Trust This Model Of Imperfection

Forget the Farmer's Almanac or analyzing chicken innards. The best way to forecast weather is through computer models that don't always work.

by Scott C. Dennstaedt

Weather prediction has come a long way in the last 50 years. This is largely a result of the advancement in computer technology and numerical weather prediction. As it stands now, our feet are firmly planted in the industrial revolution when it comes to forecasting the weather.

Tremendous progress has been made in weather forecasting and our ability to observe the atmosphere, but we are just beginning to lay the foundation for the future. There's no doubt that significant improvements will abound over the next 50 years. However, don't hold your breath; any significant leaps will take time.

Aviation has perhaps benefited the most from advances in meteorological science. Short-range forecasting and our ability to see real-time (or near real-time) weather have made our skies much safer. Airborne and ground-based radar, satellite meteorology, and surface-based observation equipment are all testaments to these desired improvements. Much of this information is now available where it counts the most—in the cockpit.

The challenge that meteorologists face is how to take a bunch of raw observations in many different flavors and turn them into a useful and timely weather picture of the future. The forecaster's crystal ball consists of a very fast computer and a number-crunching piece of software called a forecasting model. The model is initialized with the latest and greatest observations; the crank is turned, and out comes even more data or what meteorologists refer to as numerical weather guidance. Finally, the model output is post-processed and various weather maps, charts, graphs, diagrams and text are generated in a form useful to forecasters.

It goes without saying that weather forecasting isn't perfect and may never be perfect, at least not in our lifetime. Why can't the weather be predicted more accurately? As a meteorologist, I get asked this question all of the time. Perhaps a better question is: "How can we predict the weather at all?"

The Weather Underground

The process described above doesn't sound too complicated, but it's far from simple. In order to really appreciate the weather forecasts pilots routinely utilize, you must first understand what goes on behind the scenes to create your favorite weather charts, maps, and diagrams. In most instances, this means understanding the foundation of nearly all weather forecasting, that is, numerical weather prediction (NWP). Okay, who was thinking the *Farmer's Almanac*? For the purposes of this discussion we are just referring to models that forecast at a planetary, synoptic (big picture), or mesoscale (medium view) level.

> *"It goes without saying that weather forecasting isn't perfect and may never be perfect."*

As mentioned above, computers are the primary reason NWP can be done at all. Without the ability to model the atmosphere, meteorologists would rarely be able to provide an accurate short-range forecast and would never be able to provide a useful forecast several days in advance.

There are dozens of forecasting models in use today all over the world that serve a plethora of purposes to include academia, private industry, and government entities. As a result, forecasting models come in all shapes, sizes, and variants. Here are a few classifications that are important to recognize.

The first and perhaps most distinguishing classification is the model domain. Some models have a global domain and others have a regional domain. A global model normally has no defined boundary. It takes in data from all over the world to produce a forecast that spans the globe. A regional model, on the other hand, has hard boundary conditions and produces forecasts over a limited area (such as a region covering North America to include the Atlantic and Pacific Oceans).

Models can also be characterized by how many hours or days in the future the prediction is continued once the model is initialized. They are often grouped into three basic categories that include long-range, medium-range, or short-range. For example, long-range models may be run to produce a forecast valid every six hours with a range out to two weeks or longer. At the other extreme, short-range models may produce a forecast valid every hour out to a maximum of 12 or 24 hours.

Another important way to classify a forecasting model is based on its horizontal and vertical resolution. A higher resolution model may take longer to execute, but also may produce a more detailed forecast and may have a better handle on mesoscale events such as coastal fronts.

Garbage In, Garbage Out

The first and most important order of business of NWP is called data assimilation. The initial state of the atmosphere is not just described by observations alone. Instead, the initial conditions of the forecasting model are described by a previous short-range forecast that is *corrected* by real-world observations. Think about these small corrections as your nervous instrument instructor is grabbing the controls momentarily to get you back on your assigned heading or altitude.

Using a short-range forecast as the basis of the initial conditions accomplishes several things. First, it fills in the gaps between observing locations that would otherwise be unavailable. Effectively, using a short-range forecast, appropriately called the *first guess*, provides a complete and consistent set of data over the model's entire domain. A previous three-hour forecast is normally used and is assumed to be a good representation of the current conditions due to its short-range nature. Second, it provides for dynamic, physical, and numerical consistencies in the various model forecast fields. Third, it retains information about earlier observations since it was a forecast *also* corrected by its own set of earlier observations.

Real-world data sources used to correct the first guess include, but are not limited to: Observations from satellites, surface station reports, aircraft data (ACARS), wind profilers, ground-based radar, and upper-air soundings (RAOBs).

Once the raw observations are captured, a series of quality assurance checks are made to determine the viability of the initial data prior to using it in the model. The quality of data collected is important to producing an accurate forecast, especially in the long range forecasting models. Remember the saying, "Garbage in, garbage out?" Is a bad set of initial data guaranteed to produce an inaccurate forecast? Bad data alone don't necessarily cause an inaccurate forecast, although they can contribute to one. Even though low quality data can get into the system, they can normally be filtered out.

In this quality control step, meteorologically impossible data and duplicates are thrown out—called a gross check—and data are checked for consistency. That's called a buddy check. For example, assume a surface station reports a temperature of 95 F. While this is meteorologically possible, it's not likely if all of the surrounding stations are reporting temperatures in the 30s. This temperature will likely not be assimilated.

Above: *With computers we can safely put a 182RG through your windshield in IMC, but conjuring a perfect forecast takes more than cyber magic.*

Turn The Crank

When a forecaster speaks of a "run" of a forecasting model, he's referring to the starting or executing of a sophisticated computer program. The forecasting model takes the corrected first guess, that's the results of the data assimilation phase, and using the primitive equations of motion, produces gobs of data affectionately referred to as a forecast.

A forecasting model is useless if it cannot produce a forecast in a timely manner. This isn't an issue with today's forecasting models considering the advancements in computer hardware and software technology. In fact, as much as 10 times more computing resources are used in the data assimilation phase as compared to a one-day forecast.

How Good Is Good Enough?

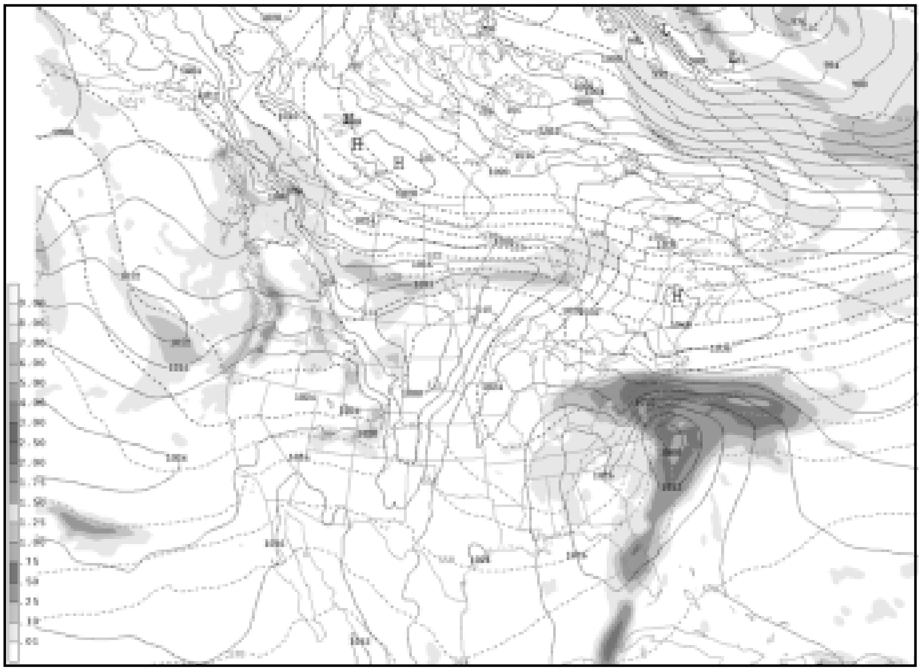
Imagine trying to understand this article by reading every tenth word? In a crude sense, this describes our data observation network. The atmosphere isn't surveyed in its entirety; the overall spatial density of the observations is extremely low. Even if you could measure the data extremely accurately, there would be many locations where you've simply captured little or no data. This is especially true of unpopulated regions or over our vast oceans.

Numerical weather guidance tends to produce the best results when the atmosphere is in an easy-to-predict state. This means that during those tough forecast situations when a forecaster needs the most help, the models tend not to rise to the occasion. Therefore, the forecaster must either accept the model guidance or, in his professional judgment, compensate for any perceived errors. I'll return to that momentarily.

The Human Element

Our inability to completely measure the atmosphere introduces random or non-systematic errors. Data are captured by a variety of instruments that inherently have some observational error. This leaves large gaps to define the initial state of the atmosphere. Even before the crank is turned, the initial state of the atmosphere is uncertain.

And if that isn't enough, models are known to exhibit certain expected behaviors or systematic errors, affectionately referred to as biases. A bias is sort of like knowing what your child will likely do when confronted with a specific situation. You can call it forecaster's intuition. As they are dis-



***Above:** Normally presented in color, this Eta model is run four times daily and produces a forecast valid at an interval of six hours through 84 hours.*

covered, biases become common knowledge to the meteorological community. Knowing about a particular model's biases gives the human an edge on the model. This is a lot like parents being one step ahead of their kids, unless of course, their kids are teenagers.

Over time, forecasters learn about these biases and can tweak a model's forecast accordingly. Knowing when and how to apply a particular bias takes a tremendous amount of skill and a little meteorological risk. Does this lead to an accurate forecast? Not always. However, it usually allows the forecaster to triumph over the model in many cases. Therefore, we can predict when a model will do poorly, hence, allowing for manual adjustment.

Model output statistics (MOS) is another tool in the forecaster's tool kit. MOS is a post-processing step that attempts to apply these known biases in an objective and systematic sense, creating even more data for the forecaster to ponder. The job of the forecaster isn't to just repeat what a given model says will be true. Instead, the forecaster must do better than the model or there's no reason for him to be part of the equation.

Numerical guidance, as it's sometimes called, represents the canvas from which the forecaster begins to prepare his forecast. In reality, the forecaster uses the guidance of many models to come up with the official forecast.

One of the first things that you'll notice when you see a graphically depicted model output of the mean sea level surface map is the lack of fronts. If you see a mean sea level surface chart with fronts depicted, it's most likely a human's forecast and not model output. Frontal analysis is subject to forecaster interpretation and may vary from one forecaster to another and from one model to another.

Post-Processing

The raw model output format isn't terribly useful to forecasters, much like binary format is typically not much help to computer programmers. Post-processing takes the raw data from the model's native forecast grid and translates them into variables that are meaningful to a forecaster. This step also serves to standardize the model output and produces forecast parameters at standard vertical levels as well as many other supplemental derived products and fields.

It's these translated products that are freely available from many academic and government web sites. The translated output is available in the form that forecasters are accustomed to using and include a few products familiar to pilots such as surface isobaric analysis maps, constant pressure charts, and thermodynamic diagrams (skew-T soundings diagrams), just to name a few.

The Big Four

While there are many NWP models and variants of these models, there are four models that stand out. These are the Eta, Rapid Update Cycle (RUC), Nested Grid Model (NGM), and Global Forecast System (GFS). Not that these models always perform beyond our wildest dreams, but they are representative of the various kinds of models available.

The RUC model is especially important because it was designed to serve users needing short-range weather forecasts, such as the aviation community. It produces an operational forecast out to 12 hours, updated every three hours, with hourly forecasts produced in the first three hours. The Forecast Systems Laboratory (FSL) runs a research counterpart of the RUC model called MAPS that forecasts out to 24 hours, updated every three hours, with hourly forecasts produced in the first six hours.

As of April 2002, the Aviation model and its close cousin the Medium Range Forecast (MRF) model were effectively combined and are now called the Global Forecast System (GFS). As the name suggests, the GFS is a global model run four times daily (0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC), and produces output at six-hour intervals out to 384 hours. Yes, you read that right, 16 days.

The Eta and NGM are regional models with a North American domain. The Eta model is run at the same interval as the GFS and produces output also at six-hour intervals out to 84 hours at the 0000 UTC and 1200 UTC runs, and out to 48 hours in the 0600 UTC and 1800 UTC runs. The NGM, on the other hand, is only run twice a day (0000 UTC and 1200 UTC) and produces output at six-hour intervals, out to 48 hours.

Cycles And Valid Times

Familiarity with how often a model is run is essential before interpreting its output. As outlined above, these models are run at different intervals depending on the model. For example, the Eta model is run four times a day and produces a forecast valid at an interval of six hours through 84 hours. The Eta forecast that's initialized at 1200 UTC is called the 12Z run. As a result, you would expect the 12Z run to produce model output for 1200 UTC (00 hour), 1800 UTC (06 hour), 0000 UTC (12 hour), and so forth. The forecast output will generally be posted on the web as the model completes the respective forecast hour.

When looking at any model output, it's vital to determine the date and time the forecast is valid. The date and time format can be somewhat cryptic and isn't necessarily presented in a standard format from one model to the next. Moreover, if you're viewing the model output on the internet, you may find that the data can indeed become stale. So, check those dates.

Who Cares About NWP?

If all of this model output eventually ends up in an official NWS forecast, why should a pilot care about forecasting models? The answer can be summed up in a single word—timeliness. While a human's representation will, on average, be more accurate than the model, it takes the forecaster many hours to put the forecast together, produce the official forecast and make it available to you and me. The official forecast can be six or more hours old for short-range forecasts, and possibly as much as 12 or 18 hours old for longer-range forecasts.

On the other hand, model output becomes available just a couple hours after the model is initialized. Having access to the latest-and-greatest guidance is a boon to the aviator. Learning how to interpret the model output can be difficult, but not impossible.

Scott C. Dennstaedt is a meteorologist, CFII, and IFR *contributing editor. He bases his Cirrus SR22 in Maryland.*

The Quiz