TOOLBOX

Visualizing uncertainty

Towards a better understandiung of weather forecasts

Uncertainty visualizations are increasingly used in communications to the general public. A well-known example is the weather forecast. Rather than providing an exact temperature value, weather forecasts often show the range in which the temperature will lie. But uncertainty visualizations are also deployed in graphical forecasts that are used for decision-making in many other different areas like agriculture, flood management, health care, and finance. Visualization appears to be an intuitive way to communicate uncertainty. In principle, uncertainty visualizations enable users to make better decisions by enhancing their awareness of the inherent uncertainty in the data. However, in practice many people (even experts) frequently misunderstand both the concept of uncertainty and its visualizations. We are currently investigating how the visual form and width of the graphical representation of uncertainty ranges affect how people interpret the underlying uncertainty distribution.

Alexander Toet, Susanne Tak en Jan van Erp

The need for uncertainty visualization

It is generally assumed that people appreciate uncertainty visualizations and use them to make their decisions. For weather forecasts, it has indeed been observed that most people infer uncertainty into forecasts anyway (even when it is not provided), and prefer forecasts that explicitly express uncertainty. Research has also shown that including uncertainty estimates in weather (and hydrological) forecasts increases trust and can in principle provide a better understanding of the possible outcomes and the amount of uncertainty in the given situation, thereby allowing people to make better decisions (Joslyn & LeClerc, 2013). Carefully designed visual representations can indeed successfully convey uncertainty information to both experts and non-experts (Nadav-Greenberg, Joslyn & Taing, 2008). However, the advantage of the availability of uncertainty estimates depends critically on how they are communicated (Ibrekk & Morgan, 1987). Communicating forecast uncertainty in an intuitive way so that the information is easily perceived and correctly interpreted still remains a challenge, especially when the information is intended for the general public (Tak, Toet & Van Erp, 2014; 2015). As a part of our ongoing work on the optimization of visualization techniques we are therefore investigating how different aspects of uncertainty visualizations affect the interpretations by non-experts.

Ways to visualize uncertainty

Before explaining the challenges in this field we first discus the different ways in which uncertainty can be visualized. There are roughly three different ways to visualize uncertainty: (1) by varying the graphical properties of the visualization, (2) by adding uncertainty information to the visualization, and (3) by animating the visualization.

The first approach deploys techniques to vary the graphical properties of depicted entities, such as size, blur, color saturation, texture and transparency (Figure 1). For example, blurring or degradation of

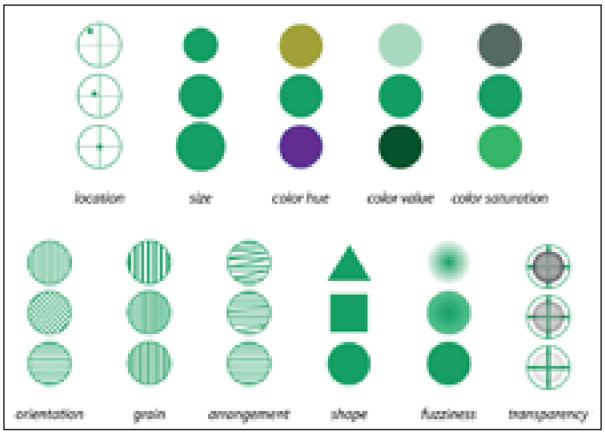

Figure 1. Different graphical properties that can be used to visualize uncertainty in data (from MacEachren, Roth, O'Brien et al., 2012).

the data has an intuitive relation with uncertainty: the harder it is to see or recognize something, the more uncertain it appears. However, blurring or degradation can also be interpreted as poor visualization quality.

The second approach is to explicitly add uncertainty information to a visualization, such as glyphs (graphical elements that can convey a number of variables through variations in their size, shape, orientation, texture, and color), geometry, labels, numbers or icons. For example, positional uncertainty can be indicated by overlaying a glyph, the size of which becomes larger the more uncertain the location is (e.g., Figure 2). Geometric techniques include contour lines and isosurfaces (i.e., surfaces of a constant value of e.g. pressure, temperature, velocity, density). Also, textual or numerical information about the magnitude of uncertainty can be added to the visualization. Adding graphical representations of uncertainty information to a data visualization may result in data obscuration (symbols may be plotted over and cover other relevant information) and user distraction. This may increase the user's response time, as it may require more cognitive effort to interpret the data.

Finally, animation of the visual data representation (e.g., turning symbols on and off at a certain rate or letting them jitter around a fixed position) can be used to indicate uncertainty. However, flickering or jittering symbols can significantly reduce the visibility of other important image details. Also, most users typically find the use of blinking and flicker annoying.

(Mis)Understanding uncertainty visualizations

As stated before, many people (even experts) frequently misunderstand visual representations of uncertainty. A notorious example of the misunderstanding of a graphical uncertainty representation is the ***deterministic*** (and not ***probabilistic*** as intended) interpretation of the well-known 'Cone of Uncertainty' graphic issued by both the US National Hurricane Center (NHC) and the media to communicate hurricane risk to the public prior to landfall (Figure 3). The main elements of this graphic are a black line representing the predicted path of the hurricane center, centered on a white 'cone' representing the potential geographic range of the track. Despite the attempt of the forecast community to make a user-friendly product this type of hurricane-warning graphics is misinterpreted by a large part of the public. Although the track line only represents the predicted (potential) track of a hurricane center, the general public typically fails to appreciate both the uncertainty about it or the statistical meaning of the wider 'cone' of uncertainty about its projected course. The black line leads many to overestimate the certainty of the projected track. The white cone is often incorrectly interpreted as the extent of the hurricane, its intensity, or the potential swath of destruction (Broad, Leiserowitz, Weinkle et al., 2007). As a result, people often fail to understand that the hurricane will potentially affect a much larger area than just the cone depicting the uncertainty about the track of the eye of the storm. People wrongly assume that only areas along the track line are at risk (over distances up to the boundary of the cone), while areas outside the cone will not be impacted. Thus, people often do not feel at risk when they do not live near the track line or outside the cone's boundaries. Another source of confusion is the fact that the white cone has been obtained by thresholding

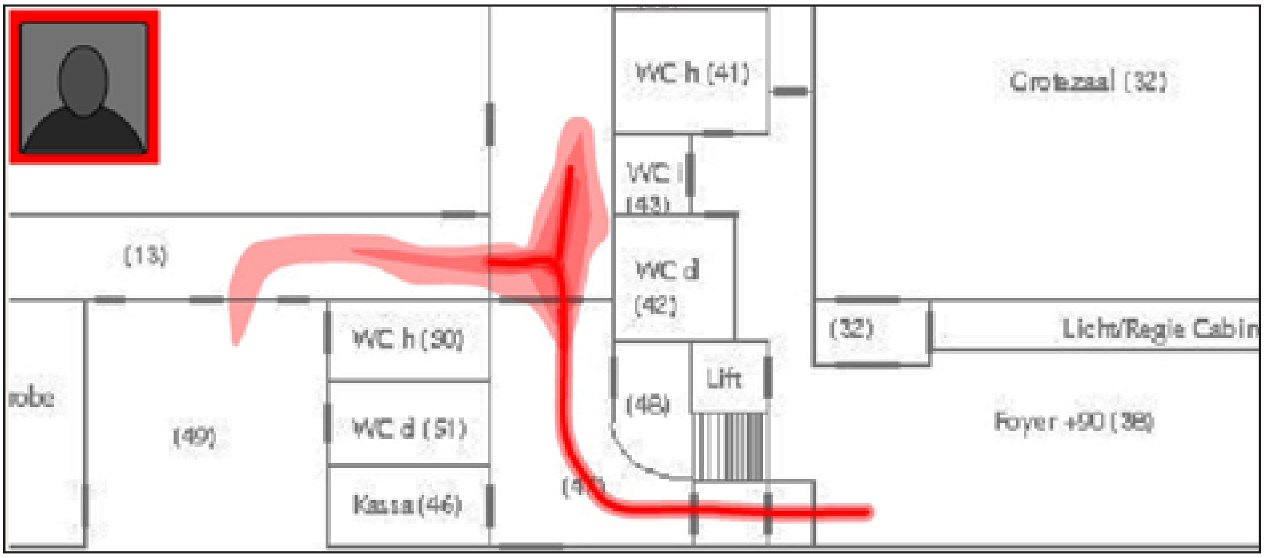

Figure 2. The red area represents the possible location of a person in a building based on previous sensor registrations. Less likely positions are plotted as more transparent and less saturated. The possibility to find the person at a given location decreases as the red area increases.

the actual overall spatial probability distribution, resulting in a loss of information (the variance in probabilities over the white area is no longer available). It has been observed that this may cause an overestimation of the probability along the centerline, even by experts (Kirlik, 2007). Sometimes an ensemble of predictions from different weather models (the data that is actually used to construct the cone) is presented to the public as a 'spaghetti plot' (Figure 4). Although this representation intuitively conveys the notion of uncertainty (there is no single 'sure' path) and thereby prevents overconfidence in a single predicted path, it has been found to confuse the public ('I have no clue which one to believe').

This example illustrates how uncertainty visualizations can easily be confusing or lead to misunderstandings. An important question is what causes the misinterpretation. In practice uncertainty visualizations are often shown without an explanation of the meaning of the uncertainty range (e.g., Figure 3 and Figure 5). In that case people need to rely on the visual form of the graphical representation to deduce the inherent uncertainty in the data. When no further information is given people sometimes assume a uniform probability distribution, both for graphical (Ibrekk & Morgan, 1987) and numerical (Rinne & Mazzocco, 2013) uncertainty representations. This means that people think that all values (including the extremes) in an uncertainty range are just as likely to occur as the mean value. It is evident that this will typically not be the case for most applications.
Especially now data science is gaining in interest in various sectors, there is a need for graphical conventions that unambiguously and intuitively convey the notion of probability as this will lead to data representations that are better and more easily comprehended by the end users. Research on visual uncertainty communication typically focuses on the development of new graphical uncertainty representations, with little attempt to evaluate their effectiveness for the end users. Also, there is still little empirical evidence to suggest that uncertainty visualization influences decision making in a robust and consistent manner (Deitrick & Edsall, 2006). It is often simply taken for granted that visual depictions of uncertainty will be useful for decision making. Until now only few studies investigated to what extent the users' interpretation of uncertainty visualizations matches the actual uncertainty distribution of the underlying data. As a result we still do not have a comprehensive understanding of the parameters that influence *successful* uncertainty visualization.

What model people exactly adopt to interpret different visualizations of uncertainty is not completely clear. Studies suggest that different visualizations (such as glyphs, color or grayscales) may result in different perceived models and may therefore induce a discrepancy between the model intended by the designer and the

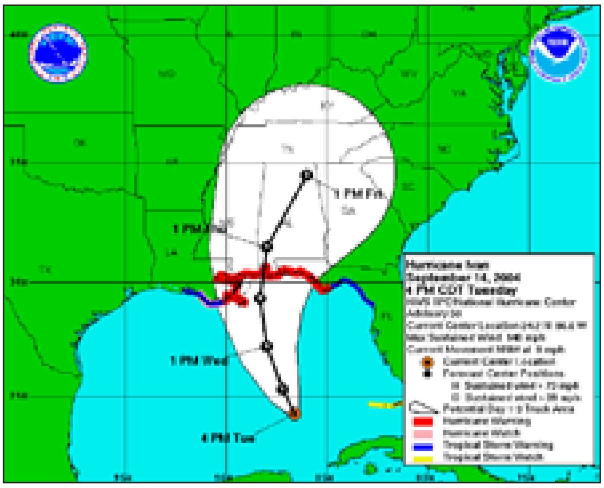

Figure 3. The (in)famous 'Cone of Uncertainty' graphic representing the area in which the path of the hurricane may lie.

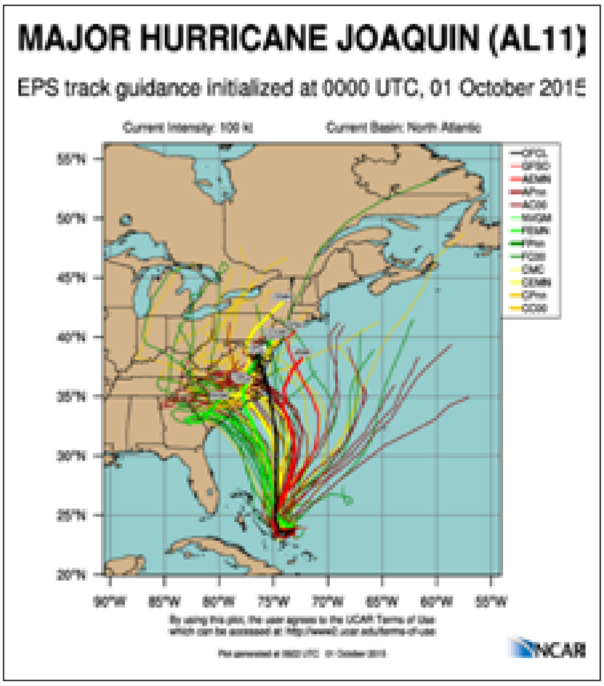

Figure 4. Spaghetti plot showing an ensemble of predictions from different hurricane forecast models.

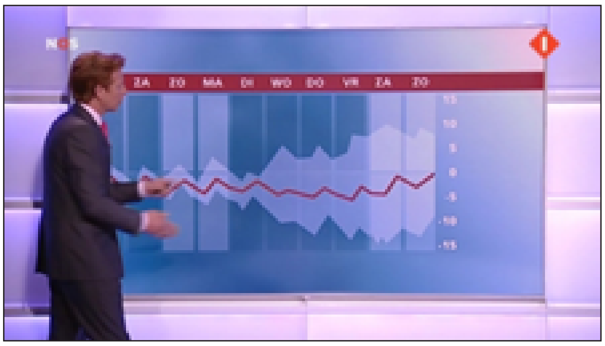

Figure 5. A visual representation of temperature forecast uncertainty in the news. Visualizations for the general public (e.g., on TV or in newspapers) often provide no explicit information on the uncertainty that is shown and the viewer needs to adopt a model to interpret the underlying distribution.

model that is actually applied by the viewer. Obviously, such discrepancies lead to incorrect interpretations, which may result in wrong decisions.
In a study on the perception of *point probability* (e.g., by asking to judge the probability that a certain predicted value will occur) in graphs with visual uncertainty bands (Tak, Toet & Van Erp, 2014; see Figure 6) it was found that observers (when given no further explanation of the 'mathematical' meaning of the uncertainty band) intuitively assume that the mean of the band is the most likely value to occur and that values further from the mean are less likely to occur. We also observed that a user's numeracy (mathematical skills) affects this intuitive model. We therefore performed an additional study to investigate the effects of type and overall width of ensemble prediction visualizations (again presented without any additional information) on range probability estimates (e.g., by asking to judge the probability that the temperature will exceed a given value or be in the higher ranges: Tak, Toet & Van Erp, 2015). More specifically, we investigated (1) the nature of the model that people assume for visual uncertainty ranges when given no additional information, (2) whether the form of visual uncertainty ranges affects this assumed model, and (3) to what extent the assumed model depends on a participants' numeracy.

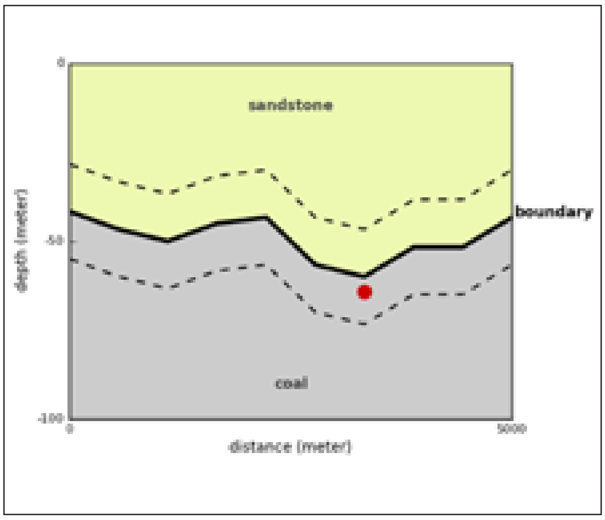

Figure 6. Stimulus from an experiment in which participants were asked to judge the probability that a certain material (coals or sandstone) was present at the location indicated by the red dot. The dashed lines represent the borders of the uncertainty region (i.e., the transition area between both materials).

Effects of graphical form and width on perceived range probability

Visualizations like those depicted in Figure 3 and Figure 5 do not provide explicit information about, for example, the (range) probability of exceeding a temperature of 5 degrees on a specific day. Taking into account the results of the aforementioned studies, we performed an experiment to investigate if users assume a consistent model to translate visualizations like the ones used in weather forecasts into probabilities. We asked participants to estimate a *range probability* by judging the probability that the afternoon temperature on a given day would exceed the temperature indicated by a red dot in a given uncertainty visualization (Figure 7).

Figure 8 shows the seven visualization forms that were investigated in this study. All seven different visualization types represented seven data points connected by a continuous black line (the center line). The data points represented the predicted temperature values for seven days ahead. The visualizations differed in the graphical representation of the uncertainty range, which was always symmetrical around the center line. To prevent stimulus familiarization the shape of the center line was varied slightly across stimuli by randomly distributing the seven predicted temperature values over the days of the week while keeping the width of the uncertainty interval fixed at each *x*-position (i.e., for any given day of the week, the *y*-value or uncertainty width was fixed, but the corresponding temperature value was randomly selected from the set of seven temperatures). This procedure yielded temperature curves with slightly varying shapes but similar and monotonously increasing uncertainty ranges.

Results

It appears that, in the absence of information about the uncertainty range, people apply a perceived model of the uncertainty distribution that closely resembles a Bell-shaped distribution, as was also found in our study about interpretation of point probability (Tak, Toet & Van Erp, 2014). In addition, we found that people typically have a bias for higher temperatures: they consistently estimate the probability of higher temperatures to be larger. The perceived probability of 'extreme values' (i.e., values far outside the uncertainty range) is affected by the visualization type, with denser fills leading to higher perceived probability of values within that area. Perceived probability also depends on the width of the uncertainty range: people judge the probability of values with the same relative distance to the centerline different for wide and for narrow uncertainty ranges (Figure 9). This means that observers take not only the relative but also the absolute distance to the center line into account and assume a model that does not simply scale with the width of the uncertainty range. Finally, the assumed model of the uncertainty distribution depends on a participant's numeracy: people with low numeracy adopt a 'flatter' (all values judged more or less equally probable) interpretation than those with high numeracy. In addition, those with low numeracy have a more pronounced bias than those with high numeracy (i.e., they more consistently judge higher values to be more likely). In practice this means that people with high numeracy have

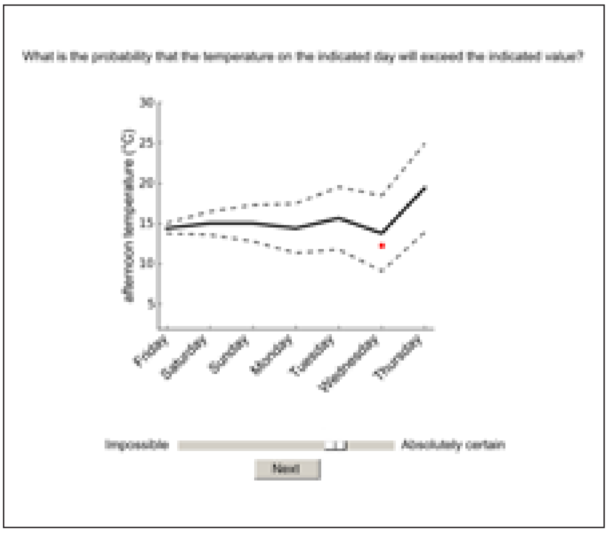

Figure 7. Screen shot of an uncertainty rating experiment.

a more realistic interpretation of uncertainty visualizations when they are presented without any further explanation. The cause of this effect (which has also been observed in several previous studies) still remains unclear. Hence, it remains a challenge to design uncertainty visualizations that will be correctly interpreted by all users, independent of their numeracy.

Conclusion

We find that the width and density of graphical representations of uncertainty ranges affect range probability estimates, and differently so for estimates relative to reference values in respectively the upper or lower part of an uncertainty range. We suspect that the effects found here are likely to hold for a wider range of visualizations. This may have practical implications for graphical forecasts used in different areas like agriculture, flood management, health care, finance, and many other decision-making contexts where incorrect inferences from range estimates may lead to suboptimal decisions. Our results suggest for instance that blurred or dashed (versus sharp) borders (like the ones shown in respectively Figure 10b and d) for the hurricane 'Cone of Uncertainty' may give the public a more realistic impression of the uncertainty in the hurricane's path, since we found that these type of borders lead to people to judge values away from the center line to be more likely (i.e., they lead to a 'flatter' interpretation of the underlying uncertainty distribution).

Our results also imply that wide uncertainty intervals are in most cases probably not be the best choice for uncertainty visualizations, since people already interpret the underlying uncertainty distributions as wider than they actually are. In terms of the different visualization types it is unclear what the 'best' choice is. However, the effects of the density of the fill on the interpretation of extreme values should be taken into account when

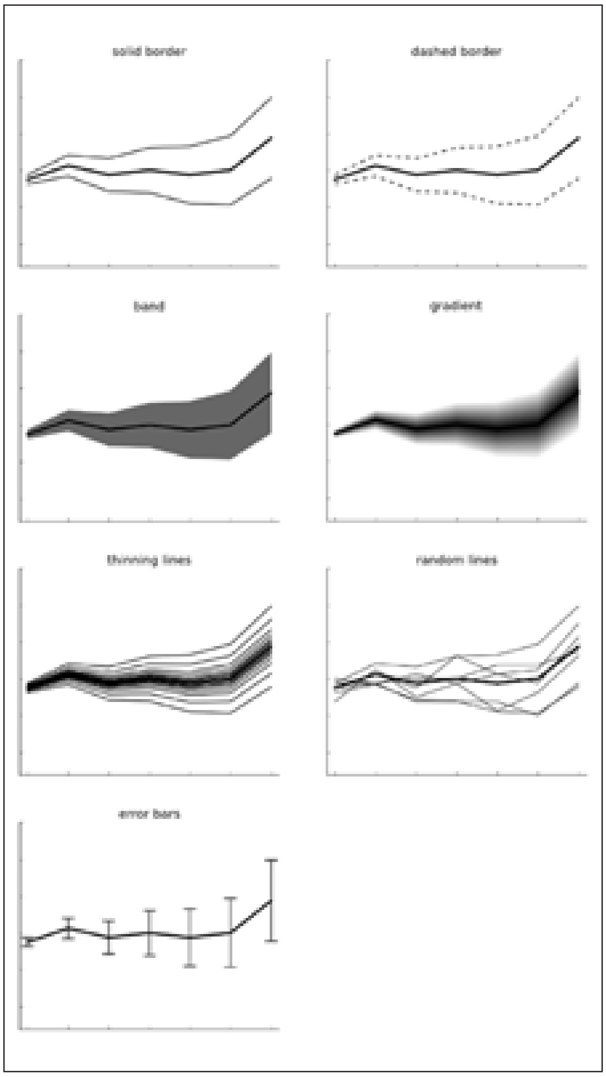

Figure 8. The seven graphical uncertainty visualization types used in our experiment to represent temperature predictions.

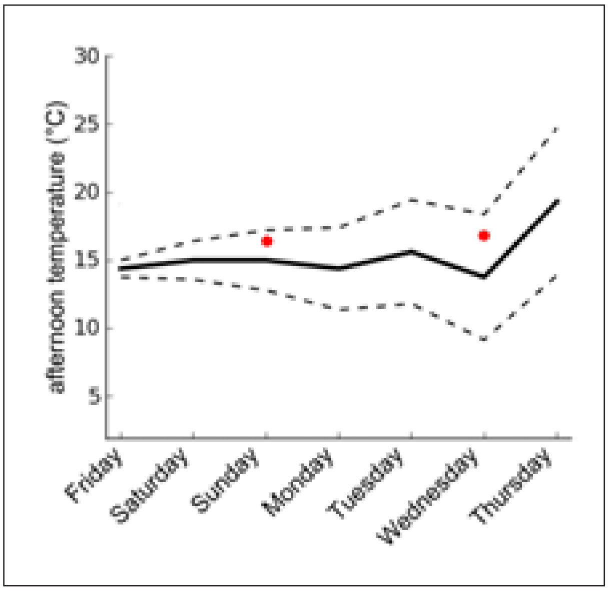

Figure 9. The temperatures indicated by the red dots on Sunday and Wednesday are at equal relative distances from the center-line (the predicted mean temperature curve) but are judged to have different probabilities.

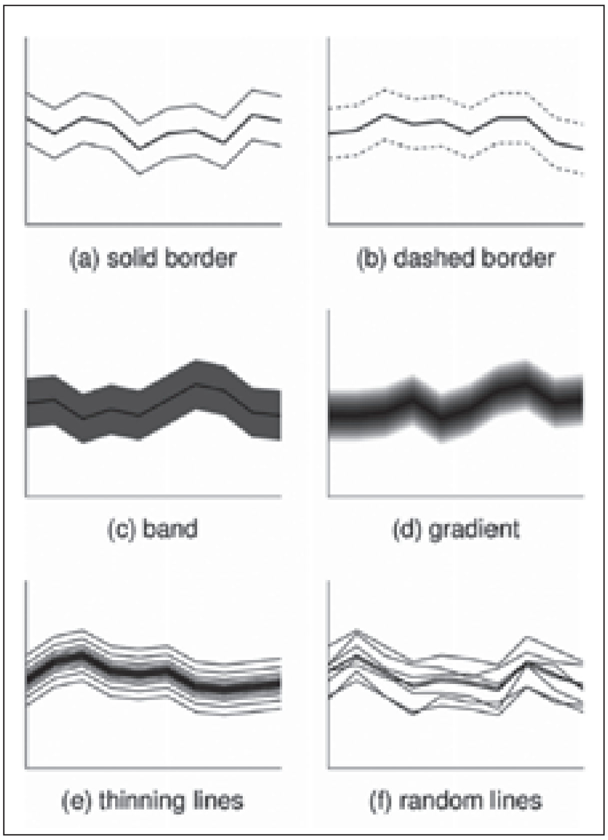

Figure 10. Graphs with different visualizations of uncertainty bands representing for instance the boundary between two different earth layers (e.g., sandstone and coal).

choosing a particular visualization, especially when these unlikely events are of the type 'low probability, high impact'. Finally, the different results on the narrow and the wide uncertainty width show that perceived uncertainty does not necessarily map linearly to visual features, and that testing of the interpretation of uncertainty visualizations prior to dissemination is important, since the intentions of the designer do not necessarily match the interpretation of the viewer.

In our previous research we used uncertainty graphs without actual underlying data. Using these graphs we established relations between the shape and width of the perceived uncertainty and the characteristics of the graphical uncertainty visualization. However, we also need to know how well perceived uncertainty corresponds to the actual uncertainty in data. We therefore started to experiment with uncertainty visualizations based on actual data for which the uncertainty is known, so that we can compare the perceived uncertainty with the actual uncertainty.

Further research can provide knowledge on the nature of the effects found here, which may in turn lead to more effective presentations of uncertainty ranges to diverse populations in a variety of judgment and decision-making contexts. This may be of crucial value in high-risk environments where people have to decide quickly, or when decisions have a high impact.

References

Broad, K., Leiserowitz, A., Weinkle, J., & Steketee, M. (2007). Misinterpretations of the 'Cone of Uncertainty' in Florida during the 2004 hurricane season. *Bulletin of the American Meteorological Society, 88*(5), 651-667.

Deitrick, S., & Edsall, R. (2006). The influence of uncertainty visualization on decision making: An empirical evaluation. In: A. Riedl, W. Kainz & G.A. Elmes (Eds.), *Progress in Spatial Data Handling*, (pp. 719-738). Berlin Heidelberg: Springer.

Gigerenzer, G., Hertwig, R., Van Den Broek, E., Fasolo, B. & Katsikopoulos, K.V. (2005). 'A 30% chance of rain tomorrow': How does the public understand probabilistic weather forecasts? *Risk Analysis, 25*(3), 623-629.

Ibrekk, H. & Morgan, M.G. (1987). Graphical communication of uncertain quantities to nontechnical people. *Risk Analysis, 7*(4), 519-529.

Joslyn, S., & LeClerc, J. (2013). Decisions with uncertainty: The glass half full. *Current Directions in Psychological Science, 22*(4), 308-315.

Joslyn, S., Nemec, L., & Savelli, S. (2013). The benefits and challenges of predictive interval forecasts and verification graphics for end users. Weather, Climate, and Society, 5(2), 133-147.

Kirlik, A. (2007). Lessons Learned from the Design of the Decision Support System Used in the Hurricane Katrina Evacuation Decision. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 51*(4), 253-257.

MacEachren, A.M., Roth, R.E., O'Brien, J., Li, B., Swingley, D., & Gahegan, M. (2012). Visual semiotics & uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics, 18*(12), 2496-2505.

Nadav-Greenberg, L., Joslyn, S.L. & Taing, M.U. (2008). The effect of uncertainty visualizations on decision making in weather forecasting. *Journal of Cognitive Engineering and Decision Making, 2*(1), 24-47.

Rinne, L.F., & Mazzocco, M.M.M. (2013). Inferring uncertainty from interval estimates: Effects of alpha level and numeracy. *Judgment and Decision Making, 8*(3), 330-344.

Tak, S., Toet, A., & Van Erp, J.B.F. (2014). The perception of visual uncertainty representation by non-experts. *IEEE Transactions on Visualization and Computer Graphics, 20*(6), 935-943.

Tak, S., Toet, A. & Van Erp, J. (2015). Public understanding of visual representations of uncertainty in temperature forecasts. *Journal of Cognitive Engineering and Decision Making, 9*(3), 241-262.

About the authors

A. Toet, PhD.
Senior Scientist
Department of Perceptual and Cognitive Systems
TNO, Soesterberg
lex.toet@tno.nl

S. Tak, PhD
Education Devoloper
Freudenthal Institute for Science and Mathematics Education
Faculty ov Science, Utrecht University, Utrecht

Prof. J.B.F. van Erp, PhD
Principal Scientist
Department of Perceptual and Cognitive Systems
TNO, Soesterberg
Professor of tangible user interaction
University of Twente, Enschede