

Atmospheric Science And Dynamics

Atmospheric Science provides the physical foundation for any artificial-intelligence system that predicts weather. Understanding the language of the atmosphere is essential before an algorithm can be trained to recognise patterns, diagnose errors, or generate forecasts. The following glossary presents the most important terms and concepts that a postgraduate student in AI-driven weather prediction must master. Each entry includes a definition, a brief example, typical applications in modelling or data-driven forecasting, and common challenges that arise when the term is incorporated into a machine-learning pipeline.

Pressure – The force per unit area exerted by the weight of the air column above a point. It is measured in pascals (Pa) or hectopascals (hPa). In practice, surface pressure maps are a primary input for numerical weather prediction (NWP) and for training deep-learning models that learn the evolution of synoptic systems. A common difficulty is the need to interpolate sparse station observations onto a regular grid while preserving the physical balance between pressure gradients and wind fields.

Temperature – The measure of the kinetic energy of air molecules, expressed in kelvin (K) or degrees Celsius (°C). Temperature gradients drive many atmospheric motions, including convection and the development of fronts. AI models often predict temperature anomalies (departures from climatology) because anomalies have reduced variance, which improves learning efficiency. However, temperature fields are strongly coupled to humidity and radiation, so models that treat temperature in isolation can generate physically inconsistent forecasts.

Humidity – The amount of water vapour present in the air. Two common descriptors are specific humidity (mass of water vapour per mass of moist air) and relative humidity (ratio of actual vapour pressure to saturation vapour pressure at a given temperature). Moist processes dominate the formation of clouds and precipitation, making accurate humidity representation crucial for AI-based rainfall prediction. The main challenge is that humidity observations are unevenly distributed, especially over oceans, leading to data-sparse regions that can bias model training.

Lapse Rate – The rate at which temperature decreases with height. The environmental lapse rate (ELR) is the actual observed profile, while the dry adiabatic lapse rate ($\approx 9.8 \text{ K km}^{-1}$) and moist adiabatic lapse rate ($\approx 6 \text{ K km}^{-1}$) describe idealised, frictionless ascent of dry or saturated parcels. In forecasting, lapse rates are used to assess atmospheric stability: If the ELR is steeper than the dry adiabatic lapse rate, the atmosphere is potentially unstable. Machine-learning classifiers that predict severe weather often include lapse-rate-derived features because they encapsulate the thermodynamic state of the column.

Stability – A qualitative description of whether a displaced air parcel will return to its original level (stable),

continue moving away (unstable), or remain neutral. Stability is quantified by comparing the ELR to the adiabatic lapse rates, or by using parameters such as Convective Available Potential Energy (CAPE). AI models that forecast convective storms frequently ingest CAPE and related stability indices as inputs; however, these indices are nonlinear functions of temperature, humidity, and pressure, requiring careful preprocessing to avoid introducing spurious correlations.

Convection – The vertical transport of heat, moisture, and momentum caused by buoyancy. Convective processes manifest as cumulus clouds, thunderstorms, and large-scale updrafts in tropical cyclones. In data-driven approaches, convection is often the most difficult process to capture because it occurs at scales smaller than the typical grid resolution of satellite or reanalysis datasets. Researchers address this by using super-resolution networks or by training sub-grid parameterisation schemes that emulate the net effect of convection on larger scales.

Advection – Horizontal transport of atmospheric properties by the wind. For example, temperature advection refers to the movement of warm or cold air masses across a region. In numerical models, advection terms are discretised using finite-difference or finite-volume methods; in AI, advection can be approximated by recurrent neural networks (RNNs) that learn temporal dependencies, or by convolutional layers that capture spatial flow patterns. A key difficulty is that advection is highly anisotropic; wind direction changes rapidly, so fixed convolution kernels may struggle to represent the true transport direction without additional attention mechanisms.

Geostrophic Wind – The theoretical wind that results from a balance between the pressure-gradient force and the Coriolis force, assuming no friction. It flows parallel to isobars (lines of constant pressure) and is a good approximation for large-scale flow in the mid-latitudes. Geostrophic wind components are often used as target variables for AI models that aim to predict the large-scale circulation, because they are smoother and less noisy than actual observations. However, near the surface where friction is significant, the geostrophic approximation breaks down, and models must learn the residual ageostrophic component.

Gradient Wind – An extension of the geostrophic wind that includes the curvature of the flow, accounting for centrifugal acceleration. Gradient wind is especially relevant around low-pressure systems such as cyclones, where the wind follows a curved path. In AI-based cyclone-track prediction, the gradient wind field can be derived from pressure data and used as a diagnostic to assess the strength of the steering flow. A challenge lies in the need for high-resolution pressure fields to compute curvature accurately, which can be limited by measurement sparsity.

Baroclinic – A state of the atmosphere where surfaces of constant pressure intersect surfaces of constant temperature. Baroclinic zones are characterized by strong temperature gradients and are the typical environment for mid-latitude cyclogenesis. AI models that predict the development of extratropical storms often include baroclinic indicators such as the thermal wind shear or the potential vorticity (PV) gradient. The non-linear interaction between temperature and momentum in baroclinic regions makes them a source of forecast error, especially for data-driven models that lack explicit physics.

Barotropic – A condition where density depends only on pressure, implying that isobaric surfaces are also isentropic (constant potential temperature). Barotropic flows lack vertical wind shear and are relatively simple to model. In practice, the atmosphere is rarely purely barotropic, but the concept is useful for diagnosing the degree of baroclinicity. Machine-learning pipelines sometimes classify flow regimes as barotropic or baroclinic to select appropriate model architectures (e.g., Simpler models for barotropic periods).

Potential Vorticity (PV) – A conserved quantity for adiabatic, frictionless flow, defined as the product of absolute vorticity and the thickness of an isentropic layer. PV is a powerful diagnostic for identifying the dynamical cores of weather systems, such as the upper-level troughs that steer cyclones. In AI, PV fields can be used as inputs or as loss-function constraints to enforce dynamical consistency. The main difficulty is that PV is derived from high-order spatial derivatives of wind and temperature, which can be noisy when computed from gridded observations, potentially degrading model performance.

Rossby Number – A dimensionless parameter that measures the relative importance of inertial forces to Coriolis forces, expressed as $Ro = U/(fL)$, where U is a characteristic velocity, f is the Coriolis parameter, and L is a characteristic length scale. Small Rossby numbers (< 1) indicate age-driven, non-rotational flow (e.g., Small convective cells). In AI, the Rossby number can guide the selection of architecture: For low- Ro regimes, convolutional networks that respect rotational symmetry may be preferred, while high- Ro regimes may benefit from more flexible, non-linear models.

Potential Temperature – The temperature an air parcel would have if brought adiabatically to a reference pressure (usually 1000 hPa). It is denoted by θ and remains constant for dry adiabatic processes, making it a convenient tracer for studying atmospheric stratification. Potential temperature profiles are used to calculate stability indices such as the lifted index (LI) or the K-index. AI models that predict vertical temperature structure often output θ rather than T to preserve physical realism, but converting back to temperature requires knowledge of the pressure field, which can introduce additional error if pressure is not accurately predicted.

Equivalent Potential Temperature – The temperature a parcel would have if all its moisture were condensed out and the latent heat released were used to warm the parcel, then brought adiabatically to 1000 hPa. It is denoted by θ_e and is conserved for moist adiabatic processes. θ_e is valuable for diagnosing the intensity of tropical cyclones and for identifying regions of high moist instability. In machine-learning, θ_e can serve as a target for networks that aim to predict the potential for deep convection, but calculating θ_e from sparse humidity observations can be a source of uncertainty.

Moist Adiabatic Lapse Rate – The rate of temperature decrease with height for a saturated parcel, which is lower than the dry adiabatic lapse rate because of latent heat release. The exact value varies with temperature and pressure, typically around 6 K km^{-1} in the lower troposphere. Moist lapse rates are central to the formulation of convective parameterisations; AI-based parameterisations often learn an effective moist lapse rate from data, thereby bypassing the need for explicit microphysics. The complexity of the

moist lapse rate, however, can make the learned representation sensitive to the training dataset's moisture distribution.

Radiative Transfer – The physical process by which electromagnetic energy is absorbed, emitted, and scattered by atmospheric constituents. Radiative transfer governs the heating and cooling rates of the atmosphere and is the basis for satellite remote-sensing retrievals. In AI, radiative-transfer models are sometimes embedded as differentiable layers, enabling end-to-end training that simultaneously learns atmospheric states and observation operators. The non-linearity of radiative transfer, especially in cloudy conditions, poses a challenge for gradient-based optimisation.

Cloud Microphysics – The set of processes that describe the formation, growth, and fallout of cloud particles (water droplets, ice crystals, snow, graupel). Microphysics schemes in NWP models include equations for condensation, evaporation, autoconversion, accretion, and sedimentation. Data-driven approaches aim to replace or augment these schemes with neural networks that predict bulk quantities such as cloud water mixing ratio or precipitation rate. A persistent challenge is that microphysical processes operate at sub-grid scales, leading to a mismatch between observable variables (e.G., Radar reflectivity) and the model state variables needed for training.

Radar Reflectivity – A measure of the power returned to a radar antenna by precipitation particles, expressed in decibels relative to Z (dBZ). Reflectivity is proportional to the sixth power of particle diameter, making it highly sensitive to large drops or hail. AI models that nowcast precipitation often ingest sequences of reflectivity fields to predict the motion and intensity of convective cells. The relationship between reflectivity and actual rain rate is non-linear and depends on drop-size distribution; consequently, converting reflectivity to quantitative precipitation can introduce systematic bias if not handled carefully.

Satellite Retrieval – The inverse problem of estimating atmospheric variables (e.G., Temperature, humidity, aerosol optical depth) from satellite radiances. Retrieval algorithms may be statistical (e.G., Regression) or physical (e.G., Optimal estimation). Recent advances use deep-learning models that map multi-spectral radiances directly to atmospheric profiles, dramatically reducing computational cost. However, satellite retrievals are subject to uncertainties due to sensor calibration, surface emissivity, and cloud contamination; these uncertainties must be propagated into downstream AI forecasting pipelines to avoid over-confident predictions.

Reanalysis – A comprehensive, gridded dataset that combines historical observations with a consistent NWP model through data assimilation, producing a continuous record of atmospheric variables. Examples include ERA5, JRA-55, and NCEP-DOE Reanalysis. Reanalyses provide the backbone for training AI models because they deliver complete, physically coherent fields. Nevertheless, reanalysis products inherit model biases and assimilation errors, which can imprint systematic distortions onto machine-learning predictions if not corrected.

Data Assimilation – The process of merging observations with a prior model state to produce an optimal

estimate of the atmospheric state, often using techniques such as variational (3D-Var, 4D-Var) or ensemble Kalman filters (EnKF). In an AI context, data assimilation can be reframed as a learning problem: Neural networks can be trained to approximate the analysis step, or to generate ensemble perturbations. The main difficulty is that assimilation requires accurate error statistics for both observations and background, which are not always readily available for data-driven methods.

Ensemble Forecasting – Running multiple model simulations with slightly varied initial conditions or physics to sample the range of possible future states. Ensembles provide probabilistic information, such as the likelihood of extreme events. AI systems can generate ensembles by perturbing inputs, using stochastic dropout, or employing generative models (e.g., GANs, VAEs). Ensuring that ensemble members remain physically plausible—respecting conservation of mass, energy, and momentum—is a current research frontier.

Bias Correction – A post-processing step that adjusts model output to reduce systematic errors relative to observations. Traditional bias correction uses statistical techniques (e.g., Linear regression, quantile mapping); modern approaches employ neural networks that learn non-linear mappings. The challenge is to avoid over-fitting to past data, which can degrade performance under novel climate conditions.

Loss Function – The mathematical expression that quantifies the discrepancy between model predictions and reference data during training. Common loss functions in weather prediction include mean squared error (MSE), mean absolute error (MAE), and more specialized metrics such as the continuous ranked probability score (CRPS) for probabilistic forecasts. Selecting an appropriate loss function is critical: A loss that penalises large errors more heavily (e.g., MSE) may encourage smoother predictions, while a loss that respects distributional shape (e.g., CRPS) can improve calibration of ensemble forecasts.

Overfitting – The situation where a model captures noise or idiosyncrasies of the training data rather than the underlying physical relationships, leading to poor generalisation on unseen data. Overfitting is particularly acute in atmospheric science because datasets are often limited in temporal extent and have strong autocorrelation, making naive random splits ineffective. Regularisation techniques such as dropout, weight decay, early stopping, and cross-validation are essential safeguards.

Cross-validation – A statistical method for assessing model performance by partitioning data into training and validation subsets multiple times. In time-series contexts, k-fold cross-validation must respect temporal ordering (e.g., Rolling-origin evaluation) to avoid leakage of future information into the training set. Proper cross-validation provides a more reliable estimate of forecast skill and helps detect over-fitting.

Training Dataset – The collection of input–output pairs used to optimise model parameters. For weather AI, the training dataset may consist of past reanalysis fields (inputs) paired with observed precipitation or temperature fields (targets). Curating a high-quality training dataset involves handling missing data, ensuring spatial-temporal consistency, and balancing representation of rare events (e.g., Tropical cyclones).

Validation Dataset – A separate subset of data used to monitor model performance during training and to tune hyper-parameters. In weather prediction, validation data may be drawn from a different year or from a distinct geographic region to test the model's ability to generalise across climate regimes.

Test Dataset – The final, unseen data on which model performance is reported. For a robust assessment, the test dataset should span multiple seasons and include extreme weather cases that were not present in the training set.

Hyper-parameter – Configuration settings that govern the behaviour of a learning algorithm but are not learned from data (e.G., Learning rate, number of layers, kernel size). Hyper-parameter optimisation can be performed manually, via grid search, or using Bayesian optimisation. Because atmospheric datasets are large and computationally intensive, efficient hyper-parameter search strategies that reduce the number of full training runs are highly valuable.

Neural Network – A computational model composed of interconnected layers of artificial neurons that learn hierarchical representations of data. In weather prediction, common architectures include convolutional neural networks (CNNs) for spatial fields, recurrent neural networks (RNNs) for temporal sequences, and transformer models for global attention. Designing a neural network that respects physical constraints (e.G., Mass conservation) is an active area of research, often addressed with physics-informed loss terms or custom layers.

Convolutional Neural Network – A class of neural networks that apply learned filters (kernels) across spatial dimensions to extract local features. CNNs excel at recognizing patterns such as fronts, cloud bands, or jet streams in gridded fields. However, standard CNNs assume translation invariance, which may conflict with the rotating Earth's geometry. Solutions include using spherical convolutions or embedding latitude-dependent weighting functions.

Recurrent Neural Network – A network architecture that processes sequences by maintaining an internal state that evolves over time. Variants such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) mitigate the vanishing-gradient problem and can capture long-range temporal dependencies, making them suitable for forecasting the evolution of atmospheric fields over several days. A drawback is that RNNs are computationally intensive for high-resolution spatial data, often requiring hybrid approaches that combine CNNs for spatial encoding with RNNs for temporal dynamics.

Transformer – A deep-learning architecture that relies on self-attention mechanisms to model relationships between all positions in a sequence, removing the need for recurrent connections. Transformers have shown promise in weather prediction by learning global dependencies across large spatial domains, such as the interaction between tropical convection and extratropical wave trains. Their main limitation is the quadratic scaling of memory with input size, which can be prohibitive for high-resolution global datasets; recent research explores sparse attention or patch-based strategies to alleviate this.

Attention Mechanism – A component that allows a model to focus on specific parts of the input when generating each output element. In atmospheric applications, attention can be used to highlight regions of high moisture or strong vorticity that are most relevant for a target forecast (e.G., Precipitation). Interpreting attention maps provides a form of explainability, yet care must be taken because attention weights do not always correspond to causal influence.

Generative Adversarial Network (GAN) – A framework consisting of a generator network that creates synthetic samples and a discriminator network that distinguishes real from fake samples. GANs have been applied to super-resolution of satellite imagery, generation of realistic precipitation fields, and stochastic downscaling of climate projections. Training GANs is notoriously unstable, requiring careful balancing of the generator and discriminator losses, and the resulting samples may lack physical consistency unless additional constraints are imposed.

Variational Autoencoder (VAE) – A probabilistic generative model that learns a low-dimensional latent representation of data while enforcing a prior distribution. VAEs can be used to produce ensembles of plausible atmospheric states by sampling the latent space, offering a computationally cheap alternative to traditional ensemble generation. The trade-off is that VAEs tend to produce blurry outputs, which can be problematic for sharp features such as convective cores.

Physics-Informed Neural Network (PINN) – A neural network that incorporates known physical equations (e.G., Navier-Stokes, thermodynamic relationships) into the loss function, guiding the learning process toward physically consistent solutions. In weather prediction, PINNs can enforce the continuity equation or the hydrostatic balance, reducing the need for massive training data. However, constructing appropriate differential-equation penalties for complex, discontinuous phenomena like fronts remains challenging.

Hybrid Model – A system that combines traditional NWP components (e.G., Dynamical core, physical parameterisations) with data-driven modules (e.G., Neural-network-based convection). Hybrid models aim to leverage the strengths of both approaches: The interpretability and stability of physics-based components and the flexibility of machine learning to capture unresolved processes. Integration challenges include ensuring consistent time stepping, handling differing spatial resolutions, and maintaining mass and energy budgets across the coupled system.

Stochastic Parameterisation – The representation of sub-grid processes using random variables to reflect uncertainty. In AI, stochastic parameterisation can be implemented by sampling from learned probability distributions for quantities such as cloud water content or turbulent fluxes. Stochastic schemes improve ensemble spread but require careful calibration to avoid generating unphysical variability.

Temporal Resolution – The time interval between successive data points (e.G., Hourly, 3-hourly). High temporal resolution provides richer information about rapid processes like squall lines, but also increases data volume and computational cost. Downsampling can be used to reduce dimensionality, yet it may discard critical signals; therefore, AI models often incorporate multi-scale temporal inputs, processing both

fine- and coarse-resolution streams.

Spatial Resolution – The distance represented by each grid cell (e.G., 0.25°, 5 Km). Finer resolution captures smaller-scale phenomena such as topographically induced rainfall, but demands more memory and longer training times. Super-resolution networks attempt to upscale coarse forecasts to higher resolutions, learning the mapping between large-scale patterns and fine-scale details. A key difficulty is that the mapping is not unique; multiple fine-scale realizations can be consistent with the same coarse field, leading to ambiguity that must be addressed through probabilistic modeling.

Vertical Coordinate – The system used to discretise the atmosphere in the vertical direction. Common choices include pressure levels, hybrid sigma-pressure coordinates, or terrain-following sigma coordinates. AI models that ingest vertical profiles must be aware of the coordinate system to correctly interpret gradients and to preserve hydrostatic balance. Converting between coordinate systems can introduce interpolation errors that propagate into the training data.

Hydrostatic Approximation – The assumption that vertical pressure gradients are balanced by the weight of the overlying air, neglecting vertical accelerations. This approximation is valid for large-scale motions but fails for deep convection. AI models that predict vertical velocity fields should therefore be aware of the regime: For hydrostatic scales, the vertical velocity is small and can be inferred from divergence, while for non-hydrostatic scales, explicit prediction of vertical motion becomes necessary.

Continuity Equation – The conservation of mass expressed as $\partial \rho / \partial t + \nabla \cdot (\rho \mathbf{v}) = 0$, where ρ is density and $\mathbf{v}$ is velocity. Enforcing the continuity equation in AI models helps maintain physically realistic wind fields. Techniques include adding a divergence-penalty term to the loss function or designing network architectures that output a streamfunction and velocity potential, guaranteeing non-divergent flow by construction.

Momentum Equation – The set of equations governing the evolution of wind, incorporating pressure-gradient force, Coriolis force, friction, and advection. In data-driven forecasting, the momentum equation can be used as a constraint to regularise the learned wind fields. However, estimating the pressure gradient from sparse observations can be noisy, limiting the effectiveness of such constraints.

Energy Conservation – The principle that total atmospheric energy (internal, kinetic, potential) is conserved in the absence of external sources or sinks. Machine-learning models that predict temperature and wind should respect energy balance to avoid drift over long integrations. One approach is to embed a prognostic energy equation into the network and penalise deviations from the conserved quantity.

Thermodynamic Diagram – Graphical representations such as the Skew-T log-P diagram or the tephigram, used to visualise temperature, moisture, and stability profiles. AI systems that ingest sounding data may convert these profiles into feature vectors derived from the diagram (e.G., CAPE, CIN) or learn to interpret the raw profile directly using sequence models. Understanding the physical meaning of these derived

features aids in interpreting model outputs and in diagnosing errors.

Convective Parameterisation – The set of empirical or semi-theoretical relationships that represent the net effect of unresolved convection on the resolved scale, typically providing heating, moistening, and momentum tendencies. AI-based convective schemes replace or augment traditional parameterisations with neural networks that predict these tendencies from large-scale variables. A persistent challenge is ensuring that the learned tendencies remain stable when coupled to the dynamical core, as small errors can amplify and cause model blow-up.

Boundary Layer – The lowest part of the atmosphere directly influenced by the surface, where turbulence and friction dominate. The planetary boundary layer (PBL) height, surface fluxes, and turbulent mixing coefficients are critical for near-surface temperature and wind forecasts. AI models that predict PBL depth often use surface observations (e.G., Temperature, humidity, wind) combined with satellite-derived land-surface temperature to capture diurnal cycles.

Land-Surface Model – A component that simulates exchange of heat, moisture, and momentum between the ground and the atmosphere, accounting for soil moisture, vegetation, and snow cover. In AI-enhanced forecasting, land-surface models can be replaced with neural-network emulators that map atmospheric forcing to surface fluxes, dramatically reducing computational cost. However, ensuring that the emulator respects soil-water balance and energy constraints remains an active research problem.

Ocean-Atmosphere Coupling – The interaction between the ocean surface and the overlying air, which influences sea-surface temperature (SST), humidity, and wind stress. Coupled forecasts require consistent exchange of fluxes at the interface. AI methods can improve coupling by learning SST-forecast corrections or by providing rapid estimators of oceanic heat uptake. The difficulty lies in the different timescales: Ocean dynamics evolve more slowly than atmospheric ones, requiring careful temporal alignment of datasets.

Sea-Surface Temperature (SST) – The temperature of the topmost ocean layer, a key driver of atmospheric circulation, especially in the tropics. SST anomalies are often used as predictors for ENSO (El Niño-Southern Oscillation) events. AI models that forecast SST typically employ recurrent architectures to capture the slow evolution of oceanic patterns, but the scarcity of high-frequency SST observations can limit training data quality.

El Niño-Southern Oscillation (ENSO) – A coupled ocean-atmosphere phenomenon characterised by periodic warming (El Niño) or cooling (La Niña) of the central-tropical Pacific. ENSO exerts a global influence on weather patterns, making its prediction a central goal of climate-scale AI models. Forecasting ENSO involves skillful modelling of oceanic subsurface temperature, thermocline depth, and atmospheric teleconnections. The challenge is the long lead times (6-12 months) and the non-linear nature of the system, which can lead to model collapse if not regularised.

Teleconnection – A climate anomaly that influences weather patterns at large distances, such as the North

Atlantic Oscillation (NAO) or the Pacific-North American (PNA) pattern. Teleconnections are identified through empirical orthogonal functions (EOFs) or similar statistical techniques. AI models may incorporate teleconnection indices as additional inputs, improving skill for regional forecasts that are modulated by large-scale modes. However, teleconnection patterns can shift under climate change, requiring models that adapt to non-stationary relationships.

Empirical Orthogonal Function (EOF) – A statistical method that decomposes a dataset into orthogonal spatial patterns (modes) and associated temporal coefficients, often used to identify dominant variability structures. EOFs provide a compact representation of high-dimensional atmospheric fields, facilitating dimensionality reduction for machine-learning. The limitation is that EOFs are linear and may not capture non-linear interactions; nonlinear alternatives such as autoencoders are sometimes employed to overcome this.

Principal Component Analysis (PCA) – A technique closely related to EOF analysis that finds orthogonal axes explaining maximum variance in the data. PCA is frequently applied to reduce the dimensionality of large reanalysis datasets before feeding them into neural networks. Care must be taken to retain sufficient components to preserve important weather signals, particularly extreme events that may be represented by low-variance modes.

Climatology – The statistical description of weather over a reference period (typically 30 years), providing baseline values such as mean temperature, precipitation frequency, and standard deviation. AI models often predict anomalies relative to climatology because the anomaly fields have lower variance and are more stationary. Nonetheless, climatological baselines themselves evolve under climate change, so models trained on historical climatology may become biased when applied to future conditions.

Bias – Systematic error in a model's output, often expressed as the mean difference between predictions and observations. Bias can arise from model physics, data preprocessing, or training-set imbalance. In AI weather prediction, bias is commonly corrected using statistical post-processing, but integrating bias correction into the training loop can improve overall skill.

Calibration – The process of adjusting forecast probabilities so that observed frequencies match predicted probabilities. Well-calibrated probabilistic forecasts are essential for decision-making (e.g., Flood warnings). Calibration techniques include isotonic regression, Bayesian model averaging, and ensemble model output statistics (EMOS). Machine-learning models may learn calibration implicitly if trained with proper scoring rules, yet explicit calibration remains valuable for operational reliability.

Verification Metric – A quantitative measure used to assess forecast quality. Common metrics include root-mean-square error (RMSE), mean absolute error (MAE), Brier score (for binary events), and the equitable threat score (ETS). In AI, selection of verification metrics guides model optimisation; for example, training with a differentiable approximation of the CRPS can directly improve probabilistic forecast skill.

Skill Score – A dimensionless indicator that compares a forecast’s performance to a reference (often climatology or persistence). The most widely used skill score is the anomaly correlation coefficient (ACC), which measures the correlation of forecast anomalies with observed anomalies. Positive skill scores indicate added value over the reference, whereas negative scores reveal degradation. Skill scores are crucial for evaluating whether an AI approach truly advances beyond traditional methods.

Persistence Forecast – A simple baseline that assumes the future state will be identical to the current observation. Persistence is surprisingly competitive for short-range forecasts (e.G., Temperature a few hours ahead). AI models must demonstrate skill beyond persistence to justify their complexity; accordingly, persistence is often used as a benchmark during model development.

Ensemble Kalman Filter (EnKF) – An algorithm that assimilates observations by updating an ensemble of model states, approximating the error covariance with the ensemble spread. EnKF can be combined with neural networks to create hybrid data-assimilation systems where the network provides a model forecast, and the EnKF updates it with observations. The challenge is maintaining ensemble diversity while preventing filter divergence, especially when the network introduces non-Gaussian errors.

Variational Data Assimilation (VarDA) – A technique that finds the model state that minimises a cost function representing the mismatch between the model forecast and observations, weighted by their error covariances. VarDA is the backbone of many operational NWP centres. Recent research explores differentiable variational assimilation, where the cost function and its gradient are computed through automatic differentiation, enabling seamless integration with deep-learning components.

Gradient Descent – An optimisation algorithm that iteratively updates model parameters in the direction of decreasing loss. Variants such as stochastic gradient descent (SGD), Adam, and RMSprop are widely used in training weather AI models. Gradient descent requires careful tuning of learning rates; too large a rate can cause divergence, while too small a rate leads to slow convergence. Adaptive optimisers like Adam alleviate some of these issues but may introduce bias in the final parameter estimates.

Learning Rate – The step size used by gradient-descent algorithms to update parameters. In atmospheric AI, learning rates are often scheduled to decay over epochs, or they are adapted per-parameter using algorithms such as Adam. Choosing an appropriate learning rate is critical: A learning rate that is too high can destabilise training, especially when the loss landscape contains steep gradients due to sharp weather fronts.

Batch Size – The number of training samples processed before the model’s parameters are updated. Larger batch sizes provide smoother gradient estimates but require more memory, whereas smaller batches introduce noise that can help escape local minima. In weather prediction, the batch size is constrained by the size of the spatial domain and the resolution of the data; practitioners often use mini-batches that contain a few days of global fields.

Normalization – The process of scaling input variables to have zero mean and unit variance (or another prescribed range). Normalisation improves training stability and speeds convergence. For atmospheric variables, it is common to normalise each field separately (e.G., Temperature, geopotential height) because they have distinct physical units and variability scales.

Data Augmentation – Techniques that artificially increase the size of the training dataset by applying transformations such as rotation, flipping, or adding noise. In atmospheric science, physical constraints limit permissible augmentations; for example, rotating a global field disrupts the relationship with latitude-dependent Coriolis force. Acceptable augmentations include adding realistic observational noise, perturbing initial conditions, or applying time-shifts that preserve the diurnal cycle.

Transfer Learning – The practice of re-using a model trained on one task (e.G., Image classification) as a starting point for a related task (e.G., Precipitation nowcasting). Transfer learning can accelerate convergence and reduce data requirements, especially when the target dataset is limited. For weather applications, pretrained CNNs on large image datasets are sometimes fine-tuned on satellite imagery, though the domain shift can be substantial, necessitating intermediate adaptation layers.

Domain Adaptation – Methods that adjust a model trained on one data distribution (source domain) to perform well on another distribution (target domain). In weather AI, domain adaptation is relevant when transferring models trained on reanalysis data to real-time observations, which may have different error characteristics. Techniques include adversarial training to align feature distributions or re-weighting loss functions based on domain-specific uncertainties.

Explainability – The set of tools and methods used to interpret the internal workings of a machine-learning model. For atmospheric forecasts, explainability helps build trust with meteorologists and end-users. Common techniques include saliency maps, layer-wise relevance propagation, and SHAP values. However, atmospheric data are high-dimensional, and interpreting why a neural network predicts a specific rainfall pattern can be ambiguous; robust explainability remains an open research frontier.

Uncertainty Quantification (UQ) – The assessment of confidence intervals or probability distributions associated with model predictions. In weather forecasting, UQ is essential for risk-based decision making. Approaches include Bayesian neural networks, Monte-Carlo dropout, and ensemble methods. Quantifying both aleatoric (inherent) and epistemic (model) uncertainty allows forecasters to distinguish between unpredictable weather and gaps in training data.

Monte-Carlo Dropout – A technique that approximates Bayesian inference by applying dropout at inference time and sampling multiple forward passes to generate a predictive distribution. This method provides a computationally cheap way to estimate uncertainty in deep-learning weather models. The drawback is that dropout-induced variance may underestimate true uncertainty, especially for extreme events where the model's epistemic uncertainty dominates.

Bayesian Neural Network (BNN) – A neural network where weights are treated as probability distributions rather than fixed values, enabling principled uncertainty estimation. BNNs have been applied to short-range temperature forecasting, yielding calibrated predictive intervals.