Bayesian Updating of Track-Forecast Uncertainty for Tropical Cyclones

NAM-YOUNG KANG AND MYEONG-SOON LIM

National Typhoon Center, Korea Meteorological Administration, Jeju, South Korea

JAMES B. ELSNER

Department of Geography, Florida State University, Tallahassee, Florida

DONG-HYUN SHIN

National Typhoon Center, Korea Meteorological Administration, Jeju, South Korea

(Manuscript received 14 October 2015, in final form 29 December 2015)

ABSTRACT

The accuracy of track forecasts for tropical cyclones (TCs) is well studied, but less attention has been paid to the representation of track-forecast uncertainty. Here, Bayesian updating is employed on the radius of the 70% probability circle using 72-h operational forecasts with comparisons made to the classical approach based on the empirical cumulative density (ECD). Despite an intuitive and efficient way of treating track errors, the ECD approach is statistically less informative than Bayesian updating. Built on a solid statistical foundation, Bayesian updating is shown to be a useful technique that can serve as a substitute for the classical approach in representing operational TC track-forecast uncertainty.

1. Introduction

Tropical cyclone (TC) activity is a major concern to a large number of people worldwide where lives and property are at risk (Mendelsohn et al. 2012; World Bank 2010). A forecast of where a TC will go is the single most important piece of information for disaster preparedness. Since forecasts are less than perfect, error distances are computed between the operational track-forecast position and the position observed as the best estimate of the TC center location. Forecast accuracy is evaluated from these error distances. Owing to improvements in numerical model guidance, the accuracy of the operational track forecasts is getting better all the time (Heming and Goerss 2010).

While forecast accuracy is well understood and studied, the representation of track-forecast uncertainty is less understood and studied (Pole et al. 1999). The tradition is to use a probability contour based on the empirical cumulative density (ECD) function of forecast-track error distances (Mannoji 2005; Kishimoto 2010). For example, a quantile value at a 0.7 probability level (70th percentile) of the ECD provides a radial error distance approximating the 70% probability circle around a forecast TC position. While useful as a historical benchmark, the ECD approach does not include information about the current forecast track error.

The problem is how to estimate track uncertainty that reflects the past statistics and current data (Elsner and Bossak 2001; Elsner and Jagger 2004). Moreover, despite the efficiency of the ECD approach, it is not clear if the uncertainties are statistically calibrated when used as forecasts. Said another way, given a forecast position and a 70% probability circle, is it the case that 30% of forecasts made under identical situations will result in a position outside the circle? The objective of this study is to demonstrate an alternative approach for representing TC track-forecast uncertainty. The approach is rooted in Bayesian inference, and the uncertainty estimates from this approach are considered to be closer to a forecaster's actual belief about the forecast spread. Comparisons are made with the classical ECD approach.

Denotes Open Access content.

Corresponding author address: Nam-Young Kang, 2 Seosung 810-gil, Namwon, Seogwipo, Jeju 699-942, South Korea.
E-mail: nkang.fsu@gmail.com

DOI: 10.1175/WAF-D-15-0140.1

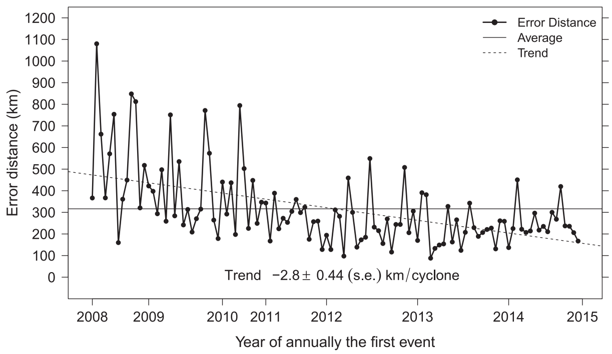

FIG. 1. Time series of the track error distances for 72-h forecasts over 7 yr (2008–14). Each value represents the average per cyclone. One-hundred and thirteen values are plotted sequentially. The solid horizontal line indicates the average over the whole period. Linear regression on the values shows a decreasing trend (dashed line).

The paper is organized as follows. Description of the data used in this study is given in section 2. The procedure of Bayesian updating is described in section 3, followed by a presentation of the results in section 4. Interpretation of the results and a discussion of the value of the Bayesian approach are given in section 5. The paper is summarized in section 6. All of the statistics and figures are created using the software R (www.r-project.org) and are available online (rpubs.com/Namyoung/P2015b).

2. Data

Operational 72-h TC track forecasts and the associated observations over a 7-yr period (2008–14) are obtained from the Korea Meteorological Administration (KMA). The consecutive annual numbers of TCs whose lifetimes exceeded 72 h are 13, 17, 10, 14, 21, 21, and 17. The error distances between the forecast and observed TC positions are averaged per cyclone and used as a sample dataset to show the Bayesian approach to estimating track-forecast uncertainty.

Figure 1 displays the 113 error values as a time series. Consecutive locations of the first TC in each year are labeled with the corresponding year number. Each value is assumed to be independent. The period-average track error is 316.2 km (horizontal line). Linear regression by ordinary least squares on the errors shows a decreasing trend of −2.8 ± 0.44 (standard error) km per cyclone (dashed line).

3. Bayesian updating

a. Gamma distribution for the error distances

Error distances x are assumed to be adequately described by a gamma distribution as the values are continuous, nonnegative, and skewed. Mathematically, the distribution is expressed as

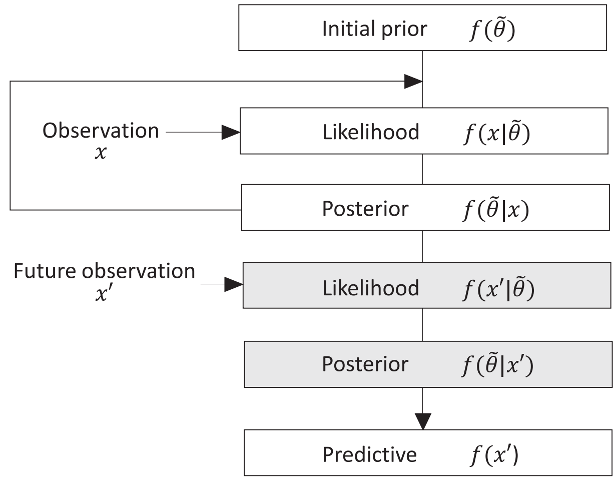

FIG. 2. Schematic flowchart of the Bayesian updating method. Given an observation x, the posterior is used as a new prior and the whole process is repeated. Predictive density is calculated for future values x' based on the posterior.

$$f(x;\alpha,\beta)=\frac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)},\quad \text{for all}\quad x>0, \qquad (1)$$

where α and β are the shape and rate (inverse scale) parameters, respectively. In addition, Γ is the gamma function defined as $\Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}e^{-x}\,dx$. In this way the probability distribution of error distance is completely described by the two parameters.

b. Prior specification

The procedure for estimating the distribution of future error distances $f(x')$ is outlined as a flowchart (Fig. 2). The approach starts with a method for assessing the initial uncertainty on two statistics that are related to the gamma parameters [here represented as a vector $\tilde{\boldsymbol{\theta}}=(\mu_x,\sigma_x^2)$]. First, a gamma distribution is fit to the first 30 error distances. Then, 1000 samples, each containing 30 error distances, are drawn from the distribution and pairs of the mean and the variance are computed from each.

To avoid complications arising from using a fully hierarchical model and a Monte Carlo Markov chain simulation, the parameter space of $\tilde{\boldsymbol{\theta}}$ is discretized. We approximate the two-dimensional density surface by configuring a grid of equally spaced values for μ_x and σ_x^2. The spacing is 2 km from 330 to 630 km for μ_x and 1000 km^2 intervals from 11 000 to 62 000 km^2 for σ_x^2, resulting in a 151 × 151 grid.

The prior density is estimated with a two-dimensional Gaussian kernel smoother over the grid. The bandwidths are set at 9 km and 3480 km^2 as a compromise between removing insignificant bumps and retaining the real peaks (Venables and Ripley 1999). The prior density gives the relative weight of x with fixed gamma

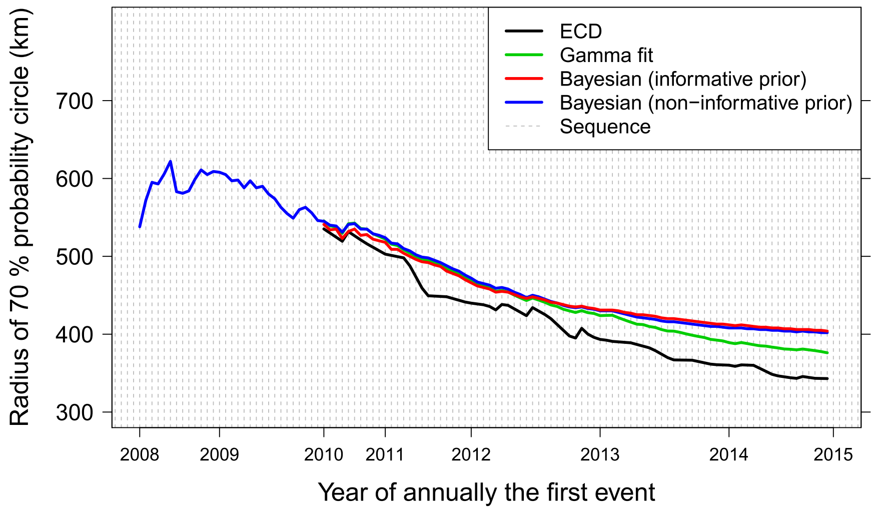

FIG. 3. Time series of the 70th-percentile probability circle radius (km). Black and green lines show the values from up-to-date ECD and theoretical gamma distribution, respectively. The red line represents the Bayesian result from an informative initial prior, which is the joint probability distribution of the gamma parameters based on the first 30 error distances. For comparison, an alternative experiment starting from a noninformative prior is shown with the blue line.

parameters. Using the method of moments, the two values composing $\tilde{\boldsymbol{\theta}}$ are translated into the parameters α and β for a gamma distribution by the relationships (Husak et al. 2007)

$$\alpha = \frac{\mu_x^2}{\sigma_x^2} \tag{2}$$

and

$$\beta = \frac{\mu_x}{\sigma_x^2}. \tag{3}$$

Finally, to test the robustness of the Bayesian updating approach for estimating track-forecast uncertainty, a noninformative prior is also specified. The noninformative prior is prepared by a specifying a uniform distribution of values on the same grid that sum to unity.

c. Likelihood and posterior

The likelihood function $f(x\,|\,\tilde{\boldsymbol{\theta}})$ is estimated using a newly computed error distance x conditional on $\tilde{\boldsymbol{\theta}}$. The likelihood has the same form as Eq. (1), but μ_x and σ_x^2 are used as the conditional parameters. Then, the posterior of $\tilde{\boldsymbol{\theta}}$ is proportional to the joint probability distribution of the likelihood and the prior as

$$f(\tilde{\boldsymbol{\theta}}\,|\,x) \propto f(x\,|\,\tilde{\boldsymbol{\theta}})f(\tilde{\boldsymbol{\theta}}), \tag{4}$$

which is summed to one. Each update uses the posterior as a new prior. The process is repeated to find a new posterior given the next computed error distance (Fig. 2).

d. Prediction from the posterior

Given the latest posterior, the marginal distribution of x' is computed by

$$f(x') = \int_{\tilde{\boldsymbol{\theta}}} f(x'\,|\,\tilde{\boldsymbol{\theta}})f(\tilde{\boldsymbol{\theta}})\,d\tilde{\boldsymbol{\theta}}, \tag{5}$$

which is a weighted convolution of multiple (151 × 151 in this experiment) gamma distributions (Di Salvo 2008), where the weights come from the posterior $f(\tilde{\boldsymbol{\theta}})$. Through the systematic use of past and recent error-distance values, the method updates the uncertainty closer to the forecaster's subjective belief in the forecast spread. A similar approach is taken to model the predicted time it takes for an article to get published in an American Meteorological Society journal (Hodges et al. 2012).

The 70th-percentile error distance of the posterior distribution is the radius of the 70% probability circle. In this study, we approximate the exact value by the 70th percentile among the error distances from 0 to 1000 km in 1-km intervals, assuming the density of the error distances beyond 1000 km is negligible.

4. Results

Radial distances of the 70th-percentile circle are plotted as a time series in Fig. 3. Curves are shown for the ECD method, a gamma fit, and for two Bayesian update methods (informative and noninformative

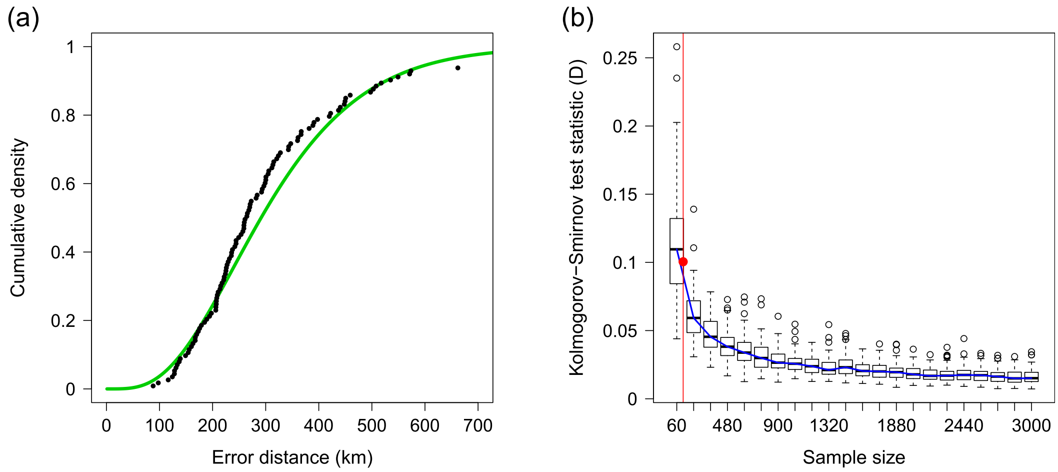

FIG. 4. Comparison of cumulative density distributions. (a) Cumulative density distributions of 113 error distances (black dot) and their fitted gamma distribution (green line), and (b) box plot of the KS test statistic by comparing the gamma distribution and its bootstrapped samples. For (b), the test statistic D is calculated 100 times for each different sample size. Each boxplot shows the median (black thick solid line), interquartile range (box range), and outliers (circles). Median values of the D samples are connected by the blue curve. The red dot represents the D calculated from the observed errors and the gamma distribution at 113 error samples (red line).

prior). Curves are shown after the 30th case except for the noninformative prior update method. As noted, the radii of the probability circle for the updating methods (blue and red lines) are estimated from the marginal distribution of x' based on the posterior of $\tilde{\boldsymbol{\theta}}$.

The radial distances show a general decreasing function over time resulting from forecast improvements as described by the raw data (Fig. 1). Case-to-case fluctuations also tend to diminish over time. The stability of the error distances is particularly evident with the update methods. The influence of the priors is seen initially but after a dozen cases or so there are only small differences between the two updating methods. The Bayesian updates show consistently larger radii likely reflecting better-calibrated forecasts.

5. Interpretation

The difference between the empirical (black) and Bayesian (red) forecast-track errors arises from the difference between the empirical and cumulative densities. For clarification, the 70th-percentile distances from updated gamma distributions are plotted with a green line in Fig. 3. Each value comes from a gamma distribution fit to an updated set of error distances. In this case the radii are closer to the Bayesian updates, but still larger than the radii from the ECD approach.

The radii from the gamma fit are similar to the radii from the Bayesian update since both are theoretical. As an example, Fig. 4 shows how the empirical and the theoretical cumulative densities can be different, where a total of 113 error samples are used to fit the theoretical density distributions. The reason for the difference lies in the shape of the distributions. Densities in both tails of a theoretical distribution are less likely to be well represented in an empirical distribution with small sample size. This makes the level of cumulative densities different still.

Values of the Kolmogorov–Smirnov (KS) goodness-of-fit test statistic for samples drawn from a gamma density as a function of sample size show how different a sample distribution can be from its parent distribution (Fig. 4b). The test statistic uses the maximum difference D between the ECD and the theoretical cumulative density.

The medians (black thick solid line in the box plots) connected by the blue curve show that the ECD from a sample can be quite different from the theoretical population according to where the samples are drawn. The red dot indicates the D of the two cumulative densities seen from Fig. 4a, confirming how the 70th-percentile value from a small number of samples can be different from that of its theoretical distribution. This also implies that ECD's deficiency might be even more exaggerated in longer lead-time forecasts such as those at 96 and 120 h with the possibility of fewer samples compared with the 72-h lead time. Assuming the error distances are described by a reasonable theoretical distribution and assuming a reasonable (not extreme) quantile level for estimating the probability of the TC position, the

ECD may not be the best approach for estimating the forecast-track error.

An additional reason for the gap between the error distances from the empirical and the Bayesian approaches comes from the fact that the Bayesian approach combines two sources of uncertainty. Practically, it is the difference between the green line and the red line in Fig. 3. Like the difference between a ''prediction interval'' and a ''confidence interval'' in a linear regression, the predictive distribution combines the uncertainty of a particular radial distance given the parameters and the uncertainty associated with the values of the unknown ''true'' parameters. The fitted gamma distribution uses fixed parameter values and so by definition lacks this latter uncertainty. Thus, by construction the Bayesian updating approach will give a larger 70% probability circle than the ECD approach.

Finally, it is noted that the ECD approach results in greater variation across the error distances. This behavior is expected since the ranked error members are sparse, and their relationship is not necessarily linear. For example, the inclusion of a new track error distance can push the quantile value to the next larger interval between ranks, leading to a big shift. This ''erratic'' behavior is absent with the Bayesian update method. Even the two results from informative and noninformative priors converge to nearly the same level after some time. Overall, these features of the Bayesian response to observed errors imply the robustness of the Bayesian approach.

6. Summary

Although typically portrayed as a cone of uncertainty by extending across consecutive forecast horizons, probability circles are widely used to represent the uncertainty of forecast TC positions. The classical approach provides a probability estimate from an empirical cumulative density function. The 70th percentile is comparable to the radius of the 70% probability circle around a TC forecast position.

This study explores Bayesian updating as an alternative representation of forecast uncertainty. The Bayesian approach has the advantage that output is in the form of a predictive distribution. The method is described and then compared with the classic ECD approach. Data for the comparison are error distances from the operational 72-h TC track forecast issued by KMA. The 70th-percentile error distances estimated from the Bayesian updating scheme are larger than the corresponding error distances from the ECD approach. In spite of the intuitive and efficient way of treating the errors, the ECD approach is shown to be statistically less informative than the Bayesian updating method for the following reasons:

- The ECD approach with a smaller sample size is more likely to under- or overestimate the 70th percentile of an assumed population following the gamma distribution.
- Error distance samples from the ECD method are assumed to be random samples from a theoretical gamma distribution with fixed parameters, which leads to a narrower uncertainty range than the Bayesian updating approach considering the uncertainty of the parameters.
- Because of the nonlinear nature of the ranked error distances, the probability circle obtained by the ECD approach is unstable with respect to newly added error samples, while the Bayesian approach returns a stable and robust result.

Consequently, it is confirmed that a smaller value of the 70th-percentile radial distance from the ECD approach does not imply a better result. The gap between the ECD and the Bayesian results reveals how much the former lacks the uncertainty information. On a firm statistical basis the Bayesian approach is considered to be a useful and statistically more consistent technique that can replace the classical ECD approach in the representation of the operational TC track-forecast uncertainty.

Acknowledgments. This work was supported by the National Typhoon Center and the National Institute of Meteorological Sciences at the Korea Meteorological Administration (NIMS-2016-3100).

REFERENCES

Di Salvo, F., 2008: A characterization of the distribution of a weighted sum of gamma variables through multiple hypergeometric functions. *Integr. Transforms Spec. Funct.*, **19**, 563–575, doi:10.1080/10652460802045258.

Elsner, J. B., and B. H. Bossak, 2001: Bayesian analysis of U.S. hurricane climate. *J. Climate*, **14**, 4341–4350, doi:10.1175/1520-0442(2001)014<4341:BAOUSH>2.0.CO;2.

——, and T. H. Jagger, 2004: A hierarchical Bayesian approach to seasonal hurricane modeling. *J. Climate*, **17**, 2813–2827, doi:10.1175/1520-0442(2004)017<2813:AHBATS>2.0.CO;2.

Heming, J., and J. Goerss, 2010: Track and structure forecasts of tropical cyclones. *Global Perspectives on Tropical Cyclones*, J. C.-L. Chan and J. D. Kepert, Eds., World Scientific Series on Asia-Pacific Weather and Climate, Vol, 4, World Scientific, 287–323.

Hodges, R. E., J. B. Elsner, and T. H. Jagger, 2012: Predictive models for time to acceptance: An example using ''hurricane'' articles in AMS journals. *Bull. Amer. Meteor. Soc.*, **93**, 879–882, doi:10.1175/BAMS-D-11-00133.1.

Husak, G. J., J. Michaelsen, and C. Funk, 2007: Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *Int. J. Climatol.*, **27**, 935–944, doi:10.1002/joc.1441.

Kishimoto, K., 2010: JMA's five-day tropical cyclone track forecast. RSMC Tokyo-Typhoon Center Technical Review, No. 12, 55–63. [Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/techrev/text12-2.pdf.]

Mannoji, N., 2005: Reduction of the radius of probability circle in typhoon track forecast. RSMC Tokyo-Typhoon Center Technical Review, No. 8, 1–9. [Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/techrev/text8-1.pdf.]

Mendelsohn, R., K. Emanuel, S. Chonabayashi, and L. Bakkensen, 2012: The impact of climate change on global tropical cyclone damage. *Nat. Climate Change*, **2**, 205–209, doi:10.1038/nclimate1357.

Pole, A., M. West, and J. Harrison, 1999: *Applied Bayesian Forecasting and Time Series Analysis*. Chapman and Hall, 409 pp.

Venables, W. N., and B. D. Ripley, 1999: *Modern Applied Statistics with S-PLUS*. Springer, 501 pp.

World Bank, 2010: *Natural Hazards, UnNatural Disasters: The Economics of Effective Prevention*. International Bank for Reconstruction and Development, 254 pp.