JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

The Gaussian Predictability of Wind Speeds

Adam H. Monahan[1]

A.H. Monahan, School of Earth and Ocean Sciences, University of Victoria PO Box 3065 STN CSC Victoria, BC, Canada, V8W 3V6. (monahana@uvic.ca).

[1]School of Earth and Ocean Sciences,
University of Victoria, Victoria, BC,
Canada.

Abstract.

The linear predictability of wind speed using Gaussian predictors, relative to the predictability of the vector wind components, is considered. Analytic expressions for the correlation-based wind speed prediction skill are obtained in terms of the prediction skills of the vector wind components and their statistical moments. This analysis is facilitated by the assumption that the vector winds are Gaussian, and by the fact that fluctuations of the wind speed are highly correlated with those of its square. It is shown that:

1. at least one of the vector wind components is generally better predicted than the wind speed

2. wind speed predictions constructed from the predictions of vector wind components are more skillful than direct wind speed predictions

3. the linear predictability of wind speed (relative to that of the vector wind components) increases as vector wind fluctuations become smaller relative to their mean value.

These model results are shown to be broadly consistent with linear predictive skills assessed using sea surface wind observations from the SeaWinds scatterometer. Biases in the model predictions are shown to be related to the degree to which vector wind fluctuations are non-Gaussian.

1. Introduction

Surface winds exert a significant influence on surface exchanges of energy, momentum, and mass, and the associated kinetic energy represents a resource of potentially global importance. Winds are highly variable; the understanding and prediction of this variability are important for the characterisation of surface fluxes and the effective utilization of wind energy. Recent years have seen a rapid increase in the number of studies examining the predictability of wind speeds and wind power [e.g. *Lange and Focken*, 2005; *Costa et al.*, 2008; *Ma et al.*, 2009; *Soman et al.*, 2010; *Giebel et al.*, 2011]. Wind power prediction models may be entirely empirical, based on statistical relationships; or they may be dynamical, making use of physically-based equations of motion. A third class of hybrid models uses statistical post-processing to correct errors in local wind predictions from dynamical models [e.g. *Thorarinsdottir and Gneiting*, 2010].

While a number of studies have addressed direct predictions of wind speeds, relatively little attention has been paid to the prediction of vector wind components. While the wind speed is of most direct interest for many applications, the vector winds contain information about both speed and direction and have the practical advantage that their joint distribution is more closely Gaussian (and therefore more amenable to the tools of classical statistical prediction) than that of speed and direction. Consideration of the vector wind components also allows for a more direct connection to dynamical models (e.g. General Circulation Models, Numerical Weather Prediction models, or mesoscale models), in which the equations of motion are normally expressed in terms of the vector winds rather than the speed and direction. The resolved vector winds in such models represent gridscale averages of the components; depending on the autocorrelation spatial scale of the winds, the norm of the mean vector wind may not represent the mean wind speed over this scale [e.g. *Mahrt and Sun*, 1995].

This study considers relationships between the predictability by Gaussian predictors of the vector wind components and of wind speed, starting with the assumption that the vector wind fluctuations can be approximated as bivariate Gaussian. It is obvious of course that exact knowledge of the vector wind components provides exact knowledge of the wind speed. What is less clear is the extent to which wind speed fluctuations are predictable, when variability in the vector wind components is only imperfectly predicted.

Of the various measures that can be used to assess predictability, the two most common are mean squared error and correlation coefficient: for a predictand y with prediction y_p, these are respectively

$$\epsilon^2 = \mathsf{E}\left\{(y - y_p)^2\right\} \tag{1}$$

$$\begin{aligned} \mathrm{corr}(y, y_p) &= \frac{\mathrm{cov}(y, y_p)}{\mathrm{std}(y)\mathrm{std}(y_p)} \\ &= \frac{\mathsf{E}\{y y_p\} - \mathsf{E}\{y\}\mathsf{E}\{y_p\}}{\mathrm{std}(y)\mathrm{std}(y_p)}, \end{aligned} \tag{2}$$

where $\mathsf{E}\{x\}$ denotes the mathematical expectation (i.e. the population mean, also denoted $\mathrm{mean}(x)$) of the variable x, $\mathrm{std}(x)$ is the standard deviation of x, and $\mathrm{cov}(x, z)$ is the covariance of x with z. These measures are not independent; in fact, it can be shown that

$$\epsilon^2 = \underbrace{[\mathrm{mean}(y) - \mathrm{mean}(y_p)]^2}_{\mathrm{I}} + \underbrace{[\mathrm{std}(y) - \mathrm{std}(y_p)]^2}_{\mathrm{II}} + \underbrace{2\mathrm{std}(y)\mathrm{std}(y_p)[1 - \mathrm{corr}(y, y_p)]}_{\mathrm{III}} \tag{3}$$

[e.g. *Lange and Focken*, 2005]. Terms I and II in this expansion are errors associated with biases in the prediction of the mean and the standard deviation, respectively. Term

III represents errors associated with mismatches of fluctuations of the prediction and the predictand, as measured by the correlation.

The focus of this study is differences in the predictability of vector wind components and wind speed inherent to the mathematical relationship between these quantities. As such, we will consider an idealised setting in which all statistical features of the vector winds and wind speeds are assumed to be known. In this case, any prediction biases can be corrected and terms I and II in the mean squared error expansion Eqn. (3) can be eliminated by rescaling the mean and standard deviation of the predicted wind speeds to match the known true values. The remaining irreducible errors are associated with imperfect correlations between prediction and predictand. For this reason, this study will focus on correlations as the key measure of predictability.

Furthermore, we will study the predictability of wind speeds in terms of the predictions of vector wind components without specifying the provenance of these predictions. These may be imagined to come from either a statistical model (e.g. multiple linear regression) or from a dynamical model. We assume only that these predictions of the vector wind components have been constructed, and that they are Gaussian.

Section 2 describes the data to be used in this study and establishes notation. Section 3 reviews the probability distribution of wind speeds associated with Gaussian vector winds and introduces an important approximation from which all subsequent results follow. The direct predictability of vector wind components and wind speed using a single Gaussian predictor is analysed in Section 4. Section 5 considers the predictability of wind speed from a single predictor when speed is not predicted directly, but rather is constructed from the vector wind components predicted from a single predictor. This analysis is extended in Section 6 to consider wind speed predictability in the situation in which the vector wind components are predicted separately. Conclusions are presented in Section 7.

2. Data and Notation

The surface wind dataset considered in this study consists of level 3.0 gridded SeaWinds scatterometer equivalent neutral 10-m zonal and meridional winds between 60°S and 60°N from the NASA QuikSCAT satellite [*Jet Propulsion Laboratory*, 2001], available twice daily at a resolution of $0.25° \times 0.25°$ from 19 July 1999 to 23 November 2009. These data are available for download from the NASA Jet Propulsion Laboratory (JPL) Distributed Active Archive Center `http://podaac-www.jpl.nasa.gov/dataset/QSCAT_LEVEL_3`. Those data points flagged as having been possibly corrupted by rain were excluded from the analysis. No further processing of the dataset was carried out.

Throughout this study, u and v will denote generic orthogonal horizontal vector wind components. The precise orientation of these components will be arbitrary, unless specifically noted. Zonal and meridional wind components will be denoted by U and V respectively. The wind speed will be denoted by $w = \sqrt{u^2 + v^2} = \sqrt{U^2 + V^2}$. For notational convenience, the squared wind speed w^2 will sometimes be referred to as kinetic energy; strictly speaking, w^2 is twice the specific kinetic energy.

3. Probability Distribution of Wind Speeds

We first review what is known about the relationship between the probability density function (pdf) of vector winds and that of wind speed, under the assumption that the vector winds are Gaussian. The approximation of the joint distribution of vector wind components as Gaussian permits analytically tractable characterisations of the probability distribution and correlation structure of surface winds. While the observed pdfs of surface vector winds display marked deviations from Gaussianity [e.g. *Monahan*, 2006], the model results that follow from the neglect of this structure are sufficiently accurate to justify the use of the Gaussian assumption as a leading order approximation. A more detailed discussion of the errors introduced into the present analysis due to vector wind non-Gaussianity appears in Section 6.

For Gaussian vector winds with statistical moments:

$$\text{mean}(u) = \overline{u} \tag{4}$$

$$\text{mean}(v) = \overline{v} \tag{5}$$

$$\text{std}(u) = \sigma_u \tag{6}$$

$$\text{std}(v) = \sigma_v \tag{7}$$

$$\text{corr}(u,v) = r, \tag{8}$$

the joint pdf of u and v is given by

$$p_{uv}(u,v) = \frac{1}{2\pi\sigma_u\sigma_v(1-r^2)^{1/2}} \exp\left(-\frac{1}{2(1-r^2)}\left[\frac{(u-\overline{u})^2}{\sigma_u^2} - \frac{2r(u-\overline{u})(v-\overline{v})}{\sigma_u\sigma_v} + \frac{(v-\overline{v})^2}{\sigma_v^2}\right]\right). \tag{9}$$

[e.g. *Wilks*, 2005]. Through an appropriate orientation of the basis vectors, any one of the following simplifications can be made:

1. $\overline{v} = 0$ (the along- and cross-mean wind basis)
2. $\sigma_u = \sigma_v = \sigma$ (the equal variance basis)
3. $r = 0$ (the principal component basis)

Starting from the relationship between vector wind components and wind speed and direction:

$$\begin{aligned} u &= w\cos\theta && (10)\\ v &= w\sin\theta, && (11) \end{aligned}$$

the joint density of w and θ is given by

$$p_{w\theta}(w,\theta) = \begin{vmatrix} \frac{\partial u}{\partial w} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial w} & \frac{\partial v}{\partial \theta} \end{vmatrix} p_{uv}(w\cos\theta, w\sin\theta) = w p_{uv}(w\cos\theta, w\sin\theta), \qquad (12)$$

so the marginal pdf for the wind speed is

$$\begin{aligned} p_w(w) &= \int_{-\pi}^{\pi} p_{uv}(w\cos\theta, w\sin\theta)d\theta \\ &= \frac{w}{2\pi\sigma_u\sigma_v(1-r^2)^{1/2}} \times \qquad (13) \\ & \int_{-\pi}^{\pi} \exp\left(-\frac{1}{2(1-r^2)}\left[\frac{(w\cos\theta-\overline{u})^2}{\sigma_u^2} - \frac{2r(w\cos\theta-\overline{u})(w\sin\theta-\overline{v})}{\sigma_u\sigma_v} + \frac{(w\sin\theta-\overline{v})^2}{\sigma_v^2}\right]\right) d\theta. \end{aligned}$$

This integral does not generally admit a closed-form solution. Neither is there in general a simple expression for the mean wind speed, mean(w), in terms of the statistical parameters of the vector wind components. However, given mean(w), an expression for the wind speed variance follows relatively simply:

$$\begin{aligned} \mathrm{var}(w) &= \mathsf{E}\left\{w^2\right\} - (\mathsf{E}\left\{w\right\})^2 \\ &= \mathsf{E}\left\{u^2\right\} + \mathsf{E}\left\{v^2\right\} - \mathrm{mean}^2(w) \\ &= \sigma_u^2 + \sigma_v^2 + \overline{u}^2 + \overline{v}^2 - \mathrm{mean}^2(w). \qquad (14) \end{aligned}$$

For the special case of equal variance ($\sigma_u = \sigma_v = \sigma$) and uncorrelated ($r = 0$) fluctuations, the pdf of w is the Rice distribution:

$$p_w(w) = \frac{w}{\sigma^2}\exp\left(-\frac{\overline{\mathcal{U}}^2 + w^2}{2\sigma^2}\right) I_0\left(\frac{\overline{\mathcal{U}}w}{\sigma^2}\right), \qquad (15)$$

where $\overline{\mathcal{U}}^2 = \overline{u}^2 + \overline{v}^2$ and $I_k(z)$ is the associated Bessel function of the first kind of order k [*Rice*, 1945]. In this case, the mean wind speed is given by

$$\begin{aligned} \text{mean}(w) &= \int_0^\infty w p_w(w) dw \\ &= \sqrt{\frac{\pi}{2}}\sigma \exp\left(\frac{-\overline{\mathcal{U}}^2}{4\sigma^2}\right)\left[\left(1+\frac{\overline{\mathcal{U}}^2}{2\sigma^2}\right) I_0\left(\frac{\overline{\mathcal{U}}^2}{4\sigma^2}\right) + \frac{\overline{\mathcal{U}}^2}{2\sigma^2} I_1\left(\frac{\overline{\mathcal{U}}^2}{4\sigma^2}\right)\right]. \end{aligned} \quad (16)$$

In the limit that $\overline{\mathcal{U}} = 0$, the Rice distribution reduces to the Rayleigh distribution [e.g. *Cakmur et al.*, 2004; *Monahan*, 2006, 2007]. It is worth emphasizing that $\overline{\mathcal{U}}$ is not the mean wind speed, but rather the amplitude of the average vector winds. In general, these two quantities will differ.

In its general form, the wind speed marginal distribution derived from the vector wind pdf (Eqn. 13) is a mathematically complicated object. In particular, statistical measures of w cannot generally be partitioned into separate statistical measures of u and v and so closed-form analytic expressions relating the statistical features of wind speed to those of the vector wind components are not available. In contrast, the statistics of the kinetic energy w^2 separate easily into those of u and v. For example, for some variable Z, we have

$$\text{cov}(Z, w^2) = \text{cov}(Z, u^2) + \text{cov}(Z, v^2). \quad (17)$$

Knowledge of the relationship between Z and the vector wind components can therefore be used to make statements about the relationship between Z and w^2.

This result is of practical utility because we find that, for observed surface winds, fluctuations of wind speed are highly correlated with those of kinetic energy (Figure 1). From this, it follows that to a first approximation

$$\text{corr}(Z, w) \simeq \text{corr}(Z, w^2). \quad (18)$$

For example, estimates of $\mathrm{corr}(u, w)$ and $\mathrm{corr}(u, w^2)$ from the SeaWinds data between 60°S and 60°N are compared in the first panel Figure 2; the second panel of this Figure similarly compares $\mathrm{corr}(v, w)$ with $\mathrm{corr}(v, w^2)$. Clearly, the linear statistical relationships between w and the vector wind components are almost indistinguishable from those between w^2 and the components - a result which is not surprising, given the high correlation between w and w^2. A similar result was discussed in *Carlin and Haslett* [1982], in which it was noted that that fluctuations of w are highly correlated with those of $\sqrt{w}$.

A simple heuristic justification of the close linear relationship between fluctuations in w and w^2 follows from writing

$$\begin{aligned} w^2 &= (\mathrm{mean}(w) + \mathrm{std}(w)w')^2 \\ &= \mathrm{mean}^2(w) + 2\mathrm{mean}(w)\mathrm{std}(w)w' + \mathrm{std}^2(w)w'^2 \\ &= \mathrm{mean}^2(w)\left[1 + 2\frac{\mathrm{std}(w)}{\mathrm{mean}(w)}w' + \frac{\mathrm{std}^2(w)}{\mathrm{mean}^2(w)}w'^2\right], \end{aligned} \tag{19}$$

where w' is the standardized wind speed anomaly with mean zero and unit variance:

$$w' = \frac{w - \mathrm{mean}(w)}{\mathrm{std}(w)}. \tag{20}$$

Fluctuations in w^2 contain contributions that are both linear and quadratic in w'; the relative size of these terms depends on the ratio $\mathrm{std}(w)/\mathrm{mean}(w)$. When this ratio is less than one, the term linear in w' is dominant. An estimate of the spatial probability distribution of this ratio from the SeaWinds observations is presented in Figure 3. The ratio is almost everywhere less than 0.5, and in many places is considerably smaller. To a first approximation, fluctuations in w^2 are therefore proportional to those in w, and Eqn. (18) follows. A more comprehensive justification of the high correlation between w

and w^2 when std(w)/mean(w) < 1 is presented in Appendix A. We will make use of the approximation Eqn. (18) throughout this study.

4. Direct Prediction of Wind Speed with a Single Gaussian Predictor

As the simplest example of the relationship between linear predictive skills of the vector wind components and wind speed, we first consider the prediction of wind speed directly from the Gaussian predictor, x, such that

$$\mathrm{corr}(x,u) = \rho_u \quad (21)$$

$$\mathrm{corr}(x,v) = \rho_v. \quad (22)$$

Without loss of generality, we can take x as being of mean zero and unit variance. Furthermore, we can always orient the basis used to define the vector wind components so that $\rho_u \geq 0, \rho_v \geq 0$. Note that x itself may be a composite of a number of different predictors (e.g., x may result from a multiple linear regression prediction of w); we require only that x be Gaussian and possess the specified correlations with the vector wind components. Note also that there is not complete freedom in specifying the parameters ρ_u, ρ_v, and r. As the correlation matrix of x, u, v must be non-negative definite (i.e. its determinant must be non-negative), we obtain:

$$1 - r^2 \geq \rho_u^2 + \rho_v^2 - 2r\rho_u\rho_v. \quad (23)$$

As discussed above, the calculation of the correlation between x and w is difficult because w is not separable into the sum of individual contributions from u and v. Using the approximation

$$\mathrm{corr}(x,w) \simeq \mathrm{corr}(x,w^2), \quad (24)$$

and the definition

$$\text{corr}(x, w^2) = \frac{\text{cov}(x, w^2)}{\text{std}(x)\text{std}(w^2)}, \tag{25}$$

we first compute

$$\text{cov}(x, w^2) = \text{cov}(x, u^2) + \text{cov}(x, v^2). \tag{26}$$

The first term is given by

$$\begin{aligned} \text{cov}(x, u^2) &= \text{cov}(x, (\sigma_u u' + \overline{u})^2) \\ &= \sigma_u^2 \text{cov}(x, u'^2) + 2\sigma_u \overline{u} \text{cov}(x, u') \\ &= 2\sigma_u \overline{u} \rho_u, \end{aligned} \tag{27}$$

where the standardized anomaly u' is defined as

$$u' = \frac{u - \overline{u}}{\sigma_u}. \tag{28}$$

It is at this step that the Gaussianity of x and u become relevant, as we have used the fact that for (x, u) bivariate Gaussian, x and u'^2 are uncorrelated (although they are not generally independent). For non-Gaussian (x, u), no such general statement can be made. Although x and u'^2 are uncorrelated, nonzero correlations between x and u^2 arise when $\overline{u} \neq 0$ because fluctuations in u^2 can then be expanded into contributions from u' and u'^2. A geometric interpretation of this result in terms of the joint pdf of x and u^2 is presented in Appendix B. A similar calculation gives:

$$\text{cov}(x, v^2) = 2\sigma_v \overline{v} \rho_v. \tag{29}$$

Finally,

$$\begin{aligned}
\mathrm{var}(w^2) &= \mathsf{E}\left\{w^4\right\} - \left(\mathsf{E}\left\{w^2\right\}\right)^2 \\
&= \mathsf{E}\left\{u^4\right\} + 2\mathsf{E}\left\{u^2v^2\right\} + \mathsf{E}\left\{v^4\right\} - \left(\mathsf{E}\left\{u^2\right\}\right)^2 - 2\mathsf{E}\left\{u^2\right\}\mathsf{E}\left\{v^2\right\} - \left(\mathsf{E}\left\{v^2\right\}\right)^2 \\
&= 2\sigma_u^4 + 4\sigma_u^2\sigma_v^2 r^2 + 2\sigma_v^4 + 4\overline{u}^2\sigma_u^2 + 8\sigma_u\sigma_v\overline{u}\overline{v}r + 4\overline{v}^2\sigma_v^2 \\
&= 2(\sigma_u^2+\sigma_v^2)^2 + 4\sigma_u^2\sigma_v^2(r^2-1) + 4(\overline{u}\sigma_u + \overline{v}\sigma_v)^2 + 8\sigma_u\sigma_v\overline{u}\overline{v}(r-1),
\end{aligned} \tag{30}$$

so

$$\mathrm{corr}(x,w) \simeq \mathrm{corr}(x,w^2) = \frac{2\overline{u}\sigma_u\rho_u + 2\overline{v}\sigma_v\rho_v}{\sqrt{2\sigma_u^4 + 4\sigma_u^2\sigma_v^2r^2 + 2\sigma_v^4 + 4\overline{u}^2\sigma_u^2 + 8\sigma_u\sigma_v\overline{u}\overline{v}r + 4\overline{v}^2\sigma_v^2}}. \tag{31}$$

As noted above, we can align the coordinate system so that $\overline{v} = 0$, obtaining

$$|\mathrm{corr}(x,w)| \simeq \left|\frac{\sigma_u\overline{u}}{\sqrt{\frac{1}{2}(\sigma_u^4+\sigma_v^4) + \sigma_u^2\sigma_v^2r^2 + \sigma_u^2\overline{u}^2}}\right|\rho_u \leq \rho_u. \tag{32}$$

We see that the linear predictability (as measured by the correlation coefficient) of w by the Gaussian predictor x is always less than that of the along-mean wind component, by a factor determined by the ratio of the mean vector wind magnitude to the size of the fluctuations and the strength of the correlation between u and v. In particular, for $\overline{u} = 0$, x carries no direct linear predictive information regarding w irrespective of how good a predictor it is of u. Similarly, the predictive skill of vector wind fluctuations in the cross mean wind direction, ρ_v, has no bearing on the linear predictability of speed. Of course, good predictions of u and v will result in good predictions of u^2 and v^2, and these can then be used to predict w. The analysis of this two-step approach to predicting wind speed will be considered in the next section.

We can also consider Eqn. (31) in the principal component basis for which $r = 0$:

$$\begin{aligned}
|\mathrm{corr}(x,w)| &\simeq \left| \frac{\rho_u \sigma_u \overline{u} + \rho_v \sigma_v \overline{v}}{\left[\frac{1}{2}(\sigma_u^4 + \sigma_v^4) + \sigma_u^2 \overline{u}^2 + \sigma_v^2 \overline{v}^2\right]^{1/2}} \right| \\
&\leq \left| \frac{\sigma_u \overline{u}}{(\sigma_u^2 \overline{u}^2 + \sigma_v^2 \overline{v}^2)^{1/2}} \right| \rho_u + \left| \frac{\sigma_v \overline{v}}{(\sigma_u^2 \overline{u}^2 + \sigma_v^2 \overline{v}^2)^{1/2}} \right| \rho_v \\
&= \delta \rho_u + \sqrt{1-\delta^2}\rho_v, \qquad (33)
\end{aligned}$$

where

$$\delta = \left| \frac{\sigma_u \overline{u}}{(\sigma_u^2 \overline{u}^2 + \sigma_v^2)^{1/2}} \right| \leq 1 \qquad (34)$$

and so

$$\mathrm{corr}(x,w) \lesssim \max(\rho_u, \rho_v). \qquad (35)$$

The linear predictability of wind speed is bounded above by that of the better predicted principal component directions (but not below by the predictability of the worse predicted component).

Finally, it is observed that over the ocean vector wind fluctuations are to a good approximation isotropic (with $\sigma_u = \sigma_v = \sigma$) and uncorrelated [e.g. *Monahan*, 2007]. In this approximation, we obtain the particularly simple expression

$$\mathrm{corr}(x,w) \simeq \frac{\overline{u}\rho_u + \overline{v}\rho_v}{\sqrt{\sigma^2 + \overline{u}^2 + \overline{v}^2}} \leq \max(\rho_u, \rho_v). \qquad (36)$$

The correlation of the zonal wind U with w is given by Eqn. (31) with $\rho_u = 1$ and $\rho_v = r$; similarly, the correlation of the meridional wind V with w is given by $\rho_u = r$ and $\rho_v = 1$. Kernel density estimates of the relationship between observed and modelled values of $\mathrm{corr}(u,w)$ and $\mathrm{corr}(v,w)$ from the SeaWinds data are presented in Figure 4. There is a good relation overall between the modelled and observed correlation values. The model results are not perfect: in particular, there is a tendency for the Gaussian model to overestimate the magnitude of the correlations. That the model results should be imperfect is consistent with the fact that the vector winds are not Gaussian. The model error is more pronounced for corr(U, w) than for corr(V, w); similarly, the zonal winds display more pronounced non-Gaussianity than the meridional winds [e.g. *Monahan*, 2004]. Despite these errors, we see that the assumption of Gaussian vector wind fluctuations provides a good first-order approximation of the correlation structure between zonal wind, meridional wind, and wind speed.

5. Prediction of Wind Speed from Components: Single Predictor

The previous section related the direct linear predictability of wind speed from a single Gaussian predictor to the linear predictability of the vector wind components. That the predictions of non-Gaussian w from Gaussian x should be less skillful than the predictions of the Gaussian vector components u and v is not in and of itself surprising. What is less clear a priori is what improvement in the predictive skill of w, if any, can be obtained by predicting speed by combining the separate predictions of u and v. It is to this analysis that we now turn.

Instead of predicting w directly from x, u and v are first predicted from x and then these predictions are used to calculate w. Denoting the respective vector wind component predictions as u_p and v_p, we have

$$u_p = \rho_u \sigma_u x + \overline{u} \quad (37)$$

$$v_p = \rho_v \sigma_v x + \overline{v}, \quad (38)$$

and we can construct the wind speed prediction

$$w_p = \left(u_p^2 + v_p^2\right)^{1/2}. \quad (39)$$

The vector wind component prediction moments are then

$$\text{mean}(u_p) = \overline{u} \tag{40}$$

$$\text{mean}(v_p) = \overline{v} \tag{41}$$

$$\text{std}(u_p) = \rho_u \sigma_u \tag{42}$$

$$\text{std}(v_p) = \rho_v \sigma_v. \tag{43}$$

Note that the variances of the predictions of u and v are generally biased low. We can write

$$u = u_p + \varepsilon_u \tag{44}$$

$$v = v_p + \varepsilon_v, \tag{45}$$

where $\varepsilon_u, \varepsilon_v$ are respectively the prediction errors of u and v. Assuming that u_p, v_p are regression-based predictions, u_p and ε_u are uncorrelated (as are v_p and ε_v). The variance of u can be decomposed as the sum of the variances of u_p and ε_u; as the error is generally nonzero, the variance of u_p is generally less than that of u. It is a standard result [e.g. *Wilks*, 2005] that the standard deviation of u_p is given by Eqn. (42) and that we can interpret ρ_u as the fraction of the variance of u accounted for by the prediction u_p. Similarly, the standard deviation of v_p is given by Eqn. (43).

Consistent with the approximations used in the previous section, we take

$$\text{corr}(w_p, w) \simeq \text{corr}(w_p^2, w^2). \tag{46}$$

The covariance between w_p^2 and w^2 is given by

$$\text{cov}(w_p^2, w^2) = \text{cov}(u_p^2, u^2) + \text{cov}(u_p^2, v^2) + \text{cov}(v_p^2, u^2) + \text{cov}(v_p^2, v^2). \tag{47}$$

Computing the first of these terms:

$$
\begin{aligned}
\mathrm{cov}(u_p^2, u^2) &= \mathsf{E}\left\{u_p^2 u^2\right\} - \mathsf{E}\left\{u_p^2\right\} \mathsf{E}\left\{u^2\right\} \\
&= \rho_u^2 \sigma_u^4 \left(2\rho_u^2 + 1\right) + 5\rho_u^2\sigma_u^2\overline{u}^2 + \sigma_u^2\overline{u}^2 + \overline{u}^4 - (\rho_u^2\sigma_u^2 + \overline{u}^2)(\sigma_u^2 + \overline{u}^2) \\
&= 2\rho_u^2\sigma_u^2(\rho_u^2\sigma_u^2 + 2\overline{u}^2). \qquad (48)
\end{aligned}
$$

where we have made use of the fact that $\mathrm{corr}(u_p'^2, u'^2) = \rho_u^2$ (Appendix C). Similarly,

$$
\begin{aligned}
\mathrm{cov}(u_p^2, v^2) &= 2\rho_u^2\rho_v^2\sigma_u^2\sigma_v^2 + 4\rho_u\rho_v\sigma_u\sigma_v\overline{uv} = \mathrm{cov}(v_p^2, u^2) \qquad (49) \\
\mathrm{cov}(v_p^2, v^2) &= 2\rho_v^2\sigma_v^2(\rho_v^2\sigma_v^2 + 2\overline{v}^2), \qquad (50)
\end{aligned}
$$

so

$$
\mathrm{cov}(w_p^2, w^2) = 2(\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2)^2 + 4(\rho_u\sigma_u\overline{u} + \rho_v\sigma_v\overline{v})^2. \qquad (51)
$$

Furthermore,

$$
\begin{aligned}
\mathrm{var}(w_p^2) &= \mathrm{cov}(w_p^2, w_p^2) \\
&= \mathrm{cov}(u_p^2, u_p^2) + 2\mathrm{cov}(u_p^2, v_p^2) + \mathrm{cov}(v_p^2, v_p^2) \\
&= 3\rho_u^4\sigma_u^4 + 6\rho_u^2\sigma_u^2\overline{u}^2 + \overline{u}^4 - (\rho_u^2\sigma_u^2 + \overline{u}^2)^2 \\
&\quad +6\rho_u^2\rho_v^2\sigma_u^2\sigma_v^2 + 2\rho_u^2\sigma_u^2\overline{v}^2 + 2\rho_v^2\sigma_v^2\overline{u}^2 + 8\rho_u\rho_v\sigma_u\sigma_v\overline{uv} + 2\overline{u}^2\overline{v}^2 - 2(\rho_u^2\sigma_u^2 + \overline{u}^2)(\rho_v^2\sigma_v^2 + \overline{v})^2 \\
&\quad +3\rho_v^4\sigma_v^4 + 6\rho_v^2\sigma_v^2\overline{v}^2 + \overline{v}^4 - (\rho_v^2\sigma_v^2 + \overline{v}^2)^2 \\
&= 2\rho_u^2\sigma_u^2(\rho_u^2\sigma_u^2 + 2\overline{u}^2) + 4\rho_u^2\rho_v^2\sigma_u^2\sigma_v^2 + 8\rho_u\rho_v\sigma_u\sigma_v\overline{uv} + 2\rho_v^2\sigma_v^2(\rho_v^2\sigma_v^2 + 2\overline{v}^2) \\
&= \mathrm{cov}(w_p^2, w^2), \qquad (52)
\end{aligned}
$$

from which it follows that

$$
\begin{aligned}
\mathrm{corr}(w_p^2, w^2) &= \frac{\mathrm{cov}(w_p^2, w^2)}{\mathrm{std}(w_p^2)\mathrm{std}(w^2)} \\
&= \left(\frac{(\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2)^2 + 2(\rho_u\sigma_u\overline{u} + \rho_v\sigma_v\overline{v})^2}{(\sigma_u^2 + \sigma_v^2)^2 + 2\sigma_u^2\sigma_v^2(r^2 - 1) + 2(\sigma_u\overline{u} + \sigma_v\overline{v})^2 + 4\sigma_u\sigma_v\overline{uv}(r - 1)}\right)^{1/2} \qquad (53)
\end{aligned}
$$

Similar results to these were obtained in *Carlin and Haslett* [1982], in which the joint distribution of wind speeds at different locations w_j was characterised based on the assumption that the distribution of $w_j^{1/2}$ is multivariate Gaussian [see also *Brown et al.*, 1984]. This assumption is distinct from those used in the present study.

Note that the prediction of w by w_p is always superior to that directly by x (Eqn. 31):

$$\frac{\mathrm{corr}(x, w^2)}{\mathrm{corr}(w_p^2, w^2)} = \frac{\rho_u \sigma_u \overline{u} + \rho_v \sigma_v \overline{v}}{\left[\frac{1}{2}(\rho_u^2 \sigma_u^2 + \rho_v^2 \sigma_v^2)^2 + (\rho_u \sigma_u \overline{u} + \rho_v \sigma_v \overline{v})^2\right]^{1/2}} = \frac{b}{\sqrt{\frac{1}{2} + b^2}} \leq 1, \tag{54}$$

with

$$b = \frac{\rho_u \sigma_u \overline{u} + \rho_v \sigma_v \overline{v}}{\rho_u^2 \sigma_u^2 + \rho_v^2 \sigma_v^2}. \tag{55}$$

We see that, with a single Gaussian predictor, it is always better to first predict u and v and then compute w, rather than predict w directly (insofar as the vector winds are Gaussian).

In the limit of strong vector wind fluctuations, $\sigma_u^2 + \sigma_v^2 >> \overline{u}^2 + \overline{v}^2$, we have

$$\mathrm{corr}(w_p, w) \simeq \frac{\rho_u^2 \sigma_u^2 + \rho_v^2 \sigma_v^2}{(\sigma_u^4 + \sigma_v^4 + 2\sigma_u^2 \sigma_v^2 r^2)^{1/2}} \leq \left(\frac{\sigma_u^2}{(\sigma_u^4 + \sigma_v^4)^{1/2}}\right) \rho_u^2 + \left(\frac{\sigma_v^2}{(\sigma_u^4 + \sigma_v^4)^{1/2}}\right) \rho_v^2, \tag{56}$$

from which it follows that

$$\mathrm{corr}(w_p, w) \leq \max(\rho_u^2, \rho_v^2) \leq \max(\rho_u, \rho_v). \tag{57}$$

It follows that in the strong fluctuation limit the linear predictability of wind speed fluctuations from w_p is bounded above by the square of the linear predictability of the best-predicted component (and not necessarily bounded below by that the worst predicted component), and so will generally be considerably more poorly predicted than the vector wind fluctuations. Also note that $\mathrm{corr}(w_p, w)$ does not vanish in the limit that $\overline{u}^2 + \overline{v}^2 \to 0$, in contrast to the direct prediction of w by x.

In the limit of weak vector wind fluctuations, $\sigma_u^2 + \sigma_v^2 << \overline{u}^2 + \overline{v}^2$, we have

$$\mathrm{corr}(w_p, w) \simeq \frac{\rho_u \sigma_u \overline{u} + \rho_v \sigma_v \overline{v}}{(\sigma_u^2 \overline{u}^2 + \sigma_v^2 \overline{v}^2 + 2\sigma_u \sigma_v \overline{u}\overline{v} r)^{1/2}}. \tag{58}$$

In the along- and cross-mean wind basis for which $\overline{v} = 0$, we find that $\mathrm{corr}(w_p, w) \leq \rho_u$, so in this limit the wind speed predictability is bounded above by the predictability of the along-mean wind component.

In the limit of isotropic, uncorrelated vector wind fluctuations (with $\sigma_u = \sigma_v = \sigma$, $r = 0$) we have

$$\begin{aligned} \mathrm{corr}(w_p, w) &\simeq \left(\frac{\sigma^2(\rho_u^2 + \rho_v^2)^2 + 2(\rho_u \overline{u} + \rho_v \overline{v})^2}{2\sigma^2 + 2\overline{u}^2 + 2\overline{v}^2} \right)^{1/2} \\ &\leq \left(\frac{\sigma^2(\rho_u^2 + \rho_v^2) + 2(\rho_u \overline{u} + \rho_v \overline{v})^2}{2\sigma^2 + 2\overline{u}^2 + 2\overline{v}^2} \right)^{1/2}, \end{aligned} \tag{59}$$

where we have used the inequality $\rho_u^2 + \rho_v^2 \leq 1$, which follows from inequality (23). Moving into the along- and cross-mean wind basis, $\overline{v} = 0$ and

$$\begin{aligned} \mathrm{corr}(w_p, w) &\lesssim \left(\frac{\sigma^2(\rho_u^2 + \rho_v^2) + 2\rho_u^2 \overline{u}^2}{2\sigma^2 + 2\overline{u}^2} \right)^{1/2} \\ &\leq \left(\frac{2\sigma^2 + 2\overline{u}^2}{2\sigma^2 + 2\overline{u}^2} \right)^{1/2} \max(\rho_u, \rho_v) \\ &= \max(\rho_u, \rho_v). \end{aligned} \tag{60}$$

For the case of isotropic, uncorrelated vector wind fluctuations, the linear predictability of the wind speed (as constructed from the predicted vector wind components) is less than the maximum linear predictability of the along- and cross-mean wind components.

The general expression for $\mathrm{corr}(w_p, w)$ in the along- and across-wind basis (Eqn. 53 with $\overline{v} = 0$) is not in fact bounded above by the predictability of either vector wind component. This can be demonstrated by sampling the ratio

$$\Gamma = \frac{\mathrm{corr}(w_p^2, w^2)}{\max(\rho_u, \rho_v)} \tag{61}$$

uniformly over the parameter ranges $0 \leq \rho_u, \rho_v \leq 1$, $-1 \leq r \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma_u) \leq 1$, $-1 \leq \log_{10}(\sigma_v/\sigma_u) \leq 1$ (Figure 5). While the great majority of parameter sets result in values of $\Gamma < 1$, this ratio exceeds unity for a small fraction (0.05%) of the parameter sets. This result indicates that it is possible for wind speeds to be better predicted than the components, but in rare circumstances that are not representative of the isotropic, uncorrelated fluctuations of observed sea surface winds.

As discussed above, for $\rho_u, \rho_v < 1$, the Gaussian predictions of u and v have a low variance bias (cf. Eqns. 42 - 43). This bias will result in biases in both the mean and standard deviation of wind speed (as well as higher-order moments). In principle, an *a posteriori* bias correction can be applied to the predicted vector wind components before these are used to construct the predicted wind speed. However, as we will now demonstrate, correcting the variance bias in u_p and v_p actually degrades the wind speed linear predictive skill $\mathrm{corr}(w_p, w)$.

Compensating for the variance biases of the predicted vector wind components, we take

$$\hat{u}_p = \sigma_u x + \overline{u} \tag{62}$$

$$\hat{v}_p = \sigma_v x + \overline{v}, \tag{63}$$

and define

$$\hat{w}_p^2 = \hat{u}_p^2 + \hat{v}_p^2. \tag{64}$$

In this case, we have

$$\mathrm{cov}(\hat{w}_p^2, w^2) = 2\rho_u\sigma_u^2(\rho_u\sigma_u^2 + 2\overline{u}^2) + 2\sigma_u^2\sigma_v^2(\rho_u^2 + \rho_v^2) + 4\sigma_u\sigma_v\overline{u}\overline{v}(\rho_u + \rho_v) + 2\sigma_v^2\rho_v(\rho_v\sigma_v^2 + 2\overline{v}^2) \tag{65}$$

and

$$\begin{aligned}\mathrm{var}(\hat{w}_p^2) &= 2\sigma_u^2(\sigma_u^2 + 2\overline{u}^2) + 4\sigma_u^2\sigma_v^2 + 8\sigma_u\sigma_v\overline{u}\overline{v} + 2\sigma_v^2(\sigma_v^2 + 2\overline{v}^2) \\ &= 2(\sigma_u^2 + \sigma_v^2)^2 + 4(\sigma_u\overline{u} + \sigma_v\overline{v})^2. \end{aligned} \tag{66}$$

Taking $\overline{v} = 0$ (without loss of generality), and defining

$$A = \frac{\mathrm{corr}(\hat{w}_p^2, w^2)}{\mathrm{corr}(w_p^2, w^2)}, \tag{67}$$

we then have:

$$\begin{aligned} A &= \frac{(\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2)(\sigma_u^2 + \sigma_v^2) + 2\rho_u\sigma_u^2\overline{u}^2}{\left[(\sigma_u^2 + \sigma_v^2)^2 + 2\sigma_u^2\overline{u}^2\right]^{1/2}\left[(\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2)^2 + 2\rho_u^2\sigma_u^2\overline{u}^2\right]^{1/2}} \\ &= \frac{(\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2)(\sigma_u^2 + \sigma_v^2) + 2\rho_u\sigma_u^2\overline{u}^2}{\left\{[(\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2)(\sigma_u^2 + \sigma_v^2) + 2\rho_u\sigma_u^2\overline{u}^2]^2 + 2\sigma_u^2\overline{u}^2[\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2 - \rho_u(\sigma_u^2 + \sigma_v^2)]^2\right\}^{1/2}} \\ &= \frac{1}{\sqrt{1 + a^2}} \\ &\leq 1, \end{aligned} \tag{68}$$

with

$$a = \frac{\sqrt{2}\sigma_u\overline{u}[\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2 - \rho_u(\sigma_u^2 + \sigma_v^2)]}{(\rho_u^2\sigma_u^2 + \rho_v^2\sigma_v^2)(\sigma_u^2 + \sigma_v^2) + 2\rho_u\sigma_u^2\overline{u}^2}. \tag{69}$$

It follows that, from the perspective of the linear predictability of wind speed, there is no benefit to removing the variance bias of the vector wind components. A more sensible strategy is to rescale the mean and variance of the predicted wind speed w_p obtained from the biased vector wind predictions u_p and v_p, using the known mean and standard deviation of the wind speed (what *Lange and Focken* [2005] call Double Bias Correction). Such rescaling will not affect the correlation skill $\mathrm{corr}(w_p, w)$. While this strategy can eliminate biases in the mean and standard deviation of the predicted wind speed, it will not generally eliminate the biases in the higher order moments.

6. Prediction of Wind Speed from Components: Separate Predictions of u and v

The previous section considered the linear predictive skill of wind speed as calculated from vector wind component predictions obtained from a single Gaussian predictor. Of course, if both u and v are to be predicted, there is no reason why the same predictor should be used for both of these. We will now consider the linear predictive skill of wind speeds obtained from separately-predicted vector wind components, continuing to assume that the vector winds and their predictors are Gaussian. Note that in principle optimized predictions of u and v separately (via e.g. regression modelling) should always improve the vector wind prediction skills over those obtained with a single predictor.

We define the new vector wind predictions u_p and v_p such that

$$\begin{aligned}
\mathrm{corr}(u_p, u) &= \rho_{11} &(70)\\
\mathrm{corr}(u_p, v) &= \rho_{12} &(71)\\
\mathrm{corr}(v_p, u) &= \rho_{21} &(72)\\
\mathrm{corr}(v_p, v) &= \rho_{22} &(73)\\
\mathrm{corr}(u_p, v_p) &= P. &(74)
\end{aligned}$$

Again, we may think of these predictions as resulting from e.g. a linear regression model. Without loss of generality we can orient the wind component axes so that $\rho_{11} \geq 0, \rho_{22} \geq 0$. As well, we can assume that

$$\begin{aligned}
\rho_{11} &> |\rho_{21}| &(75)\\
\rho_{22} &> |\rho_{12}|. &(76)
\end{aligned}$$

That is, we can assume that v_p does not yield a better prediction of u than u_p does, and that u_p does not predict v better than v_p. The requirement that the correlation matrix of u_p, v_p, u, v be non-negative definite implies the following constraint on the correlation coefficients:

$$
\begin{aligned}
(\rho_{11}\rho_{22} - \rho_{21}\rho_{12})^2 + (1-P^2)(1-r^2) \;\geq\; & (\rho_{11} - P\rho_{21})(\rho_{11} - r\rho_{12}) + (\rho_{21} - P\rho_{11})(\rho_{21} - r\rho_{22}) \\
& +(\rho_{12} - P\rho_{22})(\rho_{12} - r\rho_{11}) + (\rho_{22} - P\rho_{12})(\rho_{22} - r\rho_{21}).
\end{aligned}
\tag{77}
$$

We recover the situation considered in the previous section in the limit that $P = 1$, $\rho_{11} = \rho_{21} = \rho_u$, $\rho_{12} = \rho_{22} = \rho_v$, in which case u_p and v_p both become equivalent to the predictor x. For $\rho_{11} = \rho_{21} = \rho_u$, $\rho_{12} = \rho_{22} = \rho_v$, the inequality (77) becomes

$$
(1-r^2) \geq \frac{2}{1+P}(\rho_u^2 + \rho_v^2 - 2r\rho_u\rho_v), \tag{78}
$$

so as $P \to 1$ (and $u_p \to \rho_u\sigma_u x + \overline{u}, v_p \to \rho_v\sigma_v x + \overline{v}$) we recover the inequality Eqn. (23).

To compute $\text{corr}(w_p, w)$, we first calculate the covariance:

$$
\begin{aligned}
\text{cov}(w_p^2, w^2) \;=\; & 2\rho_{11}^2\sigma_u^2(\rho_{11}^2\sigma_u^2 + 2\overline{u}^2) + 2\rho_{11}\rho_{12}\sigma_u\sigma_v(\rho_{11}\rho_{12}\sigma_u\sigma_v + 2\overline{uv}) \\
& +2\rho_{21}\rho_{22}\sigma_u\sigma_v(\rho_{21}\rho_{22}\sigma_u\sigma_v + 2\overline{uv}) + 2\rho_{22}^2\sigma_v^2(\rho_{22}^2\sigma_v^2 + 2\overline{v}^2)
\end{aligned}
\tag{79}
$$

and the variance of the predicted wind speed:

$$
\text{var}(w_p^2) = 2\rho_{11}^2\sigma_u^2(\rho_{11}^2\sigma_u^2 + 2\overline{u}^2) + 4\rho_{11}\rho_{22}\sigma_u\sigma_v(\rho_{11}\rho_{22}\sigma_u\sigma_v P^2 + 2\overline{uv}P) + 2\rho_{22}^2\sigma_v^2(\rho_{22}^2\sigma_v^2 + 2\overline{v}^2).
\tag{80}
$$

The resulting general expression for $\text{corr}(w_p^2, w^2)$ is complicated. These results are most usefully interpreted through consideration of a set of limiting cases.

- In the limit that $P = 1$, $\rho_{11} = \rho_{21} = \rho_u$, and $\rho_{12} = \rho_{22} = \rho_v$, we recover Eqn. (53).

• In the case that the predictions u_p and v_p are uncorrelated ($P = 0$), that u_p carries no linear predictive information for v ($\rho_{12} = 0$) and that v_p carries no linear predictive information for u ($\rho_{21} = 0$), we have

$$\mathrm{corr}(w_p, w) \simeq \left(\frac{\rho_{11}^2\sigma_u^2(\rho_{11}^2\sigma_u^2 + 2\overline{u}^2) + \rho_{22}^2\sigma_v^2(\rho_{22}^2\sigma_v^2 + 2\overline{v}^2)}{\sigma_u^2(\sigma_u^2 + 2\overline{u}^2) + 2\sigma_u\sigma_v(\sigma_u\sigma_v r^2 + 2\overline{u}\overline{v}r) + \sigma_v^2(\sigma_v^2 + 2\overline{v}^2)}\right)^{1/2}. \quad (81)$$

Again, considering this result in the along- and across-wind basis:

$$\begin{aligned}
\mathrm{corr}(w_p, w) &\simeq \left(\frac{\rho_{11}^4\sigma_u^4 + \rho_{22}^4\sigma_v^4 + 2\rho_{11}^2\sigma_u^2\overline{u}^2}{\sigma_u^4 + \sigma_v^4 + 2\sigma_u^2\sigma_v^2 r^2 + 2\sigma_u^2\overline{u}^2}\right)^{1/2} && (82)\\
&\leq \left(\frac{\rho_{11}^2\sigma_u^4 + \rho_{22}^2\sigma_v^4 + 2\rho_{11}^2\sigma_u^2\overline{u}^2}{\sigma_u^4 + \sigma_v^4 + 2\sigma_u^2\sigma_v^2 r^2 + 2\sigma_u^2\overline{u}^2}\right)^{1/2} \\
&\leq \max(\rho_{11}, \rho_{22}). && (83)
\end{aligned}$$

In this limit, the linear predictability of the wind speed is bounded above by that of the larger of the linear predictabilities of the along- and cross-mean wind components.

• For isotropic, uncorrelated vector fluctuations, in the along- and cross-wind basis, we have

$$\mathrm{corr}(w_p, w) \simeq \frac{(\rho_{11}^4 + \rho_{11}^2\rho_{12}^2 + \rho_{21}^2\rho_{22}^2 + \rho_{22}^4)\sigma^2 + 2\rho_{11}^2\overline{u}^2}{[2(\rho_{11}^4 + 2\rho_{11}^2\rho_{22}^2 P^2 + \rho_{22}^4)\sigma^2 + 4\rho_{11}^2\overline{u}^2]^{1/2}[\sigma^2 + \overline{u}^2]^{1/2}}. \quad (84)$$

There are parameter sets for which the linear predictability of wind speed exceeds that of the best-predicted vector wind component. That is, the ratio

$$\Gamma = \frac{\mathrm{corr}(w_p^2, w^2)}{\max(\rho_{11}, \rho_{22})} \quad (85)$$

is not bounded above by unity for all parameter values satisfying the inequalities (75), (76), and (77). As in the previous section, the ratio Γ was explored numerically over the parameter ranges: $0 \leq \rho_{11}, \rho_{22} \leq 1$, $-1 \leq \rho_{12}, \rho_{21} \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma) \leq 1$, with uniform grid spacing. For only 1.9% of the parameter sets did Γ exceed 1 (Figure 6); these tend to be associated with larger values of the vector wind correlations ρ_{ij} (not shown). Most of the parameter sets yielding $\Gamma > 1$ are found for one (or both) of $\rho_{12}, \rho_{21} < 0$; when these cross-correlations are constrained to be non-negative, $\Gamma > 1$ occurs for only 0.03% of the parameter sets (Figure 6). Note that while the expression Eqn. (84) is insensitive to the sign of the cross-correlations, this is not true of the inequality (77).

- In the limit of perfect predictability of the vector wind components, $\rho_{11} = \rho_{22} = 1$, the constraint (77) requires that $\rho_{12} = \rho_{21} = P = r$, and so

$$\mathrm{corr}(w_p, w) \simeq \frac{2\sigma_u^2(\sigma_u^2 + 2\overline{u}^2) + 4r\sigma_u\sigma_v(r\sigma_u\sigma_v + 2\overline{u}\overline{v}) + 2\sigma_v^2(\sigma_v^2 + 2\overline{v}^2)}{2\sigma_u^2(\sigma_u^2 + 2\overline{u}^2) + 4\sigma_u\sigma_v(\sigma_u\sigma_v r^2 + 2\overline{u}\overline{v}r) + 2\sigma_v^2(\sigma_v^2 + 2\overline{v}^2)} = 1. \tag{86}$$

This otherwise obvious result is interesting for its illustration of the relevance to predictability of the constraint of non-negative definiteness of the correlation matrix of the vector winds and wind predictors.

In the previous section, we found that the linear predictability of of w was generally reduced by correcting the variance biases of u_p and v_p obtained from a single predictor. A similar analysis will now be carried out for the case of u_p and v_p predicted separately. Denoting the rescaled vector wind component predictions as $\hat{u}_p$ and $\hat{v}_p$, as in the previous section, we find:

$$\begin{aligned}\mathrm{cov}(\hat{w}_p^2, w^2) \;=\; & 2\rho_{11}\sigma_u^2(\rho_{11}\sigma_u^2 + 2\overline{u}^2) + 2\rho_{12}\sigma_u\sigma_v(\rho_{12}\sigma_u\sigma_v + 2\overline{u}\overline{v}) \\ & +2\rho_{21}\sigma_u\sigma_v(\rho_{21}\sigma_u\sigma_v + 2\overline{u}\overline{v}) + 2\rho_{22}\sigma_v^2(\rho_{22}\sigma_v^2 + 2\overline{v}^2)\end{aligned} \tag{87}$$

and

$$\mathrm{var}(\hat{w}_p^2) = 2\sigma_u^2(\sigma_u^2 + 2\overline{u}^2) + 4\sigma_u\sigma_v(\sigma_u\sigma_v P^2 + 2\overline{u}\overline{v}P) + 2\sigma_v^2(\sigma_v^2 + 2\overline{v}^2). \tag{88}$$

Once again, we define the ratio

$$A = \frac{\mathrm{corr}(\hat{w}_p{}^2, w^2)}{\mathrm{corr}(w_p^2, w^2)}. \tag{89}$$

In the case of mutually uninformative predictors ($\rho_{12} = \rho_{21} = P = 0$), we have

$$A = \frac{\rho_{11}\sigma_u^2(\rho_{11}\sigma_u^2 + 2\overline{u}^2) + \rho_{22}\sigma_v^2(\rho_{22}\sigma_v^2 + 2\overline{v}^2)}{\left[\rho_{11}^2\sigma_u^2(\rho_{11}^2\sigma_u^2 + 2\overline{u}^2) + \rho_{22}^2\sigma_v^2(\rho_{22}^2\sigma_v^2 + 2\overline{v}^2)\right]^{1/2}\left[\sigma_u^2(\sigma_u^2 + 2\overline{u}^2) + \sigma_v^2(\sigma_v^2 + 2\overline{v}^2)\right]^{1/2}}. \quad (90)$$

From this, we obtain

$$\begin{aligned} A^2 - 1 \;=\; & -\left[2\sigma_u^6\overline{u}^2\rho_{11}^2(\rho_{11}-1)^2 + 2\sigma_v^6\overline{v}^2\rho_{22}^2(\rho_{22}^2-1)^2 + \sigma_u^4\sigma_v^4(\rho_{22}-\rho_{11})^2\right. \\ & \left. +2\sigma_u^4\sigma_v^2\overline{v}^2(\rho_{11}^2-\rho_{22})^2 + 2\sigma_u^2\sigma_v^4\overline{u}^2(\rho_{22}^2-\rho_{11})^2 + 4\sigma_u^2\sigma_v^2\overline{u}^2\overline{v}^2(\rho_{11}-\rho_{22})^2\right] \\ \leq\; & 0. \end{aligned} \quad (91)$$

When the predictors of u_p and v_p are mutually uninformative, correcting for the vector wind variance bias will always reduce the wind speed predictive skill.

For isotropic, uncorrelated vector wind fluctuations (taking $\overline{v} = 0$), we have

$$A = \frac{(\rho_{11}^2 + \rho_{12}^2 + \rho_{21}^2 + \rho_{22}^2)\sigma^2 + 2\rho_{11}\overline{u}^2}{(\rho_{11}^4 + \rho_{11}^2\rho_{12}^2 + \rho_{21}^2\rho_{22}^2 + \rho_{22}^4)\sigma^2 + 2\rho_{11}^2\overline{u}^2}\left[\frac{(\rho_{11}^4 + 2\rho_{11}^2\rho_{22}^2P^2 + \rho_{22}^4)\sigma^2 + 2\rho_{11}^2\overline{u}^2}{2\sigma^2(1+P^2) + 2\overline{u}^2}\right]^{1/2}. \quad (92)$$

The ratio A was sampled numerically and uniformly over the parameter ranges used earlier in the analysis of Γ (Eqn. 85); the resulting pdf is plotted in Figure 7. We see that in general $A \not\leq 1$: for about 11% of the parameter value sets, the wind speed prediction is improved by correcting for the variance bias of u_p and v_p. The prediction skill is unchanged or worsened by this strategy for the remaining 89% of the parameter sets. Although offsetting the variance bias increases wind speed predictability for some parameter sets, it does not substantially change the predictability of wind speed relative to that of the vector wind components. The ratio

$$\hat{\Gamma} = \frac{\mathrm{corr}(\hat{w}_p^2, w^2)}{\max(\rho_{11}, \rho_{22})} \quad (93)$$

sampled over the same parameter range (Figure 7) shows little difference from the ratio Γ associated with the non-bias corrected wind speed predictions. Thus, while there are circumstances in which it is helpful to correct for the variance bias in the vector wind components, the predictability of wind speeds is still in the great majority of situations lower than that of the components.

Our analysis has focused up to this point on theoretical results for the wind speed predictability, based on the assumption of Gaussian vector winds and Gaussian predictors. To assess the relevance of these results to predicting real (non-Gaussian) winds, the following analysis was carried out. For each spatial location in the SeaWinds dataset (between 60°S and 60°N), we constructed the standardized vector wind prediction anomalies, u_p' and v_p' from the standardized anomalies of the zonal and meridional wind components:

$$u_p' = \rho U' + (1-\rho^2)^{1/2}\epsilon_u \tag{94}$$

$$v_p' = \rho V' + (1-\rho^2)^{1/2}\epsilon_v. \tag{95}$$

From these anomalies, the predictions

$$u_p = \rho\sigma_u u_p' + \overline{u} \tag{96}$$

$$v_p = \rho\sigma_v v_p' + \overline{v} \tag{97}$$

were computed using the vector wind means and standard deviations taken from observations. The random processes ϵ_u, ϵ_v are mutually uncorrelated series of independent, Gaussian random variables of unit variance and mean zero. The parameter ρ determines how well the vector wind components are "predicted" by u_p and v_p: by construction, $\rho_{11} = \rho_{22} = \rho$.

The vector wind predictions u_p and v_p were computed for $\rho = 0.25, 0.5$, and 0.75. From these, the correlations $\mathrm{corr}(w_p, w)$ between observed and predicted wind speed were computed and compared to the theoretical correlation values computed from Eqns. (30),

(79) and (80). Agreement is generally good between the observed and theoretically derived wind correlations $\mathrm{corr}(w_p, w)$ (Figure 8). As expected from the Gaussian vector wind model, $\mathrm{corr}(w_p, w) < \rho$ for the most part; sampling variability will contribute to the small number of locations for which this inequality is not satisfied. As well, the spatial joint distribution of the theoretically derived and observed values of $\mathrm{corr}(w_p, w)$ parallels the 1:1 line. The results of the theoretical model are not perfect: in particular, it predicts larger values of $\mathrm{corr}(w_p, w)$ than are observed. Much of this bias can be associated with the fact that, in contradiction to the assumptions underlying the model, variability in the vector wind components is not Gaussian [e.g. *Monahan*, 2004, 2006]. Furthermore, as u_p and v_p are themselves constructed from the non-Gaussian U and V, these predictions themselves will be non-Gaussian. That the non-Gaussianity of the vector winds is an important factor in biasing the theoretical value of $\mathrm{corr}(w_p, w)$ is demonstrated by consideration of the spatial joint distribution of this bias with the skewness of the along-mean wind component $u = (\overline{U}U + \overline{V}V)/\sqrt{U^2 + V^2}$ (Figure 9). To a first approximation, this bias varies linearly with the vector wind skewness. Note that for $\rho = 0.5$, and particularly for $\rho = 0.75$, there are nonzero offsets in the bias of $\mathrm{corr}(w_p, w)$ for $\mathrm{skew}(u) \simeq 0$. The bias is affected by factors other than non-Gaussianity of the vector winds, possibly such as the neglect of nonstationarities such as the seasonal cycle in the present analysis.

7. Conclusions

This study has considered the linear predictability of wind speed fluctuations relative to the predictability of vector wind fluctuations. Analytic expressions for the linear prediction skill of wind speed have been obtained, based on the assumption that vector wind fluctuations can be approximated as bivariate Gaussian. The following conclusions have been reached:

- Fluctuations of the wind speed w and the kinetic energy w^2 are highly correlated. This fact can be used to compare the predictability of the wind speed relative to that of the vector wind components, as $w^2 = u^2 + v^2$ is simply the sum of the squares of the components.

- We have found that there will always be some vector wind component that is better predicted by a single Gaussian predictor x than is the wind speed. The predictability of the wind speed relative to the best-predicted component decreases as vector wind fluctuations become much larger than the magnitude of the mean vector wind. In the limit that the mean vector wind vanishes, x carries no direct linear predictive information regarding w irrespective of how well the vector wind components are predicted.

- Predictions of w are always improved, relative to predicting w from x directly, by first predicting u and v from x and then determining predicted speed from these. In this approach, the predictive skill of w does not vanish in the limit that the mean vector wind goes to zero. For the vast majority of parameter sets, the wind speed prediction skill is lower than that of the vector winds in some direction (considerably so, in general) - although for a small fraction of parameter values, the wind speed is better predicted than any vector component. In the limit of isotropic and uncorrelated vector wind fluctuations, which is a good approximation to observed fluctuations of sea surface winds, the wind speed prediction skill is always less than that of the best predicted wind component. Finally, rescaling the predicted vector wind components to correct for the low variance bias always reduces the prediction skill of wind speeds.

- Predicting the vector wind components separately rather than using a single predictor always results in an improved wind speed prediction. When these predictors are mutually independent and carry no cross-predictive information, the wind speed prediction skill is always bounded above by that of the best-predicted vector wind component. For such predictions, the linear predictability of wind speed is considerably smaller than that of the vector wind components over a broad range of parameter values, and no improvement to wind speed prediction ever follows from correcting for the variance bias of the predicted vector winds.

For isotropic and uncorrelated vector wind fluctuations, it is possible through the use of separate predictions of u and v to obtain wind speed predictions that are better than those of any individual vector wind component (for about 2% of the parameter sets). Furthermore, in this limit, it is possible to improve wind speed predictions by correcting for the variance biases of the predicted vector wind components - but these improvements do not result in an increase in the parameter range over which wind speeds are better predicted than the vector winds. For the great majority of parameter sets, it is the case that wind speed predictions are worse than those of vector winds.

- The results of the theoretical analysis were broadly supported by global observations of sea-surface vector winds. Using vector wind predictions produced by corrupting observed vector wind time series with synthetic white noise, it was found that wind speed prediction skills are generally less than those of the vector winds - often substantially so. In fact, the theoretical model predicts wind speed correlation skills that are somewhat larger than those that are observed. These biases were shown to be closely related to the non-Gaussianity of the vector winds, as measured by skewness.

It may seem self-evident that the predictability of vector wind components should limit that of wind speed. However, one can imagine a situation in which the wind speed is perfectly predictable, while the wind direction has no predictability. In such a case, predictions of wind speed would be expected to be considerably better than those of any vector wind component. Such a scenario is ruled out by the results of the present study, for Gaussian vector winds and Gaussian predictors.

The influence of sampling variability on wind speed prediction was neglected in this theoretical analysis, which assumed that all statistical features of the vector winds were known with perfect accuracy. Of course, sampling errors in the vector wind statistics will result in errors in the wind speed statistics; in particular, terms I and II of the mean squared error expansion Eqn. (3) will become nonzero. These sensitivities can also be investigated using idealised models of the vector winds such as those considered in this study [e.g. *Monahan*, 2011]. The results of the present study should therefore be interpreted as representing upper limits on the predictability of wind speeds that are inherent to the relationship between speeds and vector winds and distinct from the issue of sampling error. Future studies will carry out a more thorough investigation of influence of sampling errors in the vector wind components on the wind speed predictions.

The analyses of observed winds in this study indicated that non-Gaussianity of the vector winds results in a small but systematic overestimate of wind speed predictability by the idealised model. A natural next step would be to extend the present analysis to account for the observed non-Gaussianity in u and v; such a step is facilitated by the fact that the skewness of sea surface vector winds is a feature which is well understood [e.g. *Monahan*, 2004, 2006]. Nevertheless, how best to specify a non-Gaussian distribution with given moments remains an open question [e.g. *Monahan*, 2007]. A possible approach would be to use a bivariate generalisation of the Gram-Charlier expansion [e.g. *Longuet-Higgins*, 1964; *Lokas*, 1998]; the resulting expressions for $\mathrm{corr}(w_p, w)$ would be considerably more complicated than those obtained in the present study. It is important to emphasize that the bias in the modelled value of $\mathrm{corr}(w_p, w)$ is systematically positive - that is, the idealised Gaussian model suggests higher values of wind speed predictability than are observed over the oceans. Thus, the upper bounds presented here for the Gaussian predictability of wind speeds relative to vector wind components are conservative. It is possible that the character of vector wind non-Gaussianity over land may be sufficiently different than that over oceans (due for example to the effects of complex topography) for these bounds to not hold; empirical analyses of the predictability of wind speed relative to vector wind components represent a useful direction of future study.

A crucial step in the theoretical characterisation of the predictability of wind speed was the recognition that fluctuations in speed are highly correlated with those of kinetic energy, so that for any variable x we can take $\mathrm{corr}(x, w) \simeq \mathrm{corr}(x, w^2)$ as an excellent approximation. This result will hold for any power of w, so long as the truncated Taylor series approximation

$$w^n = (w^2)^{n/2} = (\mathrm{mean}(w^2) + \mathrm{std}(w^2)w'^2)^{n/2} \simeq \mathrm{mean}^{n/2}(w^2)\left(1 + \frac{n}{2}\frac{\mathrm{std}(w^2)}{\mathrm{mean}(w^2)}w'^2\right) \quad (98)$$

is sufficiently accurate. In particular, the predictability of both $w^{1/2}$ and w^3 should be similar to that of w^2, to which these quantities are strongly correlated (Figure 10). The first of these quantities is interesting because square root transforms have been used to produce wind speed distributions that are approximately Gaussian, and therefore more amenable to prediction by the techniques of classical time series analysis [e.g. *Carlin and*

Haslett, 1982; *Brown et al.*, 1984]. The second of these quantities is proportional to the wind power density. Although the correlations between w^2 and $w^{1/2}$ are poorer than those between w and w^2, the present analysis indicates that to a first approximation this quantity and corr(w^2, w^3) should be subject to similar predictability constraints (with Gaussian predictors) as the wind speed itself.

This study has focused on the statistical prediction of vector wind components and wind speed using Gaussian predictors, such as those characteristic of the large-scale flow used for statistical downscaling [e.g. *Monahan*, 2011]. The limits to wind speed predictability that have been found could in principle be avoided through the use of strongly non-Gaussian predictors, derived from either statistical or dynamical predictions. In particular, the results obtained in this study do not exclude the possibility that wind speed predictability can be increased through the use of appropriate nonlinear statistical tools, such as neural networks or nonlinear regression [e.g. *Kretzschmar et al.*, 2004]. Furthermore, in principle excellent (non-Gaussian) predictors of u^2 and v^2 could yield an excellent prediction of w^2 without carrying any (linear) predictive information regarding u or v. The limits to wind speed predictability obtained in this study apply only to the extent that the predictors are Gaussian.

A broad range of approaches have been proposed for the prediction of wind speeds, ranging from the dynamical, to the statistical, to hybrids of these two. The results of the present study demonstrate the limits that can be achieved with Gaussian predictions. A useful direction of future study is the consideration of the extent to which non-Gaussian predictions, from either nonlinear statistical models or dynamical models, will allow the circumvention of these limits to this important prediction problem.

Appendix A: $\mathrm{corr}(w, w^2)$

Suppose that the wind speed w has mean μ, standard deviation σ_w, skewness ν and kurtosis κ. Then:

$$\begin{aligned}\mathrm{cov}(w, w^2) &= \mathsf{E}\left\{w^3\right\} - \mathsf{E}\left\{w\right\}\mathsf{E}\left\{w^2\right\} \\ &= \sigma_w^3\nu + 2\sigma_w^2\mu. \end{aligned} \quad \text{(A1)}$$

Furthermore,

$$\begin{aligned}\mathrm{var}(w^2) &= \mathsf{E}\left\{w^4\right\} - \left(\mathsf{E}\left\{w^2\right\}\right)^2 \\ &= \sigma_w^4(\kappa+2) + 4\sigma_w^3\mu\nu + 4\sigma_w^2\mu^2, \end{aligned} \quad \text{(A2)}$$

and so we have

$$\begin{aligned}\mathrm{corr}(w, w^2) &= \frac{\sigma_w\nu + 2\mu}{\sqrt{\sigma_w^2(\kappa+2) + 4\sigma_w\mu\nu + 4\mu^2}} \\ &= \frac{1+\nu\delta}{\sqrt{1 + 2\nu\delta + (\kappa+2)\delta^2}}, \end{aligned} \quad \text{(A3)}$$

where

$$\delta = \frac{\sigma_w}{2\mu}. \quad \text{(A4)}$$

This quantity is well-defined because for wind speeds $\mu > 0$. In fact, from the discussion in Section 3, we know that for observed winds $\delta < 1$ (and that for the most part $\delta << 1$; Figure 3). Approximating the denominator of Eqn. (A3):

$$\mathrm{corr}(w, w^2) \simeq (1+\nu\delta)\left(1 - \nu\delta - \frac{1}{2}(\kappa+2)\delta^2\right) \simeq 1 - \left(\nu^2 + \frac{1}{2}(\kappa+2)\right)\delta^2, \quad \text{(A5)}$$

so $\mathrm{corr}(w, w^2) \simeq 1$ to $O(\delta^2)$ accuracy.

Appendix B: The Joint Distribution of x and y^2

Suppose that the variables x and y are bivariate Gaussian, such that mean$(x) = 0$, mean$(y) = \overline{y}$, std$(x) = $ std$(y) = 1$, and corr$(x, y) = r$. Then

$$p_{xy}(x,y) = \frac{1}{2\pi(1-r^2)^{1/2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(x^2 - 2rx(y-\overline{y}) + (y-\overline{y})^2\right)\right]. \tag{B1}$$

The joint pdf of x and $z = y^2$ is given by

$$\begin{aligned} p_{xz}(x,z) &= \begin{vmatrix} \partial_x x & \partial_x z \\ \partial_y z & \partial_y z \end{vmatrix}^{-1} [p_{xy}(x,\sqrt{z}) + p_{xy}(x,-\sqrt{z})] \qquad \text{(B2)} \\ &= \frac{1}{4\pi(1-r^2)^{1/2}\sqrt{z}} \exp\left(-\frac{1}{2(1-r^2)}[x^2 + 2r\overline{y}x + \overline{y}^2 + z]\right) \cosh\left(\frac{rx+\overline{y}}{(1-r^2)}\sqrt{z}\right). \end{aligned}$$

For $\overline{y} = 0$, $p_{xz}(x,z)$ is a symmetric function of x: fluctuations of x of either sign are associated with the same fluctuations in z and corr$(x,z) = 0$ (Figure 11). As $\overline{y}$ increases, $p_{xz}(x,z)$ becomes increasingly asymmetric, so fluctuations in x and z are increasingly like-signed and corr(x,z) becomes increasingly positive. The strength of this asymmetry scales with the degree of correlation between x and y. While corr(x,y) is insensitive to the mean of y, the correlation corr(x,y^2) is proportional to $\overline{y}$.

Appendix C: corr(x^2, y^2)

Let x, y be Gaussian random variables with mean zero and unit variance, such that

$$\text{corr}(x,y) = \mathsf{E}\{xy\} = r. \tag{C1}$$

As the joint distribution of x, y is Gaussian, we can define the unit variance, mean zero, Gaussian random variable ε such that

$$y = rx + (1-r^2)^{1/2}\varepsilon, \tag{C2}$$

and

$$\text{corr}(x,\varepsilon) = 0. \tag{C3}$$

Because x and ε are Gaussian, the fact that they are uncorrelated implies that they are statistically independent.

To find $\mathrm{corr}(x^2, y^2)$ we first compute $\mathsf{E}\{x^2y^2\}$:

$$\begin{aligned}
\mathsf{E}\left\{x^2y^2\right\} &= \mathsf{E}\left\{x^2(rx + (1-r^2)^{1/2}\varepsilon)^2\right\} \\
&= \mathsf{E}\left\{r^2x^4 + 2r(1-r^2)^{1/2}x^3\varepsilon + (1-r^2)x^2\varepsilon^2\right\} \\
&= r^2\mathsf{E}\left\{x^4\right\} + 2r(1-r^2)^{1/2}\mathsf{E}\left\{x^3\right\}\mathsf{E}\left\{\varepsilon\right\} + (1-r^2)\mathsf{E}\left\{x^2\right\}\mathsf{E}\left\{\varepsilon^2\right\} \\
&= 2r^2 + 1, \qquad \text{(C4)}
\end{aligned}$$

and so

$$\mathrm{cov}(x^2, y^2) = \mathsf{E}\left\{x^2y^2\right\} - \mathsf{E}\left\{x^2\right\}\mathsf{E}\left\{y^2\right\} = 2r^2 + 1 - 1 = 2r^2. \qquad \text{(C5)}$$

We also have:

$$\mathrm{var}(x^2) = \mathrm{E}\left\{x^4\right\} - \left(\mathrm{E}\left\{x^2\right\}\right)^2 = 3 - 1^2 = 2 = \mathrm{var}(y^2). \qquad \text{(C6)}$$

It follows that

$$\mathrm{corr}(x^2, y^2) = \frac{\mathrm{cov}(x^2, y^2)}{\mathrm{std}(x^2)\mathrm{std}(y^2)} = \frac{2r^2}{2} = r^2. \qquad \text{(C7)}$$

That is, the correlation of the square of the variables is the square of their correlation.

Note that the Gaussianity of the joint distribution of x and ε was central in the above calculation in allowing uncorrelated variables to be treated as independent. The result in Eqn. (C7) will not hold in general if the joint distribution of x and y is non-Gaussian. For example, suppose that x is Gaussian with mean zero and unit variance and $y = |x|$. Then $\mathrm{corr}(x, y) = 0$ but $\mathrm{corr}(x^2, y^2) = 1$.

Acknowledgments.

The author gratefully acknowledges helpful comments on the manuscript from Charles Curry and Aaron Culver. This research was supported by the Natural Sciences and Engineering Research Council of Canada.

References

Brown, B. G., R. W. Katz, and A. H. Murphy (1984), Time series models to simulate and forecast wind speed and wind power, *J. Clim. Appl. Meteor.*, *23*, 1184–1195.

Cakmur, R., R. Miller, and O. Torres (2004), Incorporating the effect of small-scale circulations upon dust emission in an atmospheric general circulation model, *J. Geophys. Res.*, *109*, D07,201.

Carlin, J., and J. Haslett (1982), The probability distribution of wind power from a dispersed array of wind turbine generators, *J. Appl. Meteor.*, *21*, 303–313.

Costa, A., A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa (2008), A review on the young history of the wind power short-term prediction, *Ren. Sust. Energy Rev.*, *12*, 1725–1744.

Giebel, G., R. Brownsword, G. kariniotakis, M. Denhard, and C. Draxl (2011), The state-of-the-art in short-term prediction of wind power: A literature overview, *Tech. rep.*, ANEMOS.plus and SafeWind.

Jet Propulsion Laboratory (2001), SeaWinds on QuikSCAT Level 3: Daily, Gridded Ocean Wind Vectors, *Tech. Rep. Tech. Rep. JPL PO.DAAC Product 109*, California Institute of Technology.

Kretzschmar, R., P. Eckert, D. Cattani, and F. Eggimann (2004), Neural network classifiers for local wind prediction, *J. Appl. Meteor.*, *43*, 727–738.

Lange, M., and U. Focken (2005), *Physical Approach to Short-Term Wind Power Prediction*, 208 pp., Springer, Berlin.

Lokas, E. L. (1998), Evolution of peaks in weakly non-linear density field and dark halo profiles, *Mon. Not. R. Astron. Soc.*, *296*, 491–501.

Longuet-Higgins, M. (1964), Modified Gaussian distributions for slightly nonlinear variables, *Radio Sci.*, *68D*, 1049–1062.

Ma, L., S. Luan, C. Jiang, H. Liu, and Y. Zhang (2009), A review on the forecasting of wind speed and generated power, *Ren. Sust. Energy. Syst.*, *13*, 915–920.

Mahrt, L., and J. Sun (1995), The subgrid velocity scale in the bulk aerodynamic relationship for spatially averaged scalar fluxes, *Mon. Weath. Rev.*, *123*, 3032–3041.

Monahan, A. H. (2004), A simple model for the skewness of global sea-surface winds, *J. Atmos. Sci.*, *61*, 2037–2049.

Monahan, A. H. (2006), The probability distribution of sea surface wind speeds. Part I: Theory and SeaWinds observations, *J. Climate*, *19*, 497–520.

Monahan, A. H. (2007), Empirical models of the probability distribution of sea surface wind speeds, *J. Climate*, *20*, 5798–5814.

Monahan, A. H. (2011), Can we see the wind? Statistical downscaling of historical sea surface winds in the Northeast Subarctic Pacific, *J. Climate*, *in press*.

Rice, S. O. (1945), Mathematical analysis of random noise (part 2), *Bell System Tech. J.*, *24*, 46–156.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, 175 pp., Chapman and Hall.

Soman, S., H. Zareipour, O. Malik, and P. Mandal (2010), A review of wind power and wind speed forecasting methods with different time horizons, in *42nd North American Power Symposium*, Arlington, Texas, USA.

Thorarinsdottir, T., and T. Gneiting (2010), Probabilistic forecasts of wind speed: Ensemble model output statistics using heteroskedastic censored regression, *J. Roy. Stat. Soc. A*, *173*, 371–388.

Wilks, D. S. (2005), *Statistical Methods in the Atmospheric Sciences*, Academic Press.

Figure Captions

Figure 1: Kernel density estimate [e.g. *Silverman*, 1986] of the spatial probability density function of observed correlations between w and w^2. The observed sea surface winds are taken from the SeaWinds scatterometer observations, between 60°S and 60°N.

Figure 2: Kernel density estimate of the spatial joint distributions of correlations between vector wind components, wind speed and squared wind speed. Left panel: $\mathrm{corr}(U, w)$ and $\mathrm{corr}(U, w^2)$. Right panel: $\mathrm{corr}(V, w)$ and $\mathrm{corr}(V, w^2)$. Correlations are computed from SeaWinds observations between 60°S and 60°N.

Figure 3: Kernel density estimate of the spatial probability density function of the ratio of the standard deviation to the mean of wind speed, from the SeaWinds data between 60°S and 60°N.

Figure 4: Kernel density estimate of the spatial joint distribution of the observed and modelled correlations of wind speed w with vector wind components. Left panel: correlation of w with zonal wind U. Right panel: correlation of w with meridional wind V.

Figure 5: Probability density function of the ratio Γ between the modelled wind speed correlation prediction skill and that of the best-predicted vector wind component (Eqn. 61), uniformly sampled over the parameter ranges $0 \leq \rho_u, \rho_v \leq 1$, $-1 \leq r \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma_u) \leq 1$, $-1 \leq \log_{10}(\sigma_v/\sigma_u) \leq 1$. Note the logarithmic scale on the y-axis.

Figure 6: Black curve: probability density function of the ratio Γ between the wind speed prediction skill and that of the best-predicted vector wind component (Eqn. 85), obtained from uniform sampling of the parameter values over the ranges $0 \leq \rho_{11}, \rho_{12} \leq 1$, $-1 \leq \rho_{12}, \rho_{21} \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma) \leq 1$. Grey curve: as for the black curve, but sampling only positive cross-correlations $0 \leq \rho_{12}, \rho_{21} \leq 1$. Note the logarithmic scale on the y-axis.

Figure 7: Left panel: Probability distribution of A (Eqn. 92) obtained by sampling the parameter values uniformly over the range $0 \leq \rho_{11}, \rho_{22} \leq 1$, $-1 \leq \rho_{12}, \rho_{21} \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma) \leq 1$. Right panel: As in Figure 6 for the ratio $\hat{\Gamma}$ (Eqn. 93). Note the logarithmic scale on the y-axis of each panel.

Figure 8: Kernel density estimate of the relationship between the observed wind speed predictability, $\mathrm{corr}(w_p, w)$, and the theoretical value from Eqns. (30), (79) and (80). The vector wind predictors u_p and v_p were obtained from Eqns. (94)-(97) with $\rho = 0.25$ (black contours), $\rho = 0.5$ (red contours), $\rho = 0.75$ (blue contours).

Figure 9: Kernel density estimates of the spatial joint distribution of the bias (theoretical value less observed value) in $\mathrm{corr}(w_p, w)$ with the observed skewness of the along mean-wind component of the vector winds, $u = (\overline{U}U + \overline{V}V)/\sqrt{U^2 + V^2}$.

Figure 10: As in Figure 1, for the correlation between w^2 and $w^{1/2}$ (left panel) and the correlation between w^2 and w^3 (right panel).

Figure 11: Joint pdf of x and $z = y^2$ for x, y bivariate Gaussian with $\mathrm{corr}(x, y) = r$, for a range of values of $\overline{y}$ and r.

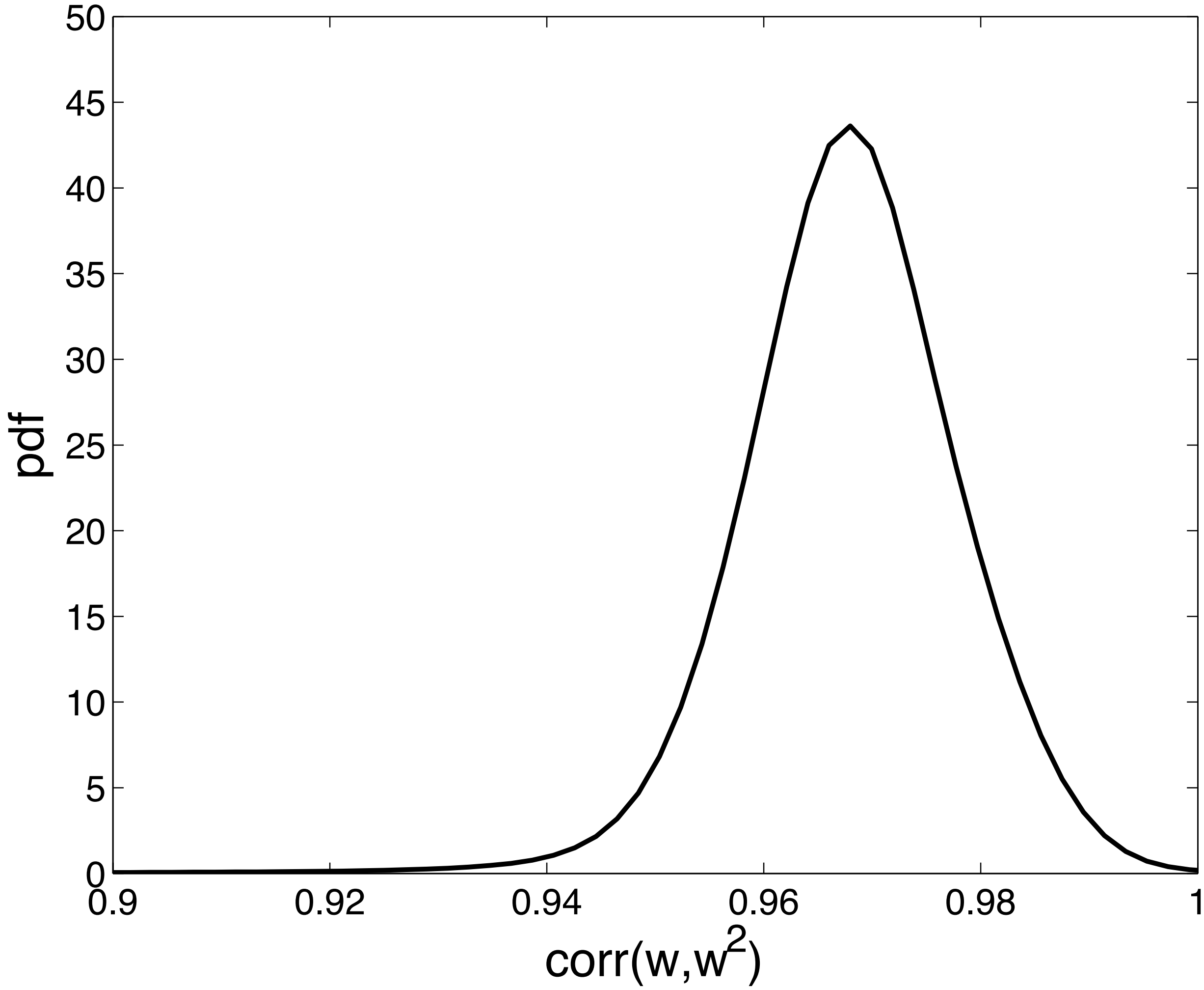

Figure 1. Kernel density estimate [e.g. *Silverman*, 1986] of the spatial probability density function of observed correlations between w and w^2. The observed sea surface winds are taken from the SeaWinds scatterometer observations, between 60°S and 60°N.

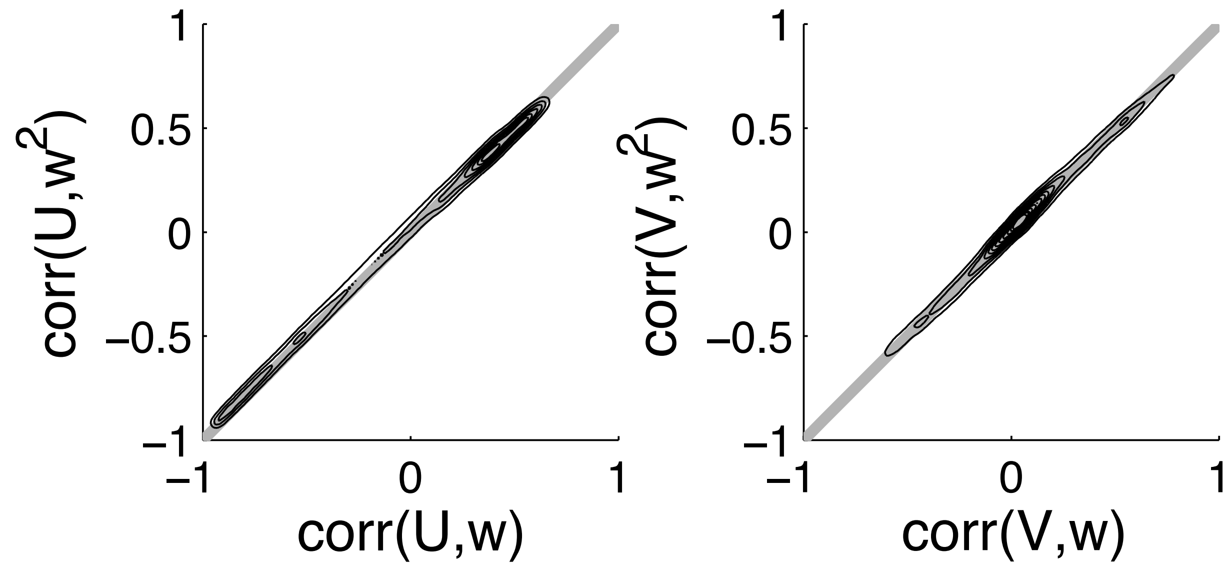

Figure 2. Kernel density estimate of the spatial joint distributions of correlations between vector wind components, wind speed and squared wind speed. Left panel: $\mathrm{corr}(U, w)$ and $\mathrm{corr}(U, w^2)$. Right panel: $\mathrm{corr}(V, w)$ and $\mathrm{corr}(V, w^2)$. Correlations are computed from SeaWinds observations between 60°S and 60°N.

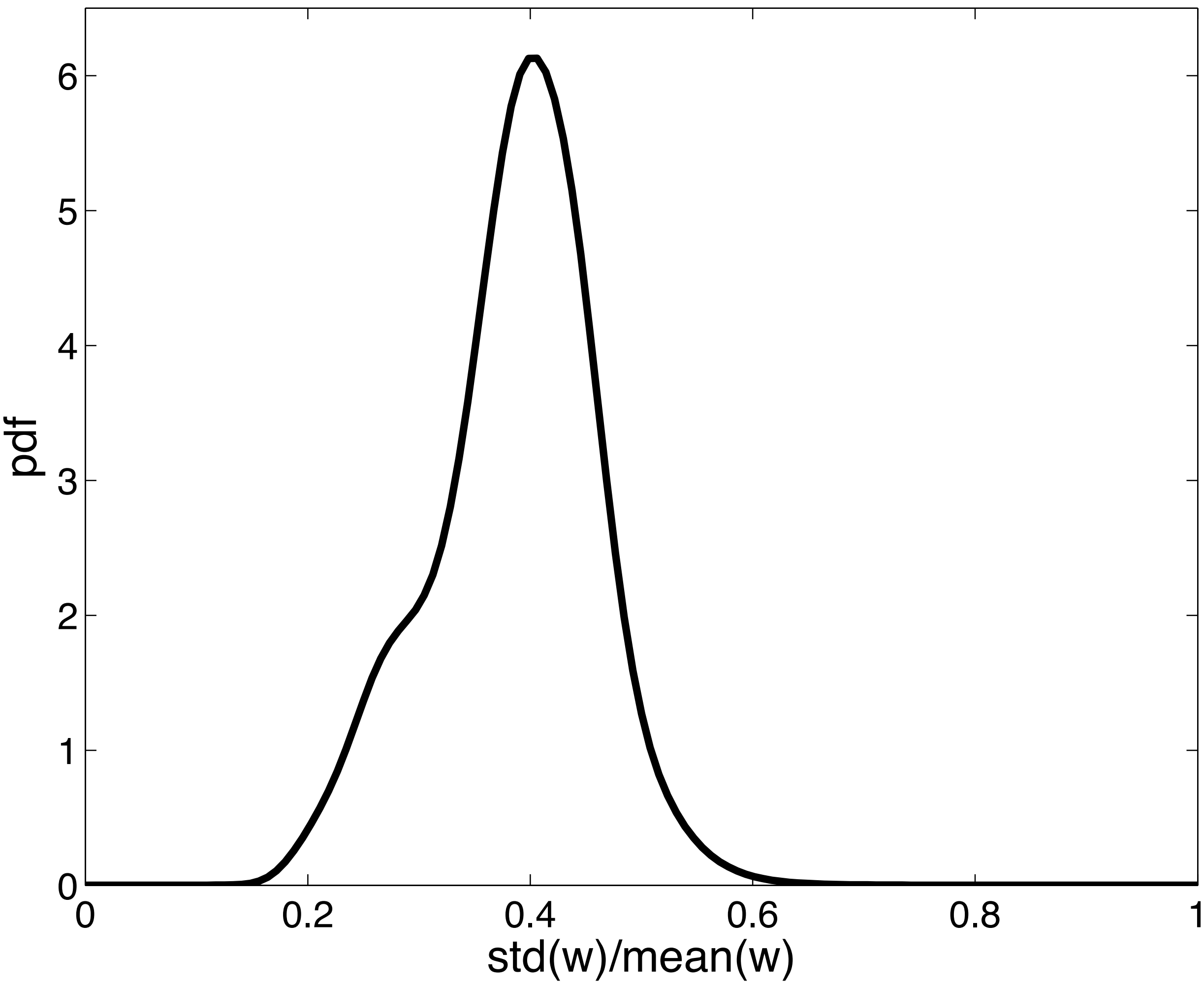

Figure 3. Kernel density estimate of the spatial probability density function of the ratio of the standard deviation to the mean of wind speed, from the SeaWinds data between 60°S and 60°N.

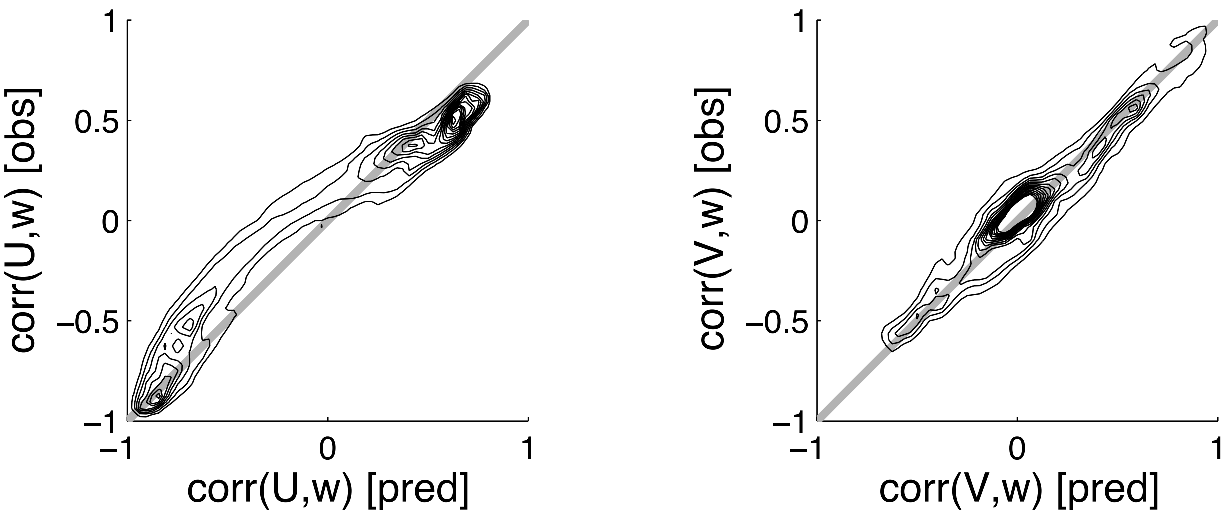

Figure 4. Kernel density estimate of the spatial joint distribution of the observed and modelled correlations of wind speed w with vector wind components. Left panel: correlation of w with zonal wind U. Right panel: correlation of w with meridional wind V.

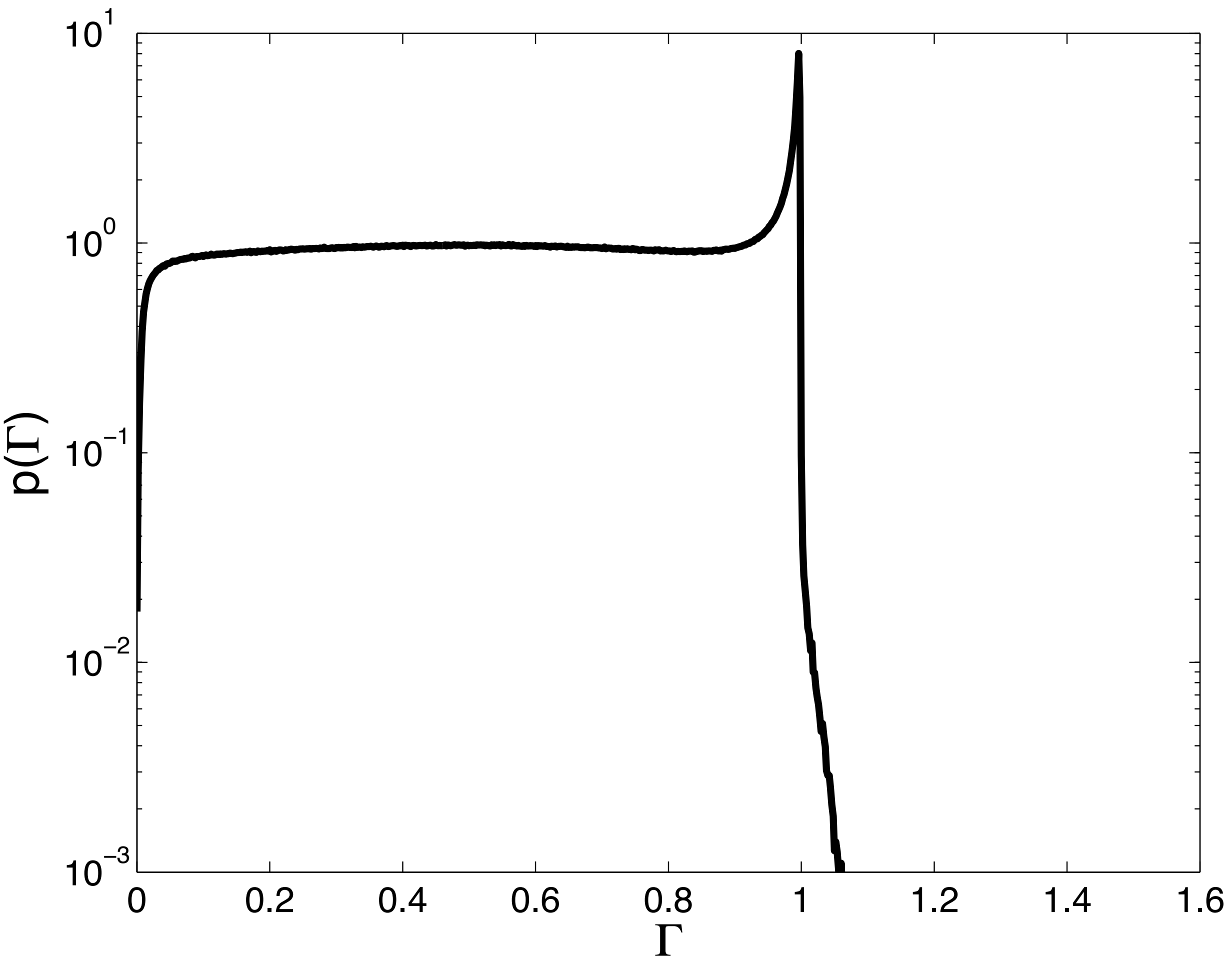

Figure 5. Probability density function of the ratio Γ between the modelled wind speed correlation prediction skill and that of the best-predicted vector wind component (Eqn. 61), uniformly sampled over the parameter ranges $0 \leq \rho_u, \rho_v \leq 1$, $-1 \leq r \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma_u) \leq 1$, $-1 \leq \log_{10}(\sigma_v/\sigma_u) \leq 1$. Note the logarithmic scale on the y-axis.

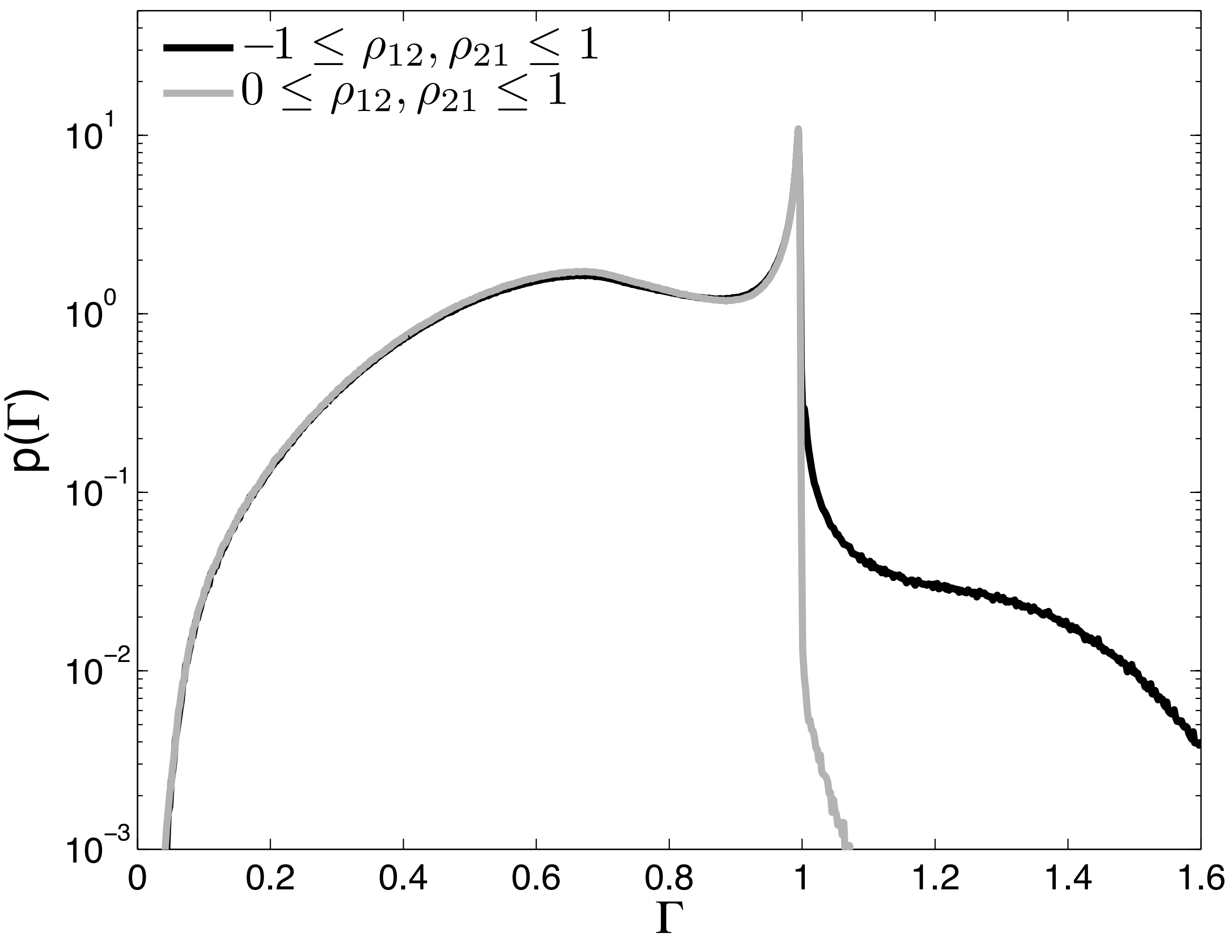

Figure 6. Black curve: probability density function of the ratio Γ between the wind speed prediction skill and that of the best-predicted vector wind component (Eqn. 85), obtained from uniform sampling of the parameter values over the ranges $0 \leq \rho_{11}, \rho_{12} \leq 1$, $-1 \leq \rho_{12}, \rho_{21} \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma) \leq 1$. Grey curve: as for the black curve, but sampling only positive cross-correlations $0 \leq \rho_{12}, \rho_{21} \leq 1$. Note the logarithmic scale on the y-axis.

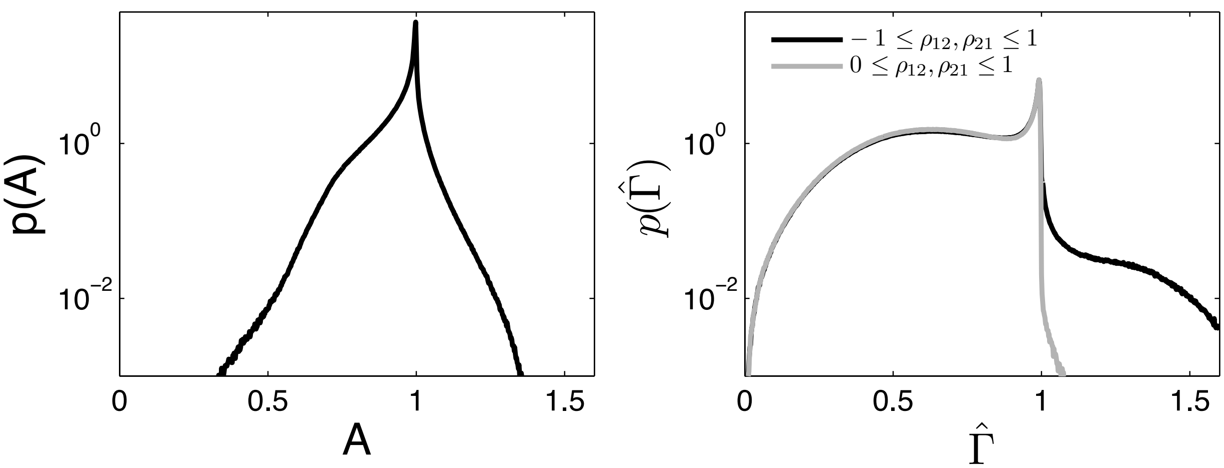

Figure 7. Left panel: Probability distribution of A (Eqn. 92) obtained by sampling the parameter values uniformly over the range $0 \leq \rho_{11}, \rho_{22} \leq 1$, $-1 \leq \rho_{12}, \rho_{21} \leq 1$, $-1 \leq \log_{10}(\overline{u}/\sigma) \leq 1$. Right panel: As in Figure 6 for the ratio $\hat{\Gamma}$ (Eqn. 93). Note the logarithmic scale on the y-axis of each panel.

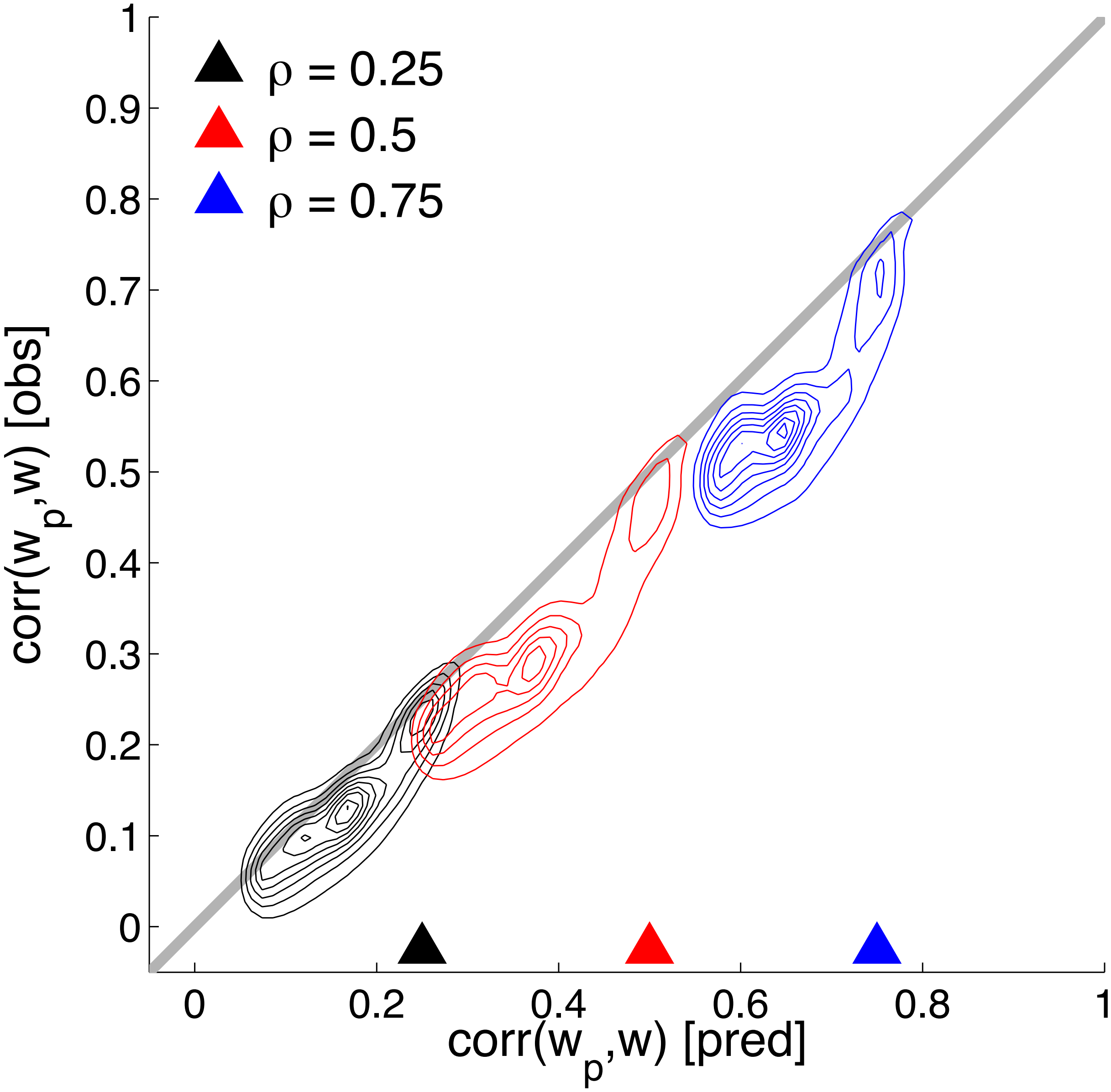

Figure 8. Kernel density estimate of the relationship between the observed wind speed predictability, $\text{corr}(w_p, w)$, and the theoretical value from Eqns. (30), (79) and (80). The vector wind predictors u_p and v_p were obtained from Eqns. (94)-(97) with $\rho = 0.25$ (black contours), $\rho = 0.5$ (red contours), $\rho = 0.75$ (blue contours).

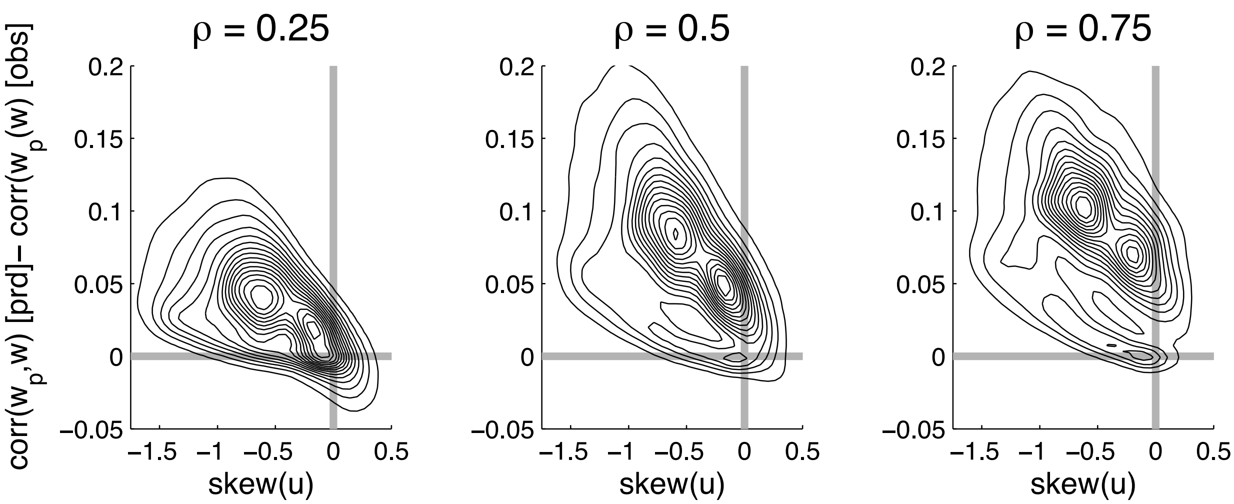

Figure 9. Kernel density estimates of the spatial joint distribution of the bias (theoretical value less observed value) in $\mathrm{corr}(w_p, w)$ with the observed skewness of the along mean-wind component of the vector winds, $u = (\overline{U}U + \overline{V}V)/\sqrt{U^2 + V^2}$.

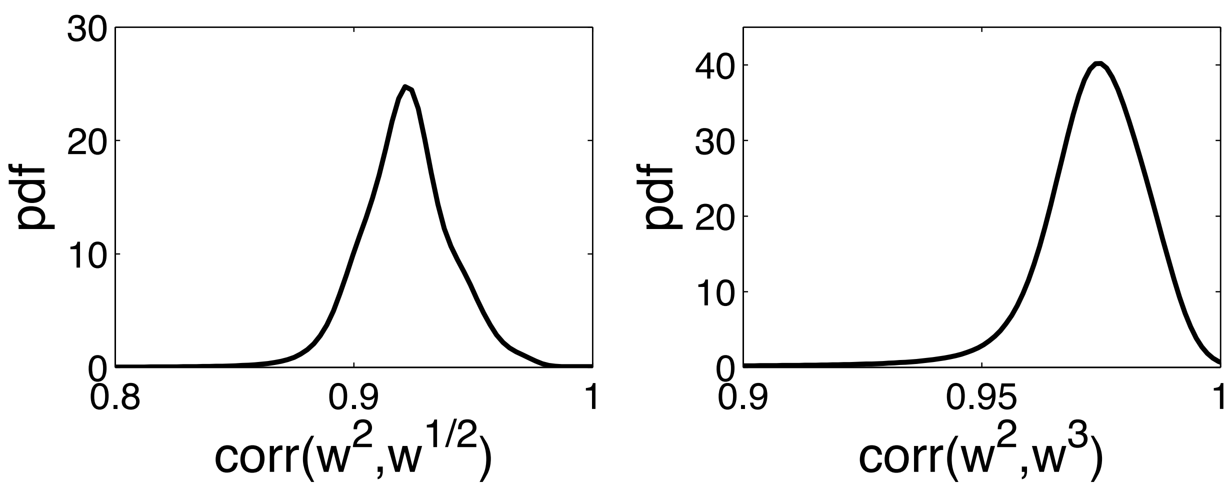

Figure 10. As in Figure 1, for the correlation between w^2 and $w^{1/2}$ (left panel) and the correlation between w^2 and w^3 (right panel).

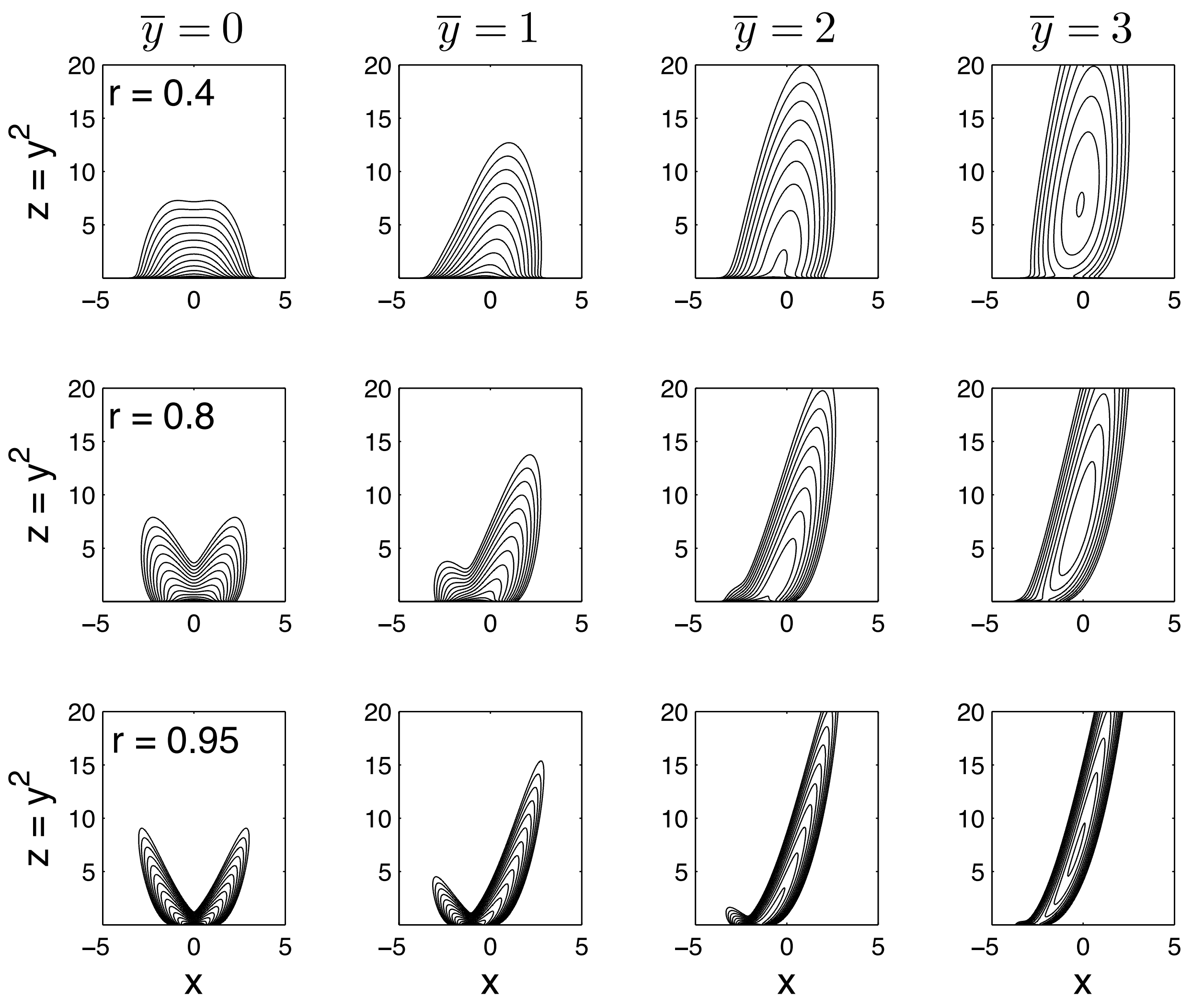

Figure 11. Joint pdf of x and $z = y^2$ for x, y bivariate Gaussian with $\mathrm{corr}(x, y) = r$, for a range of values of $\overline{y}$ and r.