Institute for Policy Research
Northwestern University
Working Paper Series

WP-14-14

Bayes and BOGSAT: Issues in When and How to Revise Earthquake Hazard Maps

Seth Stein
'William Deering Professor, Department of Earth & Planetary Sciences
Faculty Associate, Institute for Policy Research

Bruce D. Spencer
Professor, Department of Statistics
Faculty Fellow, Institute for Policy Research

Edward Brooks
Graduate Student
Department of Earth & Planetary Sciences
Northwestern University

Version: October 2014

DRAFT
Please do not quote or distribute without permission.

2040 Sheridan Rd. ◆ Evanston, IL 60208-4100 ◆ Tel: 847-491-3395 Fax: 847-491-9916
www.northwestern.edu/ipr, ◆ ipr@northwestern.edu

Abstract

Earthquake hazard maps play an important role in the formulation of building codes throughout the U.S. and much of the world. An important question is what to do after a major earthquake yielding shaking larger than anticipated in a hazard map. Common practice is to revise the map to show increased hazard in the heavily-shaken area. However, a new map that better describes the past does not necessarily better predict the future. The researchers examine the logic underlying map revision and argue that Bayesian modeling can play a useful role in deciding whether and how to revise the maps to improve forecasting the future.

Recent large earthquakes that gave rise to ground shaking larger than anticipated in an earthquake hazard map have generated interest in how to improve hazard mapping. Issues under discussion include how to evaluate the performance of maps, how to assess their uncertainties, how to make better maps, and how to best use maps given their strengths and limitations.

In this context, an important question is what to do after a major earthquake yielding shaking larger than anticipated in a hazard map. Hazard mappers have two choices. One is to regard the high shaking as a low-probability event allowed by a probabilistic seismic hazard map, which used estimates of the probability of future earthquakes and the resulting shaking to predict the maximum shaking expected with a certain probability over a given time (*Hanks et al., 2012; Frankel, 2013*). The usual choice, however, is instead to accept that high shaking was not simply a low-probability event consistent with the existing map, and revise the map to show increased hazard in the heavily-shaken area (Figure 1).

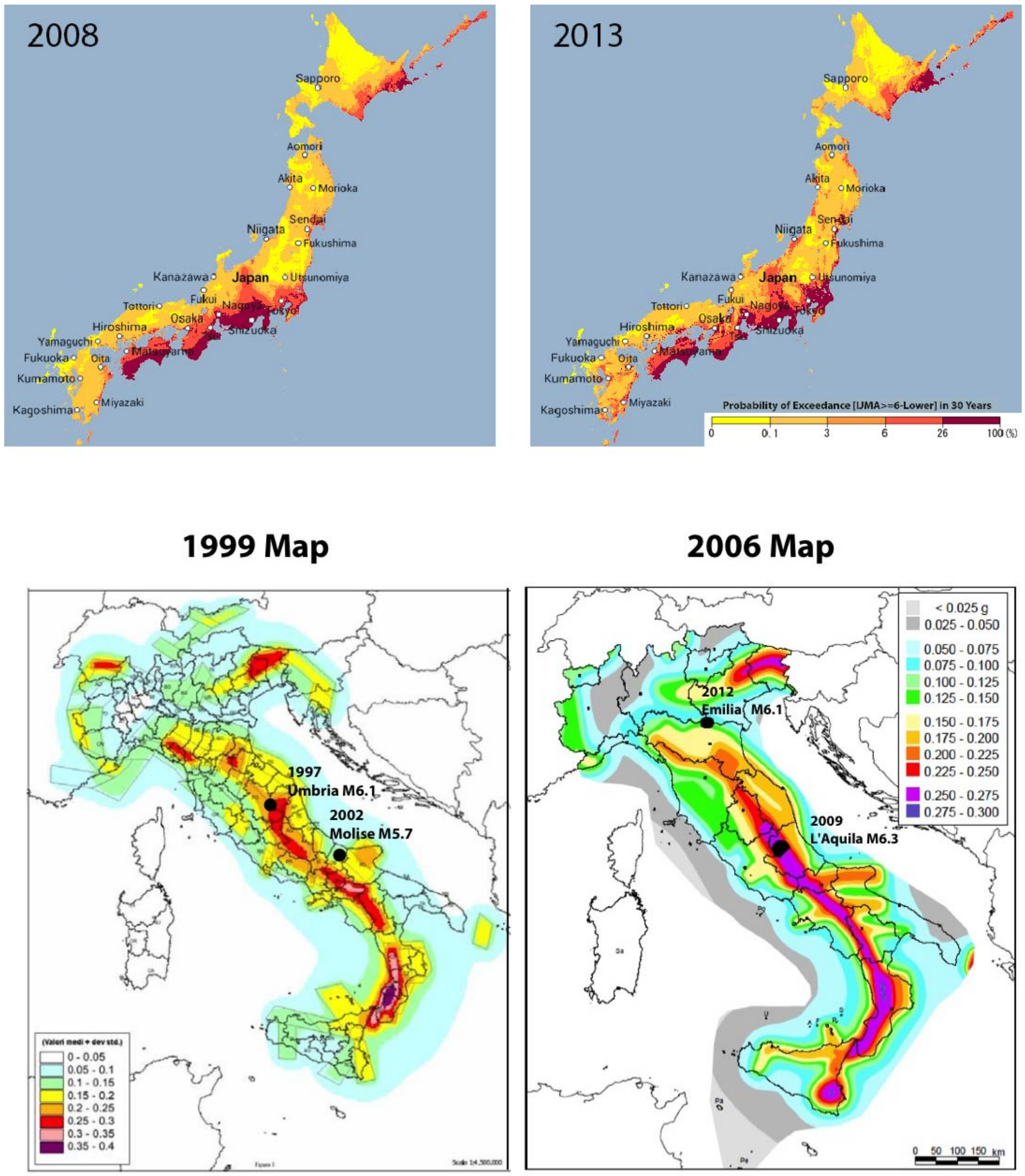

Figure 1: Top: Comparison of Japanese national seismic hazard maps before and after the 2011 Tohoku earthquake. The predicted hazard has been increased both along the east coast, where the 2011 earthquake occurred, and on the west coast. (http://www.j-shis.bosai.go.jp/map/?lang=en) Bottom: Comparison of successive Italian hazard maps (Stein et al., 2013). The 1999 map was updated to reflect the 2002 Molise earthquake and the 2006 map will likely be updated after the 2012 Emilia earthquake.

Whether and how much to revise a map is a complicated issue, because a new map that better describes the past may or may not better predict the future. For example, increasing the predicted hazard after an earthquake on a fault will make better predictions if the average recurrence time is short compared to the map’s time window, but will overpredict future shaking if the average recurrence time is much longer than the map’s time window.

Bayes’ Rule

To get insight into the question of whether and how to remake a hazard map, consider a simple analogy. Imagine tossing a coin, which comes up heads four times in a row. How likely do you think it is to come up heads on the next toss? You started off assuming that the coin is fair - equally likely to land heads or tails. Should you change that assumption after four heads?

Either choice runs a risk. If the coin is severely biased, staying with the assumption that it is fair will continue to yield poor predictions. However, if the coin is fair and the four heads were just a low-probability event, changing to the assumption that the coin is biased - heads more likely - does a better job of describing what happened in the past, but will make your prediction worse.

Your choice would likely depend on how confident you were in your initial assumption, prior to the tosses, that the coin was fair. If you were confident that the coin was fair, you would not change your model, and continue to assume that a head or tail is equally likely. However, if you were given the coin at a magic show, your confidence that the coin is fair would be lower and you would be more apt to change your model to one predicting that a head is more likely than a tail.

A statistical approach that combines preconceptions with observations to decide how to update forecasts as additional information becomes available uses Bayes' Rule (*Sivia, 2006; Rice, 2007*). In this formulation

$$\begin{array}{l} \text{revised or} \\ \text{posterior} \\ \text{probability} \end{array} \propto \begin{array}{l} \text{likelihood of} \\ \text{observations given} \\ \text{the prior model} \end{array} \times \begin{array}{l} \text{prior} \\ \text{probability,} \end{array}$$

where we omit the normalization factor. This formulation starts by assuming an initial or *prior* probability model based on information available prior to the additional observations, calculating how likely the observations were given that model, and using the product as the revised or *posterior* probability model to account for the additional observations.

For example, if we describe a coin's probability of landing heads by a parameter from 0 (always tails) to 1 (always heads), we can represent our beliefs about the parameter by a probability distribution. If, prior to observing the four heads, we are confident the coin is fair or nearly fair, our prior probability distribution is tightly clustered around 0.5 (although to allow for surprises, the distribution assigns positive probability throughout the interval). If we think the coin may be biased, our prior distribution would have a much larger spread and might be skewed towards 0 or 1.

After some tosses, the revised model depends on both the observations and the prior model. If we had high confidence that the coin was fair, a few low-probability observations would not change it much. However, if we had little confidence in the prior model, these low-probability observations change it a lot.

An important feature of the Bayesian approach is that in it “probability” represents our belief in how a system works based on the information we have. In this formulation, probability is subjective, in that given the little information we know about the coin, we have no way to know what the actual probability of a head on the next toss is. Once we have chosen a model, we can calculate precisely the probability of observing a head on the next toss. However, because this calculated probability assumes that the model is true, it also is subjective and subject to revision after the next toss.

This view is more complicated than the "frequentist" view of probability in which the probability of an event is the relative frequency in which it occurs in an indefinitely large number of trials. If we flipped the coin a thousand times independently under standard conditions, the fraction of heads would be a good estimate of the probability of a head on the next toss under the same conditions. However, because we only have the results of four tosses, we factor in our preconceptions rather than automatically assume that the four heads prove that the probability of a head in the next toss is near 1.

Although the Bayesian approach requires assuming a prior probability distribution, the effect of this assumption is reduced as more data become available, provided that the prior distribution does not assign zero probability to sets of parameter values that include the true state of nature. Once the coin has been flipped many times, there are enough observations that the posterior distribution does not depend on the assumed prior distribution as long as it assigns positive values throughout the interval from 0 to 1.

Earthquake probabilities

Seismologists' approach to estimating earthquake hazards is often in the spirit of Bayes' Rule, in that it involves assuming probability models based on limited data and then using new data to improve them (*Parvez, 2007; Marzocchi and Jordan, 2014*). To see this, consider a simple example in which we assume that the probability of a large earthquake on a fault is described by a Poisson process with parameter $\lambda = 1/T$, corresponding to an average return time of T years. Following *Cornell (1972)*, *Campbell (1982)*, and later authors, we represent our uncertainty about the value of λ by using a gamma probability distribution with mean μ and standard deviation σ as our prior probability distribution. If an earthquake occurs only 1 year after the past one, then by Bayes' Rule the prior distribution is updated to the posterior distribution, and the prior mean μ updates to the posterior mean $\mu' = \mu(1 + \sigma^2 / \mu^2) / (1 + \sigma^2 / \mu)$. (*Rice, 2007, 288*).

Consider the prior mean μ to be specified as 0.02, corresponding to T = 50 years. If we are highly confident about λ when the forecast is made, the standard deviation σ is small, so the posterior mean μ' and the prior mean μ are close. We treat the new observation that did not fit the model well as a rare event that does not change our preconception much. But, if we were uncertain that the actual value of λ would be near the prior mean μ, i.e., σ is large, then the new observation changes our view, making the posterior mean very different (larger) than the prior mean.

Figure 2 shows how the updated forecast, described by the posterior mean, increasingly differs from the initial forecast (prior mean) when the uncertainty in the prior distribution, as measured by the standard deviation σ, is larger. In other words, large changes from updating can be appropriate when the original forecast is very uncertain. The less confidence we have in the prior model, the more a new datum can change it.

This example is useful because experience shows that inferring earthquake probabilities, which are crucial inputs for hazard mapping, is very difficult given the poorly-understood faulting process and the intrinsic limitations of the earthquake record (*Sieh et al., 1989; Savage, 1992, 1994; Parsons, 2008*). It is still unclear whether to assume the recurrence of large earthquakes is described by a Poisson process that has no "memory," so the probability of an earthquake is constant with time, or by time-dependent models based on an earthquake cycle in which the probability is small shortly after the past one, and then increases with time. A numerical simulation shows that these two are difficult to distinguish even in a simple case (*Stein and Stein, 2013*). Moreover, using a time-dependent model requires choosing many parameters, because both the model used for recurrence times and its parameters are needed, but are poorly constrained by the available earthquake history.

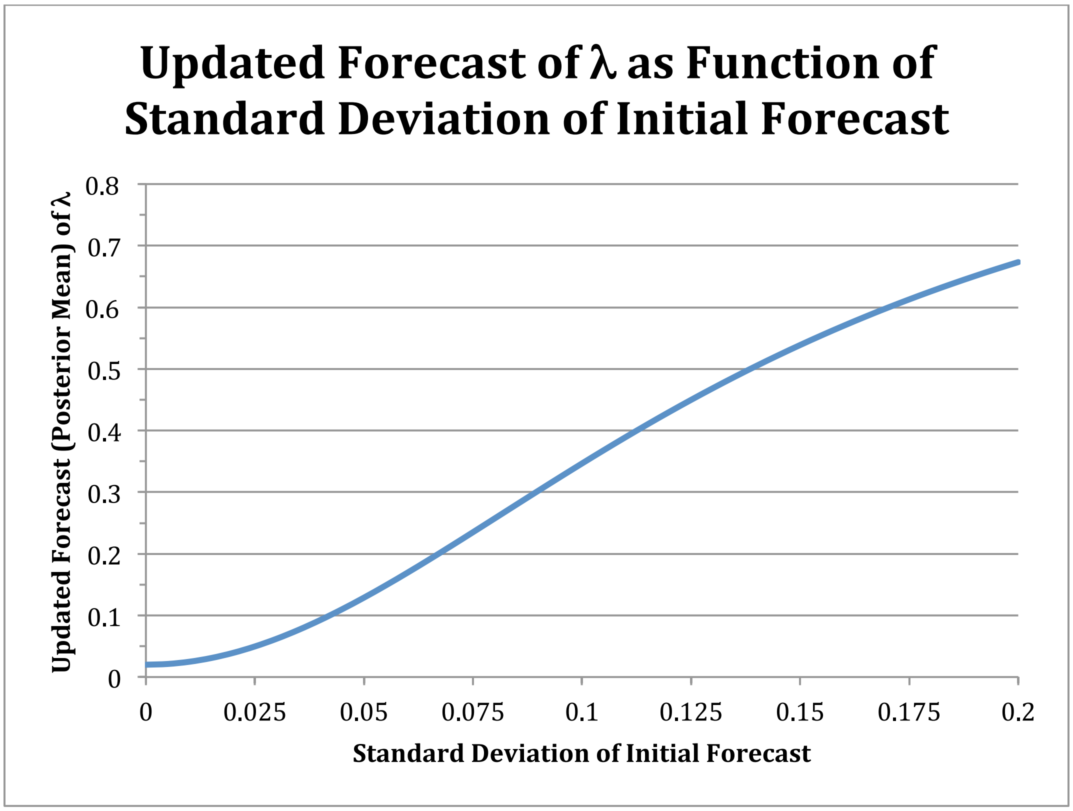

Figure 2. Sensitivity of updated forecast of λ, initially assumed to equal 0.02, to assumed prior uncertainty in the initial forecast. The less confidence we have in the initial forecast, the more the new datum changes it.

Hence from a statistical view, *Stark and Freedman (2003)* conclude that estimates of earthquake probabilities are "shaky." In their view, "the interpretation that probability is a property of a model and has meaning for the world only by analogy seems the most appropriate.... The problem in earthquake forecasts is that the models, unlike the models for coin-tossing, have not been tested against relevant data. Indeed, the models cannot be tested on a human time scale, so there is little reason to believe the probability estimate." *Savage (1991)* similarly concluded that earthquake probability estimates for California are "virtually meaningless" and that it would be meaningful only to quote broad ranges, such as low (<10%), intermediate (10–90%), or high (>90%). In other words, it seems reasonable to say that earthquakes of a given size are more likely on some faults than others, but quantifying this involves large uncertainty.

Hazard Maps

The earthquake probability example illustrates the challenge for earthquake hazard maps, namely choosing hundreds or thousands of parameters to predict the answers to four key questions over periods of 500-2500 years: Where will large earthquakes occur? When will they occur? How large will they be? How strong will their shaking be?

Some of the parameters required are reasonably well known, some are somewhat known, some are essentially unknown, and some may be unknowable (e.g., *Stein et al., 2012*). As a result, mappers combine a variety of data and models with their sense of how the earth works. In *Stark and Freedman's (2003)* words, the process involves "geological mapping, geodetic mapping, viscoelastic loading calculations, paleoseismic observations, extrapolating rules of thumb across geography and magnitude, simulation, and many appeals to expert opinion. Philosophical difficulties aside, the numerical probability values seem rather arbitrary."

Such models, which involve subjective assessments and choices among many poorly known or unknown parameters, are sometimes termed BOGSATs, from "Bunch Of Guys Sitting Around a Table" (*Kurowicka and Cooke, 2006*). Not surprisingly, sometimes the resulting maps do well at predicting what occurs in future earthquakes, and sometimes they do poorly. However, at this point, there is no way to avoid BOGSAT. Although some parameters could be better estimated, and knowledge of some will improve as new data and models become available, major uncertainties seem likely to remain (*Stein and Friedrich, 2014*).

Nonetheless, despite their large uncertainties, hazard maps have some useful information. From a mitigation policy standpoint, inaccurate hazard (and loss) estimates are still useful unless they involve gross misestimates (*Stein and Stein, 2013*). For example, a highway department would likely use its limited funds to preferentially strengthen bridges in predicted high-hazard areas.

Hence in our view, the most practical approach is to consider the BOGSAT process from a Bayesian perspective, in that the predicted hazard reflects the mapmakers' view of the world based on their assessment of diverse data and models, and that when and how maps are revised once new data become available depends significantly on the mapmakers' preconceptions. Given that this is the case, how can it be done better?

At a fundamental level, we need to learn more about when and how revising maps makes them better or worse predictors of the future. In some cases the revisions should make the map work better, and in others, worse. In particular, raising the predicted hazard where a large earthquake recently occurred may improve the match of the model to past data, but degrade its fit to future events.

On a working level, we suggest several changes to current procedures.

First, maps should specify what they seek to predict and how their future performance should be measured. Various metrics can be used, so users can know what the mappers' goals are and be able at later time to assess how well the map met them. For example, how well did the map perform compared to one that assumed a much smoother variation in the predicted hazard (*Geller, 2011*)?

Second, documentation of hazard maps should explicitly list the parameters used and estimates of their uncertainties. Often much of this information is available in the documentation for maps (e.g., *Field et al., 2008*). In particular, the weights assigned to various branches in a logic tree are a discretized version of the prior probability density function assumed for that parameter. It would be useful to have model assumptions listed in a consistent form to make changes between successive maps easier to identify and discuss.

Third, estimates of the expected uncertainty in the resulting map should be presented and explained. Such uncertainties are typically presented in other complicated forecasts that have significant economic and policy implications (Figure 3). Although the forecasts sometimes miss their targets (Figure 4), uncertainty estimates are still useful. These would illustrate the range of predicted hazard at points in the mapped area. This would involve generating hazard curves and thus maps for different parameter values within their assumed uncertainties. The resulting large number of estimates could be presented in various ways, including maps of uncertainty or tabulations at sites, that would give users a sense of the uncertainties the mapmakers ascribe to predictions in different areas of the map. These uncertainties could then be factored in the policy making process, as is done for most other forecasts.

Fourth, changes in the parameters between successive generations of maps should be explicitly listed and explained. Some changes will likely reflect what happened in earthquakes after the map was made (Figure 1), whereas others will reflect data not used in the earlier map, because they were not recognized, not appreciated, or not available. The criteria used to decide when parameters were changed should be defined *(Ramsey 1926, 180)*.

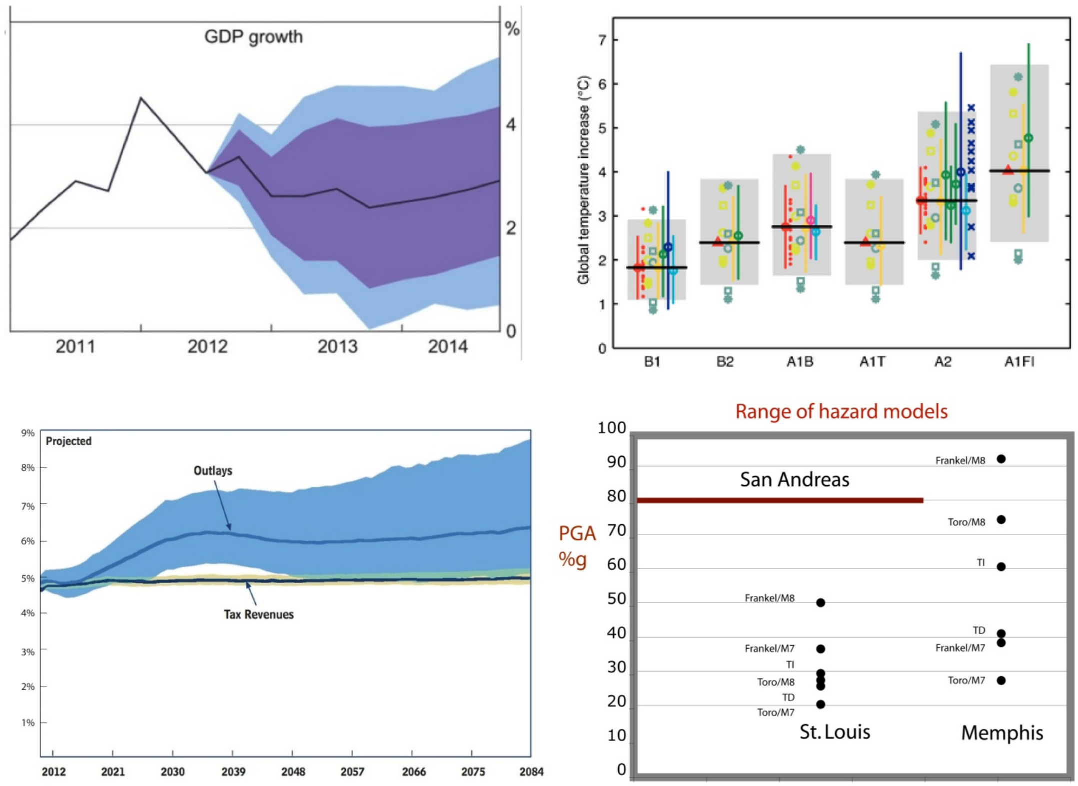

Figure 3: Presenting forecast uncertainties. a) Forecast of Australian gross domestic product growth. Uncertainty bounds are 70% and 90% (*Reserve Bank of Australia*, 2013). b) Forecast of US Social Security expenditure as percentage of Gross Domestic Product (*Congressional Budget Office*, 2010) c) Comparison of the rise in global temperature by the year 2099 predicted by various climate models. For various scenarios of carbon emissions, e.g., B1, the vertical band shows the predicted warming (*IPCC*, 2007). d) Comparison of earthquake hazard, described as peak ground acceleration (PGA) as a percentage of the acceleration of gravity expected with 2% risk in 50 years, predicted by various assumptions for two sites in the central U.S. (*Stein et al.*, 2012).

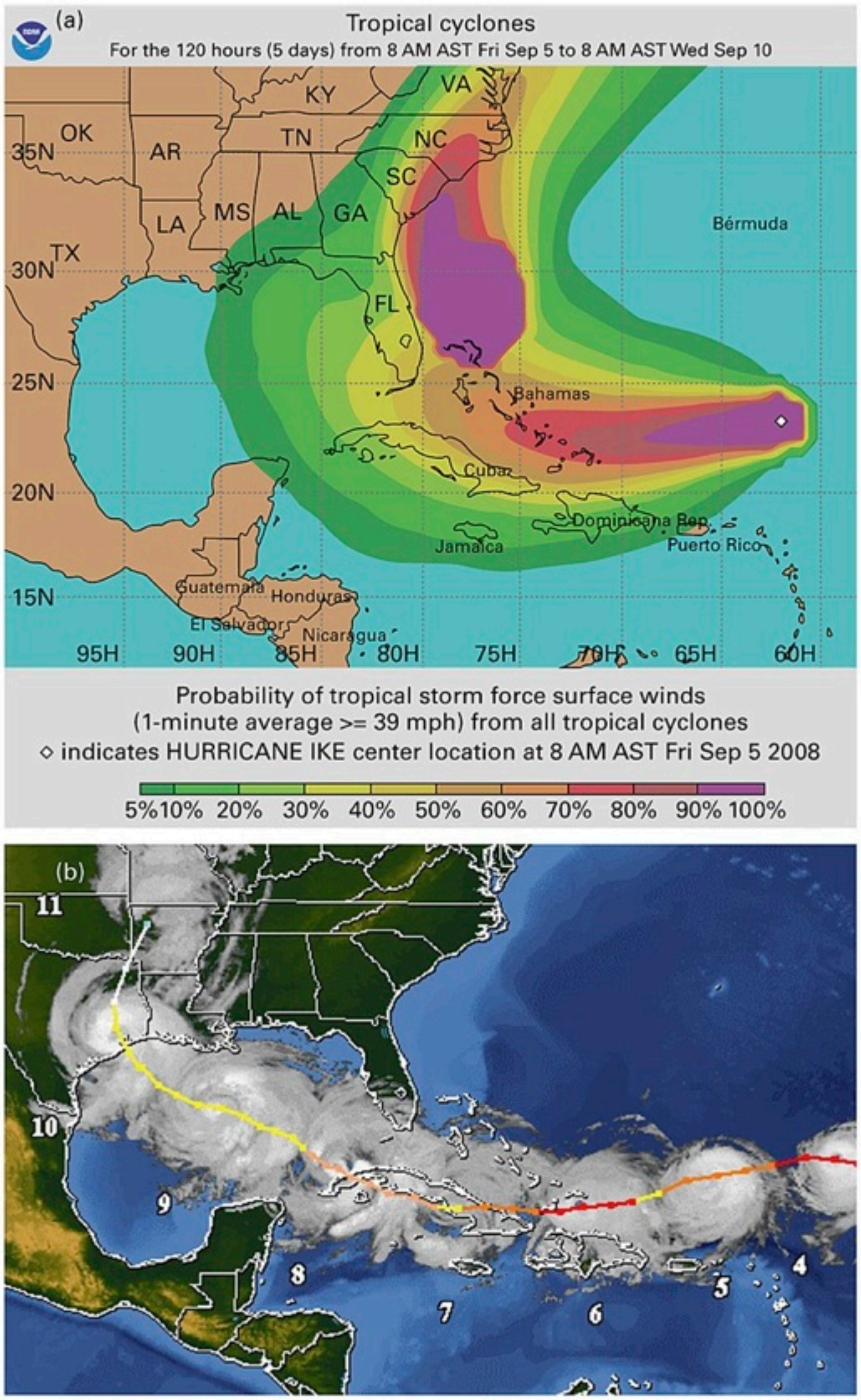

Figure 4: Comparison of the predicted (top) and actual (bottom) tracks of Hurricane Ike in December 2008. The storm was predicted to continue westward, and then turn north along the Florida coast, but instead followed a track outside the 95% uncertainty cone that headed into the Gulf of Mexico, striking the Texas coast. (Stein and Stein, 2014).

In this approach, deciding when and how to revise hazard maps would involve combining Bayes and BOGSAT. Conceptually, the process of changing parameters would reflect Bayes' Rule, in that those previously thought to have greater uncertainty would be most easily changed by new data or ideas. Operationally, because most parameters are estimated via a combination of data, models, and assumptions, the actual values would still come from BOGSAT rather than explicit calculation. Even so, the Bayesian approach can add value because it is systematic. If BOGSAT leads to big changes in the map, one can assess what that implies about prior confidence in the forecasts, and possibly revise BOGSAT.

The real beneficiaries of posing the process in this combined form would be map hazard map users, who would gain information about the uncertainties involved and thus could make better decisions. The meteorological community (*Hirschberg et al., 2011*) has adopted a goal of "routinely providing the nation with comprehensive, skillful, reliable, sharp, and useful information about the uncertainty of hydrometeorological forecasts." Although seismologists have an even tougher challenge and a longer way to go, we should try to do the same.

Acknowledgements

Stein thanks the USGS John Wesley Powell Center for Analysis and Synthesis for hosting a working group under auspices of the Global Earthquake Model project, whose stimulating discussions inspired this work. Spencer thanks the Institute for Policy Research at Northwestern for supporting his research. We also thank Sandy Zabell for helpful comments.

References

Campbell, K. W. (1982), Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model. *Bull. Seismol. Soc. Amer.*, **72**, 1689-1705.

Congressional Budget Office (2010), *CBO's 2010 Long-Term Projections for Social Security*.

Cornell, C. A. (1972), Bayesian statistical decision theory and reliability-based design, *Proceedings of the International Conference on Structural Safety and Reliability*, A. M. Freudenthal, Editor, April 9-11, 1969, Washington, D.C., Smithsonian Institute, 47-66.

Field, E., Dawson, T., Felzer, K., Frankel, A., Gupta, V., Jordan, T., Parsons, T., Petersen, M., Stein, R., Weldon, R., Wills, C. (2008). *The Uniform California Earthquake Rupture Forecast, Version 2,* U.S. Geol. Surv. Open File Report 2007-1437.

Frankel, A. (2013). Comment on "Why earthquake hazard maps often fail and what to do about it," by S. Stein, R.J. Geller, and M. Liu. *Tectonophysics*. **592**, 200–206.

Geller, R.J (2011). Shake-up time for Japanese seismology, *Nature*, 472, 407-409.

Hanks, T. C., G. C. Beroza, and S. Toda (2012). Have recent earthquakes exposed flaws in or misunderstandings of probabilistic seismic hazard analysis?, *Seismol. Res. Lett.* **83**, 759–764.

Hirschberg, P., et al. (2011). A weather and climate enterprise strategic implementation plan for generating and communicating forecast uncertainty information, *Bull. Am. Meteorol. Soc.*, **92**, 1651–1666, doi:10.1175/BAMS-D-11-00073.1.

Intergovernmental Panel on Climate Change (IPCC) (2007), *Climate Change 2007: The Physical Science Basis: Contribution of Working Group I to the Fourth Assessment*

Report of the Intergovernmental Panel on Climate Change, Cambridge Univ. Press, New York.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modeling*, John Wiley & Sons.

Parsons, T. (2008). Earthquake recurrence on the south Hayward fault is most consistent with a time dependent, renewal process, *Geophys. Res. Lett.*, **35**, doi:10.1029/2008GL035887.

Parvez, I.A. (2007), On the Bayesian analysis of the earthquake hazard in the North-East Indian peninsula, *Nat. Hazards,* **40**, 397–412, DOI 10.1007/s11069-006-9002-4.

Ramsey, F.P. (1926) Truth and probability. Pp. 156 – 198 in R.B. Braithwaite (editor) (1965) *The Foundations of Mathematics and other Logical Essays.* Totowa, N.J.: Littlefield, Adams & Co.

Reserve Bank of Australia (2013), *Statement on Monetary Policy.*

Rice, J. (2007). *Mathematical Statistics and Data Analysis,* Duxbury Advanced, India.

Savage, J. C., 1991. Criticism of some forecasts of the national earthquake prediction council, *Bull. Seismol. Soc. Am.,* **81**, 862-881.

Savage, J. C. (1992). The uncertainty in earthquake conditional probabilities, *Geophys. Res. Lett,* **19**, 709-712.

Savage, J. C. (1994). Empirical earthquake probabilities from observed recurrence intervals, *Bull. Seismol. Soc. Am.,* **84**, 219-221.

Sieh, K., Stuiver, M., Brillinger, D. (1989). A more precise chronology of earthquakes produced by the San

Andreas fault in southern California, *J. Geophys. Res.*, **94**, 603-624.

Sivia, D.S. (2006). *Data Analysis: A Bayesian Tutorial*, Oxford.

Stark, P.B. and D.A. Freedman (2003). What is the chance of an earthquake? in Mulargia, F. and Geller, R.J. (eds.), *Earthquake Science and Seismic Risk Reduction*, NATO Science Series IV: Earth and Environmental Sciences 32 201–213. Kluwer, Dordrecht, The Netherlands.

Stein, S., and A. Friedrich (2014). How much can we clear the crystal ball? *Astronomy and Geophysics* **55**, 2.11-2.17.

Stein, S., R. J. Geller, and M. Liu (2012). Why earthquake hazard maps often fail and what to do about it. *Tectonophysics* **562–563**, 1–25.

Stein, S., R. J. Geller, and M. Liu (2013). Reply to comment by Arthur Frankel on “Why Earthquake Hazard Maps Often Fail and What to do About It”, *Tectonophysics* **592**, 207–209.

Stein, S. and J.L. Stein (2013). Shallow versus deep uncertainties in natural hazard assessments, *EOS, 94,* 4, **133-140**.

Stein, S. and J. Stein (2014). *Playing Against Nature: Integrating Science and Economics to Mitigate Natural Hazards in an Uncertain World,* Wiley/AGU.

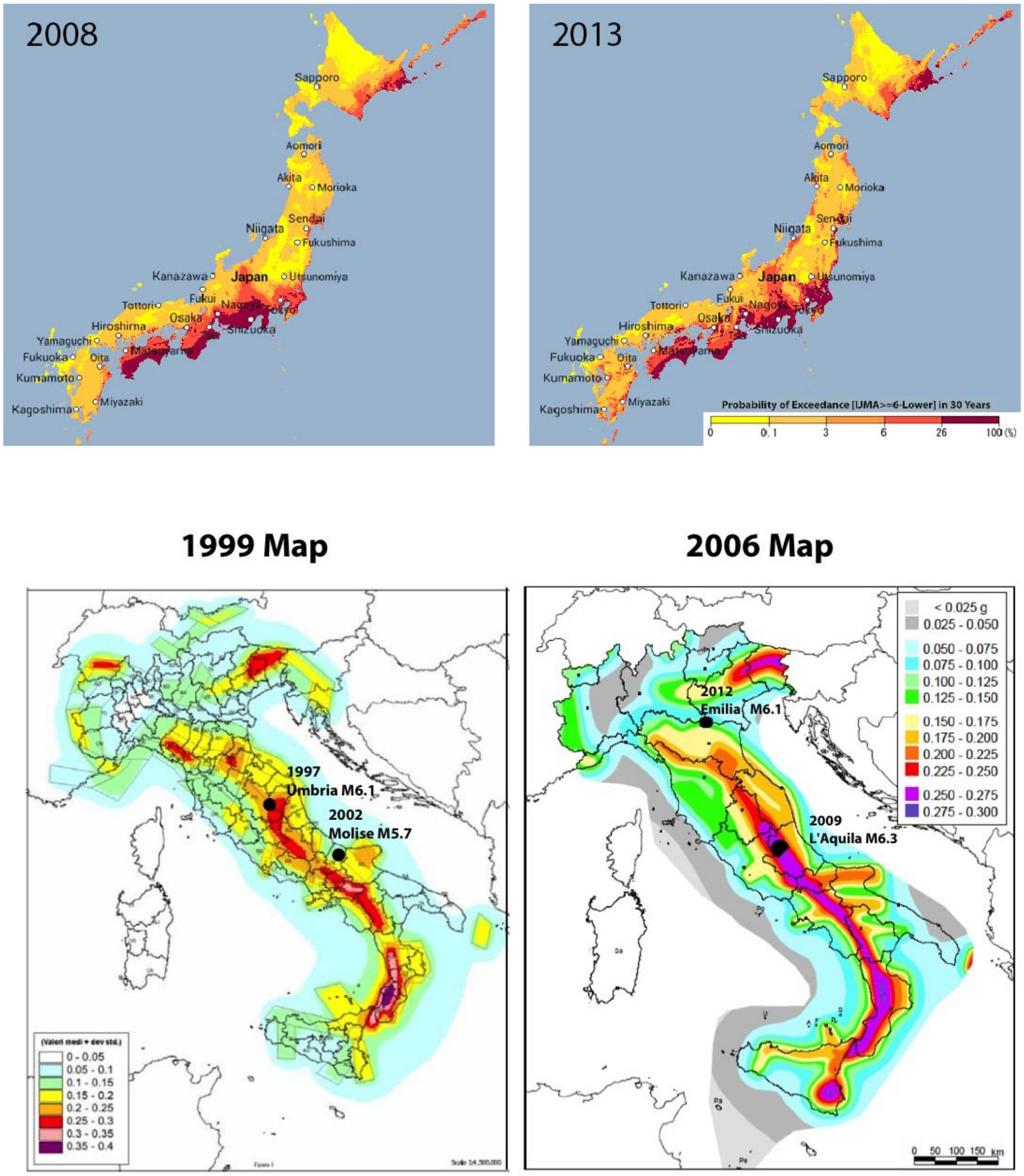

Figure 1: Top: Comparison of Japanese national seismic hazard maps before and after the 2011 Tohoku earthquake. The predicted hazard has been increased both along the east coast, where the 2011 earthquake occurred, and on the west coast. (http://www.j-shis.bosai.go.jp/map/?lang=en) Bottom: Comparison of successive Italian hazard maps (Stein et al., 2013). The 1999 map was updated to reflect the 2002 Molise earthquake and the 2006 map will likely be updated after the 2012 Emilia earthquake.

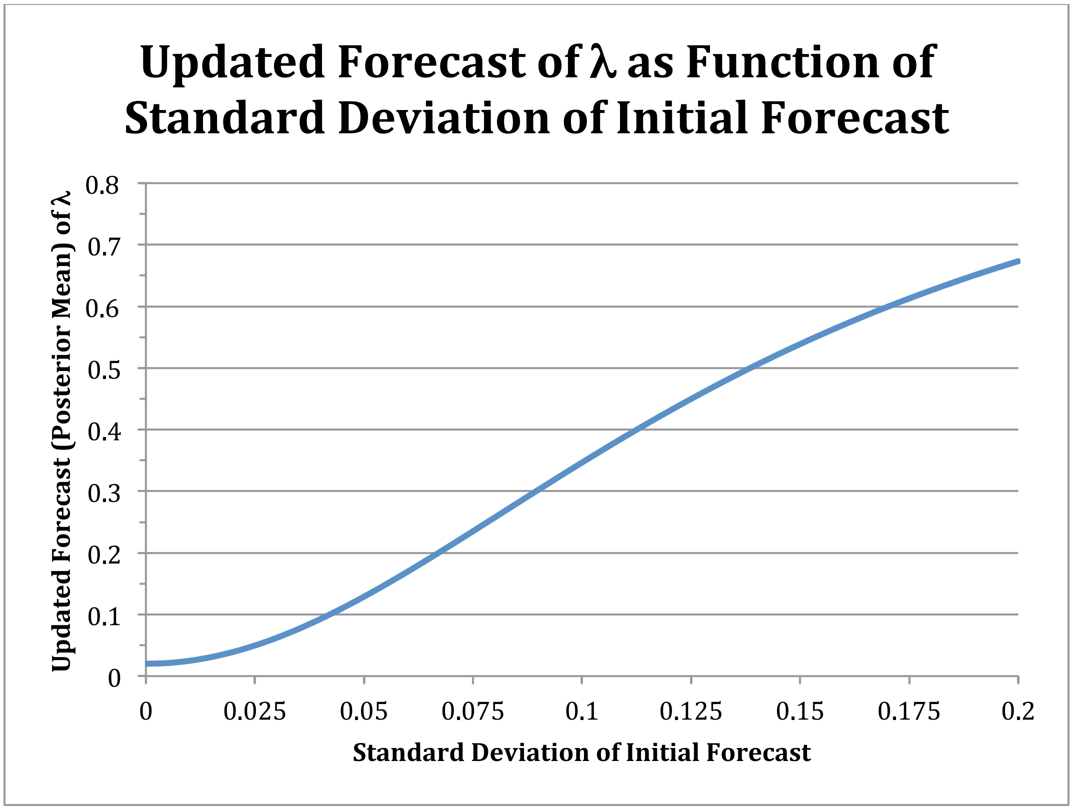

Figure 2. Sensitivity of updated forecast of λ, initially assumed to equal 0.02, to assumed prior uncertainty in the initial forecast. The less confidence we have in the initial forecast, the more the new datum changes it.

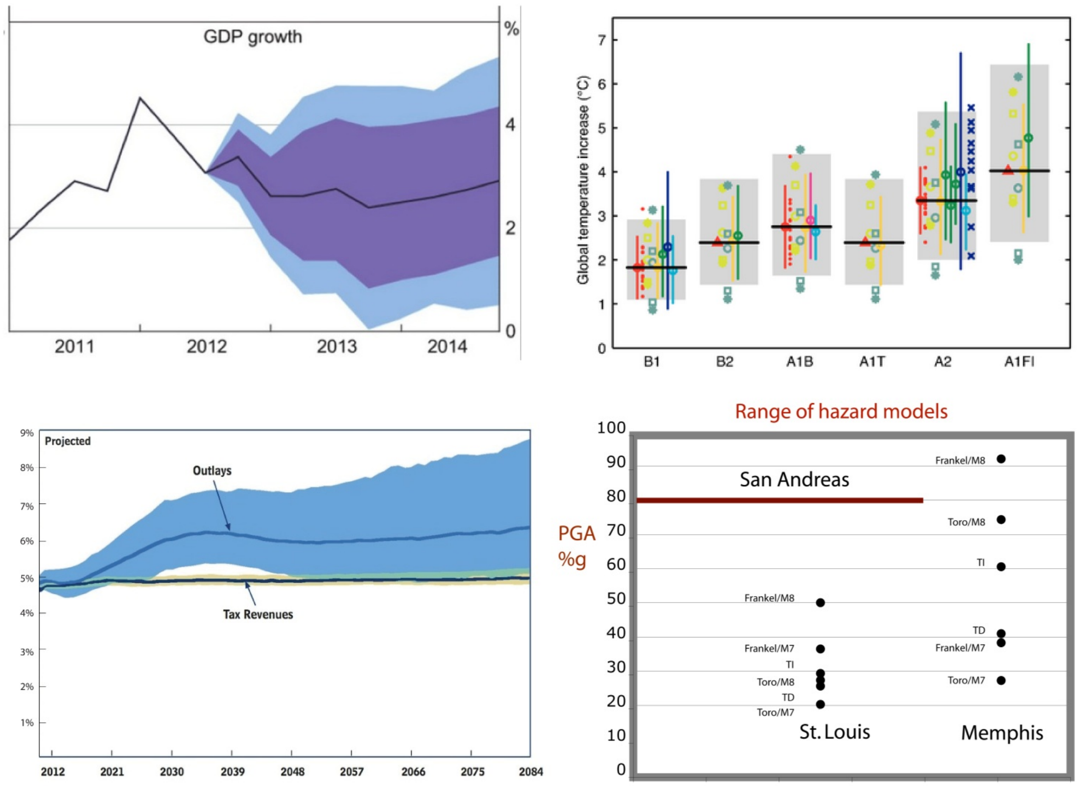

Figure 3: Presenting forecast uncertainties. a) Forecast of Australian gross domestic product growth. Uncertainty bounds are 70% and 90% (*Reserve Bank of Australia*, 2013). b) Forecast of US Social Security expenditure as percentage of Gross Domestic Product (*Congressional Budget Office*, 2010) c) Comparison of the rise in global temperature by the year 2099 predicted by various climate models. For various scenarios of carbon emissions, e.g., B1, the vertical band shows the predicted warming (*IPCC*, 2007). d) Comparison of earthquake hazard, described as peak ground acceleration (PGA) as a percentage of the acceleration of gravity expected with 2% risk in 50 years, predicted by various assumptions for two sites in the central U.S. (*Stein et al.*, 2012).

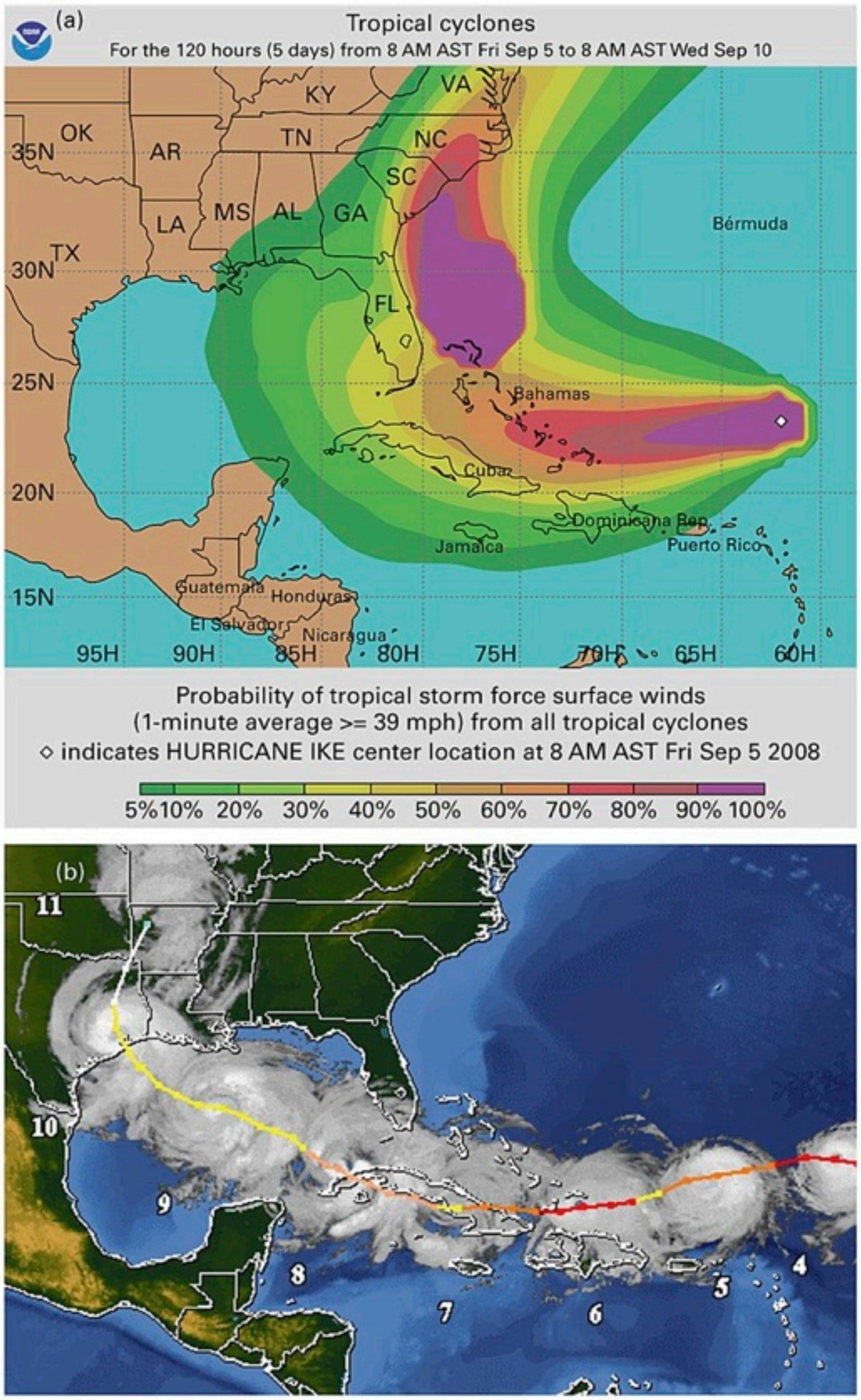

Figure 4: Comparison of the predicted (top) and actual (bottom) tracks of Hurricane Ike in December 2008. The storm was predicted to continue westward, and then turn north along the Florida coast, but instead followed a track outside the 95% uncertainty cone that headed into the Gulf of Mexico, striking the Texas coast. (Stein and Stein, 2014).