

Effects of post-fire vegetation regrowth on wind fields over complex terrain

R. Quill^{a b}, J.J. Sharples^{a b} and L.A. Sidhu^a

^a*School of Physical, Environmental and Mathematical Sciences, UNSW Canberra, Australia.*

^b*Bushfire and Natural Hazards Cooperative Research Centre, Australia.*

Email: rachael.quill@student.adfa.edu.au

Abstract: The spread of bushfire is highly sensitive to wind speed and direction. Consequently, strong variation of wind fields over areas of complex terrain, with multi-scale changes in topography and surface roughness, significantly complicates the prediction of bushfire behaviour. In the current suite of operational fire modelling schemes, mesoscale variations in wind fields are often over-simplified, thereby reducing the accuracy of fire behaviour modelling in areas of potentially volatile or dangerous fire behaviour. To address this issue, emerging fire modelling frameworks are using ensemble-based approaches to accommodate for the inherent uncertainties in the factors influencing fire behaviour. To account for the effects of wind variability within ensemble approaches it is necessary to recast wind field information in probabilistic terms.

As part of an investigation into a probabilistic representation of wind field information, the mountainous region west of Canberra is used as a case study. In particular, analysis is focussed on Flea Creek Valley within the Brindabella Ranges. The valley runs in a north-south direction, approximately perpendicular to the dominant west-north-westerly (WNW) prevailing wind direction. Flea Creek Valley, and much of the surrounding region, was heavily burnt during the 2003 Canberra bushfires but the region has since experienced no major fire activity. In 2007 and 2014, wind data were collected across a 3-4km east-west transect of Flea Creek Valley.

To statistically characterise wind fields across complex terrain, the directional response of surface winds to changes in prevailing winds is considered as a toroidal, or bivariate circular, probability distribution. To construct these distributions, wind direction measured on the ridge top (indicative of the prevailing winds) is plotted against the concurrent wind direction measured within the landscape. Discrete observed directional response distributions are taken as noisy realisations of continuous underlying distributions of wind direction response. These continuous distributions can be estimated using a number of mathematical techniques including cubic or thin plate smoothing splines. Both the observed discrete and estimated continuous distributions highlight the modal nature of wind direction response across the landscape. Understanding the impacts of variables such as surface roughness on this directional response is an important step towards spatially extending a statistical characterisation of wind fields across complex terrain.

To discern any changes in the probabilistic response caused by changes in surface roughness due to vegetation regrowth, a number of mathematical and statistical comparison techniques are available. In this paper, techniques from astronomy, biometrics and statistics are employed to investigate the effects of seven years of post-fire vegetation regrowth on the directional response of surface winds across Flea Creek Valley.

The findings of this study suggest that the choice of statistical test as well as smoothing technique can have a significant impact on the results. Despite this, there is also evidence that the wind response across Flea Creek Valley may have been significantly altered by regrowth in some areas, but in other areas no significant difference is found. There are important implications here for wind and fire modelling, and it is clear that there is much more work to be done to better understand the impacts of physical variables on the probabilistic characterisation of wind fields.

Keywords: *Wind modelling, vegetation regrowth, complex terrain, extended Kolmogorov-Smirnov test, non-parametric surface comparison*

1 INTRODUCTION

The spread of bushfire is highly sensitive to wind flow. Sudden changes in wind speed and direction can cause unpredictable changes in fire behaviour. Over areas of complex terrain, multi-scale variations in topography and surface roughness significantly complicate the accurate prediction of wind fields (e.g. Finnigan, 2000). In current operational fire modelling schemes, mesoscale variations in wind fields are often over-simplified, reducing the accuracy of modelling in areas of potentially volatile or dangerous fire behaviour (Tolhurst *et al.*, 2008). To address this issue, emerging fire modelling frameworks are using ensemble-based approaches to understand the inherent variabilities in fire behaviour (French *et al.*, 2013). A probabilistic approach to wind modelling would be well suited to these new fire modelling frameworks, and complement current models to enhance discussions of uncertainty and ‘what-if’ or ‘worst-case’ bushfire scenarios.

To develop a probabilistic understanding of wind variability over complex terrain, the response of surface wind direction within complex terrain to changes in the prevailing wind direction is represented as a toroidal, or bivariate circular, probability distribution (Gross and Wippermann, 1987; Whiteman and Doran, 1993; Sharples *et al.*, 2010). Furthermore, understanding the impacts of variables such as surface roughness on these wind direction response distributions becomes an important step in developing a statistical characterisation of wind flow using physical parameters, and extending this characterisation across the landscape. To capture the changes in distribution caused by changes in surface roughness, a number of comparison techniques are available from the mathematical and statistical literature (Peacock, 1983; Bowman, 2006; Wang and Ye, 2010).

For the case study outlined in Section 2, wind direction response distributions are constructed for four points across a valley transect. Two sets of these distributions are constructed; the first using 2007 data and the second using data from 2014, after seven years of post-fire vegetation regrowth. At each of the four points in the valley, the response distributions from each year are compared using three non-parametric comparison techniques to test for statistically significant differences. This comparison rigorously builds evidence to support whether or not post-fire regrowth has significantly impacted wind direction response across the case study region.

2 CASE STUDY

In 2003, much of the region west of Canberra was heavily burnt by devastating bushfires; since this time there have been no major fires in the area. In 2007 and 2014, wind direction data were collected in Flea Creek Valley, which runs north-south through the Brindabella Ranges, west of the ACT. The 2007 data were analysed by Sharples *et al.* (2010) and the data collected in 2014 are presented here for comparison. Wind direction response distributions were constructed using data from four points across a 3-4km east-west transect of the valley for both years. Table 1 describes the location of each of the four points across the valley as well as the differences in vegetation between the two sample years. Figure 1 permits a visual assessment of the differences in vegetation across the valley between the two years.

Table 1. Weather stations in Flea Creek Valley, with vegetation descriptions from 2007 (Sharples *et al.*, 2010) and 2014.

	Location	2007 Vegetation	2014 Vegetation
Point 1	East-facing, steep slope. Lee-ward to the prevailing WNW winds.	Sparse Eucalypt regrowth. Some taller trees, some with cambial growth. Partially intact canopy overhead.	Intermediate foliage. Some large trees.
Point 2	Valley floor, flat terrain.	Some burnt trees. Scattered canopy overhead.	Medium density canopy surrounding station. Dense scrub up to 1m.
Point 3	West-facing, mid slope. Stations were in similar but not identical positions between years.	Very dense (dead) bracken. Some larger trees. Some with cambial growth. Sparse canopy overhead.	Little to no scrub. Some tall trees. Areas of dense cambial growth.
Point 4	West-facing, steep slope. Windward to prevailing WNW winds.	Very sparse acacia regrowth and bracken. Some large eucalypts. Relatively sparse canopy overhead.	Scrub up to 2m high. Large trees surrounding station with cambial growth.

The wind direction response distributions show the densities of concurrent prevailing wind directions measured on the ridge top and surface wind directions measured within the valley. Wind direction was recorded in 22.5° bins, corresponding to the 16 points of the compass, so the observed wind response distributions were represented as 16×16 discrete bivariate distributions. These discrete distributions were considered to be

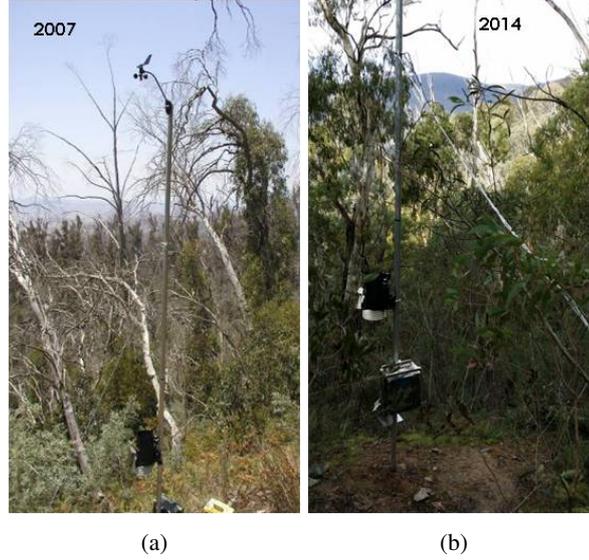


Figure 1. Photographs of vegetation in Flea Creek Valley in (a) 2007 and (b) 2014.

noisy realisations of underlying smooth continuous distributions. For the set of observed wind direction pairs, $X_{ij} = (\phi_i, \psi_j)$, where ϕ_i is the wind direction measured on the ridge top and ψ_j is the wind direction measured in the valley, a model was proposed;

$$Y_k(X_{ij}) = m_k(X_{ij}) + \epsilon_k(X_{ij}). \quad (1)$$

For each year $k = \{2007, 2014\}$, Y_k is the observed wind direction distribution, m_k is some underlying continuous distribution and ϵ_k is a random noise component. A thin plate smoothing (TPS) spline method was used to estimate the underlying continuous surface, m_k . The TPS spline method allowed for a trade-off between data fidelity and smoothness of the surface to account for the inherent noise in the observations. A cubic spline was also used to estimate m_k . This exact interpolation method was more computationally efficient than the thin plate smoothing spline, but could exaggerate noise from the data. Both methods ensured that the toroidal nature of the surface was maintained during estimation (Quill *et al.*, 2015).

The estimated continuous distributions were realised on a 360×360 grid, equivalent to 1° intervals around the compass. Figure 2 shows the estimated continuous wind direction response distributions using the TPS spline method for the four points across Flea Creek Valley in both 2007 and 2014. It is apparent that there are some similarities between the surface pairs, such as the number of modes across each surface. However, there also appear to be some differences, including the shape and density of some modes as well as shifts in modal locations. Seven years of undisturbed post-fire regrowth has occurred in Flea Creek Valley between the two sample years but, for all but one pair, the sampling locations have remained identical. The investigation in this paper addresses whether the observed differences in wind direction distributions are caused by these vegetation changes.

3 METHODOLOGY

In order to understand the impacts of the post-fire vegetation regrowth on the wind direction response distributions from Flea Creek Valley, statistical comparison techniques are used to determine whether the apparent differences observed between the two years (Figure 2) are statistically significant. Given the model in (1) for $k = \{2007, 2014\}$, the following hypotheses are proposed:

$$H_0 : m_{07}(X_{ij}) = m_{14}(X_{ij}),$$

$$H_1 : m_{07}(X_{ij}) \neq m_{14}(X_{ij}).$$

The null hypothesis assumes equality between the two distributions, suggesting that the vegetation regrowth has had no significant impact on the directional response of wind flow. While the alternative hypothesis posits inequality between distributions, and a significant impact on wind direction response across the valley caused by regrowth.

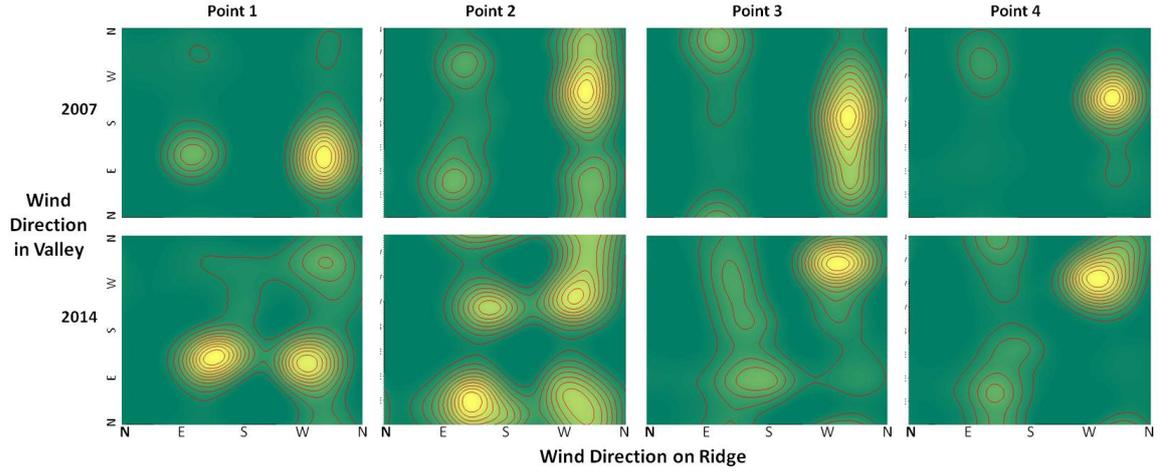


Figure 2. Estimated continuous wind response distributions for four points along the transect of Flea Creek Valley in 2007 and 2014.

If the wind direction responses are considered as a probability distribution function, there are a number of parametric and non-parametric tests available to test the above hypotheses. The wind direction response distributions can also be thought of as surfaces or images, where each cell has a value. Non-parametric techniques from areas such as bioinformatics can be used to measure the average difference between the values across each surface. For this study, the wind response distributions are compared using an extension of the Kolmogorov-Smirnov test for equality between probability distributions (Section 3.1) and two non-parametric comparison tests for equality between surfaces (Section 3.2).

3.1 Extended Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (KS) test considers the difference between an observed dataset and a known probability density function. The test takes the maximum difference between the empirical distribution function (EDF) of the observed dataset and the theoretical cumulative distribution function (CDF) of the known distribution. The test can also be used to compare two observed datasets by considering the two empirical cumulative distributions. The univariate KS test is well-known and commonly used in fields of research such as astronomy (Peacock, 1983). The test is highly efficient, distribution-free and invariant to the ordering of the CDF.

Peacock (1983) extended the univariate KS test to the bivariate case, and showed it to again be efficient and sufficiently distribution-free for practically useful cases. The extended KS test becomes invariant to the ordering of the CDF by constructing the test statistic based on the maximum differences between all possible CDFs over the bivariate space. In the bivariate space, the CDF can be defined in four directions; $Q1 = (x \leq X, y \leq Y)$, $Q2 = (x \leq X, y \geq Y)$, $Q3 = (x \geq X, y \leq Y)$ and $Q4 = (x \geq X, y \geq Y)$. The test statistic for the extended KS test becomes the maximum of the maximum difference between CDFs defined in each of these directions.

For a comparison of a single sample, with n data points, to a known distribution of the variables X and Y , the test statistic is defined as (Gosset, 1987);

$$D_n = \max(D_n^{Q1}, D_n^{Q2}, D_n^{Q3}, D_n^{Q4}), \quad (2)$$

where D_n^Q defines the maximum absolute difference between the CDFs of the sample and the known distribution constructed over each direction, for example

$$D_n^{Q1} = \sup_{(x,y)} |F_n^{Q1}(x,y) - F_{X,Y}^{Q1}(x,y)|, \quad (3)$$

where $F_{X,Y}^{Q1}(x,y) = P(x \leq X, y \leq Y)$. Since this statistic, D_n , is assumed to be proportional to \sqrt{n} , an adjusted statistic is used;

$$Z_n = \sqrt{n}D_n. \quad (4)$$

The asymptotic behaviour of Z_n is given by Peacock (1983);

$$P(Z_\infty > z) \simeq 2 \exp(-2(z - 0.5)^2). \quad (5)$$

To compare two sample datasets with sizes $n_1, n_2 \gtrsim 10$, Peacock (1983) defines;

$$n = \frac{n_1 n_2}{n_1 + n_2}. \quad (6)$$

In this study, two observed datasets are compared, and so the Kolmogorov-Smirnov test statistic is calculated using the empirical distribution functions (EDFs) from the observed discrete datasets. The p-values for the test are derived from (5). For each test, the value of Z_n is also compared to the critical values derived by Peacock (1983) and given in Table 2.

Table 2. Relevant critical values of Z_n given in Peacock (1983).

n	Significance Levels			
	0.20	0.10	0.05	0.01
...
50	1.57	1.71	1.83	2.06

3.2 Non-parametric Surface Comparison Tests

The wind direction response distributions can also be thought of as surfaces or images, where each pixel or cell has a value. In this case, two non-parametric surface comparison tests are used to compare the estimated continuous surfaces from 2007 and 2014. The first test is developed using the methodology described by Wang and Ye (2010). This was developed to compare medical images and test for significant changes due to treatment. The test statistic is derived from the mean squared difference between the sample values across the surfaces;

$$T_W = \frac{1}{n} \sum_{ij} \{\hat{m}_{07}(X_{ij}) - \hat{m}_{14}(X_{ij})\}^2 \quad (7)$$

where n is the number of data points.

The second test statistic is defined by Bowman (2006) and was used to test the difference between two air quality datasets. The test statistic takes the difference between each surface \hat{m}_k and the surface constructed using both datasets under the null hypothesis of equality, \hat{m} ;

$$T_B = \frac{1}{2} \sum_{ijk} \{\hat{m}_k(X_{ij}) - \hat{m}(X_{ij})\}^2 \quad (8)$$

for $k = \{2007, 2014\}$.

To calculate the p-values for these tests, the distributions of the statistics under the null hypothesis are constructed using Monte Carlo simulations. The errors, $\epsilon_k(X_{ij})$, are assumed to be homoscedastic across the wind response distributions so bootstrap re-sampling is used within each Monte Carlo simulation.

4 RESULTS AND DISCUSSION

Table 3 shows the p-values given by each comparison test for each pair of wind direction response distributions between 2007 and 2014. Table 2 shows the critical values derived by Peacock (1983) for comparison with the observed Z_n statistics also given in Table 3. Peacock (1983) only empirically derived the critical values up to a sample size of 50, so these values only serve as an indication since the sample sizes here are $n_1 = n_2 = 256$.

It is clear from Table 3 that according to the extended KS test all of the pairs are significantly different, with very small p-values. Although the comparison of the Z_n statistics with the critical values in Table 2 can not be absolute due to the differing sample sizes, the statistic values are significantly larger than those given and so the null hypothesis of equality between surfaces is rejected for all points across the valley.

Table 3. Results of KS test between wind response surfaces from 2007 and 2014. For the KS test, both the p-values and observed test statistics are shown. For the non-parametric comparison tests, the p-values are shown for both surface estimation techniques.

	KS Test		<i>Cubic</i>	T_W		T_B	
	P_f	Z_n		<i>TPS</i>	<i>TPS</i>	<i>Cubic</i>	<i>TPS</i>
Point 1	1.46×10^{-9}	3.7434	0.0140	0.0000	0.1070	0.0000	
Point 2	2.57×10^{-7}	3.3166	0.0001	0.0000	0.3820	0.0020	
Point 3	2.06×10^{-19}	5.1752	0.0020	0.0000	0.0770	0.0000	
Point 4	8.28×10^{-20}	5.2239	0.0380	0.0010	0.1490	0.0010	

For both non-parametric tests applied to the TPS spline, very small p-values are shown across all points. This again suggests that all pairs are significantly different. However, the results of both non-parametric tests on the cubic spline surfaces provide a different story. For the T_W test, Point 1 and Point 4 show the highest p-values, suggesting less evidence against equality, and for the T_B test, the same points also show high p-values. Across both test statistics, Point 3 shows small p-values providing evidence against equality. This is to be expected since the two stations of this point were not in the same location in each year and the distributions in Figure 2 are visually very different.

The findings from both tests on the TPS surfaces could be a result of disproportionate weight being given to the small differences between surfaces away from the modes of the distributions. Under the cubic spline, these areas are mostly zero so the differences are also zero.

For the cubic spline surfaces, there are similar patterns shown between the results from both tests. However, there is a distinct difference between the magnitude of p-values given by each test, which may be caused by their constructions and applications. For instance, T_W was constructed to compare medical images and so be highly sensitive to changes in pixel values, while T_B may not need to be as sensitive when considering environmental variables.

The results for Point 2 also highlight the influence of statistic construction on the results. Under T_W , Point 2 shows the smallest p-value, while under T_B the point gives the highest p-value. The two surfaces from 2007 and 2014 exhibit significant variation and less distinctive modal patterns than the other three pairs (Figure 2). The differences taken between the two surfaces, i.e. for T_W , would therefore be relatively large, producing a small p-value and suggesting inequality. Whereas, the difference taken between each surface and their combined surface, i.e. for T_B , would be smaller across the entire surface than if the modes showed less variation. In this case, the p-value is therefore large. The large p-value suggests there is not enough evidence against the null hypothesis to reject the assumption of equality between the two wind response surfaces.

5 CONCLUSIONS

Initial results from three different statistical comparison tests show that the choice and construction of the tests themselves, as well as the surface estimation techniques, can have a significant impact on the findings. The small p-values shown across the KS test suggest that all surface pairs are not equal and thus the post-fire vegetation regrowth has had a significant impact on wind direction response in Flea Creek Valley. However, further investigation into the critical values of the test under large sample sizes is required to validate this conclusion.

The results of the non-parametric tests applied to continuous surfaces estimated by the thin plate smoothing spline method also suggest that the post-fire vegetation regrowth has had a significant impact on wind direction response across Flea Creek Valley. These results are, however, likely due to a disproportionate weighting towards extremely small differences across the estimated surfaces as opposed to the zero differences shown when applied to the cubic spline. This issue could be investigated further through applying some weighting to the differences around modes of the distributions, or defining a cut-off where extremely small surface values are taken to be zero.

The results of the non-parametric tests applied to the cubic spline surfaces provide a more interesting story; suggesting that at places across the landscape, particularly on the steep lee and windward slopes, the increased vegetation has had no significant impact on the directional response of wind flow. The choice of test statistic still significantly influences these results and further investigation is required to consider the mechanics of all three tests. However these initial results should serve to indicate the importance of physical parameters in the development of statistical models of wind direction response distributions for use in fire modelling.

6 FURTHER WORK

It is clear that further investigation into the application of these statistical techniques to wind direction response distributions is required to verify these initial findings. In addition, there are a number of ways this study can progress towards the statistical characterisation of wind fields over complex terrain;

- Since all the tests used in this study consider the differences across the entire surface or distribution, the results give a view of the broad scale differences in wind direction response distributions without providing any further details on how and where the surfaces or distributions differ. Analysis of the sensitivity of the tests to known changes in modal location and spread will help to understand the mechanics of each test in the context of wind response distributions as well as enable better interpretation of the results.
- To further investigate the differences in wind direction response distributions, parametric tests will also be appropriate. For instance, it is meaningful to define the location, density and spread of each mode on the surface, and compare how these parameters have changed between the two years. This will identify how growth in vegetation has shifted the response of wind flow in the valley to specific prevailing winds. The shifts in modal spread or variability will also help to indicate potential changes in predictability or uncertainty with increased vegetation growth.
- In the context of fire modelling, a quantification of the impacts of vegetation regrowth on wind direction response will build into the probabilistic characterisation of wind fields across complex terrain. Further studies have begun to enable some quantification of vegetation across a landscape and so analyse the quantitative relationship between vegetation cover and wind direction response distributions.

ACKNOWLEDGEMENTS

The authors would like to thank NSW National Parks & Wildlife Services for allowing the research to be conducted within Brindabella National Park, NSW. The first author, R. Quill, is supported by a UNSW Canberra Scholarship and by the Bushfire and Natural Hazards Cooperative Research Centre.

REFERENCES

- Bowman, A. W. (2006). Comparing nonparametric surfaces. *Statistical Modelling* 6(4), 279–299.
- Finnigan, J. (2000). Turbulence in plant canopies. *Annual Review of Fluid Mechanics* 32(1), 519–571.
- French, I., B. Cechet, T. Yang, and A. Sanabria (2013). FireDST: Fire Impact and Risk Evaluation Decision Support Tool - model description. In J. Piantadosi, R. Anderssen, and J. Boland (Eds.), *MODSIM2013, 20th International Congress on Modelling and Simulation*, Adelaide, Australia. Modelling and Simulation Society of Australia and New Zealand Inc.
- Gosset, E. (1987). A three-dimensional extended Kolmogorov-Smirnov test as a useful tool in astronomy. *Astronomy and Astrophysics* 188(1), 258–264.
- Gross, G. and F. Wippermann (1987). Channeling and countercurrent in the Upper Rhine Valley: Numerical simulations. *Journal of Climate and Applied Meteorology* 26(10), 1293–1304.
- Peacock, J. A. (1983). Two-dimensional goodness-of-fit testing in astronomy. *Monthly Notices of the Royal Astronomical Society* 202(Feb 1983), 615–627.
- Quill, R., J. J. Sharples, and L. Sidhu (2015). Approximation of toroidal surfaces arising from noisy bivariate circular data. *In Prep.*
- Sharples, J. J., R. H. D. McRae, and R. O. Weber (2010). Wind characteristics over complex terrain with implications for bushfire risk management. *Environmental Modelling and Software* 25(10), 1099–1120.
- Tolhurst, K. G., B. Shields, and D. Chong (2008). Phoenix: development and application of a bushfire risk management tool. *Australian Journal of Emergency Management* 23(4), 47–54.
- Wang, X.-F. and D. Ye (2010). On nonparametric comparison of images and regression surfaces. *Journal of Statistical Planning and Inference* 140(10), 2875–2884.
- Whiteman, C. D. and J. C. Doran (1993). The relationship between overlying synoptic-scale flows and winds within a valley. *Journal of Applied Meteorology* 32(11), 1669–1682.