Cornell University
Center for Advanced Computing

CAC
we enable your success

CAC, the Northeast Regional Climate Center, and Cornell's Computational Agriculture Initiative

Improving Weather Data Accuracy and Accessibility

How can the accuracy of weather data be improved to help farmers?

Finding the Answer

As part of Cornell's Computational Agriculture Initiative with the College of Agriculture and Life Sciences, CAC worked with the Northeast Regional Climate Center (NRCC) to improve the accuracy and accessibility of weather data. Weather data impacts the quality of crop models that are used by farmers to determine precisely when to apply fertilizers, herbicides, and pesticides. Easy access to temperature data and the extensive weather data collections at NRCC helps farmers make better crop selection and management decisions.

Improving Weather Data Accuracy

Accurate weather data at biologically meaningful spatial and temporal resolutions are important inputs for the crop models used by farmers. Unfortunately, it is not possible to erect a weather station in every farmer's field to gather temperature, wind, and precipitation data; interpolation of existing weather station data from radar, satellites, weather balloons, and ground stations is, therefore, necessary.

Improved Research

Research Metrics

- Cost: Reduce data handling by storing data in databases rather than flat files
- Accessibility: Provide fast data access to public and researcher requests
- Insight: Create tools to generate contour maps
- Analysis: Enable data queries and analysis with standard tools such as Excel

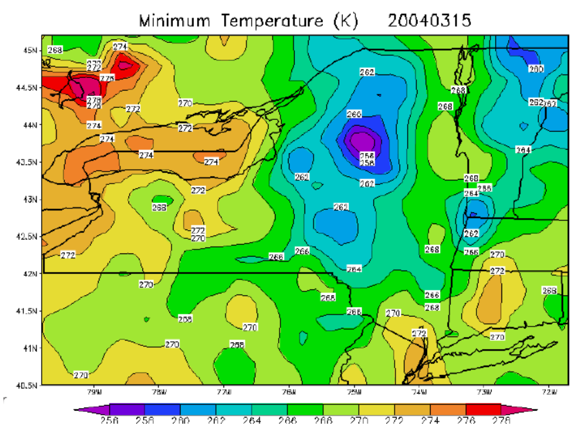
Research Challenge

To make temperature estimates as accurate as possible, researchers at Cornell's Northeast Regional Climate Center are inputting data from the hourly Rapid Update Cycle (RUC) Model Analyses, data obtained from the National Climatic Data Center, and observed temperatures from weather stations across the Northeast into an interpolation program. The interpolation program is a work in progress, as new algorithms are continuously improving interpolation routines.

Researchers use RUC data with a horizontal resolution of 40 km, however, model resolutions of 20 or even 12 km are also used. The ultimate goal is interpolation to 5 km resolution or better via a digital elevation model. Finer scale (1 km) interpolation over limited area is also possible. Pat Sullivan, Associate Professor of Natural Resources, is researching ways to make the interpolation process more accurate.

Each day the NRCC receives hourly temperature and pressure level height data from the RUC model analyses for five levels of atmospheric pressure at 1322 grid points covering the Northeast. After adjusting for elevation differences and correcting for actual weather station observations, daily minimum and maximum temperatures are output for each grid point. While each hourly input file at the 40 km scale is small, the data at the 20 km scale will be much larger, not just because of the finer resolution, but because the recipient can no longer specify the region of the country or the variables needed. As a consequence, each hourly input file will be over 50 MB. At this size, managing data as flat files becomes problematic.

One file of daily minimum and maximum temperatures for the Northeast region at 5 km resolution is over 4 MB. The concern is not only the size of the data, but getting access to the data in its raw form. It is imperative that a seamless way to query this data be developed to maximize its utility. Responding to individual requests for data is labor intensive.



Solution

Staff members of the Cornell University Center for Advanced Computing (CAC) developed a database schema in SQL Server to contain both the hourly input files and the daily minimum and maximum temperature output from the interpolation programs. CAC also developed a Web service that queries the daily minimum, maximum, or average temperatures or cumulative growing degree days and returns a contour map for the Northeast region.

Microsoft Excel can also be used to query the SQL database and return the data as Excel columns which can then be used to generate graphs or perform statistical analysis, either within Excel or in other applications.

The Clients

Northeast Regional Climate Center (NRCC)

- Located in the Dept. of Earth and Atmospheric Sciences in Cornell's College of Agriculture and Life Sciences; directed by Art DeGaetano
- Serves New York State and 11 additional Northeastern states

- Funded by the National Oceanic and Atmospheric Administration (NOAA)
- Facilitates the collection, dissemination, and use of climate data; assesses regional climate conditions and their impacts; and, conducts applied research

Cornell Computational Agriculture Initiative

- Funded by the Co-operative State Research, Education, and Extension Service of the United States Department of Agriculture
- Interdisciplinary research in crop and soil sciences, economics, climate modeling, data-mining, and high-performance computing

The Collaborative Relationship

CAC expertise in database and Web services helped the Northeast Regional Climate Center provide easy access to its immense weather data collection. This data helps farmers with crop selection and is also valuable to researchers interested in determining how climate conditions impact factors such as pollutant transportation. The data has been used as input to a nitrogen model as well. Soon, a farmer will be able to enter information about the farm on a Web page and receive recommendations for how much nitrogen fertilizer to apply to a corn crop, tailored to the farm's location within 5 kilometers or less. "This is a fairly significant leap forward in information for farmers," says Harold van Es, Cornell Professor of Crop and Soil Sciences. Eventually, says Art DeGaetano, the database could be used to refine other applications, including turfgrass management and mosquito control.

Cornell, the land grant college for New York State, is uniquely position to offer such a service, van Es points out, through collaboration between the College of Agriculture and Life Sciences, CAC, and the Northeast Regional Climate Center.

In another number-crunching project, van Es' group will process data from optical spectroscopy, which provides a quick analysis of soil using reflected light. The data from thousands of readings will be entered into a statewide database of soil properties, and the group will validate spectral analysis against chemical, physical and biological analyses of the 2000 or so soil samples sent every year to Cornell's soil health test laboratory.

- This case study is largely based on a 5/3/07 article in the *Cornell Chronicle* written by Bill Steele -