Bayesian Typhoon Track Prediction Using Wind Vector Data

Minkyu Han[1,a], Jaeyong Lee[a]

[a]Department of Statistics, Seoul National University, Korea

Abstract

In this paper we predict the track of typhoons using a Bayesian principal component regression model based on wind field data. Data is obtained at each time point and we applied the Bayesian principal component regression model to conduct the track prediction based on the time point. Based on regression model, we applied to variable selection prior and two kinds of prior distribution; normal and Laplace distribution. We show prediction results based on Bayesian Model Averaging (BMA) estimator and Median Probability Model (MPM) estimator. We analysis 8 typhoons in 2006 using data obtained from previous 6 years (2000–2005). We compare our prediction results with a moving-nest typhoon model (MTM) proposed by the Korea Meteorological Administration. We posit that is possible to predict the track of a typhoon accurately using only a statistical model and without a dynamical model.

Keywords: Bayesian principal component regression, wind field data, typhoon track prediction, Laplace distribution, Haversine formula

1. Introduction

1.1. Motivation

A typhoon is a tropical depression originating in the southwestern ocean of northern Pacific, whose maximum wind velocity at center is over 17 m/s (about 33 knot). Since typhoons were recorded, about 30 typhoons have arisen every year. Most typhoons develop in the latitude 5°–20° and move in a northwest direction. Some typhoons reach Taiwan or the Chinese mainland and disappear, but others make a turn to the northeast and reach northeastern Asia. These typhoons cause damage to South Korea and Japan. Establishing forecasting system through typhoon analysis is essential to minimize direct and indirect damage caused by typhoons. There are 4 prediction problems in typhoon analysis: predictions of track, intensity, storm surge and rainfall. Among them, the most important problem is the track prediction.

While some typhoons move zigzag, most typhoons are influenced by basic atmospheric flow due to pressure patterns. Figure 1 shows the track of Typhoon Man-yi (the fourth typhoon in 2007), over the pressure field. The typhoon appears to move alongside the edge of anticyclone.

In this paper, we cast the typhoon track prediction problem into the regression prediction problem with wind field around the center of the typhoon used as predictors.
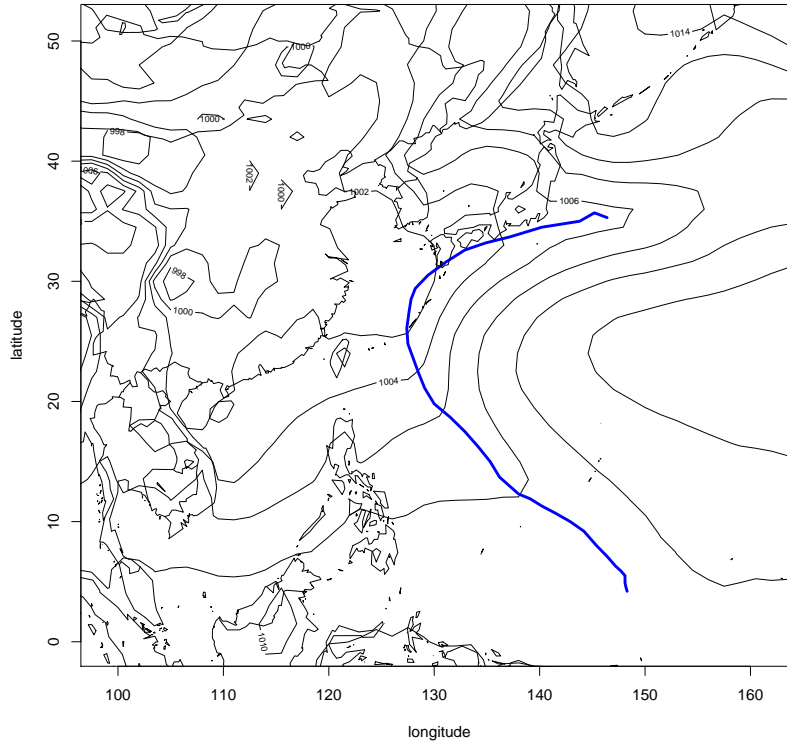
Figure 1: *Typhoon Man-yi track (blue) on the pressure field.*

1.2. Previous works

Generally, there are two kinds of typhoon track prediction models. One is the dynamical models used by most countries' meteorological administration. Dynamical models are composed of differential equations verified as mesoscale phenomena in the atmospheric science field. The other is the statistical model that utilizes statistical methods such as regression using past data; however, statistical models often play a peripheral role to support dynamical models.

Among dynamic models, European Centre of Medium-Range Weather Forecasts (ECMWF), National Centers for Environmental Prediction (NCEP), Japan Meteorological Agency (JMA) have a higher prediction accuracy as a Global model or typhoon-focused model. The moving-nest typhoon model (MTM) proposed by Kim *et al.* (2007) constitutes an interesting example of a typhoon-only model. This model is based on the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) mesoscale model (MM5) and has applied bogusing method suggested by the Geophysical Fluid Dynamics Laboratory (GFDL). When we simulate typhoons, there is a drawback that air pressure in the center of typhoon is overcharged in general models. In order to compensate the defect, we can put stronger artificial typhoon in the atmospheric field based on a bogusing method. It has auto-moving function through the typhoon track using the MM5 moving nest system. When typhoon moves, nest grids in the system follow the typhoon and conduct integration. They used Global Data Assimilation and Prediction System (GDAPS) results, surface sea temperature (SST) from NCEP and best track data (longitude, latitude, maximum velocity) from the Joint Typhoon
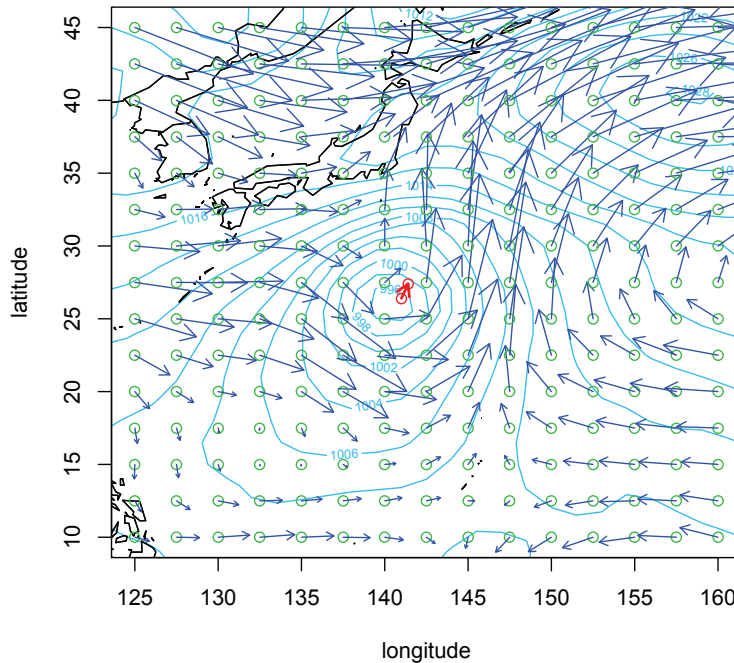
26th location of 18th typhoon, 2006

Figure 2: *Wind vectors on 225 grid points in pressure field.*

Warning Center (JTWC).

Sohn and Kim (2001) suggest a statistical model with numerical model output. They developed a dynamic linear model based on Kalman filter model with Barotropic Adaptive-grid Typhoon Simulation (BATS) results as explanatory variables and estimated coefficient parameters for predicting a typhoon track. However, this model requires a numerical model output.

Rozanova *et al.* (2010) verified explicit coordinate form through time using a Navier Stokes equation with parameters set for various cases (baroclinic, barotropic and abnormal). Their results were acceptable for each case, but it is difficult to classify typhoons into appropriate cases in the early phase.

In general, dynamical models are more accurate when they are more complicated; subsequently, they are get larger over time and require huge simulations and computations. Therefore if possible, it is necessary for prediction to approach intuitively. There are rarely prediction methods using statistical model only.

Song *et al.* (2005) shows that it is possible to deal with a typhoon track using data obtained eidetically. They suggest a prediction model using raw wind field data. They provided a support vector machine regression model with covariates obtained by a data reduction method. However, It has limitations because they focus on reducing the computation load and use only one type of typhoon which is obtained by a fuzzy c-means clustering method.

In this paper, we suggest a prediction method based on statistical models. Contrary to previous

works except Song *et al.* (2005), our model has a simpler concept by using the data intuitively. We use 2-dimensional wind field data obtained by averaging vertical values of 3-dimensional wind field data from NCEP. In particular, wind field data of 225 grid points (15 longitude points × 15 latitude points) around the eye of typhoon are used as covariates. Figure 2 shows a wind field pattern at certain time point. Grid points are formed by $2.5° \times 2.5°$.

The rest of the paper is organized as follows. In Section 2, we present regression model, data description and appropriate prior distribution and parameters for the model. In Section 3, posterior computation is presented. We display some results with actual typhoon track prediction problem comparing with Kim *et al.* (2007). Finally, we discuss some limitation and future works in Section 5.

2. Bayesian High-Dimensional Regression Model

We fit two predictions (latitude and longitude) in parallel since the prediction is enacted on the surface of the earth. We set the response variable y to be the differences between the present point and the next point of latitude or longitude. We arrange u-component and v-component from 225 grid points in order and we also set a series of the 450×1 vector as covariates.

Consider a simple regression model,

$$y = X\beta + \epsilon, \quad \epsilon \sim N\left(0, \sigma^2 I\right), \tag{2.1}$$

where the data y is an $n \times 1$ response vector, X is a $n \times p$ design matrix and β is a $p \times 1$ coefficient vector. When p is large, it is known that many problems such as multicollinearity and sparsity arise. We use the model suggested by Lee and Oh (2013). $X^T X$ can be decomposed into eigenvectors and eigenvalues. We set a matrix A having eigen vectors that correspond to nonzero eigenvalues. They can be expressed as follows.

$$X^T X = \begin{bmatrix} A & * \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A \\ * \end{bmatrix}.$$

A diagonal matrix which have nonzero eigen values is set to Λ. We reparametrize the Equation (2.1) by multiplying both sides of (2.1) by $(1/\sqrt{n})\Lambda_1^{-1/2}A_1^T X_n^T$ as follows.

$$y^* = \eta_n + \epsilon^*, \quad \xi_n \sim N\left(0, \frac{\sigma^2}{n}I_p\right),$$

where

$$y^* = \frac{1}{\sqrt{n}}\Lambda^{-\frac{1}{2}}A^T X^T y,$$

$$\eta_n = \frac{1}{\sqrt{n}}\Lambda^{-\frac{1}{2}}A^T X^T X\beta = \frac{1}{\sqrt{n}}\Lambda^{\frac{1}{2}}A^T\beta,$$

$$\epsilon^* = \frac{1}{\sqrt{n}}\Lambda^{-\frac{1}{2}}A^T X^T \epsilon.$$

If some eigenvalues are 0, it means that reparametrization make the number of covariates small because n is larger than p. It is effective to simply computation since it is a kind of dimension reduction.

On the other hand, if all eigenvalues are nonzeros, the design matrix is regarded as principal components of X_n. It is sufficient for this model to use the principal components as a new design matrix. For some cases, it is proved that Bayes estimator and posterior of parameters are consistent in Lee and Oh (2013). We estimate η_n since there is a 1-1 correspondence between β_n and η_n.

2.1. Covariate selection prior

We consider the variable selection prior as:

$$\eta_{ni} \sim \pi_0\delta_0 + (1 - \pi_0)h_\sigma, \quad i = 1, 2, \ldots, p_n,$$

where π_0 is the prior probability that η_{ni} equals 0, and h_α is a distribution where α is the precision parameter that controls degree of spreadness. We assume that each η_{ni}, $(i = 1, \ldots, p_n)$ are mutually independent. A common choice for h_α is a normal distribution with mean 0, variance σ^2/α. We also consider the Laplace or T distribution for h_α.

2.2. Estimation of prior parameter

For the fast computation, we estimate the parameters π_0, σ^2 and α of prior using the data. The empirical Bayes estimators, $\hat{\pi}_0, \hat{\sigma}^2, \hat{\alpha}$, are obtained by maximizing $\prod_{i=1}^{n} m_i(t_{ni}|\pi_o, \sigma^2, \alpha)$, where

$$m_i(t_{ni}|\pi_0, \sigma^2, \alpha) = \int p\left(\eta_{ni}|\pi_0, \sigma^2, \alpha\right) f\left(t_{ni}|\eta_{ni}, \sigma^2\right) d\eta_{ni}.$$

We consider following two distributions as priors and the calculation for each cases are shown as below. First, we consider the normal distribution with probability $1 - \pi_0$ as the prior distribution as below.

$$\eta_{ni} \overset{indep.}{\sim} \pi_0\delta_0 + (1 - \pi_0)N\left(0, \frac{\sigma^2}{\alpha}\right), \quad i = 1, 2, \ldots, p_n.$$

In the following manner, we derive m_i functions for the normal case. In normal case, m_i are weighted sum of two probability density values and it can be easily calculated.

$$
\begin{aligned}
&m_i\left(t_{ni}|\pi_0, \sigma^2, \alpha\right) \\
&= \int f\left(t_{ni}|\eta_{ni}\right) p\left(\eta_{ni}|\pi_0, \alpha\right) d\eta_{ni} \\
&= \int \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{n(t_{ni} - \eta_{ni})^2}{2\sigma^2}\right] \times \left(\pi_0 I(\eta_{ni} = 0) + (1 - \pi_0)\frac{\sqrt{\alpha}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\alpha\eta_{ni}^2}{2\sigma^2}\right]\right) d\eta_{ni} \\
&= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right] + (1 - \pi_0)\frac{\sqrt{n\alpha}}{2\pi\sigma^2} \int \exp\left[-\frac{1}{2\sigma^2}\left\{n\eta_{ni}^2 - 2nt_{ni}\eta_i + nt_{ni}^2 + \alpha\eta_{ni}^2\right\}\right] d\eta_{ni} \\
&= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right] + (1 - \pi_0)\frac{\sqrt{n\alpha}}{2\pi\sigma^2} \int \exp\left[-\frac{1}{2\sigma^2}\left\{\frac{\left(\eta_{ni} - \frac{nt_{ni}}{\alpha+n}\right)^2}{\frac{1}{\alpha+n}} - \frac{n^2 t_{ni}^2}{\alpha + n} + nt_{ni}^2\right\}\right] d\eta_{ni} \\
&= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right] + (1 - \pi_0)\frac{\sqrt{n\alpha}}{\sqrt{2\pi}\sigma\sqrt{\alpha + n}} \exp\left[-\frac{n\alpha}{2\sigma^2(\alpha + n)}t_{ni}^2\right].
\end{aligned}
\tag{2.2}
$$

Second, we consider the Laplace distribution with probability $1 - \pi_0$ as the prior distribution as follows. Laplace distribution is also sometimes called the double exponential distribution.

$$\eta_{ni} \sim \pi_0\delta_0 + (1 - \pi_0)\text{Laplace}(0, b), \quad i = 1, 2, \ldots, p_n.$$

We derive m_i function for Laplace case in the following. Contrary to normal case, there is a separation, whether η_{ni} is positive or not.

$$m_i(t_{ni}|\pi_0, \sigma, b)$$

$$= \int f(t_{ni}|\eta_{ni}, \sigma)p(\eta_{ni}|\pi_0, b)d\eta_{ni}$$

$$= \int \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{n(t_{ni} - \eta_{ni})^2}{2\sigma^2}\right] \times \left(\pi_0 I(\eta_{ni} = 0) + (1 - \pi_0)\frac{1}{2b} \exp\left[-\frac{|\eta_{ni}|}{b}\right]\right) d\eta_{ni}$$

$$= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right] + (1 - \pi_0)\frac{\sqrt{n}}{2\sqrt{2\pi}\sigma b} \int \exp\left[-\frac{1}{2\sigma^2}\left\{n\eta_{ni}^2 - 2nt_{ni}\eta_i + nt_{ni}^2\right\} - \frac{|\eta_{ni}|}{b}\right] d\eta_{ni}$$

$$= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right]$$

$$\quad + (1 - \pi_0)\frac{\sqrt{n}}{2\sqrt{2\pi}\sigma b} \int_{-\infty}^{0} \exp\left[-\frac{1}{2\sigma^2}\left\{n\eta_{ni}^2 - 2nt_{ni}\eta_i + nt_{ni}^2\right\} + \frac{\eta_{ni}}{b}\right] d\eta_{ni}$$

$$\quad + (1 - \pi_0)\frac{\sqrt{n}}{2\sqrt{2\pi}\sigma b} \int_{0}^{\infty} \exp\left[-\frac{1}{2\sigma^2}\left\{n\eta_{ni}^2 - 2nt_{ni}\eta_i + nt_{ni}^2\right\} - \frac{\eta_{ni}}{b}\right] d\eta_{ni}$$

$$= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right]$$

$$\quad + (1 - \pi_0)\frac{1}{2b} \exp\left[-\frac{t_{ni}}{b} + \frac{\sigma^2}{2nb^2}\right] \times \Phi_+(0)$$

$$\quad + (1 - \pi_0)\frac{1}{2b} \exp\left[\frac{t_{ni}}{b} + \frac{\sigma^2}{2nb^2}\right] \times (1 - \Phi_-(0)), \tag{2.3}$$

where $\Phi_+(\cdot)$ is the cdf of normal distribution with mean $(t_{ni} + \sigma^2/(nb))$ and variance σ^2/n and $\Phi_-(\cdot)$ is the cdf of normal distribution with mean $(t_{ni} - \sigma^2/(nb))$ and variance σ^2/n. In contrast with normal case, the m_i function is the weighted sum of three values: normal density and two terms with cumulative distributions.

For both cases, we can find the values π_0, σ^2 and α or b maximizing the product of function $m_i(\cdot)$. They are called empirical Bayes estimators. We use the estimators in this paper.

3. Posterior Computation

In this section, posterior calculations are accompanied with the priors based on Section 2 with π_0, σ^2 and α or b at their fixed empirical Bayes estimates. For the normal prior case, Lee and Oh (2013) showed the necessary derivation given below:

$$\eta_{ni}|t_n \sim \pi_0(t_{ni})\delta_0 + \pi_1(t_{ni})N\left(\frac{n}{\alpha + n}t_{ni}, \frac{\sigma^2}{\alpha + n}\right), \quad i = 1, 2, \ldots, p_n,$$

where

$$\pi_0(t) = \frac{\pi_0 \frac{\sqrt{n}}{\sigma}\phi\left(\frac{\sqrt{n}}{\sigma}t\right)}{\pi_0 \frac{\sqrt{n}}{\sigma}\phi\left(\frac{\sqrt{n}}{\sigma}t\right) + \pi_1 \frac{1}{\sigma\sqrt{1/n+1/\alpha}}\phi\left(\frac{1}{\sigma\sqrt{1/n+1/\alpha}}t\right)},$$

$\phi(\cdot)$ is the probability density function of standard normal distribution.

For the Laplace distribution prior case, we can derive the posterior distribution of η_{ni} as below. The posterior distribution is divided into three parts. For the first term, we draw 0 for η_{ni} with weight obtained from the first term of (2.3). For the second and third terms, their posterior distributions are distributed as truncated normal distribution. When η_{ni} is less than 0, the posterior for η_{ni} are drawn from truncated normal distribution with mean $t_{ni} + \sigma^2/(nb)$ and variance σ^2/n with weight obtained from the second term of (2.3). Otherwise, the posterior for η_{ni} are drawn a truncated normal distribution with mean $t_{ni} - \sigma^2/(nb)$ and variance σ^2/n with weight obtained from the last term of (2.3).

$$
\begin{aligned}
\pi(&\eta_{ni}|t_{ni}, \pi_0, \sigma, b) \\
&\propto f(t_{ni}|\eta_{ni}, \sigma)\, p(\eta_{ni}|\pi_0, b) \\
&= \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{n(t_{ni} - \eta_{ni})^2}{2\sigma^2}\right] \times \left(\pi_0 I(\eta_{ni} = 0) + (1 - \pi_0)\frac{1}{2b}\exp\left[-\frac{|\eta_{ni}|}{b}\right] I(\eta_{ni} \neq 0)\right) \\
&= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right] I(\eta_{ni} = 0) \\
&\quad + (1 - \pi_0)\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\frac{1}{2b}\exp\left[-\frac{n(t_{ni} - \eta_{ni})^2}{2\sigma^2} - \frac{|\eta_{ni}|}{b}\right] I(\eta_{ni} \neq 0) \\
&= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right] I(\eta_{ni} = 0) \\
&\quad + (1 - \pi_0)\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\frac{1}{2b}\exp\left[-\frac{n(t_{ni} - \eta_{ni})^2}{2\sigma^2} + \frac{\eta_{ni}}{b}\right] I(\eta_{ni} < 0) \\
&\quad + (1 - \pi_0)\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\frac{1}{2b}\exp\left[-\frac{n(t_{ni} - \eta_{ni})^2}{2\sigma^2} - \frac{\eta_{ni}}{b}\right] I(\eta_{ni} > 0) \\
&= \pi_0 \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{nt_{ni}^2}{2\sigma^2}\right] I(\eta_{ni} = 0) \\
&\quad + (1 - \pi_0)\frac{1}{2b}\exp\left[\frac{t_{ni}}{b} + \frac{\sigma^2}{2nb^2}\right]\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\exp\left[-\frac{n}{2\sigma^2}\left(\eta_{ni} - \left(t_{ni} + \frac{\sigma^2}{nb}\right)\right)^2\right] I(\eta_{ni} < 0) \\
&\quad + (1 - \pi_0)\frac{1}{2b}\exp\left[-\frac{t_{ni}}{b} + \frac{\sigma^2}{2nb^2}\right]\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\exp\left[-\frac{n}{2\sigma^2}\left(\eta_{ni} - \left(t_{ni} - \frac{\sigma^2}{nb}\right)\right)^2\right] I(\eta_{ni} > 0).
\end{aligned}
$$

3.1. BMA and MPM estimators

We consider two estimators of η_n. The Bayesian model averaging (BMA) estimator is a mixture of posterior mean vector and the weights are derived from (2.2). In normal case, the BMA estimator is

$$
\hat{\eta}_n^{BMA} = \pi_1(t_n)\frac{n}{\alpha + n}t_n.
$$

The median probability model (MPM) estimator is a vector where elements are 0 when $\pi_1(t_n)$ is less than 1/2. When $\pi_1(t_n)$ is bigger than or equal to 1/2, its elements are exactly the same as BMA

estimator. In the same manner, the MPM estimator for normal case is

$$\hat{\eta}_n{}^{MPM} = \frac{n}{\alpha + n} t_n I\left(\pi_1(t_n) > \frac{1}{2}\right).$$

On the other hand, in the Laplace case, both estimators are calculated more complicate because of truncation. The BMA estimator is

$$\hat{\eta}_n{}^{BMA} = \pi_1(t_n)\mu_n,$$

where $t_n = (t_{n1}, \ldots, t_{n,n_i})$ and $\mu_n = (\mu_{n1}, \ldots, \mu_{n,n_i})$.

For each i, when the t_{ni} is less than 0, μ_{ni} is defined as

$$t_{ni} + \frac{\sigma^2}{nb} + \frac{\phi\left(-\frac{t_{ni} + \frac{\sigma^2}{nb}}{\sigma/\sqrt{n}}\right)}{\Phi\left(-\frac{t_{ni} + \frac{\sigma^2}{nb}}{\sigma/\sqrt{n}}\right)}.$$

On the other hand, when the elements of t_n is larger than or equal to 0, μ_{ni} is defined as

$$t_{ni} - \frac{\sigma^2}{nb} - \frac{\phi\left(-\frac{t_{ni} - \frac{\sigma^2}{nb}}{\sigma/\sqrt{n}}\right)}{\Phi\left(\frac{t_{ni} - \frac{\sigma^2}{nb}}{\sigma/\sqrt{n}}\right)}.$$

We aggregate the two values with probabilities calculated above.

Posterior median of the parameter was calculated explicitly in the Johnstone and Silverman (2005), especially for the case when σ and n are 1. The median estimator of η_n is

$$\hat{\eta}_n{}^{MED} = \left(\hat{\eta}_{n1}^{MED}, \ldots, \hat{\eta}_{np}^{MED}\right).$$

For $t_{ni} > 0$, $\hat{\eta}_{ni}{}^{MED}$, the ith element of $\hat{\eta}_n{}^{MED}$ is

$$t_{ni} - \frac{\sigma^2}{bn} + \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left[1 - \left[\Phi\left\{\frac{\sqrt{n}}{\sigma}\left(t_{ni} - \frac{\sigma^2}{bn}\right) + \exp\left(\frac{2t_{ni}}{b}\right)\left(1 - \Phi\left(\frac{\sqrt{n}}{\sigma}\left(t_{ni} + \frac{\sigma^2}{bn}\right)\right)\right)\right\}/2(1 - \pi_0)\right]\right]$$

and for negative cases, we can derive the median by using formula $\mathrm{med}(x) = -\mathrm{med}(-x)$.

3.2. Prediction

Prediction error for new observation is derived from $\hat{\eta}_n$. Prior to that, we estimate the prediction value for new observation as below.

$$\begin{aligned}
E(y_{n+1}|y_1, \ldots, y_n) &= E\left[E\left(y_{n+1}|y_1, \ldots, y_n, \beta_{n+1}\right)|y_1, \ldots, y_n\right] \\
&= E\left[X_{n+1}\beta_{n+1}|y_1, \ldots, y_n\right] \\
&= X_{n+1}\hat{\beta}_n = X_{n+1}\left(\sqrt{n}A_1\Lambda_1^{-\frac{1}{2}}\hat{\eta}_n\right),
\end{aligned} \tag{3.1}$$

where $\hat{\eta}_n$ can be chosen $\hat{\eta}_n{}^{BMA}$ or $\hat{\eta}_n{}^{MPM}$ and X_{n+1} is a design matrix that new observations are added to $(n+1)$th row. The difference between true observation and (3.1) is prediction error at a given time.
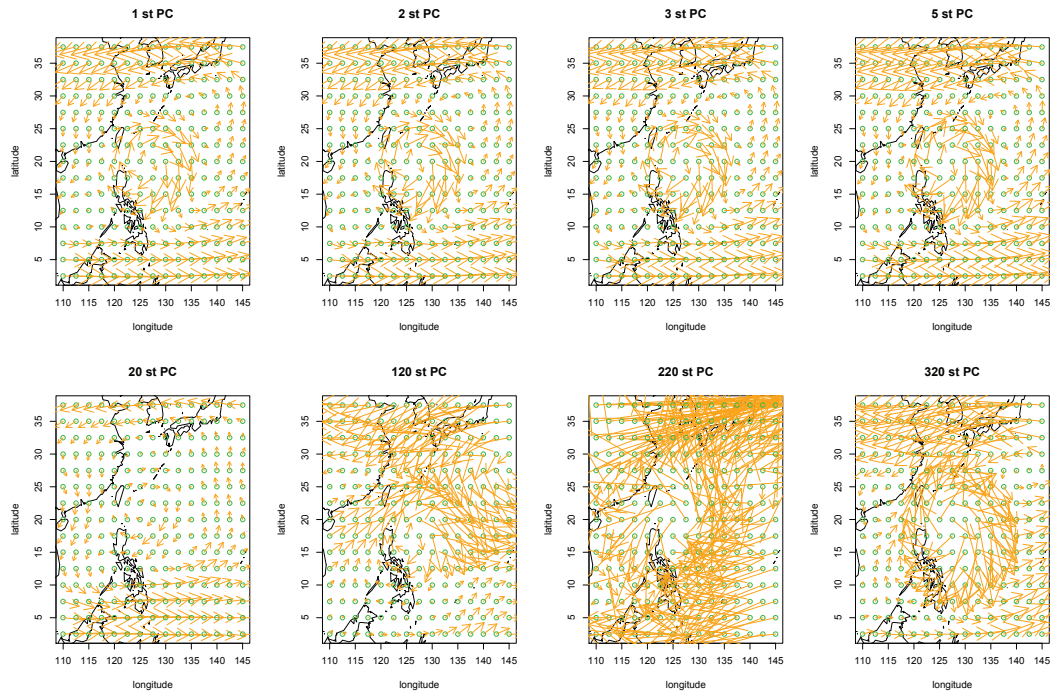
Figure 3: *PC vector (1st, 2nd, 3rd, 5th PC, upper row), (20th, 120th, 220th, 320th PC, lower row).*

4. Results

4.1. Track prediction

The Joint Typhoon Warning Center (JTWC) managed by U.S. Navy Observatory is an organization that analyzes and forecasts tropical depression. They provide annually the best track data. The best track data consist of the position and intensity of typhoons for the duration of its existence. We obtained the data from 1950 to 2013 and we obtained the wind field data from the NCEP. There are 225 points (including wind vector) and each vector can be divided into two components: east-west direction (u-component) and south-north direction (v-component). The ith row of X_n is made up for 225 u-components and 225 v-components of the ith time point. We set y_n be the vector of positions changes of typhoon in longitude or latitude. We use data of 3 years (2006–2008) and y_n has a total 887 time points.

PC vectors in the upper row in Figure 3 show a similar pattern that general flow rotates on center clockwise and the rest rotate reversely. On the other hand, certain PC vectors in the lower row show various patterns. The former have posterior probabilities over 0.5 and the latter have significantly lower probabilities. We use the whole PC in the analysis in this paper in order to lessen prediction error.

4.1.1. Comparison to Kim et al. (2007)

There are three lines in Figure 4: original typhoon track (black), predicted typhoon track by BMA (red) and predicted typhoon track by MPM (blue). The figure shows that BMA and MPM estimators work similarly and both predicted tracks do not follow the original track in the last stage of the
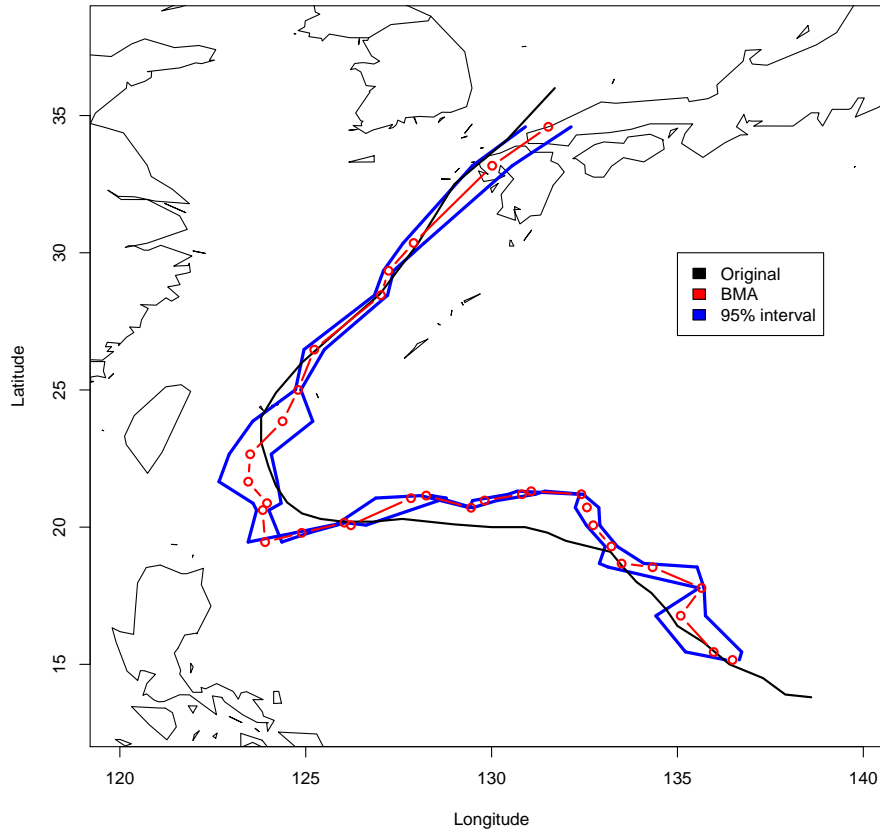
Figure 4: *Predicted typhoon track for 13th typhoon in 2006.*

typhoon. Detecting a curving point is very important for typhoon preparation. In this case, it is quite similar to the original curving point.

There are four graphs in Figure 5: BMA prediction error for longitudinal case, MPM prediction error for longitudinal case, MPM prediction error for latitudinal case and BMA prediction error for latitudinal case from upper left clockwise. For all cases, prediction errors are large in the endpiece because the typhoon disappeared irregularly in the last stage.

Figure 6 corresponds with lower-left in Figure 4. At that time point, wind vector at 225 points around the typhoon eye are presented in the figure and the prediction result is presented in the real scale (red line); consequently, there are no strong movement toward higher latitude and the red line lies in the same latitude. Therefore, it is reasonable to predict the typhoon track by using wind vector.

Kim *et al.* (2007) shows their prediction results for 14 typhoons in 2006. They give 24 hour ahead typhoon track forecasts. Table 1 shows the prediction errors of three forecasting methods and that of the prediction method given in Kim *et al.* (2007). A total of eight typhoons typhoons which correspond with JTWC data are included in the results of Table 1.

We divide into two parts: longitude and latitude. Prediction using the method in Section 3.2 is applied to each part separately. In order to aggregate the results from both parts, We calculate the
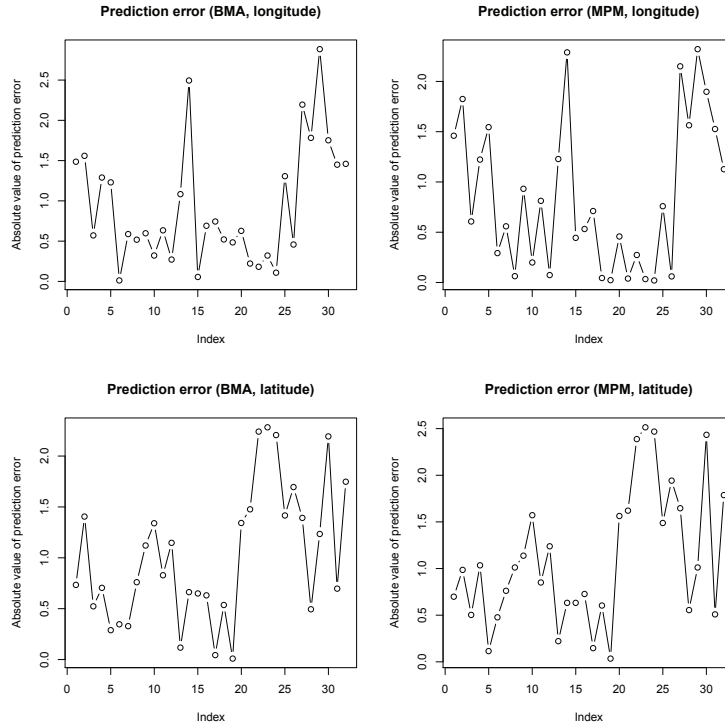
Figure 5: *Prediction errors for 13th typhoon in 2006.*

difference d between true track and prediction track by using Haversine formula as follows.

$$d = 2r \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right),$$

where r is a radius of the earth, ϕ_1, ϕ_2 are a true latitude and a prediction latitude and λ_1, λ_2 are a true longitude and a prediction longitude respectively. It is necessary to use this formula because the distance between two points becomes different when the latitude is higher.

First, we consider the case that all parameters are fixed. π_0 is fixed with 0.5 since it means that the probability to choose degenerate distribution or normal distribution is same as 0.5. σ is fixed with the least square estimator of the regression model for longitude and latitude, respectively. α is fixed with 1 and it means that the variance of each y_i is the same as η_i's

Second, the values for all parameters are fixed with empirical Bayes estimates. For both cases, π_0 are so high that the probability of choosing degenerate distribution is larger than the other.

Third, for Laplace distribution case, we set the values for all parameters with empirical Bayes estimates in the same manner.

In most cases, BMA estimator predict better than MPM estimator because the latter shrink many parameters to 0. When σ is fixed with least square estimator of general regression model, normal prior works better than Laplace prior; however, Laplace prior works better than Normal prior when σ is fixed with empirical Bayes estimator. Actually, under the appropriate conditions, the prediction error in this paper is smaller than the one calculated by MTM in Kim *et al.* (2007).
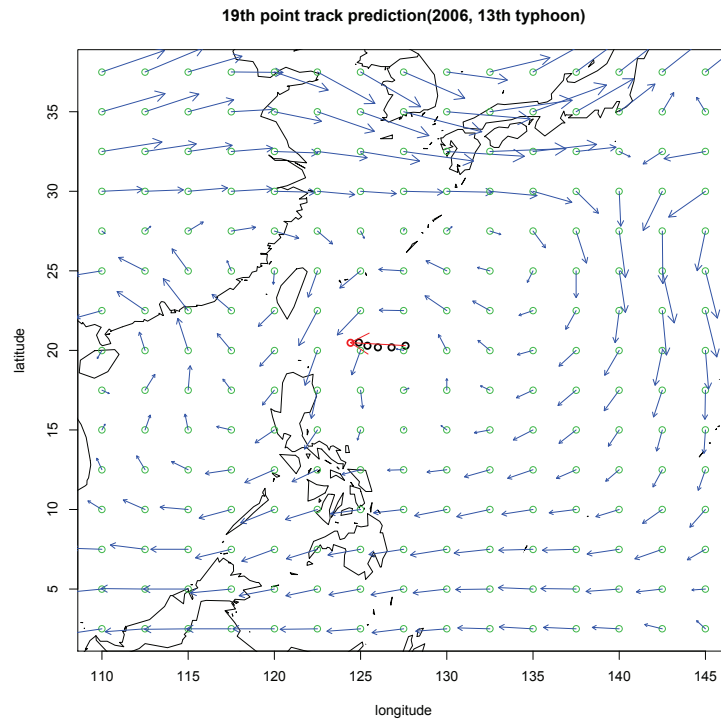
19th point track prediction(2006, 13th typhoon)

Figure 6: *Wind vector (19th point in the 13th typhoon in 2006).*

Table 1: Prediction errors for BMA, MPM and MTM cases

TP ID	BMA			MPM			Median	MTM
	(1)	(2)	(3)	(1)	(2)	(3)	(3)	(4)
0606	183.3	205.2	193.7	191.2	204.6	197.4	179.3	186.5
0607	186.8	207.6	203.6	193.5	207.6	217.3	202.3	221.0
0608	106.8	209.9	108.3	104.5	209.8	109.5	115.2	152.1
0609	141.1	223.4	136.8	152.1	224.0	144.7	143.3	143.8
0610	192.0	244.7	191.7	190.1	244.2	199.8	192.0	182.5
0613	163.4	241.3	152.8	162.7	241.3	149.1	155.6	180.1
0614	195.0	228.8	191.5	201.6	228.6	193.9	200.8	166.8
0618	199.6	246.5	207.2	202.7	246.6	213.9	209.9	186.5
Ave	172.8	227.2	174.4	176.3	227.1	178.8	176.2	178.2
sd	29.8	16.0	28.9	31.4	16.1	35.7	31.1	31.8

(1) $\hat{\sigma}$ = 1.26 (lon), 0.95 (lat), empirical Bayes estimates of σ when π_0 and α are fixed with 0.5 and 1 respectively with normal prior, (2) $\hat{\pi}_0$ = 0.97 (lon), 0.96 (lat), $\hat{\sigma}$=0.06 (lon), 0.04 (lat), $\hat{\alpha}$ = 0.003 (lon), 0.004 (lat), empirical Bayes estimates of π_0, σ, and α with normal prior, (3) $\hat{\pi}_0$ = 0.96 (lon), 0.94 (lat), $\hat{\sigma}$ = 1.6 (lon), 1.1 (lat), $\hat{b}$ = 0.4 (lon), 0.3 (lat), empirical Bayes estimates of π_0, σ and α with Laplace prior, (4) MTM = moving-nest typhoon model.

5. Concluding Remarks

We applied empirical Bayes estimates that made computation fast. Within reasonable limits, we will give a prior distribution for some parameters in place of empirical Bayes estimates.

In this paper, we mainly use 2-dimensional wind field data from averaging 3-dimensional data vertically. There is loss of information using 2-dimensional data because the original data has 3-

dimensional information; consequently, we should check the gap between the information.

In addition to above, we consider a peculiarity in the typhoon. Flow near the eye of typhoons have very small effect because it just rotates on the center. The wind near the eye of typhoon is very strong; however, it rarely influences the track of typhoon. In this paper, we use all the wind field data including flow near the eye of typhoons and it may cause a larger prediction error.

In the future works, we will use all the 3-dimensional data without loss of information and it yields heavy computation load. We should suggest a method that control a huge dataset with fast computation without loss of accuracy. Furthermore, we should distinguish the flow near the eye of the typhoon and the flow moving the typhoon in order to lose the error generated by the flow near the eye of the typhoon.

We will face a sparsity problem when the number of variables increases. In some cases, the elements of MPM estimator will be all 0's. We should control the condition adjust to situation because it will be the most important problem.

Finally, we can give the prior distribution for some parameters in place of using empirical Bayes estimators.

References

Johnstone, I. M. and Silverman, B. W. (2005). Empirical Bayes selection of wavelet thresholds, *The Annals of Statistics*, **33**, 1700–1752.

Kim, J. H., Choo, G. M., Kim, B. J., Won, S. H. and Kwon, H. J. (2007). Performance of MTM in 2006 typhoon forecast, *Atmosphere*, **17**, 207–216.

Lee, J. and Oh, H. S. (2013). Bayesian regression based on principal components for high-dimensional data, *Journal of Multivariate Analysis*, **117**, 175–192.

Rozanova, O. S., Yu, J. L. and Hu, C. K. (2010). Typhoon eye trajectory based on a mathematical model: Comparing with observational data, *Nonlinear Analysis: Real World Applications*, **11**, 1847–1861.

Sohn, K. T. and Kim, S. D. (2001). Typhoon track prediction using dynamic linear model, *Atmosphere*, **11**, 125–128.

Song, H. J., Huh, S. H., Kim, J. H., Ho, C. H. and Park, S. K. (2005). Typhoon track prediction by a support vector machine using data reduction methods. In Y. Hao, *et al.* (Eds.), *Lecture Notes in Computer Science: Vol. 3801. Computational Intelligence and Security* (pp. 503–511), Springer, Berlin.