Lecture (4)

Data Assimilation and Initialization

4.1 Data assimilation

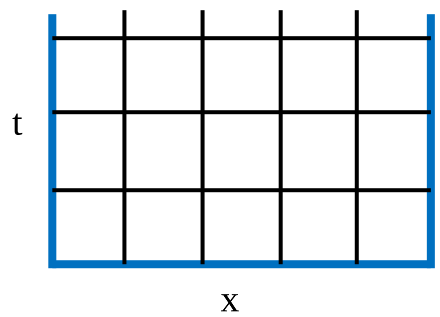
It is a mathematical discipline that seeks to optimally combine *theory* (usually in the form of a numerical model) with *observations*. There may be a number of different goals, for example, to determine initial conditions for a numerical forecast model, to interpolate sparse observation data.

Data assimilation initially developed in the field of numerical weather prediction models. In order to use these models to make forecasts, initial conditions are needed for the model that closely resemble the current state of the atmosphere. Simply inserting point-wise measurements into the numerical models did not provide a satisfactory solution. Real world measurements contain errors both due to the quality of the instrument and how accurately the position of the measurement is known. These errors can cause instabilities in the models that lessen the forecast skill. Thus, more sophisticated methods were needed in order to initialize a model using all available data while making sure to maintain stability in the numerical model.

4.2 Initialization

Mathematically, numerical weather prediction (NWP) can be viewed as solving an initial-boundary value problem in which the governing equations of atmosphere are integrated forward in time in a finite domain.

- Therefore, in addition to the boundary conditions, we must also provide suitable initial conditions for the model.
- For idealized numerical simulations, the initial conditions may be prescribed by known functions or values.



- If the Coriolis force is included in the model, then the initial basic state should be in geostrophic balance. Otherwise, the initial state will be adjusted to reach a new balanced state by the model.

- For numerical weather prediction, the observational data must be modified to be accurate and dynamically consistent with the model's governing equations.

  - Strictly speaking, the process in producing initial conditions may be classified as the following four components:

(i) *quality control*, (ii) *objective analysis*, (iii) *initialization*, and
(iv) *initial guess from a short-range forecast by an NWP model*.

  - Reason for quality control: The errors associated with the data may be misrepresented and amplified by the model.

  - To reduce the errors in the sounding data (observations), the following steps of quality control have been taken in numerical weather prediction:

(a) *plausibility check, (b) contradiction check, (c) gross check, and*
(d) buddy check.

 In ***plausibility check***, data values that cannot possibly occur in the real atmosphere or extremely exceed climatological mean are rejected.

For example, positive temperatures in Celsius at 300 hPa are rejected.

 In ***contradiction check***, data values of two or more parameters at the same location contradicting to each other are removed.

For example, the occurrence of rain in the absence of clouds is removed.

 In ***gross check***, observations with large deviations from the first guess field forecast by an operational model are removed.

 In ***buddy check***, observations not agreeing with neighboring observations are removed.

- Observational data are often not regularly spaced, which are not ready for use as initial fields for a NWP model because they do not match the model grid mesh.
- In some areas, such as over ocean, observational data are sparse.

Therefore, in order to use the observational data as initial fields for NWP model, we need to interpolate or extrapolate the data to fit into the grid mesh of the model and to apply some balance relations, such as geostrophy and mass continuity, to make the data dynamically consistent. This procedure is called ***objective analysis***.

- In an objective analysis, it is desirable to:

(1) Filter out scales of motion that cannot be resolved by the model,

(2) Use a first guess field or background field provided by an earlier forecast from the same model, and

(3) Make use of our knowledge of the probable errors associated with each observation, which may be weighted based on past records of accuracy.

- When the maximum information from data sources, including the observations, climatological records, space correlation among the meteorological variables, etc., are extracted statistically, the approach is called ***optimal interpolation***.

- This often requires knowledge of the statistical structure of the fields of the variables. The variables may be analyzed separately or simultaneously, which is referred to as ***univariate analysis*** *or* ***multivariate analysis***, respectively.

Thus, the use of such objectively analyzed data to initialize an NWP model may generate large, spurious inertial-gravity wave modes. Theoretically, these inertial-gravity wave modes will be dispersed, dissipated or propagate out of the domain due to redistribution of mass and wind fields. However, these noises, as often referred to by NWP modelers, need to be removed.

Therefore, an additional procedure, called ***initialization***, is required to force the data after objective analysis to be dynamically consistent with the model dynamics, and to allow the model to integrate forward in time with a minimum of noise and maximum accuracy of the forecasts.

Historically, a number of initialization techniques have been developed and used in NWP models, such as:

(a) *damping method*,

(b) *static initialization*,

(c) *variational method*,

(d) *normal mode initialization*, and

(e) *dynamic initialization.*

Homework: Plot a suitable graph summarizes the above lecture.