners understand the robustness of grid operations and investment strategies. Forecast models can be used to generate realistic scenarios, but the selection of representative scenarios is critical.

Sensitivity Analysis examines how changes in input variables affect output forecasts. For example, varying cloud cover input by ±10 % may reveal the sensitivity of PV output to cloud uncertainties. Sensitivity analysis guides data collection priorities and model refinement by highlighting the most influential factors.

Data Latency describes the delay between the occurrence of an event (e.g., a wind measurement) and the availability of that data for processing. Low latency is crucial for nowcasting and real-time dispatch. High latency can render forecasts obsolete, especially for fast-changing conditions. Reducing latency often requires investment in communication infrastructure and edge computing.

Telemetry refers to the automated transmission of measurement data from remote sensors to a central system. In wind farms, telemetry may include turbine status, wind speed, and power output. Reliable telemetry is essential for real-time model updates and for detecting anomalies. Telemetry failures can lead to gaps in the data record, degrading forecast quality.

Data Integration involves combining heterogeneous data sources such as satellite images, ground stations, and SCADA (Supervisory Control and Data Acquisition) systems into a unified dataset for modeling. Effective integration requires handling differing formats, time stamps, and spatial resolutions. Incomplete or mismatched integration can introduce errors that propagate through the forecasting pipeline.

SCADA System collects high-frequency operational data from power plants, including turbine rotational speed, blade pitch, and generated power. SCADA data provide a rich source of real-time information for model calibration and validation. Access to SCADA data may be restricted due to proprietary concerns, posing a barrier to model development.

Big Data describes extremely large and complex datasets that exceed traditional processing capabilities. Renewable forecasting increasingly relies on big data from high-resolution satellite imagery, extensive sensor networks, and historical market data. Big data analytics require scalable storage, parallel processing, and advanced algorithms to extract meaningful patterns.

Cloud Computing offers on-demand computational resources that can scale with the workload of forecasting models. Cloud platforms enable rapid deployment of machine-learning pipelines, large-scale NWP processing, and storage of massive datasets. Cost management and data security are key considerations when leveraging cloud services.

Edge Computing processes data close to the source, reducing latency and bandwidth usage. In renewable forecasting, edge devices may perform preliminary data cleaning or generate local forecasts that are later refined centrally. Edge computing is valuable for remote wind sites with limited connectivity, but hardware constraints may limit model complexity.

Model Drift occurs when a forecasting model's performance degrades over time due to changes in underlying data distributions, such as climate shifts or equipment upgrades. Detecting model drift involves monitoring error metrics and retraining the model when performance declines. Continuous model monitoring is essential for maintaining forecast reliability.

Model Retraining updates a forecasting model using recent data to adapt to new patterns. Retraining frequency depends on the rate of change in the environment; for solar PV, weekly or monthly retraining may suffice, while wind farms exposed to seasonal shifts may require quarterly updates. Automated retraining pipelines reduce manual effort but must be carefully validated to avoid regressions.

Explainable AI (XAI) seeks to make machine-learning models transparent and interpretable. In renewable forecasting, XAI techniques such as SHAP (SHapley Additive exPlanations) can reveal which weather variables most influence a prediction, aiding stakeholder trust. Balancing model performance with interpretability is a common challenge.

Regulatory Compliance mandates adherence to standards set by authorities such as the Energy Regulatory Commission of Thailand. Forecasting models may need to meet accuracy thresholds, reporting formats, and auditability requirements. Non-compliance can result in penalties or loss of market participation rights.

Data Governance defines policies for data ownership, quality, security, and usage. Effective governance ensures that forecast inputs are reliable, that privacy concerns are addressed, and that data lineage is traceable. Poor governance can lead to inconsistent data, legal exposure, and reduced confidence in forecast outcomes.

Cybersecurity protects forecasting infrastructure from malicious attacks that could disrupt data integrity or model operation. Threats include ransomware targeting SCADA data, spoofing of sensor measurements, and denial-of-service attacks on cloud services. Implementing robust authentication, encryption, and monitoring safeguards the forecasting pipeline.

Renewable Integration Study assesses the impact of adding renewable generation to an existing grid, evaluating factors such as voltage stability, congestion, and reserve requirements. Forecasting forms a core component of these studies by providing projected generation profiles. Accurate forecasts reduce the uncertainty in integration assessments, facilitating smoother project approvals.

Transmission Constraint refers to the physical limitation of power lines to carry electricity without violating thermal or stability limits. Renewable generation located far from load centers may be limited by transmission constraints, influencing the value of forecasts for congestion management. Forecasts can be used to schedule power flows that respect line limits.

Congestion Management involves actions taken to alleviate overloads on transmission lines, such as re-dispatching generators, curtailing renewables, or activating demand-response. Forecasts predict when and where congestion may occur, allowing proactive mitigation. Inaccurate forecasts can exacerbate congestion, leading to costly real-time interventions.

Demand Response programs incentivize consumers to adjust their electricity usage in response to price signals or grid conditions. Renewable forecasts enable more precise demand-response scheduling by indicating periods of high renewable availability. Coordination between forecast providers and demand-response aggregators is essential for maximizing benefits.

Market Clearing Price is the price at which supply equals demand in an electricity market. Forecast errors can cause mismatches that lead to price volatility. Accurate renewable forecasts contribute to a more stable market clearing price by reducing unexpected supply fluctuations.

Capacity Expansion Planning determines the future build-out of generation assets to meet projected demand. Renewable forecasts inform planners about expected generation trends, influencing decisions on new wind farms, solar parks, or storage installations. Uncertainty in forecasts must be accounted for in scenario-based planning.

Resource Adequacy assesses whether enough generation capacity, including renewables, is available to meet peak demand with an acceptable reliability level. Forecasting plays a central role by estimating the contribution of variable resources under different weather conditions. Inadequate forecasts can lead to over-estimation of resource adequacy, risking supply shortages.

Renewable Energy Certificate (REC) is a tradable instrument representing the environmental attributes of one megawatt-hour of renewable electricity generation. Forecasting helps generators estimate future REC production, influencing market participation strategies. Accurate generation forecasts ensure compliance with REC obligations and optimize revenue.

Forecast Verification is the systematic assessment of forecast performance using statistical metrics and visual tools. Verification processes include calculating RMSE, MAE, skill scores, and reliability diagrams. Continuous verification supports model improvement and provides transparency to stakeholders.