NEURAL COMPUTING
NCRG
RESEARCH GROUP

Neural Computing Research Group
Dept of Computer Science & Applied Mathematics
Aston University
Birmingham B4 7ET
United Kingdom
Tel: +44 (0)121 333 4631
Fax: +44 (0)121 333 4586
http://www.ncrg.aston.ac.uk/

NEUROSAT: An Overview

Dan Cornford and Ian Nabney
d.cornford@aston.ac.uk, i.t.nabney@aston.ac.uk

Technical Report NCRG/98/011 May 1, 1998

Project Report Version 1.0

1 Introduction

This report gives an overview of the work being carried out, as part of the NEUROSAT project, in the Neural Computing Research Group at Aston University. The aim is to give a general review of the work and methods, with reference to other documents which provide the detail. The document is ongoing and will be updated as parts of the project are completed. Thus some of the references have yet to be completed.

In the broadest sense, the Aston part of NEUROSAT is about using neural networks (and other advanced statistical techniques) to extract wind vectors from satellite measurements of ocean surface radar backscatter. The work involves several phases, which are outlined below. A brief summary of the theory and application of satellite scatterometers forms the first section. The next section deals with the forward modelling of the scatterometer data, after which the inverse problem is addressed. Dealiasing (or disambiguation) is discussed, together with proposed solutions. Finally a holistic framework is presented in which the problem can be solved.

2 Satellite Scatterometer Theory

Obtaining wind vectors over the ocean is important to Numerical Weather Prediction (NWP) since the ability to produce a forecast of the future state of the atmosphere depends critically on knowing the current state accurately (Haltiner and Williams, 1980). However, the observation network over the oceans (particularly in the southern hemisphere) is very limited (Daley, 1991). Thus it is hoped that the global coverage of ocean wind vectors provided by satellite borne scatterometers will improve the accuracy of weather forecasts by providing better initial conditions (Harlan and

O'Brien, 1986; Lorenc *et al.*, 1993). The scatterometer data also offers the ability to improve wind climatologies over the oceans (Levy, 1994) and the possibility of studying, at high resolution, interesting meteorological features such as cyclones (Dickinson and Brown, 1996).

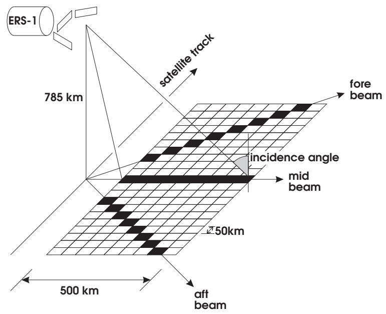

Figure 1: The ERS-1 scatterometer geometry.

The ERS-1 satellite was launched in July 1991 by the European Space Agency. Many instruments were carried by the satellite (Offiler, 1987), including the Advanced Microwave Instrument which is capable of indirectly measuring both ocean surface waves and winds. The on-board microwave radar operates at 5.3 GHz and measures the amount of backscatter generated by small ripples on the ocean surface of around 5 cm wavelength. Measured backscatter from the ocean surface is given as the Normalised Radar Cross Section, and generally denoted by σ^o, which has units of decibels. A 500 km wide swathe is swept by the satellite along the track of its polar orbit, with nineteen cells sampled across the swathe, each cell having dimensions of roughly 50 by 50 km (Figure 1). Thus there is some overlap between cells. Also, each cell is sampled from three different directions by the fore, mid and aft beams respectively giving a triplet of σ^o's, $(\sigma^o_1, \sigma^o_2, \sigma^o_3)$. This $\boldsymbol{\sigma^o}$ triplet, together with the incidence angle of the beams (which varies across the swathe) can be used to determine the average wind vector within the cell (Offiler, 1994).

Many methods to compute wind vectors from scatterometer data exist. Most have considered model based techniques (Wentz, 1991; Stoffelen and Anderson, 1992; Offiler, 1994) where a physically based mapping from wind vectors to $\boldsymbol{\sigma^o}$ is formulated. Thiria *et al.* (1993) modelled the mapping from $\boldsymbol{\sigma^o}$ to wind vectors using simulated data, and a neural network based classifier, which gave probabilities[1] of the wind direction being in each of thirty-six intervals. Simulated data was used since real $\boldsymbol{\sigma^o}$ measurements were not available at the time the work was undertaken. This group (Sylvie Thiria, Michel Crepon, Carlos Badran and Phillipe Richaume) are also involved in NEUROSAT, enhancing their model. To our knowledge, no other published work has considered the prediction of wind vectors directly from $\boldsymbol{\sigma^o}$.

2.1 The geophysical scatterometer model

Much effort has been put into understanding the theoretical relationship between σ^o and wind direction (Wentz, 1991; Stoffelen and Anderson, 1992; Offiler, 1994). This has been based on studies of the physical processes that govern backscattering from water surfaces (Ebuchi *et al.*, 1993) together with analysis of the relationship between wind vectors (both buoy observed and NWP derived) and scatterometer measurements (Offiler, 1994). From these studies empirical forward models between single σ^o's and relative wind direction (ϑ) have been established of the

[1] Strictly, a classification problem was solved, interpretting the network outputs as direction-class conditional probabilities, although these could be negative and need not sum to one.

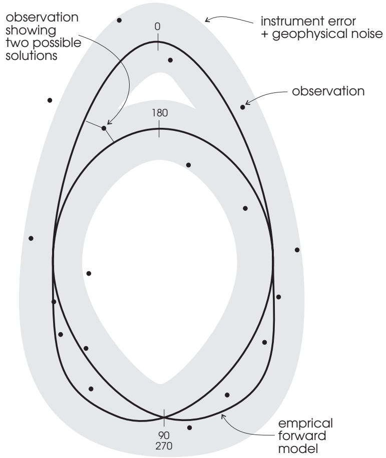

Figure 2: A theoretical sketch of the relationship between backscattered radiation and wind direction, for a fixed wind speed. This corresponds to a slice across the 'cone-like' manifold. The height up the cone defines the wind speed.

general form

$$\sigma^o \sim b_0 + b_1 cos(\vartheta) + b_2 cos(2\vartheta) \tag{1}$$

where the coefficients are complicated functions of the scatterometer incidence angle and the wind speed. The most widely used and currently operational forward model is known as CMOD4 (Offiler, 1994; Stoffelen and Anderson, 1997). We have three $\boldsymbol{\sigma^o}$ measurements for each cell and these together define a cone-like manifold in 3 dimensional space. For most $\boldsymbol{\sigma^o}$ triples, which are observed with noise, there is ambiguity over the optimal direction to select (Figure 2).

This is typical of many inverse problems in the geophysical sciences, where the forward model output is uni-valued for a given set of inputs (e.g. $\boldsymbol{\sigma^o}$ as a function of wind direction) but the inverse model is multi-valued (e.g. wind direction as a function of $\boldsymbol{\sigma^o}$). The relationship between wind speed and σ^o is, however, known to be uni-valued (Thiria *et al.*, 1993). Since the wind speed is largely uncorrelated with the *relative* wind direction, the problem of modelling a wind vector can be split into modelling the speed and direction separately. We should note that in general we are considering *relative* wind direction - that is wind direction which is relative to the satellite azimuth angle[2] - rather than absolute wind direction.

It is worth noting the distinction between the forward model:

$$(u, v) \rightarrow \boldsymbol{\sigma^o} \tag{2}$$

and the inverse problem:

$$\boldsymbol{\sigma^o} \rightarrow (u, v) \tag{3}$$

The use of the term inverse problem here is rather loose. In the engineering literature an inverse problem is generally concerned with the estimation of model parameters given some observations (Cornford, 1998b). Here we use the term inverse problem, in the sense of statistical terminology, to denote the fact that the problem is multi-valued.[3]

Operationally, the problem of obtaining wind directions from scatterometer data is resolved using the CMOD4 forward model (at present) and minimising some cost function (which is typically

[2]The azimuth angle gives the clockwise angle from north of the scatterometer beam incident on the cell.

[3]Taking a Bayesian perspective leads to the conclusion that model parameters and variables can be *treated* in exactly the same way, thus no distinction would be necessary.

a sum of squares error) between the observed $\boldsymbol{\sigma^o}$ triplets and the manifold defined by CMOD4. A look up table is used and NWP forecasts improve the chances of finding the correct solution (Offiler, 1994). Up to 4 possibly valid solutions are generally obtained (although there are often only 2 solutions with 180 degree ambiguity — the true and alias solutions) and some other method is applied to decide which direction is to be selected. Two techniques have been employed based on local median filters (Schultz, 1990; Shaffer *et al.*, 1991) and the use of background wind vectors from NWP forecasts (Chelton *et al.*, 1989). The ambiguity removal problem is discussed later.

3 The NEUROSAT Approach at Aston

The work carried out in this project is performed within a pragmatic Bayesian (O'Hagan, 1994) framework. It is felt that probability theory is one of the most powerful ways to deal with uncertainty in natural systems. For instance when we measure backscatter from the surface of the ocean we are not measuring the 'true' value (if this even exists). There are a multitude of factors from transmission errors, instrument errors and errors derived from the transmissivity properties of the atmosphere (which vary in space and time) which prevent us measuring the backscatter itself exactly. Thus the probabilistic framework (Cornford, 1997a) provides a natural framework with which we can address, and cope with, these errors. Once one has accepted that the probabilistic framework is the correct way to proceed, one naturally arrives at a Bayesian conclusion, since Bayes theorem is nothing more than a definition of the rules of conditional probability.

We adopt a pragmatic framework because we recognise that while the Bayesian solution is often the 'correct' solution (given your modelling assumptions) it is often desirable to take into account the computational feasibility. Bayesian methods tend to be based on Monte Carlo methods which often makes them very computationally demanding. While this is no argument for not applying Bayesian methods the results of this project might not be greatly appreciated if it took one week on an 'average' workstation to derive one wind field from a single scatterometer swathe. Thus we are pragmatic (rather than dogmatic) and are happy to use other principled methods along side the Bayesian ones and justified simplifications within the Bayessian approach.

The most general information we can have about the scatterometer data and wind vectors is given by the joint probability $P(\boldsymbol{u_i}, \boldsymbol{\sigma_i^o})$ where we use $\boldsymbol{u_i} = (u_i, v_i)$ to represent the wind vector in the i'th cell. Note that we can also write the vector $\boldsymbol{u_i}$ as the speed and direction $(\|\boldsymbol{u_i}\|, \vartheta_i)$. Throughout this document we assume that $\boldsymbol{\sigma_i^o}$ is a vector of observations of backscatter triple together with other information necessary such as the incidence angle of the beam, all from the i'th cell of the scene[4]. We also assume that we are working in a reference frame relative to the satellite azimuth angle, which allows us notational simplicity since we can neglect the azimuth angle and never explicitly mention the incidence angle unless specifically required.

The joint probability can be written:

$$P(\boldsymbol{u_i}, \boldsymbol{\sigma_i^o}) = P(\boldsymbol{u_i}|\boldsymbol{\sigma_i^o})P(\boldsymbol{\sigma_i^o}) = P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i})P(\boldsymbol{u_i}) \tag{4}$$

The conditional probability $P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i})$ can be interpreted as the forward model (2) while $P(\boldsymbol{u_i}|\boldsymbol{\sigma_i^o})$ describes the inverse problem (3). What we are really after is:

$$P(\boldsymbol{u}|\boldsymbol{\sigma^o}) = \frac{\prod_i P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i})P(\boldsymbol{u})}{\prod_i P(\boldsymbol{\sigma_i^o})} \propto \prod_i P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i})P(\boldsymbol{u}) \tag{5}$$

where $P(\boldsymbol{u}|\boldsymbol{\sigma^o})$ is the conditional probability density of $\boldsymbol{u}$ and $\boldsymbol{\sigma^o}$ the wind vectors and backscatter over the whole scene. This is just Bayes theorem (in one form) and forms the basis of the

[4] Cell is used to describe the $50 \times 50\ km$ region which forms one observation by the satellite. A group of cells, measured almost simultaneously as the satellite over-passes, is called a scene.

NEUROSAT approach at Aston. Note that:

$$P(\boldsymbol{\sigma^o}|\boldsymbol{u}) = \prod_i P(\boldsymbol{\sigma^o_i}|\boldsymbol{u_i}) \tag{6}$$

because given the wind vector in each cell, theoretically there is a one to one mapping to the backscatter values. Given a good model, there is no reason for the conditional probabilities (errors) to be spatially correlated[5]. However:

$$P(\boldsymbol{u}|\boldsymbol{\sigma^o}) \neq \prod_i P(\boldsymbol{u_i}|\boldsymbol{\sigma^o_i}) \tag{7}$$

because this time even if we have the backscatter measurements for each cell the ambiguity (one to many mapping) means that we cannot uniquely determine the probability of the wind field[6].

We have proposed several algorithms to retrieve wind vectors from the scatterometer data and these are outlined below.

In the first approach we attempt to directly estimate $P(\boldsymbol{u_i}|\boldsymbol{\sigma^o_i})$ using mixture density networks. We can then use a number of heuristics to choose the correct solution from those possible (ambiguity removal). This is developed in section 4. We shall also examine the use of the so called *scaled likelihood trick* to use the inverse model with prior models over wind fields (Williams, 1997)

In the second approach we model the forward problem, $P(\boldsymbol{\sigma^o_i}|\boldsymbol{u_i})$ using neural networks and then use Bayes theorem (5) together with a (spatial) prior field wind model $P(\boldsymbol{u})$ to determine the posterior $P(\boldsymbol{u}|\boldsymbol{\sigma^o})$ which should be unimodal, or nearly so. This is developed in section 5.

4 Solving the Inverse Problem

Recall that we seek to model $P(\boldsymbol{u_i}|\boldsymbol{\sigma^o_i})$. We will use the framework of mixture density networks (Bishop, 1994; Williams, 1996) to estimate the conditional probability distributions as functions of the input variables, $\boldsymbol{\sigma^o}$. There are several ways that one could attempt to perform this estimation.

We have already done some work examining the possibility of the splitting the wind vector into speed and direction component. A multi-layer perceptron (Bishop, 1995) with linear output units is used to estimate the wind speed - which is a single valued function and can thus be trained using a sum of squares error function (which corresponds to maximum likelihood estimation under the assumption of Gaussian errors). The wind direction is the predicted using a mixture density network which includes wind speed as an input. Care must be taken to ensure the periodic nature of wind direction is taken into account. We use mixtures of circular normal distributions to do this (Bishop and Nabney, 1996). The results are promising, however it is still necessary to select the correct directions (speed is uniquely predicted). We must also note at this point we do not have $P(\boldsymbol{u_i}|\boldsymbol{\sigma^o_i})$ rather we have $E[||\boldsymbol{u_i}|||\boldsymbol{\sigma^o_i}]$ and $P(\vartheta_i|\boldsymbol{\sigma^o_i}, ||\boldsymbol{u_i}||)$ where E denotes the expectations. If we had a density model for $||\boldsymbol{u_i}||$ then we could theoretically combine the two probabilities to obtain $P(\boldsymbol{u_i}|\boldsymbol{\sigma^o_i})$.[7]

[5] Actually the conditional probabilities may be correlated since satellite borne sensors often have correlated errors. This would be due to the same sensor being used to observe every cell.

[6] In this we are assuming perfect observations and models. When observation and model errors are considered then strictly speaking neither (6) or (7) are likely to have equality. Also, surrounding observations may impart useful information to the individual cell probabilities. It then becomes a modelling descision as to whether the spatial context is given in $\boldsymbol{u}$ or $\boldsymbol{\sigma^o}$. We choose to put priors over $\boldsymbol{u}$ because wind fields are reasonable well understood.

[7] The density model is emplicitly defined to be a fixed variance Gaussian in the wind speed network. This is probably not a very reasonable model - the noise might be expected to increase with $\boldsymbol{u_i}$.

We assume $\|\boldsymbol{u_i}\|$ to be the unique (correct) value and choose the four most likely directions ϑ_i from $P(\vartheta_i|\boldsymbol{\sigma_i^o}, \|\boldsymbol{u_i}\|)$. We have then attempted to use several ambiguity removal procedures, which are all rather ad-hoc, yet fast and potentially very useful. One powerful method for defining a wind field is to use div-curl splines (Wahba, 1982; Amodei and Benbourhim, 1991; Cornford, 1997b). The div-curl spline is fitted to all the wind vectors and then the 'correct' directions are chosen from the possible directions on the basis of the minimum error between the spline and chosen direction. This procedure can be iterated, but clearly requires at least 50% of the most likely solutions to be correct. There are many different heuristics and models available here - see Cornford (1997b) for methodological details and Cornford (1998a) for implementation heuristics and results.

We are also using mixture density networks to directly model $P(\boldsymbol{u_i}|\boldsymbol{\sigma_i^o})$ together with MSc student David Evans. This approach mitigates the need for combining distributions at the end, and allows us to apply the *scaled likelihood trick* of Williams (1997). This will require priors over possible wind fields and will be dealt with in the next section. We can also apply the set of heuristic techniques (such as div-curl splines) to the two or four most likely wind vectors extracted from $P(\boldsymbol{u_i}|\boldsymbol{\sigma_i^o})$.

In addition to the considerations outlined above we shall assess the role of the neural network model which forms the basis of the mixture density network (that is the multi-layer perceptron or radial basis function network) investigate the role of regularisation (Bishop, 1995) and assess the effect of changing the inputs to the networks (for instance the possibility of having one model for each of the satellite tracks and the use of information from surrounding scatterometer observations - cf. (Thiria *et al.*, 1993)).

5 Solving the Forward Problem

Although there are a number of forward models currently in existence, such as CMOD4, none are probabilistic. In order to progress we needed a probabilistic forward model: $P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i})$. This is being developed using the mixture density network framework of Williams (1996) by another MSc student Guillaume Ramage. Once we have this forward model we can use Bayes theorem to get:

$$P(\boldsymbol{u}|\boldsymbol{\sigma^o}) \propto \int \left(\prod_i P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i}) \right) P(\boldsymbol{u}|\boldsymbol{\theta_u}) P(\boldsymbol{\theta_u}) d\boldsymbol{\theta_u} \tag{8}$$

where $P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i})$ is a local model for each scatterometer cell, while $P(\boldsymbol{u}|\boldsymbol{\theta_u})$ is a spatial (prior) model for wind fields defined by parameters $\boldsymbol{\theta_u}$. In this case we are going to use random field based priors (Cornford, 1997a). We may be able to evaluate this integral analytically (which will be a good thing) and thus compute the posterior analytically since $P(\boldsymbol{\sigma_i^o}|\boldsymbol{u_i})$ is Gaussian. If we cannot evaluate the posterior analytically then we will be able to sample from it using Markov Chain Monte Carlo (MCMC) methods (Smith and Robserts, 1993; Neal, 1997). This will be more fully discussed elsewhere.

In order to enhance the prior model $P(\boldsymbol{u}|\boldsymbol{\theta_u})$ we have investigated wind field models based on the flexible modified Bessel function based covariance (Cornford, 1997a; Cornford, 1998c) and models which include discontinuities (Cornford, 1998d) which allow us to represent fronts (Cornford, 1997c). The model becomes considerably more complex, but may help greatly to resolve fronts - which are important features.

6 Ambiguity Removal

The problem of ambiguity removal has been largely addressed in the preceding two sections. We shall be investigating two main methods. The first involves heuristically based ('quick and dirty') methods based on models for wind fields. Div-curl splines (Cornford, 1997b) are suitable for this section although random field model priors are also possible (Cornford, 1997a). The four (or two) most likely solutions are chosen from $P(\boldsymbol{u_i}|\boldsymbol{\sigma_i^o})$ and heuristics are used to select those which best fit the adaptive wind field model. These methods will be fast but sensitive to the heuristics used. The methods are unlikely to perform well unless 75 percent of the most likely solutions are the correct ones (this is based on recent results yet to be published).

The second method uses Bayes theorem together with the forward (probabilistic) model and the random field wind field prior. As outlined previously, this is likely to be computationally intensive because of the need to use MCMC techniques. We hope to use Hybrid Monte Carlo (Duane *et al.*, 1987) methods to speed up the convergence of the Markov Chain and ensure we are sampling from regions of high posterior density.

6.1 Final Processing

It is envisaged that once the correct ambiguous solution has been selected, an accurate forward model, developed by Carlos Badran in France, will be used to produce the final wind vectors. It is important that this forward model is initialised near the true (as given by the scatterometer measurement) wind vector, thus the previous modelling steps are crucial. It is hoped to use the Jacobian of the forward model to find the 'true' value. It remains to be seen whether this step will improve the quality of the wind fields produced.

7 Conclusion

The above report has given an overview of what work is being carried out in the NCRG as part of the NEUROSAT program. The intention is not to completely specify all parts of the project, rather to provide pointers to relevant sources of information. This is an ongoing project and thus some of the reports may not yet be available. It is envisaged that this will be updated as progress is made.

References

Amodei, L. and M. N. Benbourhim 1991. A Vector Spline Approximation. *Journal of Approximation Theory* **67**, 51–79.

Bishop, C. M. 1994. Mixture Density Networks. Technical Report NCRG/94/004, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Bishop, C. M. 1995. *Neural Networks for Pattern Recognition.* Oxford: Oxford University Press.

Bishop, C. M. and I. T. Nabney 1996. Modelling Conditional Probability Distributions for Periodic Variables. *Neural Computation* **8**, 1123–1133.

Chelton, D. B., M. H. Freilich, and J. R. Johnson 1989. Evaluation of Unambiguous Vector Winds from the Seasat Scatterometer. *Journal of Atmospheric and Oceanic Technology* **6**, 1024–1039.

Cornford, D. 1997a. Random Field Models and Priors on Wind. Technical Report NCRG/97/023, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Cornford, D. 1997b. Splines and Vector Splines. Technical Report NCRG/97/024, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Cornford, D. 1997c. Surface Wind Fields (on Earth). Technical Report NCRG/97/022, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Cornford, D. 1998a. Ambiguity Removal: Heuristically Based Methods. Technical Report NCRG/98/???, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Cornford, D. 1998b. Inverse Problem Theory. Research Note NCRG/98/???, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Cornford, D. 1998c. Modified Bessel Function Based Covariances. Research Note NCRG/98/???, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Cornford, D. 1998d. Random Field Models with Fronts. Research Note NCRG/98/???, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK.

Daley, R. 1991. *Atmospheric Data Analysis.* Cambridge: Cambridge University Press.

Dickinson, S. and R. A. Brown 1996. A Study of Near-Surface Winds in Marine Cyclones Using Multiple Satellite Sensors. *Journal of Applied Meteorology* **35**, 769–781.

Duane, S., A. D. Kennedy, B. J. Pendleton, and D. Roweth 1987. Hybrid Monte Carlo. *Physics Letters B* **195**, 216–222.

Ebuchi, N., H. Kawamura, and Y. Toba 1993. Physical Processes of Microwave Backscattering from Laboratory Wind Wave Measurements. *Journal of Geophysical Research* **98**, 14669–14681.

Haltiner, G. J. and R. T. Williams 1980. *Numerical Prediction and Dynamic Meteorology.* Chichester: John Wiley.

Harlan, J. and J. J. O'Brien 1986. Assimilation of Scatterometer Winds into Surface Pressure Fields Using a Variational Method. *Journal of Geophysical Research* **91**, 7816–7836.

Levy, G. 1994. Southern-Hemisphere Low-Level Wind Circulation Statistics from the SeaSat Scatterometer. *Annales Geophysicae - Atmospheres, Hydroshperes and Space Sciences* **12**, 65–79.

Lorenc, A. C., R. S. Bell, S. J. Foreman, C. D. Hall, D. L. Harrison, M. W. Holt, D. Offiler, and S. G. Smith 1993. The Use of ERS-1 Products in Operational Meteorology. *Advances in Space Research* **13**, 19–27.

Neal, R. M. 1997. Monte Carlo Implementation of Gaussian Process Models for Bayesian Regression and Classification. Technical Report 9702, Department of Statistics, University of Toronto.

Offiler, D. 1987. Wind Measurements from the Earth Remote-sensing Satellite (ERS-1). *Meteorological Magazine* **116**, 279–285.

Offiler, D. 1994. The Calibration of ERS-1 Satellite Scatterometer Winds. *Journal of Atmospheric and Oceanic Technology* **11**, 1002–1017.

O'Hagan, A. 1994. *Kendall's Advanced Theory of Statistics, Volume 2B: Bayesian Inference.* London: Edward Arnold.

Schultz, H. 1990. A Circular Median Filter Approach for Resolving Directional Ambiguities in Wind Fields Retrieved from Spaceborne Scatterometer Data. *Journal of Geophysical Research* **95**, 5291–5303.

Shaffer, S. J., R. S. Dunbar, S. V. Hsiao, and D. G. Long 1991. A Median-Filter-Based Ambiguity Removal Algorithm for NSCAT. *IEEE Transactions on Geoscience and Remote Sensing* **29**, 167–174.

Smith, A. F. M. and G. O. Robserts 1993. Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods. *Journal of the Royal Statistical Society- Series B* **55**, 2–23.

Stoffelen, A. and D. Anderson 1997. Scatterometer Data Interpretation: Estimation and Validation of the Transfer Function CMOD4. *Journal of Geophysical Research* **102**, 5767–5780.

Stoffelen, A. and D. L. T. Anderson 1992. ERS-1 Scatterometer Data Characteristics and Wind Retrieval Skill. In B. Kaldeich (Ed.), *Proceedings of the First ERS-1 Symposium - Space at the Service of the Environment*, pp. 41–47. Cannes, France, ESA.

Thiria, S., C. Mejia, F. Badran, and M. Crepon 1993. A Neural Network Approach for Modeling Nonlinear Transfer Functions: Application for Wind Retrieval from Spaceborne Scatterometer Data. *Journal of Geophysical Research* **98**, 22827–22841.

Wahba, G. 1982. Vector Splines on the Sphere, with Application to the Estimation of Vorticity and Divergence from Discrete, Noisy Data. In W. Schempp and K. Zeller (Eds.), *Multivariate Approximation Theory, Vol. 2*, pp. 407–429. Basel: Birkhauser Verlag.

Wentz, F. J. 1991. A Simplified Vector Algorithm for Satellite Scatterometers. *Journal of Atmospheric and Oceanic Technology* **8**, 697–704.

Williams, C. K. I. 1997, September. Combining Spatially Distributed Predictions from Neural Networks. Research Note.

Williams, P. M. 1996. Using Neural Networksto Model Conditional Multivariate Densities. *Neural Computation* **8**, 843–854.