A web application for visualizing uncertainty in numerical ensemble models

Koko Alberti, Paul Hiemstra, Kor de Jong, and Derek Karssenberg

Department of Physical Geography
Faculty of Geosciences
Utrecht University

Heidelberglaan 2.
PO Box 80115
3508TC Utrecht, The Netherlands.

E-mail: k.alberti@students.uu.nl

Universiteit Utrecht

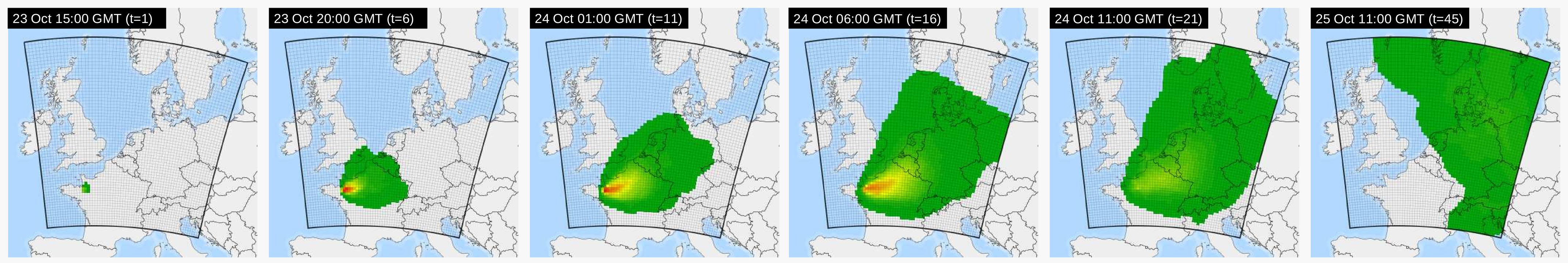
There are no best practices for visualizing attribute value and uncertainty in maps. How would you do it? In this approach:

- Spread amongst ensembles is characterized using *uncertainty metrics*: standard deviation, interquartile range, the median of absolute deviations, or a probability derived from a cumulative probability density function;
- Attribute values are upscaled before metrics are calculated. This allows decluttering of the map and makes uncertainty information available which is upscaled to the resolution of the visualization (Fig. 2);
- A web application was developed (Fig. 5 and 6) that allows users to investigate uncertainty in an atmospheric dispersion model (Fig. 1) using visualizations based on dynamic circular glyphs.

Quantifying Uncertainty

An ensemble dataset created by an atmospheric dispersion model (Hiemstra et *al*., 2011) was used as a demo dataset (Figure 1). The sample values at a particular timestep and support are considered to be a numerical estimation of a continuous random variable with an unknown distribution. Using the ordered sample values, a discretized empirical cumulative distribution function (ecdf) is constructed, and various uncertainty metrics (which are indicative of the uncertainty at a particular location) are calculated. Implemented metrics are the standard deviation, interquartile range, width of the 95% confidence interval, median of absolute deviations, legend quality, and user-defined interval probabilities derived from the ecdf.

Figure 1: The ensemble mean of the atmospheric dispersion model



Visualizing Uncertainty

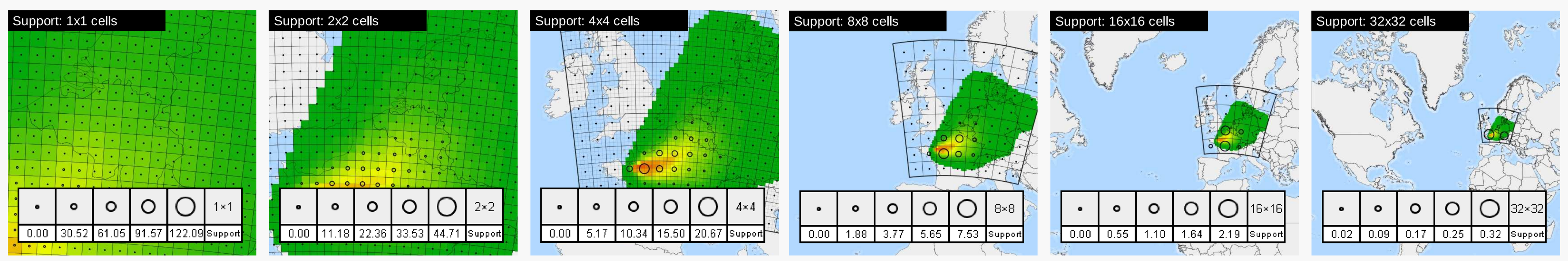
The visualization scripts import a dataset from the PCRaster-Python modeling framework (Karssenberg et *al*., 2009) and apply an upscaling algorithm (Bierkens et *al*., 2000) using Equation 1:

(Eq. 1) $$z(s_2) = \frac{1}{n}\sum_{i=1}^{n} z(s_1;i)$$

- Calculates a new value of z at support s_2 by aggregating multiple values at s_1
- Upscaling is done for each timestep, for each attribute, for each support size

The uncertainty metrics are calculated for the original and for the upscaled data, and saved as point features in a PostGIS database. In the map display these points are represented as circular glyphs and allow the uncertainty information to be adjusted to the zoom level, resulting in a visually pleasing bivariate display in which both attribute value and attribute uncertainty are embedded (Figure 4).

Figure 2: Upscaling the uncertainty information to match zoom level



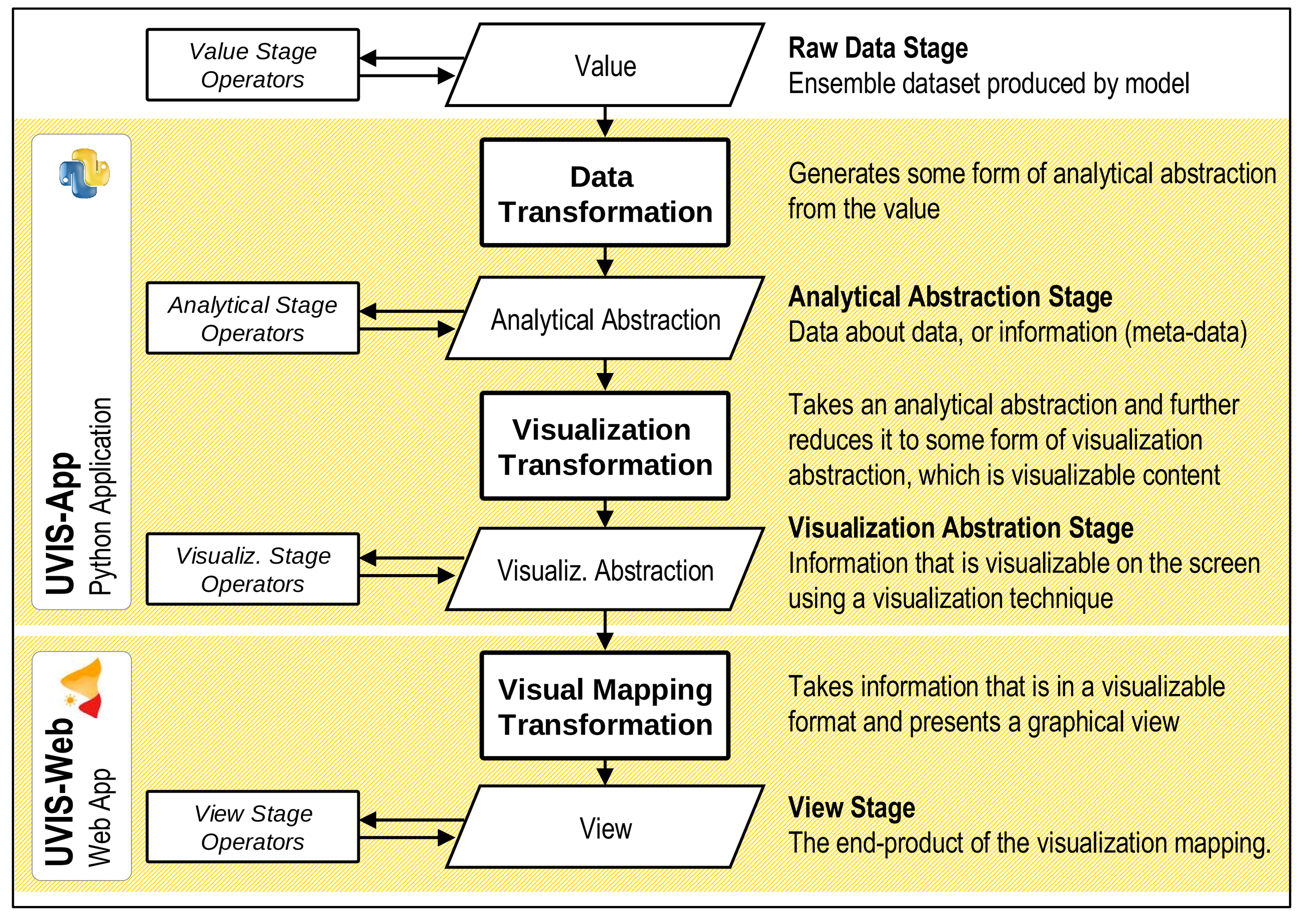
The Web Application

The data processing in the web application is in line with Chi's (2000) *data state reference model* (Figure 3). A Python application (UVIS-App) manages the analytical and visualization abstraction stages, and a web application (UVIS-Web) presents an interactive graphical view of the data produced in the higher level stages.

The web application (Figure 4) is responsible for the visual mapping transformation and for the view stage in the data state model. Functionality is present for selecting model attributes (Fig. 4A), uncertainty visualizations (Fig. 4B), model timesteps (Fig. 4E), and zooming and panning within the map (Fig 4C, 4D).

Dynamic circular glyphs are overlaid on the map display to represent uncertainty (Figure 5). The legend shows the glyph values (Fig. 5A) and a control pane appears where the user can further customize the visualization to select a different metric or an attribute value interval (Fig. 5B). Hovering over the glyphs will show its value (Fig. 5C), and clicking it opens a popup which shows more information about the region represented by the glyph (Fig. 5D).

Figure 3: The Data State Reference Model



Source: Chi, 2000

Figure 4: Navigation and Control Panes

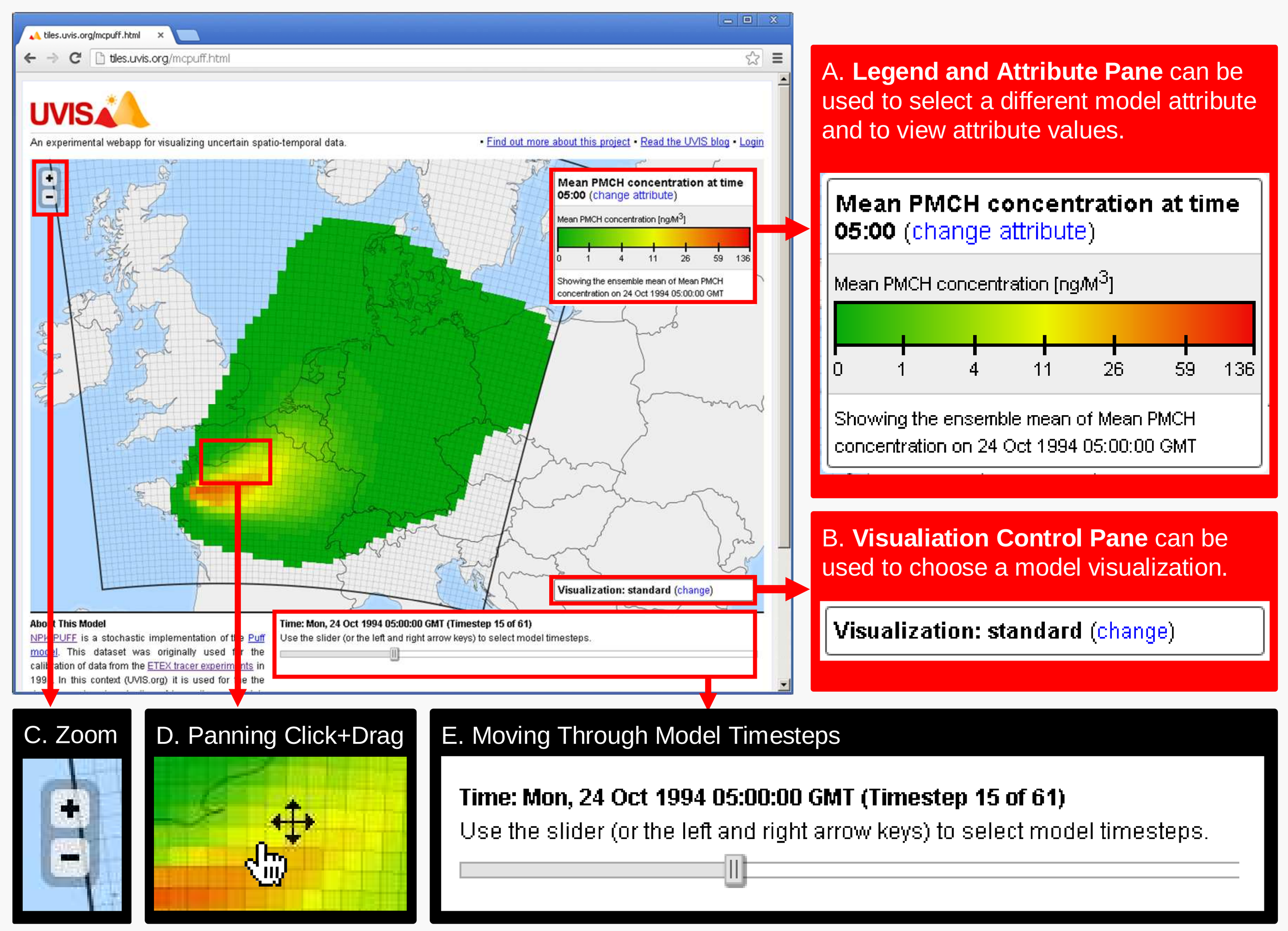
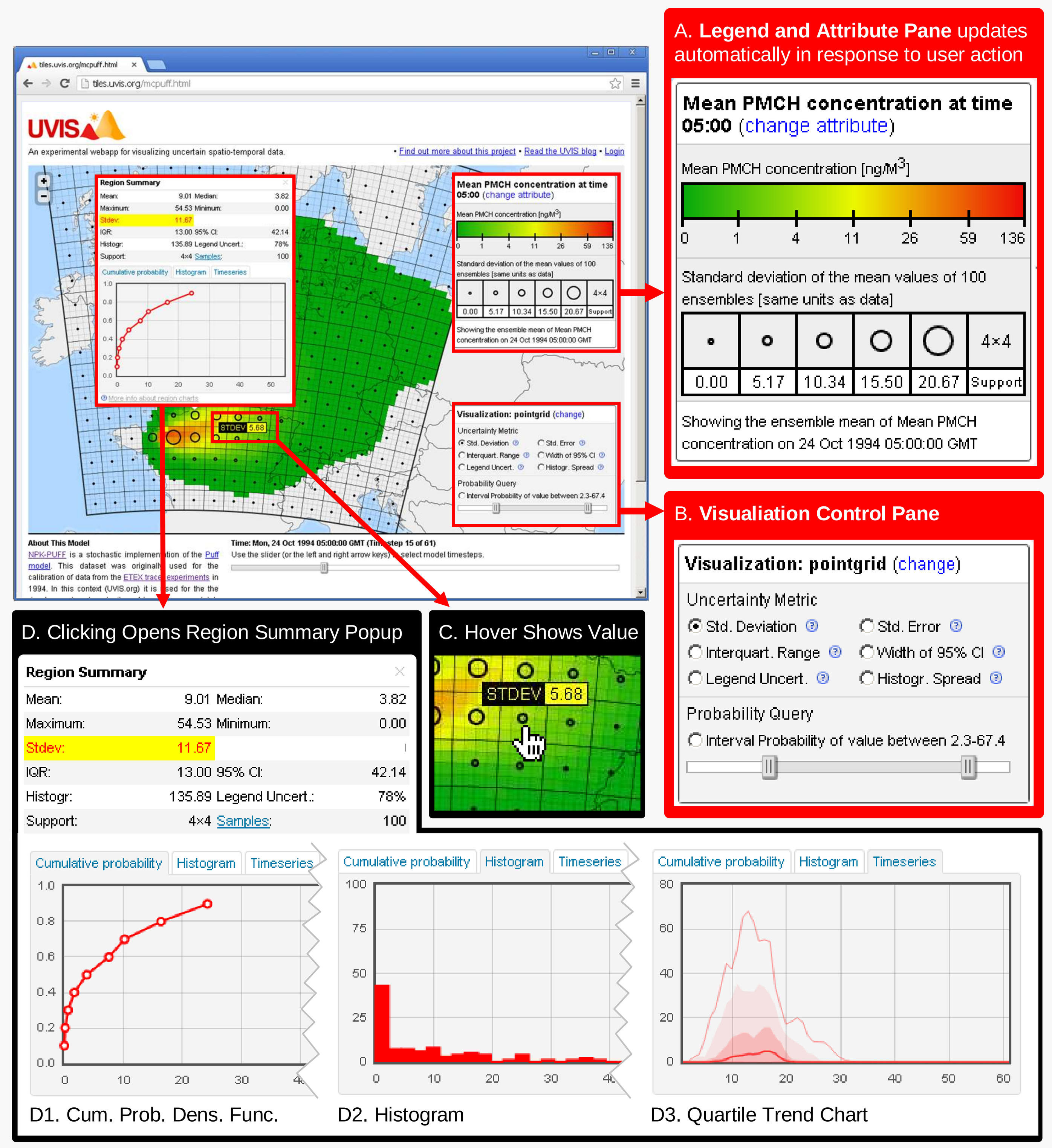


Figure 5: Uncertainty Visualization View



Discussion & Conclusions

With the aid of a user experience survey, dynamic circular glyphs, as implemented in the UVIS web application, were found to be an effective way to visualize spatial, quantative, and probabilistic aspects of uncertainty in an ensemble dataset. The upscaling of uncertainty information to the resolution of the visualization reduces the risk of a cognitive overload, and is a promising technique for creating comprehensible bivariate maps in which attribute value and uncertainty are embedded.

Other methods of quantifying uncertainty in ensemble datasets can also be explored further. Some metrics assume that the sample values are normally distributed and they are therefore sensitive to outliers. More robust methods of quantifying uncertainty, such as simulation based resampling (bootstrapping) can be implemented in future versions of the UVIS web application.

References

Bierkens, M.F.P., Finke, P.A., and Willigen, P. de (eds) (2000) Upscaling and Downscaling methods for Environmental Research. Kluwer Academic Publishers, Dordrecht.

Chi, E.H. (2000) A taxonomy of visualization techniques using the data state reference model. *IEEE Symposium on Information Visualization* (InfoVis 2000).

Hiemstra, P.H., Karssenberg, D., Dijk, A. van (2011) Assimilation of observations of radiation level into an atmospheric transport model: A case study with the particle filter and the ETEX tracer dataset. *Atmospheric Environment* 45(34): 6149-6157.

Karssenberg, D., Schmitz, O., Salamon, P., De Jong, K., Bierkens, M.F.P. (2009) A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environmental Modelling and Software* 25(2009): 489-502.