METHODOLOGY TO VISUALIZE AND VALIDATE MODELLING DATA WITH OPEN SOURCE SOFTWARE

Hamish R. Corbett-Hains, M.A.Sc., EIT*
Novus Environmental, Guelph, ON, Canada

1. INTRODUCTION

Atmospheric pollutant modelling has evolved in both accuracy and complexion with advances in computing ability. Each iteration of a particular model brings with it new capabilities which allow for more accurate and refined predictions. Ostensibly, modelling capabilities have out-paced the development or implementation of the graphic displays used to illustrate these results, although some software such as VERDI is successfully used by the CMAS community.

Without a robust way to communicate model results, many of the advances made by modellers may fail to be accepted by the broader scientific and engineering community.

The open source movement of recent years has developed free, easy to use software packages which may be used to generate high-quality data-rich images which can be used both to communicate model results and provide analysis. This poster illustrates some of the ways which open-sourced software - for example, QGIS (http://www.qgis.org[1]) - can be used to easily communicate model results and provides professional-grade graphic capabilities to the scientific community.

2. SOFTWARE DESCRIPTION

Many GIS packages are user-friendly, open source offerings supported by members of the scientific community. Anyone may develop plugins which are freely available and provide a wide variety of functionality. Many of these plugins perform complex analysis which was previously performed manually, improving the efficiency of these types of analysis.

From a modelling perspective, using GIS software to communicate results also allows the user to perform spatial queries through a graphic user interface. Additionally, the prevalence of open source mapping data allows the user to compare results against relevant datasets (e.g. compare ozone levels against the location of sensitive land uses, i.e. hospitals).

An added benefit to using open source software is the ability for the user to develop scripts or plugins which work natively in the program. This ability allows for a high degree of customization within an already robust graphics package.

Additionally, in recent years, the availability of free data published through both public and private sectors has vastly expanded the capabilities of this software. For instance, files containing regional boundaries, transportation corridors, tree counts, or political tendencies are all freely available for the entirety of North America.

3. APPLICATIONS

There are countless uses for spatial software within the modelling communities; this poster covers several broad applications and is intended to showcase the GIS software's capabilities rather than be used as a step-by-step instruction manual. Indeed, with an understanding of data structure and the capabilities of spatial software, applications are limited only by the user's imagination.

3.1 Displaying Spatial Data

Local meteorological conditions and ambient pollutant concentrations, for example, are used or generated by atmospheric models. Often the selection of monitoring stations is dictated by data availability; in other cases the modeller may have to choose between several monitoring stations near to the study area. In either case, being able to visualize the location of a monitoring station along with relevant local features (e.g. unique topography for met stations, large industrial complexes for ambient monitoring stations) can save time and provide greater insight into a dataset. For instance, figure 1 illustrates the location of all the available meteorological stations in Ontario, Canada, within a particular dataset. This alone is useful in trying to find the location of a nearby station to a modelling domain. Figure 2

**Corresponding author:* Hamish R. Corbett-Hains, Novus Environmental, Guelph, ON, Canada; tel. 226-706-8080 ext. 215, fax. 226-706-8081, e-mail: hamishh@novusenv.com

illustrates the type of analysis which is easily performed using GIS software; this figure shows the location of meteorological stations scaled to represent a theoretical value such as wind speed over a certain period. These methods can be extended to provide very robust analysis and visualization tools which can inform tasks such as selecting appropriate datasets or understanding regional variations in environmental data.

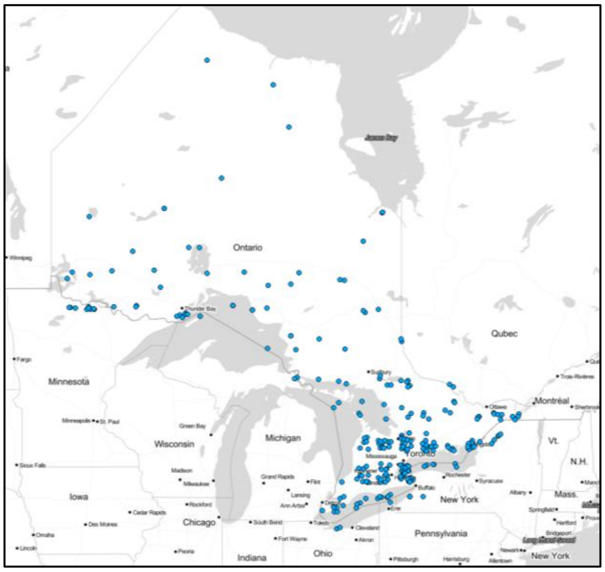

Fig. 1. Location of Meteorological Monitoring Stations within the Province of Ontario

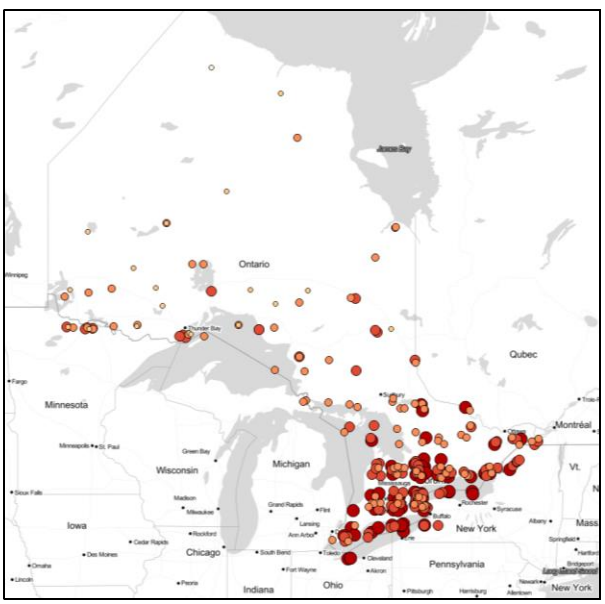

Fig. 2. Location of Meteorological Monitoring Stations within the Province of Ontario Scaled to Reflect Measured Values such as Wind Speed

Spatially plotting results can also provide better insight into their meaning. For example, Figure 3 shows wind roses at the three major airports surrounding Manhattan (LaGuardia, Newark and JFK). When viewed in this fashion, the high degree of spatial variation in the wind patterns becomes clear. The wind rose centered over Manhattan was modelled using the WRF model. This illustrates that none of the airports surrounding Manhattan give realistic representations of the wind patterns on the island. Results of this nature are not unfamiliar to modellers; the ability to communicate these findings to a broader audience, however, is essential. Spatial analysis provides a quick interface for developing these types of plots and adds to the value of a dataset.

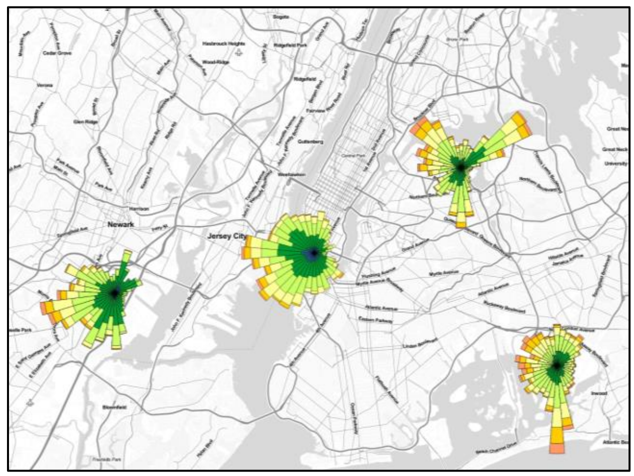

Fig. 3. Measured vs. Modelled Wind Data in the Manhattan Area

3.2 Plotting Modelled Data

Many modelling packages come with embedded visualization methods. Some models, however, do not include this ability; and some packages are too constrained to allow users to display a result set in a meaningful way. Using spatial software can allow users to plot location-based results with ease. This is especially important when the application of the results is spatially variable. For example, some municipalities require that air contaminants be assessed against different standards; spatial analysis allows the user to easily plot results in comparison to regional standards. Figure 3 shows CMAQ modelled ozone concentrations as a percentage of the regional standards over Lake Erie and Lake Ontario. The United States limit for ozone is 70 ppb averaged over 8 hours while the

Canadian limit is 65 ppb averaged over 8 hours. The image illustrates that the difference between the two regional standards is significant when assessing compliance and demonstrates the power of applying a spatial filter to a dataset.

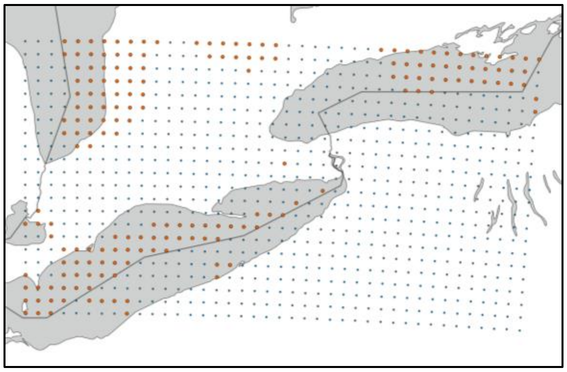

Fig. 3. Cross-Border Modelled Ozone Concentrations. Red Icons Represent Results Greater than 100% of Regional Standard

3.3 Validating Data

Another strength of using spatial software for data visualization is the ability to validate datasets. Quick checks can be performed by plotting results where discrepancies will usually become readily apparent.

Furthermore, statistical comparison of nearby points can easily be performed, allowing users to compare their results with measured data or other modelled points. For instance, forecasting models can be compared against measured values at monitoring stations to assess a model's reliability. While this can be done manually, spatial software allows the user to expedite this analysis by comparing nearby data points automatically. Furthermore, since data is already loaded into the software for this verification, visualizing the data becomes much easier.

4. CONCLUSIONS

As modellers develop more complex tools, the ability to communicate the results becomes increasingly important. Although spatial software should not be used ubiquitously, it provides tools which can vastly improve both the visualizations and validation of data. Additionally, the ability to develop specialized tools which make use of an already robust visualization package makes open source software packages extremely versatile. In short, open source GIS software should be tool which every environmental modeller has, at the least, a cursory knowledge of.

1. Quantum GIS, *Open Source Geospatial Foundation*, 2013, http://www.qgis.org