



Earth system data assimilation

What is Earth system data assimilation?

Earth system data assimilation combines the latest observations with a short-range forecast constrained by previous observations to obtain the best possible estimate of the current state of the Earth system. That estimate, called the analysis, is used as the initial conditions on which weather forecasts are based. The quality of forecasts crucially depends on the accuracy of the analysis. If analyses are produced for extended periods reaching back decades, but using one and the same data assimilation system, they become a 'reanalysis' of past weather and climate.

Why do we need to combine observations with short-range forecasts?

Observations are unevenly distributed and come with errors. They cannot provide a complete and accurate picture of the state of the Earth system across the globe at a given point in time. Fortunately, we know a lot about that state from a highly skilful short-range forecast based on the previous analysis and thus on previous observations. Data assimilation adjusts that forecast slightly in a physically consistent manner so that it matches the latest observations closely, taking into account uncertainties in the observations and the short-range forecast.

How are the observations combined with short-range forecasts?

Different mathematical techniques can be used to combine Earth system observations with short-range forecasts. The method used at ECMWF for the atmosphere is called four-dimensional variational data assimilation (4D-Var). 4D-Var iteratively adjusts the initial conditions of a short-range forecast to bring it into closer agreement with meteorological observations in space and time. The adjustments are made in a manner which respects the physical laws that govern the Earth system. The adjusted forecast then serves to define the initial conditions of a new forecast. ECMWF implemented 4D-Var in 1997. This quickly led to a dramatic improvement in forecast quality. Since then,

continued enhancements of the 4D-Var system and the introduction of a probabilistic approach through an Ensemble of Data Assimilations system have improved forecasts further.

What kind of observations are used?

Some observations reflect conditions at the location of the instrument. Such in-situ measurements come for example from weather stations, aircraft, radiosondes, ships or buoys. Other observations are based on remote sensing, for example from satellites or ground-based radar. Most satellite instruments measure electromagnetic radiation emitted from the Earth's surface and the atmosphere. The signals received can be used to infer conditions at the surface or in the atmosphere. Passive sensors measure radiation that is emitted naturally, while active sensors measure the backscatter of radiation emitted from the instrument.

How important are different kinds of observations?

Remote sensing, especially using satellite instruments, has become ever more important because of the global coverage it provides. As we became better at extracting satellite data information and the amount of global weather data increased, the skill of forecasts in areas with few in-situ observations, such as the southern hemisphere, increased sharply. In-situ weather data continue to be very important, and they have a key role in calibrating remote sensing instruments, but overall today satellite data make the biggest contribution to global forecast skill.

What is coupled data assimilation?

ECMWF's Integrated Forecasting System (IFS) uses separate data assimilation systems for the atmosphere, the ocean, ocean waves, the land surface and sea ice. This may produce an internally inconsistent analysis if the data assimilation systems are independent from each other. Coupled data assimilation aims to ensure that the analysis of different Earth system components is consistent. In the IFS, this is achieved by allowing the observations of one Earth system component to influence the analysis in other components and changing the initial conditions of each Earth system component in a way which is physically consistent with the other components.

What is the Ensemble of Data Assimilations?

The analysis on which forecasts are based comes with uncertainties. This is inevitable due to the errors that are present in the observations and the short-range forecasts. While we cannot reduce those errors to zero, we can quantify the uncertainties. To this end, multiple 4D-Var data assimilations are run in parallel using slightly perturbed observations, models and atmospheric boundary conditions, such as sea-surface temperature. ECMWF has run such an Ensemble of Data Assimilations (EDA) since 2010 to help define the initial conditions for ensemble forecasts as well as for high-resolution deterministic forecasts.