PRELIMINARY RESULTS FROM THE ANALYSIS OF WIND COMPONENT ERROR

Donald P. Gaver
Patricia A. Jacobs

September 1991

Approved for public release; distribution is unlimited.

Prepared for:
Naval Oceanographic and Atmospheric Research Laboratory
Monterey, CA 93943-5006
This report was prepared in conjunction with research funded by the Naval Oceanographic and Atmospheric Research Laboratory, Monterey, California.

This report was prepared by:

DONALD P. GAVER, JR.
Professor of Operations Research

PATRICIA A. JACOBS
Professor of Operations Research

Reviewed by:

PETER PURDUE
Professor and Chairman
Department of Operations Research

Released by:

PAUL J. MARTO
Dean of Research
Preliminary Results from the Analysis of Wind Component Error

Estimation of mean square prediction error of wind components is required in the optimal interpolation (OI) process in numerical prediction of atmospheric variables. Statistical models with log-linear scale parameters which include covariates are described for the prediction error. Data from February and April of 1991 are used to fit the model parameters and to study the predictive ability of the models. This preliminary investigation indicates that observational and first guess wind components can be helpful in predicting mean square prediction error for wind components.
0. INTRODUCTION

Numerical meteorological models are used to assist in the prediction of weather. Each run of a numerical model produces forecasts of meteorological variables which are used as preliminary predictions of the future values of these variables. These initial predictions are referred to as first-guess values. In this paper first-guess values will refer to the most recent 12 hour forecasts.

In certain areas of the world observations of the values of forecasted variables become available, in our case the observations become available 12 hours after the first-guess values are computed. Prior to the next run of the numerical model a multivariate optimal interpolation analysis updates a first-guess value of a variable by adding to it a weighted observed value of the variable if it is available. The weight multiplying the observed value depends on estimates of the squared error of the first-guess value and the squared error of the observation; cf. Goerss et al. [1991, a, b]. Thus it is of importance to predict such first-guess squared errors.

The general problem of modeling and predicting mean square errors is important but not widely studied; see Efron (1986) and Jorgenson (1987). In the next section statistical models for the error of the first-guess are introduced. The models assume the error of the first-guess has mean 0 but has a scale parameter that is log-linear with suitable covariates, i.e. explanatory or regression variables.
Results are reported concerning the estimation of model parameters, and model cross-validation and predictive ability for $u, v$ wind component data from the months of February and April 1991. The data consist of measurements and 12 hour forecasts (first-guess values) from 93 stations in North America, 25N–75N. The forecasts are produced using the NOGAPS Spectral Forecast Model; cf. Hogan et al. Each station has measurement and first-guess values for every 12 hours; there are some missing observations. The first-guess values are subtracted from measurement values (if available) to obtain observations of the error of the first-guess. The results appear in Sections 3 and 4 and in Appendices B, C and D.

The results indicate that estimates of the variance of the error of first-guess wind components can be improved by using covariates which are functions of the wind components. Covariates using observed values of the wind components appear to have more predictive ability than those using first-guess values. Further exploratory work is needed to determine the degree with which these statistical results can be used to improve the forecasting ability of the numerical model.

1. THE MODELS

Let

\begin{align*}
U_0(t) &= \text{observed } u\text{-wind component at time } t \\
U_f(t) &= \text{first-guess } u\text{-wind component at time } t \\
V_0(t) &= \text{observed } v\text{-wind component at time } t \\
V_f(t) &= \text{first-guess } v\text{-wind component at time } t \\
\end{align*}

\begin{align*}
\tau(t) &= \left( (U_0(t) - U_0(t - 1))^2 + (V_0(t) - V_0(t - 1))^2 \right)^{1/2} \\
s(t) &= \left( U_0(t)^2 + V_0(t)^2 \right)^{1/2} \\
Y(t) &= U_0(t) - U_f(t) \quad \text{or} \quad Y(t) = V_0(t) - V_f(t) \\
\end{align*}

The models considered are as follows:
NORMAL MODELS:

One Variable Models

1. \{Y(t)\} are independent normally distributed random variables with mean 0 and variance

\[ \sigma^2_1(t; t) = \exp\{\alpha_1(1) + \beta_1(1)r(t)\}. \]  

(1)

2. \{Y(t)\} are independent normally distributed random variables with mean 0 and variance

\[ \sigma^2_2(t; t) = \exp\{\alpha_1(2) + \beta_1(2)s(t)\}. \]  

(2)

Two Variable Model

3. \{Y(t)\} are independent normally distributed random variables with mean 0 and variance

\[ \sigma^2_2(t) = \exp\{\alpha + \beta_1r(t) + \beta_2s(t)\}. \]  

(3)

CAUCHY MODELS:

While many measurement errors of physical quantities are approximately normal, especially "in the middle" of their distribution, there can well be thicker-than-normal tails and occasional extreme outliers. These attributes can have seriously degrading effects in regression-like problems; cf. Mosteller and Tukey (1977), Huber (1981) and Hampel (1986). The Cauchy distribution is a symmetric distribution with thicker tails than those of the normal distribution. Distributions with long straggling tails have the tendency to produce outlying values. The following models use the Cauchy distribution to represent and suitably compensate for more-thick-tailed measurement error than that of the Normal distribution.
One Variable Models

4. \( \{Y(t)\} \) are independent Cauchy random variables with scale parameter

\[
\sigma_1^2(t) = \exp\{\alpha_1(1) + \beta_1(1)r(t)\}.
\] (4)

5. \( \{Y(t)\} \) are independent Cauchy random variables with scale parameter

\[
\sigma_1^2(2; t) = \exp\{\alpha_1(2) + \beta_1(2)s(t)\}.
\] (5)

Two Variable Model

6. \( \{Y(t)\} \) are independent Cauchy random variables with scale parameter

\[
\sigma_2^2(t) = \exp\{\alpha + \beta_1 r(t) + \beta_2 s(t)\}.
\] (6)

The form of the Cauchy density function with scale parameter \( \sigma \) that is used is

\[
f(y) = \frac{1}{\pi\sigma} \left[ 1 + \frac{y^2}{\sigma^2} \right]^{-1} \quad \text{for} \quad -\infty < y < \infty.
\]

2. ESTIMATION OF PARAMETERS

For both normal and Cauchy models, the model parameters are estimated by maximum likelihood. A system of equations is obtained by setting the first partial derivative with respect to each parameter of the ln likelihood function equal to zero. The system of equations is solved numerically using Newton’s method to obtain the maximum likelihood estimates. The procedure for the normal models is given in Appendix A.
3. THE DATA ANALYSIS—FEBRUARY DATA

3.1 Observed Wind Covariate Models

In this subsection we report an assessment of the goodness of fit and cross-validation for the normal models (1)–(3) using observational wind components as covariates. There are six analyses; one for the $u$-wind component (respectively $v$-wind component) for each pressure level height. Each analysis proceeds along the same lines. In what follows by data we mean triples $(y(t), r(t), s(t))$.

In each analysis the data are randomly divided into two sets called DA and DB without regard to the values of the data.

The maximum likelihood parameter estimates for each model (1)–(3) are obtained for each set DA and DB and for all the data. The estimated values appear in Table 1. The estimated variances $\sigma_1^2(t), \sigma_2^2(t)$, are computed for the parameters estimated from DA and DB using (1)–(3)' for each data point in DA and DB.

The models are for the variances of the observations rather than the observations themselves. One possible procedure to assess goodness-of-fit and cross-validate the models is by binning the data. To assess models (1) and (3) the data $(y(t), r(t), s(t))$ are binned into 10 bins based on ordering the values of $r(t)$ from smallest to largest. The data in the first bin correspond to the smallest values of $r(t)$; the data in the 10th bin correspond to the largest values of $r(t)$. Each bin contains about $\frac{1}{10}$ of the data with the 10th bin containing a few more data. The averages of the estimated variances for models (1) and (3) are computed for each bin. The average $y(t)^2$ is also computed for each bin.
To assess models (2) and (3) the same procedure is used but the binning is based on the values of \( s(t) \).

Figures 1-24 present graphs of the \( \ln \) \([\text{average } y(t)^2]\) in each bin versus \( \ln \) \([\text{average estimated variance}]\) in each bin for models (1) and (3) and models (2) and (3). Figures 1, 5, 9, 13, 17, 21 (respectively 2, 6, 10, 14, 18, 22) show the logarithm of the average of the \( y(t)^2 \) values of DA (respectively DB) versus the logarithm of the average of the estimated variances for each bin using the estimated parameters from DA (respectively DB). If a model were perfect, a point should be close to the 45° line shown.

Figures 3, 7, 11, 15, 19, 23, (respectively 4, 8, 12, 16, 20, 24) present graphs of \( \ln \) \([\text{average } y(t)^2]\) of DA (respectively DB) versus \( \ln \) \([\text{average estimated variances}]\) using parameters estimated using data DB (respectively DA). Once again if the model were perfect, the points would be close to the 45° line.

Since the two-variate model (3) is shown with both one-variate models, it is possible to obtain some idea of the effect of the two different sets of bins on the \( \ln \) averages. In particular, the graphs corresponding to the 500 Mb height winds, Figures 9-16, show that the display of \( \ln \) averages can be quite sensitive to which variate is used to do the binning.

Keeping this binning sensitivity in mind, the figures suggest the following concerning the models using observed winds as covariates. It appears that of the two one-variate models, model (1) which uses \( r(t) \) as the covariate is the better. The two-variate model (3) appears not much better than model (1). If wind speed is used as the single covariate, it appears to overstate the variance; the addition of the second covariate \( r(t) \) in this case seems to tend to make the estimated variance smaller and bring the \( \ln \) average predicted variance in a bin closer to the \( \ln \) average \( y^2 \) in the bin.
Preliminary examination of ln average $y^2$ in bins and ln average model variances in bins for the Cauchy models suggests that the Cauchy models result in little or no improvement over the results of the normal model. The results of the Cauchy models will not be reported here.

Another way to assess goodness of fit and to cross validate is to evaluate the ln-likelihood for the different models at the parameter estimates. Larger values of the ln-likelihood suggest better model fit; cf. Cox and Hinkley [1974].

Table 2 presents the values of the ln-likelihood up to addition and multiplication of constants for the parameter estimates of Table 1; the function being evaluated is

$$\tilde{\ell} = -n\alpha - \sum_{i=1}^{n} x_i \beta - \sum_{i=1}^{n} y_i^2 \exp\{-\alpha + x_i \beta\}. \quad (7)$$

where $x_i \beta = \sum_{j} x_{ij} \beta_j$. The values of $\tilde{\ell}$ are presented for data DA (respectively DB) using the parameters fit using DA (respectively DB); these are values assessing goodness of fit; since maximum likelihood is the estimation procedure, the largest value of $\tilde{\ell}$ in each of these two rows is the one corresponding to the two-variate model. Values of $\tilde{\ell}$ are also presented for data DA (respectively DB) using the parameters fit using DB (respectively DA); these are values assessing cross-validation. The underlined value in each row is the maximum value in that row; the corresponding model provides the best model fit. The bold italicized value in each row is the maximum value for the two one-variate models; the corresponding one-variate model provides the best model fit between the two one-variate models.
### TABLE 1. NORMAL MODELS
PARAMETER ESTIMATES
OBSERVED WIND COVARIATES

<table>
<thead>
<tr>
<th>Pressure Height</th>
<th>Wind Comp.</th>
<th>Data Set</th>
<th>One-Variate Models</th>
<th>Two-Variate Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$r(t)$</td>
<td>$s(t)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\alpha$ $\beta$</td>
<td>$\alpha$ $\beta$</td>
</tr>
<tr>
<td>850</td>
<td>$u$</td>
<td>A</td>
<td>2.02 0.054</td>
<td>1.94 0.050</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>2.09 0.050</td>
<td>1.76 0.066</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>2.06 0.052</td>
<td>1.85 0.058</td>
</tr>
<tr>
<td>850</td>
<td>$v$</td>
<td>A</td>
<td>2.19 0.040</td>
<td>1.59 0.080</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>2.05 0.051</td>
<td>1.68 0.071</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>2.12 0.045</td>
<td>1.64 0.076</td>
</tr>
<tr>
<td>500</td>
<td>$u$</td>
<td>A</td>
<td>2.29 0.045</td>
<td>2.45 0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>2.18 0.054</td>
<td>2.19 0.029</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>2.23 0.050</td>
<td>2.32 0.024</td>
</tr>
<tr>
<td>500</td>
<td>$v$</td>
<td>A</td>
<td>2.31 0.039</td>
<td>2.27 0.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>2.24 0.042</td>
<td>2.14 0.028</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>2.28 0.041</td>
<td>2.21 0.025</td>
</tr>
<tr>
<td>250</td>
<td>$u$</td>
<td>A</td>
<td>3.12 0.034</td>
<td>2.48 0.034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>2.95 0.039</td>
<td>2.45 0.032</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>3.04 0.036</td>
<td>2.46 0.033</td>
</tr>
<tr>
<td>250</td>
<td>$v$</td>
<td>A</td>
<td>3.01 0.031</td>
<td>2.36 0.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>2.97 0.032</td>
<td>2.28 0.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>2.98 0.031</td>
<td>2.31 0.034</td>
</tr>
</tbody>
</table>

$r(t) = \left[\left((u(t) - u(t-1))^2 + (v(t) - v(t-1))^2\right)\right]^{1/2}$

$s(t) = \left[u(t)^2 + v(t)^2\right]^{1/2}$

NOTE: Data are divided into two sets randomly without regard to data values. One set is called A; the other is called B.
<table>
<thead>
<tr>
<th>Pressure Height</th>
<th>Wind Comp.</th>
<th>Data Set</th>
<th>Model</th>
<th>Constant</th>
<th>( r(t) )</th>
<th>( s(t) )</th>
<th>Two-Variate Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>850 u</td>
<td>A</td>
<td>A</td>
<td>-7695.9</td>
<td>-7596.5</td>
<td>-7591.6</td>
<td>-7538.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>-7746.9</td>
<td>-7661.8</td>
<td>-7560.5</td>
<td>-7540.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>-7747.5</td>
<td>-7663.9</td>
<td>-7571.1</td>
<td>-7553.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>-7696.4</td>
<td>-7598.5</td>
<td>-7601.9</td>
<td>-7551.7</td>
<td></td>
</tr>
<tr>
<td>850 v</td>
<td>A</td>
<td>A</td>
<td>-7759.1</td>
<td>-7703.4</td>
<td>-7505.5</td>
<td>-7498.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>-7707.6</td>
<td>-7614.0</td>
<td>-7512.8</td>
<td>-7489.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>-7708.2</td>
<td>-7620.6</td>
<td>-7515.4</td>
<td>-7497.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>-7759.7</td>
<td>-7710.3</td>
<td>-7508.2</td>
<td>-7507.4</td>
<td></td>
</tr>
<tr>
<td>500 u</td>
<td>A</td>
<td>A</td>
<td>-9454.3</td>
<td>-9314.6</td>
<td>-9405.3</td>
<td>-9299.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>-9518.7</td>
<td>-9296.8</td>
<td>-9376.7</td>
<td>-9239.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>-9519.5</td>
<td>-9303.2</td>
<td>-9397.9</td>
<td>-9257.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>-9455.2</td>
<td>-9320.4</td>
<td>-9425.2</td>
<td>-9315.8</td>
<td></td>
</tr>
<tr>
<td>500 v</td>
<td>A</td>
<td>A</td>
<td>-9317.7</td>
<td>-9317.1</td>
<td>-9243.0</td>
<td>-9174.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>-9258.3</td>
<td>-9140.7</td>
<td>-9161.1</td>
<td>-9091.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>-9259.1</td>
<td>-9142.8</td>
<td>-9165.1</td>
<td>-9094.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>-9318.4</td>
<td>-9219.2</td>
<td>-9247.5</td>
<td>-9177.5</td>
<td></td>
</tr>
<tr>
<td>250 u</td>
<td>A</td>
<td>A</td>
<td>-11265.4</td>
<td>-10657.4</td>
<td>-10431.5</td>
<td>-10344.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>-10782.7</td>
<td>-10389.8</td>
<td>-10249.0</td>
<td>-10149.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>-10829.6</td>
<td>-10403.9</td>
<td>-10261.7</td>
<td>-10162.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>-11319.3</td>
<td>-10673.4</td>
<td>-10445.1</td>
<td>-10358.6</td>
<td></td>
</tr>
<tr>
<td>250 v</td>
<td>A</td>
<td>A</td>
<td>-10417.8</td>
<td>-10259.4</td>
<td>-10094.9</td>
<td>-10032.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>-10783.1</td>
<td>-10181.5</td>
<td>-10050.1</td>
<td>-9960.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>-10814.9</td>
<td>-10182.5</td>
<td>-10051.8</td>
<td>-9961.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>-10446.4</td>
<td>-10260.3</td>
<td>-10096.3</td>
<td>-10033.2</td>
<td></td>
</tr>
</tbody>
</table>
The models considered are \( \{Y_i\} \) are independent normal with mean 0 and variance

\[
\sigma_i^2(t) = e^{\alpha} \quad \text{(Constant variance)} \tag{8}
\]

and models (1)–(3).

The two-variate model (3) maximizes the cross-validation value of \( \tilde{l} \) for data DA (respectively DB) with a model using parameters fit using DB (respectively DA). This suggests that both \( r(t) \) and \( s(t) \) together have predictive ability.

For the one-variate models (1) and (2) the cross-validation value of \( \tilde{l} \) for DA (respectively DB) using the parameters fit using DB (respectively DA) are equally divided as to whether \( r(t) \) by itself or \( s(t) \) by itself produces the higher value of \( \tilde{l} \). This suggests that neither variate by itself has obviously better predictive value than the other. The goodness of fit values of \( \tilde{l} \) for the one-variate models using DA (respectively DB) have a higher value of \( \tilde{l} \) associated with \( s(t) \) the majority of the time. This suggests that \( s(t) \) by itself provides a better description of the data than \( r(t) \) by itself.

Comparing the value of \( \tilde{l} \), \( \tilde{l}_c \), for DA (respectively DB) using the constant variance model (8) fit using DA (respectively DB) with the corresponding cross-validation value of \( \tilde{l} \) for DA (respectively DB) using models (2), (3) fit using DB (respectively DA) indicates the following. The values of \( \tilde{l} \) for models (2) and (3) fit with the other half of the data are larger than the corresponding value \( \tilde{l}_c \) for the constant variance model fit using the data to be modeled. This indicates that both models (2) and (3) fit with the other half of the data describe the data better than the best constant variance model (8) fit with the same data it is used to summarize.
3.2 First Guess Wind Covariate Models

In this section we report the results of using models (1)–(3) and (8) with first guess winds as covariates; the two covariates considered are

\[ r_f(t) = \left[ (U_f(t) - U_f(t-1))^2 + (V_f(t) - V_f(t-1))^2 \right]^{1/2} \]

and

\[ s_f(t) = \left[ U_f(t)^2 + V_f(t)^2 \right]^{1/2}. \]

The analysis is the same as in the previous subsection. The data sets DA and DB are the same as those in the previous subsection in each case.

The values of the parameter estimates appear in Table 3. The corresponding values of \( \tilde{c} \) appear in Table 4. Once again the underlined value of \( \tilde{c} \) is the largest value in each row; the bold italicized value \( \tilde{c} \) is the largest value between the two one-variate models.

In all but two cases the values of \( \tilde{c} \) for the observed wind covariates are larger than those for the first-guess wind covariates. This suggests that the observed wind components have better predictive and descriptive value than the first guess wind components.

Table 4 also indicates the following results concerning models using first guess wind covariates. Between the two one-variate models (1) and (2) the one-variate model using first guess wind speed always has the greater \( \tilde{c} \)-value. This suggests that first guess wind speed alone has better predictive and descriptive value than \( r_f(t) \) alone. The cross-validation values of \( \tilde{c} \) for data DA (respectively DB) using parameters fit with DB (respectively DA) are maximized about half the time using the one-variate model with \( s_f(t) \). The other times the maximal \( \tilde{c} \) is associated with the two-variate model.

10
<table>
<thead>
<tr>
<th>Pressure Heights</th>
<th>Wind Comp.</th>
<th>Data Set</th>
<th>One-variate Models $r_j(t)$</th>
<th>Two Variate Models $s_j(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>850 $u$</td>
<td>A</td>
<td>2.52</td>
<td>-0.006</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2.48</td>
<td>0.004</td>
<td>2.21</td>
</tr>
<tr>
<td></td>
<td>ALL</td>
<td>2.50</td>
<td>-0.0007</td>
<td>2.22</td>
</tr>
<tr>
<td>850 $v$</td>
<td>A</td>
<td>2.54</td>
<td>-0.004</td>
<td>2.34</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2.41</td>
<td>0.013</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>ALL</td>
<td>2.47</td>
<td>0.005</td>
<td>2.31</td>
</tr>
<tr>
<td>500 $u$</td>
<td>A</td>
<td>2.61</td>
<td>0.023</td>
<td>2.35</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2.68</td>
<td>0.019</td>
<td>2.48</td>
</tr>
<tr>
<td></td>
<td>ALL</td>
<td>2.65</td>
<td>0.021</td>
<td>2.42</td>
</tr>
<tr>
<td>500 $v$</td>
<td>A</td>
<td>2.71</td>
<td>0.006</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2.76</td>
<td>-0.002</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td>ALL</td>
<td>2.73</td>
<td>0.002</td>
<td>2.34</td>
</tr>
<tr>
<td>250 $u$</td>
<td>A</td>
<td>4.02</td>
<td>-0.005</td>
<td>2.66</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>3.46</td>
<td>0.021</td>
<td>3.01</td>
</tr>
<tr>
<td></td>
<td>ALL</td>
<td>3.74</td>
<td>0.009</td>
<td>2.79</td>
</tr>
<tr>
<td>250 $v$</td>
<td>A</td>
<td>3.49</td>
<td>0.007</td>
<td>3.00</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>3.53</td>
<td>0.016</td>
<td>2.87</td>
</tr>
<tr>
<td></td>
<td>ALL</td>
<td>3.50</td>
<td>0.012</td>
<td>2.92</td>
</tr>
<tr>
<td>Pressure Height</td>
<td>Wind Comp.</td>
<td>Data Set</td>
<td>Model</td>
<td>Constant</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>850 u</td>
<td>A A</td>
<td>-7695.9</td>
<td>-7695.1</td>
<td>-7667.6</td>
</tr>
<tr>
<td></td>
<td>B B</td>
<td>-7746.9</td>
<td>-7746.6</td>
<td>-7708.8</td>
</tr>
<tr>
<td></td>
<td>B A</td>
<td>-7747.5</td>
<td>-7749.5</td>
<td>-7709.8</td>
</tr>
<tr>
<td></td>
<td>A B</td>
<td>-7696.4</td>
<td>-7697.7</td>
<td>-7668.6</td>
</tr>
<tr>
<td>850 v</td>
<td>A A</td>
<td>-7759.1</td>
<td>-7758.7</td>
<td>-7745.4</td>
</tr>
<tr>
<td></td>
<td>B B</td>
<td>-7707.6</td>
<td>-7703.7</td>
<td>-7685.5</td>
</tr>
<tr>
<td></td>
<td>B A</td>
<td>-7708.2</td>
<td>-7711.3</td>
<td>-7687.3</td>
</tr>
<tr>
<td></td>
<td>A B</td>
<td>-7759.7</td>
<td>-7765.3</td>
<td>-7747.3</td>
</tr>
<tr>
<td>500 u</td>
<td>A A</td>
<td>-9454.3</td>
<td>-9433.7</td>
<td>-9391.1</td>
</tr>
<tr>
<td></td>
<td>B B</td>
<td>-9518.7</td>
<td>-9505.7</td>
<td>-9481.9</td>
</tr>
<tr>
<td></td>
<td>B A</td>
<td>-9519.5</td>
<td>-9507.5</td>
<td>-9486.2</td>
</tr>
<tr>
<td></td>
<td>A B</td>
<td>-9455.2</td>
<td>-9435.5</td>
<td>-9395.4</td>
</tr>
<tr>
<td>500 v</td>
<td>A A</td>
<td>-9317.7</td>
<td>-9316.0</td>
<td>-9281.2</td>
</tr>
<tr>
<td></td>
<td>B B</td>
<td>-9258.3</td>
<td>-9258.2</td>
<td>-9202.6</td>
</tr>
<tr>
<td></td>
<td>B A</td>
<td>-9259.1</td>
<td>-9261.5</td>
<td>-9205.6</td>
</tr>
<tr>
<td></td>
<td>A B</td>
<td>-9318.4</td>
<td>-9319.7</td>
<td>-9284.3</td>
</tr>
<tr>
<td>250 u</td>
<td>A A</td>
<td>-11265.4</td>
<td>-11263.9</td>
<td>-10907.3</td>
</tr>
<tr>
<td></td>
<td>B B</td>
<td>-10782.7</td>
<td>-10745.7</td>
<td>-10684.5</td>
</tr>
<tr>
<td></td>
<td>B A</td>
<td>-10829.6</td>
<td>-10846.5</td>
<td>-10739.9</td>
</tr>
<tr>
<td></td>
<td>A B</td>
<td>-11319.3</td>
<td>-11371.6</td>
<td>-10987.7</td>
</tr>
<tr>
<td>250 v</td>
<td>A A</td>
<td>-10417.8</td>
<td>-10414.2</td>
<td>-10349.3</td>
</tr>
<tr>
<td></td>
<td>B B</td>
<td>-10783.1</td>
<td>-10758.9</td>
<td>-10622.4</td>
</tr>
<tr>
<td></td>
<td>B A</td>
<td>-10814.9</td>
<td>-10796.9</td>
<td>-10658.7</td>
</tr>
<tr>
<td></td>
<td>A B</td>
<td>-10446.4</td>
<td>-10446.4</td>
<td>-10379.2</td>
</tr>
</tbody>
</table>
Comparing the value of $\hat{\ell}, \hat{\ell}_c$, for DA (respectively DB) using the constant variance model (8) fit using DA (respectively DB) with the cross-validation value of $\hat{\ell}$ for DA (respectively DB) using models (2), (3) fit using DB (respectively DA) indicates the following. The values of $\hat{\ell}$ for models (2) and (3) fit with the other half of the data are always larger than the corresponding value $\hat{\ell}_c$ for the constant variance model fit using the data to be modeled. This suggests that both models (2) and (3) fit with the other half of the data describe the data somewhat better than the best constant variance model (8) fit with the data to be described.

In summary, based on values of $\hat{\ell}$, when first guess winds are used as covariates it appears that the one-variate model using first guess wind speed is an attractive choice for predictive purposes. When observational winds are used as covariates, the two-variate model appears to have the best predictive value.

Assessing goodness of fit and cross-validation using values of $\hat{\ell}$ has the advantage of not being sensitive to binning. However, $\hat{\ell}$ may be sensitive to data sets DA and DB. Further work needs to be done to develop procedures to assess goodness of fit and for cross-validation. Procedures based on bootstrapping or jackknifing hold some promise.

4. THE DATA ANALYSIS—APRIL AND FEBRUARY DATA

In this section we report results of an assessment of goodness of fit for the normal models (1)–(3) for April data. We also report results concerning using a model whose parameters are fit using February data (respectively April) data to model April data (respectively February) data.
4.1 Observed Wind Covariate Models

In this subsection we report results for normal models (1)–(3) using observed wind components as covariates. There are six analyses; one for the $u$-wind component (respectively $v$-wind component) for each pressure height.

Table 5 shows the values of the parameter estimates for both the February data and April data. Table 6 shows the values of $\tilde{L}$ for February data (respectively April data) using parameters fit using February data (respectively April data). Values of $\tilde{L}$ are also presented for February data (respectively April data) using parameters fit using April data (respectively February data). Once again, larger values of $\tilde{L}$ indicate better model fit. The underlined value in each row is the maximum value in that row. The bold italicized value in each row is the maximum value of $\tilde{L}$ for the two one-variate models.

The values of $\tilde{L}$ for February data (respectively April data) using parameters fit using April data (respectively February data) are maximized by the two-variate model in all but one case; between the two one-variate models $\tilde{L}$ is the maximized half the time for the model involving $s(t)$.

Comparing the value of $\tilde{L}$, $\tilde{L}_c$, for the model of constant variance (8) for February (respectively April) data fit using February (respectively April) data with that for the prediction value of $\tilde{L}$ for the models (2)–(3) for February (respectively April) data fit using April (respectively February) data indicate the following. The values of $\tilde{L}$ for models (2) and (3) fit with data from the other month are always larger than the corresponding values of $\tilde{L}_c$ fit with the data of the same month. This suggests that models (2) and (3) fit using data from the other month have predictive value over a model of constant variance fit using the data that is to be modeled.
### Table 5. Normal Models Parameter Estimates

<table>
<thead>
<tr>
<th>Pressure Height</th>
<th>Wind Comp.</th>
<th>Data Set</th>
<th>( r(t) )</th>
<th>( s(t) )</th>
<th>In MSE = ( \alpha + \beta_1 r(t) + \beta_2 s(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>850 u</td>
<td>Feb.</td>
<td>( \alpha )</td>
<td>( \beta )</td>
<td>( \alpha )</td>
<td>( \beta_1 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.06</td>
<td>0.052</td>
<td>1.85</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>1.86</td>
<td>0.084</td>
<td>1.69</td>
<td>0.086</td>
</tr>
<tr>
<td>850 v</td>
<td>Feb.</td>
<td>2.12</td>
<td>0.045</td>
<td>1.64</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>1.83</td>
<td>0.089</td>
<td>1.69</td>
<td>0.090</td>
</tr>
<tr>
<td>500 u</td>
<td>Feb.</td>
<td>2.23</td>
<td>0.050</td>
<td>2.32</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>2.12</td>
<td>0.058</td>
<td>2.20</td>
<td>0.030</td>
</tr>
<tr>
<td>500 v</td>
<td>Feb.</td>
<td>2.28</td>
<td>0.041</td>
<td>2.21</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>2.02</td>
<td>0.065</td>
<td>1.97</td>
<td>0.041</td>
</tr>
<tr>
<td>250 u</td>
<td>Feb.</td>
<td>3.04</td>
<td>0.036</td>
<td>2.46</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>2.77</td>
<td>0.044</td>
<td>2.69</td>
<td>0.027</td>
</tr>
<tr>
<td>250 v</td>
<td>Feb.</td>
<td>2.98</td>
<td>0.031</td>
<td>2.31</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>2.73</td>
<td>0.041</td>
<td>2.61</td>
<td>0.027</td>
</tr>
</tbody>
</table>

\[
  r(t) = \left( (u(t) - u(t-1))^2 + (v(t) - v(t-1))^2 \right)^{1/2}
\]

\[
  s(t) = \left( u(t)^2 + v(t)^2 \right)^{1/2}
\]
<table>
<thead>
<tr>
<th>Pressure Height</th>
<th>Wind Comp.</th>
<th>Data Set</th>
<th>Model</th>
<th>Constant</th>
<th>( r(t) )</th>
<th>( s(t) )</th>
<th>Two-Variate Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>850 ( u )</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-15443.1</td>
<td>-15259.3</td>
<td>-15157.3</td>
<td>-15084.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-15709.1</td>
<td>-15311.6</td>
<td>-15169.9</td>
<td>-15043.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-15713.2</td>
<td>-15368.2</td>
<td>-15238.7</td>
<td>-15117.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-15447.0</td>
<td>-15321.2</td>
<td>-15240.3</td>
<td>-15171.9</td>
<td></td>
</tr>
<tr>
<td>850 ( v )</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-15467.0</td>
<td>-15320.8</td>
<td>-15019.6</td>
<td>-14992.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-15819.1</td>
<td>-15320.6</td>
<td>-15307.2</td>
<td>-15106.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-15827.8</td>
<td>-15429.9</td>
<td>-15386.7</td>
<td>-15251.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-15475.3</td>
<td>-15428.6</td>
<td>-15105.9</td>
<td>-15121.1</td>
<td></td>
</tr>
<tr>
<td>500 ( u )</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-18973.4</td>
<td>-18614.5</td>
<td>-18792.2</td>
<td>-18547.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-18504.1</td>
<td>-18083.1</td>
<td>-18270.7</td>
<td>-17969.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-18528.9</td>
<td>-18093.4</td>
<td>-18280.6</td>
<td>-17989.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-18999.9</td>
<td>-18625.3</td>
<td>-18804.3</td>
<td>-18576.6</td>
<td></td>
</tr>
<tr>
<td>500 ( v )</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-18576.4</td>
<td>-18358.8</td>
<td>-18406.2</td>
<td>-18267.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-18698.2</td>
<td>-17869.0</td>
<td>-18014.4</td>
<td>-17733.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-18699.0</td>
<td>-17938.0</td>
<td>-18086.9</td>
<td>-17796.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-18577.1</td>
<td>-18421.3</td>
<td>-18480.2</td>
<td>-18345.2</td>
<td></td>
</tr>
<tr>
<td>250 ( u )</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-22073.2</td>
<td>-21054.7</td>
<td>-20687.1</td>
<td>-20500.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-22712.2</td>
<td>-20364.3</td>
<td>-20658.3</td>
<td>-20195.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-22712.4</td>
<td>-20439.9</td>
<td>-20703.3</td>
<td>-20276.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-22073.4</td>
<td>-21127.5</td>
<td>-20726.0</td>
<td>-20591.9</td>
<td></td>
</tr>
<tr>
<td>250 ( v )</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-21216.0</td>
<td>-20441.4</td>
<td>-20145.7</td>
<td>-19992.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-21205.7</td>
<td>-20066.9</td>
<td>-20274.8</td>
<td>-19840.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-21240.2</td>
<td>-20065.6</td>
<td>-20338.7</td>
<td>-19924.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-21252.5</td>
<td>-20498.8</td>
<td>-20190.2</td>
<td>-20070.5</td>
<td></td>
</tr>
</tbody>
</table>
4.2 First Guess Wind Covariate Models

In this section we report results for normal models (1)–(3) using first guess wind components as covariates.

Table 7 shows the values of the parameter estimates for both February data and April data. Table 8 shows the values of $i$ for February data (respectively April data) using parameters fit using February data (respectively April data). Values of $i$ are also presented for February data (respectively April data) using parameters fit using April data (respectively February data). The underlined value in each row is the maximum value in that row. The bold italicized value in each row is the maximum value of $i$ for the two one-variate models.

The values of $i$ for the observed wind covariates are larger than those for the first guess wind covariates except for two values associated with the one-variate model using $s(t)$ to model $u$-wind component error at the 250 mb height for the model using parameters fit with the same data. This suggests that the observed wind covariates provide a better model of the data both in terms of goodness-of-fit and prediction.

The values of $i$ for February data (respectively April data) using parameters fit using April data (respectively February data) are maximized about half the time by the two-variate model and the other half the time by the one-variate model using the first guess wind speed $s(t)$.
TABLE 7. NORMAL MODELS
PARAMETER ESTIMATES
FIRST GUESS WIND COVARIATES

<table>
<thead>
<tr>
<th>Pressure Heights</th>
<th>Wind Comp.</th>
<th>Data Set</th>
<th>One-variate Models</th>
<th>Two-Variate Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>( r_\alpha(t) )</td>
<td>( s_\alpha(t) )</td>
</tr>
<tr>
<td>850</td>
<td>u</td>
<td>Feb.</td>
<td>2.50</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>2.42</td>
<td>0.022</td>
</tr>
<tr>
<td>850</td>
<td>v</td>
<td>Feb.</td>
<td>2.47</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>2.46</td>
<td>0.019</td>
</tr>
<tr>
<td>500</td>
<td>u</td>
<td>Feb.</td>
<td>2.65</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>2.50</td>
<td>0.031</td>
</tr>
<tr>
<td>500</td>
<td>v</td>
<td>Feb.</td>
<td>2.73</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>2.30</td>
<td>0.061</td>
</tr>
<tr>
<td>250</td>
<td>u</td>
<td>Feb.</td>
<td>3.74</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>4.00</td>
<td>-0.010</td>
</tr>
<tr>
<td>250</td>
<td>v</td>
<td>Feb.</td>
<td>3.50</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>3.26</td>
<td>0.025</td>
</tr>
<tr>
<td>Pressure Height</td>
<td>Wind Comp.</td>
<td>Data Set</td>
<td>Model</td>
<td>Constant</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>----------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>850</td>
<td>$u$</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-15443.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-15709.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-15713.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-15447.0</td>
</tr>
<tr>
<td>850</td>
<td>$v$</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-15467.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-15819.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-15827.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-15475.3</td>
</tr>
<tr>
<td>500</td>
<td>$u$</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-18973.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-18504.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-18528.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-18999.9</td>
</tr>
<tr>
<td>500</td>
<td>$v$</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-18576.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-18698.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-18699.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-18577.1</td>
</tr>
<tr>
<td>250</td>
<td>$u$</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-22073.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-22712.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-22712.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-22073.4</td>
</tr>
<tr>
<td>250</td>
<td>$v$</td>
<td>Feb.</td>
<td>Feb.</td>
<td>-21216.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Apr.</td>
<td>-21205.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apr.</td>
<td>Feb.</td>
<td>-21240.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb.</td>
<td>Apr.</td>
<td>-21252.5</td>
</tr>
</tbody>
</table>
A comparison of the value of $\hat{i}$, $\hat{i}_c$, for the constant variance model of February (respectively April) data fit using the same month February (respectively April) data and the prediction values of $\hat{i}$ for models (1)–(3) of February (respectively April) data fit using the other month of April (respectively February) indicate the following. A little fewer than half the time $\hat{i}_c$ is smaller than the corresponding values of $\hat{i}$ for models (1)–(3) fit with the other month’s data. This suggests that the first-guess wind speed models fit using the other month’s data may not describe the data as well as a constant variance model fit using the data being modeled. This may be an indication that models fit using first-guess February wind (respectively April wind) data are not good predictors of April (respectively February) wind component error.

4.3 Conclusions

Models (2) and (3) using observed wind components as covariates and fit using February (respectively April) data appear to have predictive value for April (respectively February) data. It is less clear if models (1)–(3) using first-guess wind components as covariates and fit using February (respectively April) data have predictive value for April (respectively February) wind component error data. It might be that models (1)–(3) fit with first-guess data from other Aprils (respectively Februaries) are better predictors of April (respectively February) wind component error. Alternatively, if first-guess winds are to be used as predictors, it might be worthwhile to develop a procedure to update the fitted model parameters using new data as it comes in.
REFERENCES


APPENDIX A
MAXIMUM LIKELIHOOD ESTIMATION FOR THE NORMAL MODEL

Let $Y_1, Y_2, \ldots, Y_n$ be independent normal random variables with mean 0 and variances

$$
\sigma_i^2 = \exp \left\{ \alpha + \sum_{j=1}^{p} x_{ij} \beta_j \right\} = \exp \{ \alpha + x_i \beta \} \quad i = 1, \ldots, n
$$

(A.1)

where $(x_{i1}, \ldots, x_{ip})$ are fixed explanatory variables associated with $Y_i$.

The likelihood function for this model is

$$
L(\alpha, \beta; y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} (\alpha + x_i \beta) \right\} \exp \left\{ -\frac{1}{2} y_i^2 \exp \{ -(\alpha + x_i \beta) \} \right\}.
$$

(A.2)

Hence, the ln-likelihood function is

$$
\ell(\alpha, \beta; y) = \frac{1}{2} \left[ -n \alpha - \sum_{i=1}^{n} x_i \beta - \sum_{i=1}^{n} y_i^2 \exp \{ -(\alpha + x_i \beta) \} \right] - n \frac{1}{2} \ln 2\pi.
$$

(A.3)

Computing partial derivatives of $\ell$ with respect to $\alpha$ and $\beta_j$ results in

$$
\frac{\partial}{\partial \alpha} \ell(\alpha, \beta; y) = -\frac{1}{2} \left[ -n + \sum_{i=1}^{n} y_i^2 \exp \{ -(\alpha + x_i \beta) \} \right]
$$

(A.4)

$$
\frac{\partial}{\partial \beta_j} \ell(\alpha, \beta; y) = \frac{1}{2} \left[ -\sum_{i=1}^{n} x_{ij} \sum_{i=1}^{n} y_i^2 \exp \{ -(\alpha + x_i \beta) \} x_{ij} \right]
$$

(A.5)

Setting $\frac{\partial}{\partial \alpha} \ell = 0$ results in the equation

$$
e^\alpha = \frac{1}{n} \sum_{i=1}^{n} y_i^2 \exp \{ x_i \beta \}.
$$

(A.6)
Setting \( \frac{\partial}{\partial \beta_j} \ell = 0 \) and replacing \( e^a \) by (A.6) yields the equation

\[
0 = f_j(\beta) = -\bar{x}_j \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta_j} + \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta_j} x_{ij} \tag{A.7}
\]

where \( \bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij} \).

Further,

\[
\frac{\partial}{\partial \beta_k} f_j(\beta) = +\bar{x}_j \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta} x_{ik} - \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta} x_{ik} x_{ik}. \tag{A.8}
\]

If \( f_k(\beta) = 0 \), then

\[
\bar{x}_k \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta} = \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta} x_{ik}. \tag{A.9}
\]

Substituting (A.9) into (A.8) yields

\[
\frac{\partial}{\partial \beta_k} f_j(\beta) = -\sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta} (x_{ij} x_{ik} - \bar{x}_j \bar{x}_k). \tag{A.10}
\]

An iteration of a Newton procedure to solve the system of equations \( 0 = f_j(\beta) \), \( (j = 1, \ldots, p) \) yields the system of linear equations

\[
0 = f_j(\beta) = f_j(\beta^0) + \sum_{k=1}^{p} \left[ \frac{\partial}{\partial \beta_k} f_j(\beta^0) \right] \beta_k - \beta_k^0 \tag{A.11}
\]

\[
= \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta^0} (x_{ij} - \bar{x}_j) - \frac{p}{\sum_{k=1}^{p} \sum_{i=1}^{n} y_i^2 e^{-\bar{x}_i \beta^0} [x_{ij} x_{ik} - \bar{x}_j \bar{x}_k] (\beta_k - \beta_k^0), \tag{A.12}
\]

where \( \beta^0 \) is the current value for \( \beta \). This system of linear equations is solved for \( \{\beta_k\} \). The Newton procedure is iterated until it converges. The resulting
Setting $\frac{\partial}{\partial \beta_j} \ell = 0$ and replacing $e^a$ by (A.6) yields the equation

$$0 = f_j(\beta) = -x_j \sum_{i=1}^n y_i^2 e^{x_i \beta} + \sum_{i=1}^n y_i^2 e^{x_i \beta} x_{ij}$$

(A.7)

where $x_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$.

Further,

$$\frac{\partial}{\partial \beta_k} f_j(\beta) = +x_j \sum_{i=1}^n y_i^2 e^{x_i \beta} x_{ik} + \sum_{i=1}^n y_i^2 e^{x_i \beta} x_{ij} x_{ik}.$$  

(A.8)

If $f_k(\beta) = 0$, then

$$x_k \sum_{i=1}^n y_i^2 e^{x_i \beta} = \sum_{i=1}^n y_i^2 e^{x_i \beta} x_{ij} x_{ik}.$$  

(A.9)

Substituting (A.9) into (A.8) yields

$$\frac{\partial}{\partial \beta_k} f_j(\beta) = -x_j \sum_{i=1}^n y_i^2 e^{x_i \beta} (x_{ij} x_{ik} - x_j x_k).$$  

(A.10)

An iteration of a Newton procedure to solve the system of equations $0 = f_j(\beta)$, $(j = 1, ..., p)$ yields the system of linear equations

$$0 = f_j(\beta) = f_j(\beta^0) + \sum_{k=1}^P \left[ \frac{\partial}{\partial \beta_k} f_j(\beta^0) \right] \beta_k - \beta_k^0$$

(A.11)

$$= \sum_{i=1}^n y_i^2 e^{x_i \beta^0} (x_{ij} - x_j) - \sum_{k=1}^P \sum_{i=1}^n y_i^2 e^{x_i \beta^0} (x_{ij} x_{ik} - x_j x_k) \left[ (\beta_k - \beta_k^0) \right]$$

(A.12)

where $\beta^0$ is the current value for $\beta$. This system of linear equations is solved for $\{\beta_k\}$. The Newton procedure is iterated until it converges. The resulting
Figure 2
850 MB U WIND; MODEL B ON DATA A; FEB OBS WIND

1VAR=R[T]=o; 2VAR=+; BIN ON R[T]

Figure 3
850 MB U WIND; MODEL A ON DATA B; FEB OBS WIND

Figure 4
850 MB V WIND; MODEL B ON DATA B; FEB OBS WIND

$1\text{VAR}=R[T]=0; 2\text{VAR}=+; \text{BIN ON } R[T]$
850 MB V WIND; MODEL A ON DATA B; FEB OBS WIND

1VAR = R[T] = o; 2VAR = +; BIN ON R[T]

LN AV MSE PER BIN

LN AV PRED MSE PER BIN

1VAR = WS[T] = o; 2VAR = +; BIN ON WS[T]

LN AV MSE PER BIN

LN AV PRED MSE PER BIN

Figure 8
500 MB U WIND; MODEL A ON DATA A; FEB OBS WIND

1VAR = R[T] = 0; 2VAR = +; BIN ON R[T]

Figure 9
500 MB U WIND; MODEL B ON DATA B; FEB OBS WIND

1VAR=R(T)=*; 2VAR=++; BIN ON R(T)

Figure 10
500 MB U WIND; MODEL B ON DATA A; FEB OBS WIND

1VAR=R[T]=*; 2VAR=++; BIN ON R[T]

Figure 11
Figure 12
500 MB V WIND; MODEL B ON DATA A; FEB OBS WIND

1VAR=R[T]=o; 2VAR=+; BIN ON R[T]

1VAR=WS[T]=o; 2VAR=+; BIN ON WS[T]

Figure 15
500 MB V WIND; MODEL A ON DATA B; FEB OBS WIND

1VAR=R[T]=•; 2VAR=++; B1N ON R[T]

![Graph 1: LN AV MSE PER BIN vs LN AV PRED MSE PER BIN](image)

1VAR=WS[T]=•; 2VAR=++; B1N ON WS[T]

![Graph 2: LN AV MSE PER BIN vs LN AV PRED MSE PER BIN](image)

Figure 16
250 MB U WIND: MODEL A ON DATA A; FEB OBS WIND

1VAR=R[T] = + ; 2VAR=+; BIN ON R[T]

Figure 17

LN AV PRED MSE PER BIN

LN AV MSE PER BIN
250 MB U WIND, MODEL B ON DATA B, FEB OBS WIND

\[ \text{VAR} = \text{T} \times \text{VAR} = \frac{\text{BIN} + \text{BIN}}{\text{BIN}} \]

Figure 18
250 MB U WIND; MODEL A ON DATA B; FEB OBS WIND

1VAR=R[T]=o; 2VAR=++; BIN ON R[T]

LN AV MSE PER BIN

LN AV PRED MSE PER BIN

1VAR=WS[T]=o; 2VAR=++; BIN ON WS[T]

LN AV MSE PER BIN

LN AV PRED MSE PER BIN

Figure 20
250 MB V WIND; MODEL B ON DATA B: FEB OBS WIND

$1 \text{VAR} = R[T] = 0.2 \text{VAR} = +\text{BIN ON R[T]}$

Figure 22
250 MB V WIND; MODEL B ON DATA A; FEB OBS WIND

1 VAR = R[T] = o; 2 VAR = +; BIN ON R[T]

Figure 23
250 MB V WIND; MODEL A ON DATA B; FEB OBS WIND

1VAR=R[T]=o; 2VAR=+; BIN ON R[T]

Figure 24
APPENDIX B. A GRAPHICAL ASSESSMENT OF GOODNESS OF FIT AND CROSS-VALIDATION OF MODELS OF FEBRUARY WIND COMPONENT MEAN SQUARE ERROR USING FIRST-GUESS WIND COVARIATES

In this appendix we present figures assessing goodness of fit and cross-validation of the normal models (1)–(3) with first-guess wind covariates fit to February data. As in subsection (3.2) the data is randomly divided into two sets called DA and DB without regard to the values of the data; these sets are the same as those in that section.

The maximum likelihood parameter estimates for each model (1)–(3) are obtained for each set DA and DB and appear in Table 3. The estimated variances $\sigma^2_1(1, t), \sigma^2_1(2, t), \sigma^2_2(t)$ are computed for the parameters estimated from DA and DB using (1)-(3) for each data point in DA and DB.

To assess models (1) and (3) the data $(y(t), r(t), s(t))$ are binned into 10 bins based on ordering the values of $r(t)$ from smallest to largest. The data in the first bin correspond to the smaller values of $r(t); the data in the 10^{th}$ bin correspond to the larger values of $r(t). Each bin contains about $\frac{1}{10}$ of the data with the 10^{th} bin containing a few more data. The averages of the estimated variances for models (1) and (3) are computed for each bin. The average $y(t)^2$ is also computed for each bin.

To assess models (2) and (3) the same procedure is used but the binning is based on values of $s(t)$.

Figures 1B-24B present graphs of the $\ln[\text{average } y(t)^2]$ in each bin versus $\ln[\text{average estimated variance}]$ in each bin for models (1) and (3) and models (2) and (3). Figures 1B, 5B, 9B, 13B, 17B, 21B (respectively 2B, 6B, 10B, 14B, 18B
22B) show the logarithm of the average of the $y(t)^2$ values of DA (respectively DB) versus the logarithm of the average of the estimated variances for each bin using the estimated parameters from DA (respectively DB). If a model were perfect, a point should be close to the 45° line shown. These figures assess goodness of fit.

Figures 3B, 7B, 11B, 15B, 19B, 23B (respectively 4B, 8B, 12B, 16B, 20B, 24B) present graphs of ln average $y(t)^2$ of DA (respectively DB) versus ln average estimated variances using parameters estimated using data DB (respectively DA). Once again if the model were perfect, the points would be close to the 45° line.

As suggested by the values of the ln-likelihood $\hat{L}$ in Tables 2 and 4, the figures for models using first guess covariates indicate weaker goodness of fit and weaker cross-validation than Figures 1-24 for models with observed wind speed covariates. Both goodness-of-fit and cross-validation appear to improve somewhat for higher pressure height levels; Figures 17B-24B. This suggests that models using first guess covariates have greater predictive and descriptive value at 250mb height levels. However, they appear to be not as good as models using observed wind speed as covariates.
850 MB U WIND: MODEL A ON DATA A; FIRST-GUESS WINDS

1 VAR=R[T]=*; 2 VAR=+; BIN ON R[T]

Figure 1B
850 MB U WIND; MODEL B ON DATA B; FIRST-GUESS WINDS

1 VAR=\text{R}[T]=\circ; 2 \text{VAR}=\text{+}; \text{BIN ON } \text{R}[T]

Figure 2B
850 MB U WIND;MODEL A ON DATA B;FIRST-GUESS WINDS

1 VAR=R[T]=o; 2 VAR=+; BIN ON R[T]

1 VAR=WS[T]=o; 2 VAR=+; BIN ON WS[T]

Figure 4B
850 MB V WIND; MODEL A ON DATA A; FIRST-GUESS WINDS

1 \text{VAR}=R[T]=\ast; 2 \text{VAR}+=; \text{BIN ON } R[T]

\text{LN AV MSE PER BIN}

\text{LN AV PRED MSE PER BIN}

1 \text{VAR}=W_S[T]=\ast; 2 \text{VAR}+=; \text{BIN ON } W_S[T]

\text{Figure 5B}
850 MB V-WIND; MODEL B ON DATA B; FIRST-GUESS WINDS

- Figure 6B

VAR[R] = 0.2, VAR[+] = BIN ON R[T]

VAR[WS] = 0.2, VAR[+] = BIN ON WS[T]

LN AV PRED MSE PER BIN

LN AV MSE PER BIN
850 MB V WIND; MODEL B ON DATA A; FIRST-GUESS WINDS

1 VAR=R[T]=--; 2 VAR=++; BIN ON R[T]

Figure 7B
850 MB V WIND: MODEL A ON DATA B: FIRST-GUESS WINDS

1 VAR=\text{R}^{[T]} = \text{e}^{0.2} \text{VAR} = +\text{BIN ON R}^{[T]}

Figure 8B
500 MB U WIND; MODEL A ON DATA A; FIRST-GUESS WINDS

1 VAR = R[T] = *; 2 VAR = +; BIN ON R[T]

Figure 9B
500 MB U WIND; MODEL B ON DATA B; FIRST-GUESS WINDS

1 VAR=R[T]=+; 2 VAR=-; BIN ON R[T]

Figure 10B
500 MB U WIND; MODEL B ON DATA A; FIRST-GUESS WINDS

1 VAR=R[T]=*; 2 VAR=+; BIN ON R[T]

Figure 11B
500 MB V WIND; MODEL A ON DATA A; FIRST-GUESS WINDS
1 VAR=R[T]; 2 VAR=WS[T]; BIN ON R[T]

Figure 13B
Figure 14B
500 MB V WIND; MODEL B ON DATA A; FIRST-GUESS WINDS

1 VAR=R[T]=+; 2 VAR=+; BIN ON R[T]

Figure 15B
500 MB V WIND: MODEL A ON DATA B; FIRST-GUESS WINDS

1 VAR = R[T] = o; 2 VAR = +; BIN ON R[T]

Figure 16B
250 MB U WIND; MODEL B ON DATA B; FIRST-GUESS WINDS

1 VAR = R[T] = 0; 2 VAR = +; BIN ON R[T]

Figure 18B
250 MB U WIND; MODEL B ON DATA A; FIRST-GUESS WINDS

1 VAR = R[T] = *; 2 VAR = +; BIN ON R[T]

Figure 19B
250 MB U WIND; MODEL A ON DATA B; FIRST-GUESS WINDS

Figure 20B
250 MB V WIND; MODEL A ON DATA A; FIRST-GUESS WINDS

1 VAR=R[T]=; 2 VAR=+; BIN ON R[T]

Figure 21B
250 MB V WIND; MODEL B ON DATA B; FIRST-GUESS WINDS

1 VAR=R[T] ; 2 VAR=+; BIN ON R[T]

![Graph showing a linear relationship between LN AV MSE PER BIN and LN AV PRED MSE PER BIN for first-guess winds.](image)

Figure 22B
250 MB V WIND; MODEL B ON DATA A; FIRST-GUESS WINDS

1 VAR=R[T]=○; 2 VAR=+; BIN ON R[T]

LN AV MSE PER BIN

FIGURE 23B
250 MB V WIND; MODEL A ON DATA B; FIRST-GUESS WINDS

1 VAR=R[T]=@; 2 VAR=+; BIN ON R[T]

Figure 24B
APPENDIX C. GRAPHICAL ASSESSMENT OF GOODNESS OF FIT AND CROSS-VALIDATION OF MODELS FOR FEBRUARY AND APRIL WIND COMPONENT MEAN SQUARE ERROR USING OBSERVED WIND COVARIATES

In this appendix we present graphs assessing goodness of fit and predictive ability of the normal models (1)–(3) with observed wind covariates fit to April and February data.

The maximum likelihood parameter estimates for each model (1)–(3) are obtained for both February and April data and are displayed in Table 5. The estimated variances $\sigma_1^2(1,t)$, $\sigma_1^2(2,t)$, $\sigma_2^2(t)$ are computed for the parameters estimated from February and April data using (1)-(3) for each data point in February and April.

To assess models (1) and (3) the data $(y(t), r(t), s(t))$, for each data set are binned into 10 bins based on ordering the values of $r(t)$ from smallest to largest. The data in the first bin correspond to the smaller values of $r(t)$; the data in the $10^{th}$ bin correspond to the larger values of $r(t)$. Each bin contains about $\frac{1}{10}$ of the data with the $10^{th}$ bin containing a few more. The averages of the estimated variances for models (1) and (3) are computed for each bin. The average $y(t)^2$ is also computed for each bin.

To assess models (2) and (3) the same procedure is used but the binning is done using $s(t)$.

Figures 1C-24C present graphs of the $\ln[\text{average } y(t)^2]$ in each bin versus $\ln[\text{average estimated variance}]$ in each bin for models (1) and (3) and models (2) and (3). Figures 1C, 5C, 9C, 13C, 17C, 21C (respectively 2C, 6C, 10C, 14C, 18C
22C) show the logarithm of the average of the $y(t)^2$ values for February (respectively April) versus the logarithm of the average of the estimated variances for each bin using the estimated parameters from February (respectively April). If a model were perfect, a point should be close to the 45° line shown. These figures assess goodness of fit.

Figures 3C, 7C, 11C, 15C, 19C, 23C (respectively 4C, 8C, 12C, 16C, 20C, 24C) present graphs of ln average $y(t)^2$ of February (respectively April) versus ln average estimated variances using parameters estimated using April (respectively February) data. Once again if the model were perfect, the points would be close to the 45° line. These figures assess the ability of models fit using February (respectively April) observed data to predict April (respectively February) wind component mean square error.

The figures indicate once again that the display of ln averages can be quite sensitive to which variate is used to do the binning.

Keeping this binning sensitivity in mind, the figures suggest the following. The two-variate model (3) appears to best describe and predict the mean square component wind error. Of the two one-variable models, model (1) which uses $r(t)$ as the covariate appears to be better. The one-variate model using $s(t)$ appears to tend to overstate the predicted mean square error. The addition of the second covariate $r(t)$ to the one-variate model using $s(t)$ appears to tend to decrease the predicted mean square error.
850 MB APR MODEL ON FEB DATA; OBS WIND

1VAR=R[T]=; 2VAR=++; BIN ON R[T]

Figure 3C
850 MB FEB MODEL ON APR DATA: OBS WIND
1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

Figure 4C
850 MB V WIND; APR MODEL ON FEB DATA; OBS WIND

1VAR=R[T]=•; 2VAR=+; BIN ON R[T]

Figure 7C
850 MB V WIND; FEB MODEL ON APR DATA; OBS WIND

1VAR=R[T]=o; 2VAR=+; BIN ON R[T]

Figure 8C
500 MB U WIND; APR 'ODEL ON APR DATA; OBS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

Figure 10C
500 MB U WIND; APR MODEL ON FEB DATA; OBS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T].

Figure 11C
500 MB V WIND; FEB MODEL ON FEB DATA; OBS WIND

1VAR = R[T] = o; 2VAR = +; BIN ON R[T]

1VAR = WS[T] = o; 2VAR = +; BIN ON WS[T]

Figure 13C
500 MB V WIND; APR MODEL ON APR DATA; OBS WIND

1VAR=R[T]=; 2VAR=; BIN ON R[T].

Figure 14C
500 MB V WIND: APR MODEL ON FEB DATA; OBS WIND
1VAR=R(T)++:2VAR=+;BIN ON R(T)

Figure 15C
500 MB V WIND; FEB MODEL ON APR DATA; OBS WIND

1VAR=R[T]=*;2VAR=+;BIN ON R[T]

1VAR=WS[T]=*;2VAR=+;BIN ON WS[T]

Figure 16C
250 MB U WIND; FEB MODEL ON FEB DATA; OBS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

\[ \text{LN AV MSE U WIND PER BIN} \]

\[ \text{LN AV PRED MSE PER BIN} \]

1VAR=WS[T]=*; 2VAR=+; BIN ON WS[T]

Figure 17C
250 MB U WIND; APR MODEL ON APR DATA; OBS WIND

1VAR=R[T]=; 2VAR=+; BIN ON R[T]

LN AV MSE U WIND PER BIN

LN AV PRED MSE PER BIN

1VAR=WS[T]=; 2VAR=+; BIN ON WS[T]

LN AV MSE U WIND PER BIN

LN AV PRED MSE PER BIN

Figure 18C
250 MB V WIND; FEB MODEL ON FEB DATA; OBS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

Figure 21C
250 MB V WIND; APR MODEL ON APR DATA; OBS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

Figure 22C
250 MB V WIND; APR MODEL ON FEB DATA; OBS WIND

1VAR=R[T]=.; 2VAR=+; BIN ON R[T]

Figure 23C
APPENDIX D. GRAPHICAL ASSESSMENT OF GOODNESS OF FIT AND CROSS-VALIDATION OF MODELS FOR FEBRUARY AND APRIL WIND COMPONENT MEAN SQUARE ERROR USING FIRST-GUESS WIND COVARIATES

In this appendix we present graphs assessing goodness of fit and predictive ability of the normal models (1)–(3) with first-guess wind covariates fit to April and February data.

The maximum likelihood parameter estimates for each model (1)–(3) are obtained for both February and April data and are displayed in Table 7. The estimated variances $\sigma^2_1(1,t)$, $\sigma^2_1(2,t)$, $\sigma^2_2(t)$ are computed for the parameters estimated from February and April data using (1)-(3) for each data point in February and April.

To assess models (1) and (3) the data $(y(t), r(t), s(t))$ for each data set are binned into 10 bins based on ordering the values of $r(t)$ from smallest to largest. The data in the first bin correspond to the smaller values of $r(t)$; the data in the 10th bin correspond to the larger values of $r(t)$. Each bin contains about $\frac{1}{10}$ of the data with the 10th bin containing a few more. The averages of the estimated variances for models (1) and (3) are computed for each bin. The average $y(t)^2$ is also computed for each bin.

To assess models (2) and (3) the same procedure is used but the binning is done using $s(t)$.

Figures 1D-24D present graphs of the ln[average $y(t)^2$] in each bin versus ln[average estimated variance] in each bin for models (1) and (3) and models (2) and (3). Figures 1D, 5D, 9D, 13D, 17D, 21D (respectively 2D, 6D, 10D, 14D,
18D 22D) show the logarithm of the average of the $y(t)^2$ values for February (respectively April) versus the logarithm of the average of the estimated variance for each bin using the estimated parameters from February (respectively April). If a model were perfect, a point should be close to the 45° line shown. These figures are an indication of goodness of fit.

Figures 3D, 7D, 11D, 15D, 19D, 23D (respectively 4D, 8D, 12D, 16D, 20D, 24D) present graphs of In average $y(t)^2$ of February (respectively April) versus In average estimated variances using parameters estimated using April (respectively February) data. Once again if the model were perfect, the points would be close to the 45° line. These figures assess the ability of models fit using February (respectively April) first-guess data to predict April (respectively February) wind component mean square error.

The figures indicate once again that the display of In averages can be quite sensitive to which variate is used to do the binning.

The figures indicate the following. As suggested by comparison of the In likelihood values, $\hat{l}$, of Tables 6 and 8 for models with observed wind covariates and first guess wind covariates, the figures suggest that models using first guess wind covariates do not describe or predict mean square error for wind components as well as models using observed wind components. The two-variate model appears to tend to produce smaller mean square errors than the one-variate models; this tendency is most striking in the figure with first guess wind speed being used as the single covariate.

The models fit using April first guess data appear to tend to be better descriptive and predictive models than those fit using February first guess data.
The figures indicating predictive ability (3D, 4D, 7D, 8D, 11D, 15D, 19D, 20D, 23D and 24D) correspond fairly well to the differences between the minimizing value of $\tilde{L}$ for the models with covariates and the value of $\tilde{L}$ for the constant model (no covariates) in the corresponding rows of Table 8. If the value of $\tilde{L}$ for the constant model is larger than any other values in the row, the corresponding figure for that row shows no association.
850 MB U WIND; APR MODEL ON APR DATA; 1ST GUESS WIND
1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

Figure 2D
850 MB U WIND; FEB MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=•; 2VAR=+; BIN ON R[T]

1VAR=WS[T]=•; 2VAR=+; BIN ON WS[T]

Figure 4D
850 MB V WIND; FEB MODEL ON FEB DATA; 1ST GUESS WIND

Figure 5D
850 MB V WIND; APR MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

Figure 6D
850 MB V WIND; APR MODEL ON FEB DATA; 1ST GUESS WIND

1VAR = R[T] = ○; 2VAR = +; BIN ON R[T]

Figure 7D
850 MB V WIND; FEB MODEL ON APR DATA; 1ST GUESS WIND

1VAR = R[T] = *; 2VAR = +; BIN ON R[T]

Figure 8D
500 MB U WIND; FEB MODEL ON FEB DATA; 1ST GUESS WIND

1VAR = R[T]; 2VAR = +; BIN ON R[T]

1VAR = WS[T]; 2VAR = +; BIN ON WS[T]

Figure 9D
500 MB U WIND; APR MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

1VAR=WS[T]=*; 2VAR=+; BIN ON WS[T]

Figure 10D
500 MB U WIND; APR MODEL ON FEB DATA; 1ST GUESS WIND

1VAR = R[T] = °; 2VAR = +; BIN ON R[T]

1VAR = WS[T] = °; 2VAR = +; BIN ON WS[T]

Figure 11D
500 MB V WIND; FEB MODEL ON FEB DATA; 1ST GUESS WIND
1VAR=\( R[T] \)=•; 2VAR=+; BIN ON \( R[T] \)

\[ \text{LN MSE V WIND PER BIN} \]

\[ \text{LN AV PRED MSE PER BIN} \]

\[ 1VAR=\text{WS}[T]=\text{•}; 2VAR=+; \text{BIN ON WS}[T] \]

Figure 13D
500 MB V WIND; APR MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=.; 2VAR=++; BIN ON R[T]

1VAR=WS[T]=.; 2VAR=++; BIN ON WS[T]

Figure 14D
500 MB V WIND; APR MODEL ON FEB DATA; 1ST GUESS WIND

1VAR = R[T] = o; 2VAR = +; BIN ON R[T]

1VAR = WS[T] = o; 2VAR = +; BIN ON WS[T]

Figure 15D
500 MB V WIND; FEB MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=*; 2VAR=+: BIN ON R[T]

Figure 16D
250 MB U WIND; APR MODEL ON FEB DATA; 1ST GUESS WIND

1VAR = R[T] = x; 2VAR = +; BIN ON R[T].

1VAR = WS[T] = x; 2VAR = +; BIN ON WS[T].

Figure 19D
250 MB U WIND; FEB MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=*; 2VAR=+; BIN ON R[T]

LN AV PRED MSE PER BIN

1VAR=WS[T]=*; 2VAR=+; BIN ON WS[T]

Figure 20D
250 MB V WIND; FEB MODEL ON FEB DATA; 1ST GUESS WIND

1VAR=R[T]=--; 2VAR=+; BIN ON R[T]

1VAR=WS[T]=--; 2VAR=+; BIN ON WS[T]

Figure 21D
250 MB V WIND; APR MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=.; 2VAR=+; BIN ON R[T]

Figure 22D
Figure 24D

250 MB V. WIND; FEBO MODEL ON APR DATA; 1ST GUESS WIND

1VAR=R[T]=0; 2VAR=+BIN ON R[T]

LN AV PRED MSE PER BIN

LN MSE V. WIND PER BIN

127
<table>
<thead>
<tr>
<th>No.</th>
<th>Distribution List</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Library (Code 0142)</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, CA 93943-5000</td>
</tr>
<tr>
<td>2.</td>
<td>Defense Technical Information Center</td>
</tr>
<tr>
<td></td>
<td>Cameron Station</td>
</tr>
<tr>
<td></td>
<td>Alexandria, VA 22314</td>
</tr>
<tr>
<td>3.</td>
<td>Office of Research Administration (Code 012)</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, CA 93943-5000</td>
</tr>
<tr>
<td>4.</td>
<td>Prof. Peter Purdue</td>
</tr>
<tr>
<td></td>
<td>Code OR-Pd</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, CA 93943-5000</td>
</tr>
<tr>
<td>5.</td>
<td>Department of Operations Research (Code 55)</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, CA 93943-5000</td>
</tr>
<tr>
<td>6.</td>
<td>Prof. Donald Gaver, Code OR-Gv</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, CA 93943-5000</td>
</tr>
<tr>
<td>7.</td>
<td>Prof. Patricia Jacobs</td>
</tr>
<tr>
<td></td>
<td>Code OR/Jc</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, CA 93943-5000</td>
</tr>
<tr>
<td>8.</td>
<td>Center for Naval Analyses</td>
</tr>
<tr>
<td></td>
<td>4401 Ford Avenue</td>
</tr>
<tr>
<td></td>
<td>Alexandria, VA 22302-0268</td>
</tr>
<tr>
<td>9.</td>
<td>Dr. David Brillinger</td>
</tr>
<tr>
<td></td>
<td>Statistics Department</td>
</tr>
<tr>
<td></td>
<td>University of California</td>
</tr>
<tr>
<td></td>
<td>Berkeley, CA 94720</td>
</tr>
</tbody>
</table>

128
10. Prof. W. R. Schucany
    Dept. of Statistics
    Southern Methodist University
    Dallas, TX 75222

11. Prof. D. C. Siegmund
    Dept. of Statistics
    Sequoia Hall
    Stanford University
    Stanford, CA 94305

12. Prof. H. Solomon
    Department of Statistics
    Sequoia Hall
    Stanford University
    Stanford, CA 94305

13. Dr. Ed Wegman
    George Mason University
    Fairfax, VA 22030

14. Dr. P. Welch
    IBM Research Laboratory
    Yorktown Heights, NY 10598

15. Dr. Neil Gerr
    Office of Naval Research
    Arlington, VA 22217

16. Prof. Roy Welsch
    Sloan School
    M.I.T.
    Cambridge, MA 02139

17. Dr. J. Abrahams, Code 1111, Room 607
    Mathematical Sciences Division, Office of Naval Research
    800 North Quincy Street
    Arlington, VA 22217-5000

18. Prof. J. R. Thompson
    Dept. of Mathematical Science
    Rice University
    Houston, TX 77001
19. Dr. P. Heidelberger
IBM Research Laboratory
Yorktown Heights
New York, NY 10598

20. Prof. J. B. Kadane
Dept. of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

21. Prof. J. Lehoczky
Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

22. Prof. M. Mazumdar
Dept. of Industrial Engineering
University of Pittsburgh
Pittsburgh, PA 15235

23. Prof. H. Chernoff
Department of Statistics
Harvard University
1 Oxford Street
Cambridge, MA 02138

24. Prof. Frank Samaniego
Statistics Department
University of California
Davis, CA 95616

25. Dr. Shlomo Halpin
Bellcore,
445 South Street
Morristown, NJ 07962-1910
(MRE 2L309)

26. Prof. Tom A. Louis
School of Public Health
University of Minnesota
Mayo Bldg. A460
Minneapolis, MN 55455
27. Dr. Nan Laird 
   Biostatistics Dept. 
   Harvard School of Public Health 
   677 Huntington Ave. 
   Boston, MA 02115

28. Dr. Marvin Zelen 
   Biostatistics Department 
   Harvard School of Public Health 
   677 Huntington Ave. 
   Boston, MA 02115

29. Dr. John Orav 
   Biostatistics Department 
   Harvard School of Public Health 
   677 Huntington Ave. 
   Boston, MA 02115

30. Prof. W. Stuetzle 
   Department of Statistics 
   University of Washington 
   Seattle, WA 98195

31. Prof. F. W. Mosteller 
   Department of Statistics 
   Harvard University 
   1 Oxford St. 
   Cambridge, MA 02138

32. Dr. D. C. Hoaglin 
   Department of Statistics 
   Harvard University 
   1 Oxford Street 
   Cambridge, MA 02138

33. Prof. N. D. Singpurwalla 
   George Washington University 
   Washington, DC 20052

34. Prof. George S. Fishman 
   Curr. in OR & Systems Analysis 
   University of North Carolina 
   Chapel Hill, NC 20742
35. Dr. Alan F. Petty .............................................................. 1
   Code 7930
   Naval Research Laboratory
   Washington, DC 20375

36. Prof. Bradley Efron ....................................................... 1
   Statistics Dept.
   Sequoia Hall
   Stanford University
   Stanford, CA 94305

37. Prof. Carl N. Morris ..................................................... 1
   Statistics Department
   Harvard University
   1 Oxford St.
   Cambridge, MA 02138

38. Dr. John E. Rolph .......................................................... 1
   RAND Corporation
   1700 Main St.
   Santa Monica, CA 90406

39. Prof. D. C. Siegmund .................................................. 1
   Dept. of Statistics, Sequoia Hall
   Stanford University
   Stanford, CA 94305

40. Operations Research Center, Rm E40-164 .......................... 1
    Massachusetts Institute of Technology
    Attn: R. C. Larson and J. F. Shapiro
    Cambridge, MA 02139

41. Arthur P. Hurter, Jr .................................................... 1
    Professor and Chairman
    Dept. of Industrial Engineering and Management Sciences
    Northwestern University
    Evanston, IL 60201-9990

42. Institute for Defense Analysis ........................................ 1
    1800 North Beauregard
    Alexandria, VA 22311

43. Prof. J. W. Tukey ...................................................... 1
    Statistics Dept., Fine Hall
    Princeton University
    Princeton, NJ 08540
44. Dr. Daniel H. Wagner
Station Square One
Paoli, PA  19301

45. Dr. Colin Mallows
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ  07974

46. Dr. D. Pregibon
AT&T Bell Telephone Laboratories
Mountain Avenue
Murray Hill, NJ  07974

47. Dr. Jon Kettenring
Bellcore
445 South Street
Morris Township, NJ  07962-1910

48. Prof. David L. Wallace
Statistics Dept., University of Chicago
5734 S. University Ave.
Chicago, IL  60637

49. Dr. S. R. Dalal
Bellcore
445 South Street
Morristown, NJ  07962-1910

50. Dr. D. F. Daley
Statistic Dept. (I.A.S.)
Australian National University
Canberra, A.C.T.  2606
AUSTRALIA

51. Koh Peng Kong
OA Branch, DSO
Ministry of Defense
Blk 29 Middlesex Road
SINGAPORE  1024

52. Professor Sir David Cox
Nuffield College
Oxford, OX1 INF
ENGLAND
53. Dr. A. J. Lawrence ................................................................. 1  
Dept. of Mathematics,  
University of Birmingham  
P. O. Box 363  
Birmingham B15 2TT  
ENGLAND

54. Dr. John Copas ................................................................. 1  
Dept. of Mathematics,  
University of Birmingham  
P. O. Box 363  
Birmingham B15 2TT  
ENGLAND

55. Dr. D. Vere-Jones ................................................................. 1  
Dept. of Math, Victoria Univ. of Wellington  
P. O. Box 196  
Wellington  
NEW ZEALAND

56. Dr. J. Goerss ................................................................. 50  
Naval Oceanographic and Atmospheric Laboratory  
Monterey, CA 93943-5006