Modelling Frontal Discontinuities in Wind Fields

Dan Cornford   **Ian T. Nabney**   **Christopher K. I. Williams***
d.cornford@aston.ac.uk   i.t.nabney@aston.ac.uk   ckiw@dai.ed.ac.uk

Abstract

A Bayesian procedure for the retrieval of wind vectors over the ocean using satellite borne scatterometers requires realistic prior near-surface wind field models over the oceans. We have implemented carefully chosen vector Gaussian Process models; however in some cases these models are too smooth to reproduce real atmospheric features, such as fronts. At the scale of the scatterometer observations, fronts appear as discontinuities in wind direction. Due to the nature of the retrieval problem a simple discontinuity model is not feasible, and hence we have developed a constrained discontinuity vector Gaussian Process model which ensures realistic fronts. We describe the generative model and show how to compute the data likelihood given the model. We show the results of inference using the model with Markov Chain Monte Carlo methods on both synthetic and real data.

Keywords: Gaussian Processes, Discontinuities, Wind Fields, Fronts.

*Now at: Division of Informatics, University of Edinburgh, 5 Forrest Hill, Edinburgh EH1 2QL, Scotland, UK

1 Introduction

Knowledge of the current state of the atmosphere is critical for the prediction of its future evolution. The problem of estimating the current state of the atmosphere is important in meteorology (Daley, 1991). Observations of the current state of the atmosphere are sparse, particularly over the oceans. Recently scatterometers have been used to infer wind vectors over the ocean (Offiler, 1994). A scatterometer is a satellite-borne microwave radar that measures the ratio of the power transmitted (toward the Earth's surface) to the power received (from scattering at the Earth's surface). The retrieval of wind fields from scatterometer observations over the oceans is non-trivial since the physics of the scattering mechanism by small waves at the ocean surface is not perfectly understood. Thus statistical models, such as the current operational model CMOD4 (Stoffelen and Anderson, 1997), are used for wind retrieval. The wind retrieval is often an ad-hoc procedure and frequently referred to as disambiguation. This term is used because there are typically two to four valid local wind vector solutions given a single noisy scatterometer observation. This ambiguity is the source of the requirement to impose constraints on the fronts it is possible to identify.

We propose a Bayesian method for retrieval, which can be seen as an application of the classical methods for the solution of inverse problems (Tarantola, 1987). The aim is to obtain $p(U, V \mid \Sigma^o)$, the conditional probability of the wind vector field, (U, V), given the satellite observations, Σ^o. Using Bayes' theorem:

$$p(U, V \mid \Sigma^o) = \frac{p(\Sigma^o \mid U, V) p(U, V)}{p(\Sigma^o)}. \tag{1}$$

Once the Σ^o have been observed, $p(\Sigma^o)$ is a constant and thus we can write:

$$p(U, V \mid \Sigma^o) \propto p(\Sigma^o \mid U, V) p(U, V), \tag{2}$$

and $p(\Sigma^o | U, V)$ gives the probability of the scatterometer observations given a proposed wind field. Further details of the approach can be found in (Nabney *et al.*, 1998), the overarching approach being similar to that in (Long, 1993). In this paper the prior model for wind fields, $p(U, V)$, is of concern.

2 Prior Wind Field Models

The prior wind field models, $p(U, V)$, should be able to represent all the possible wind fields which can be observed in the atmosphere, however it should also penalise very unlikely wind fields, and minimise the possibility of retrieving physically unreasonable wind fields. The main constraint that we can use to specify the prior model is a smoothness constraint.

2.1 Smooth wind fields

Figure 1 shows two examples of our best estimates of what the real wind field is like at the scale of our scatterometer observations. The wind vectors are taken from a Numerical Weather Prediction (NWP) model (Haltiner and Williams, 1980) which uses a set of simplified differential equations and parameterisations to forecast the future evolution of the atmosphere given an initial state. Due to the nature of the model, which is discretised (either over a grid or in terms of spectral coefficients), these wind fields are over-smooth with respect to reality, but form our best estimates. Since the main use of scatterometer winds will be for initialising NWP models, it seems appropriate that our prior model should reflect the scales represented in NWP models.

We use a vector Gaussian Process model to represent the smooth wind fields. The model is based on the decomposition of a vector field into purely divergent flow (from the velocity potential) and purely rotational flow (from the stream function), known as Helmholtz' theorem (Daley, 1991). We
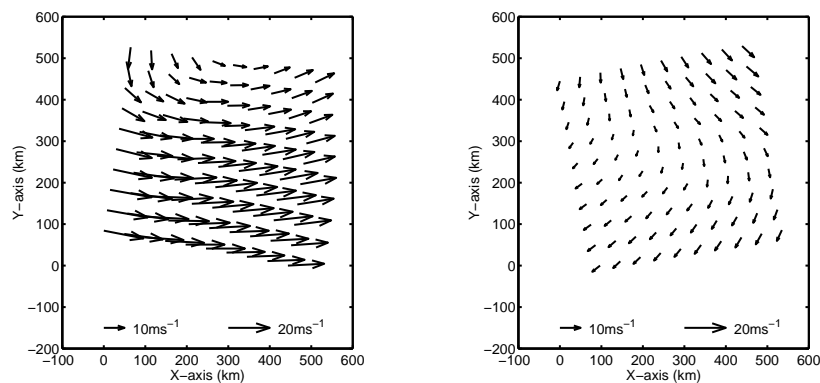
Figure 1: Two typical smooth wind fields from February 1998.

do not give details here, the full development can be found in (Cornford, 1998). Essentially a scalar modified Bessel covariance function (Abrahamsen, 1997) based Gaussian Process was applied to both the stream function (Ψ) and velocity potential (Φ) rather than directly to the wind vector components. The wind vector components can be written in terms of derivatives of the stream function and velocity potential, and thus the covariance functions for the wind vector components are computed in terms of second order derivatives of the stream function and velocity potential covariances (Cornford, 1998). This allows control over the ratio of divergence to vorticity in the resulting wind field, and automatically produces valid, positive definite, joint covariance matrices for the wind vector components. By controlling the length scales, process variance, smoothness and noise variance a very flexible model for wind fields is produced with the additional benefit of being able to control the ratio of divergence to vorticity through the relative magnitudes of the variances of the stream function and velocity potential.
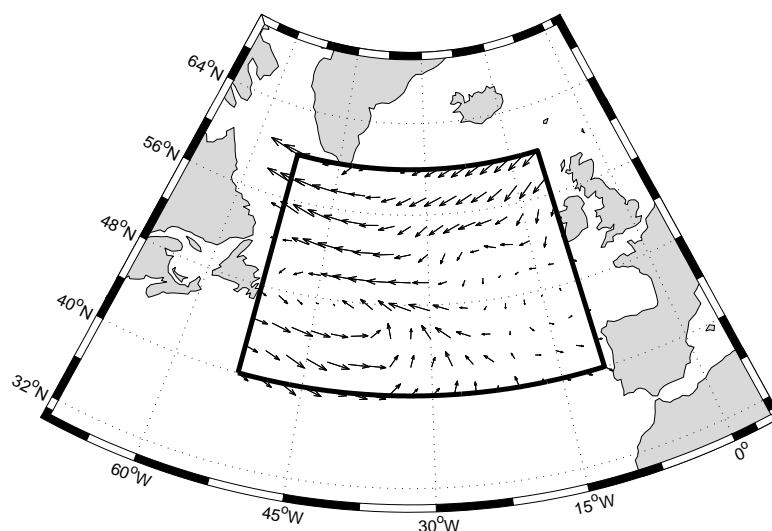

Figure 2: The area of the North Atlantic over which the prior model is defined together with an example of the ECMWF model winds for midnight 25/04/95

In this work we consider a sector of the North Atlantic from $52.5^\circ W, 40^\circ N$ to $10^\circ W, 60^\circ N$ (Figure 2) since this is the most reliably observed ocean region. In order to set the parameters in the wind field prior a large amount of data from the European Centre for Medium range Weather Forecasting (ECMWF) was used; this consists of gridded analysis data on a regular 2.5 degree latitude-longitude grid. The surface wind field[1] for a small region of the North Atlantic is extracted from the global,

---

[1]The standard vector component wind field measured at 10 m above the ocean surface

gridded data set. The British Atmospheric Data Centre[2] performed the interpolation from the model grid to the regular grid using software provided by ECMWF. An example wind field is shown in Figure 2. In order to obtain reliable climatological estimates of the parameters in the wind field model (in particular the mean parameter) three complete years of data from 1995, 1996 and 1997 were used.

Data before 1995 was not used in order to take advantage of the revisions made to the ECMWF model surface wind parameterisations in 1994, which increased model $10\ m$ wind speeds. The wind patterns in the North Atlantic show strong seasonal periodicity. Generally, winds are stronger during the winter season than during the summer season. It is also possible that the characteristic length scales of features change through the seasons. In order to account for this a separate prior wind field model was developed for each month, using the three years data. Since ECMWF analyse the incoming synoptic data (including wind observations) every six hours, there are roughly 360 sample wind fields for each month, which were used to estimate (or sample) the parameters of the prior wind field models.
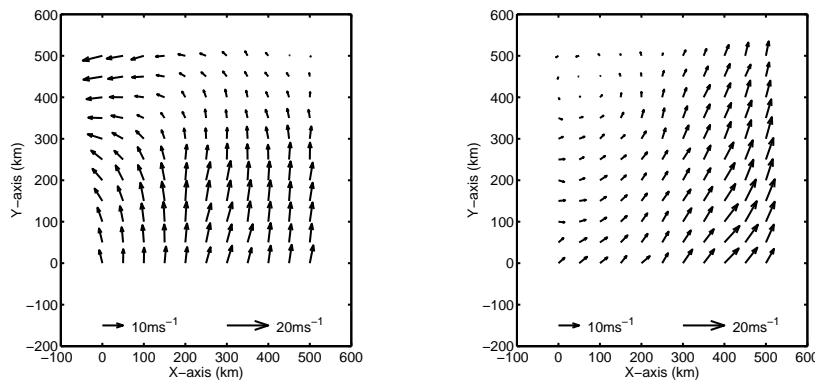

Figure 3: Two simulated wind fields using Gaussian Process parameters for January.

In this work we consider the maximum *a posteriori* probability values of the parameters of the wind field model (that is the parameters of the covariance functions on the stream function and velocity potential). We placed very loose priors over the parameters based on expert knowledge of the atmospheric system. This is possible because our model has physically interpretable parameters. The smoothness parameter in the modified Bessel covariance function was found to be ~ 2.5 for all months, suggesting the wind fields were once differentiable (Adler, 1981). With this value, the modified Bessel covariance function simplifies to a polynomial-exponential covariance function which has the form:

$$C(r) = E^2 \left(1 + \frac{r}{L} + \frac{r^2}{3L^2}\right) \exp\left(-\frac{r}{L}\right) + \eta^2 \tag{3}$$

where r is the separation distance of two points, L is a characteristic length scale parameter, E^2 is the energy (process variance) and η^2 is the noise variance. This form is much quicker to compute. Under this model we can write:

$$p(U, V) = p(U, V \mid \boldsymbol{\xi}) = \frac{1}{(2\pi)^{\frac{n}{2}} \det (K_{uv})^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(U, V)' K_{uv}^{-1}(U, V)\right) \tag{4}$$

where (U, V) is an observed wind field, K_{uv} is the covariance matrix defined by (Equation 3) and the dependence on the parameters of the Gaussian Process, $\boldsymbol{\xi} = \{L, E^2, \eta^2\}$ is made explicit.

Simulations from this prior model can be seen in Figure 3. In most cases this prior model is sufficient. However it does not capture the full range of possible wind fields. There is a class of atmospheric feature which are not well represented by this model.

---

[2] See http://www.badc.rl.ac.uk/

2.2   Frontal wind fields
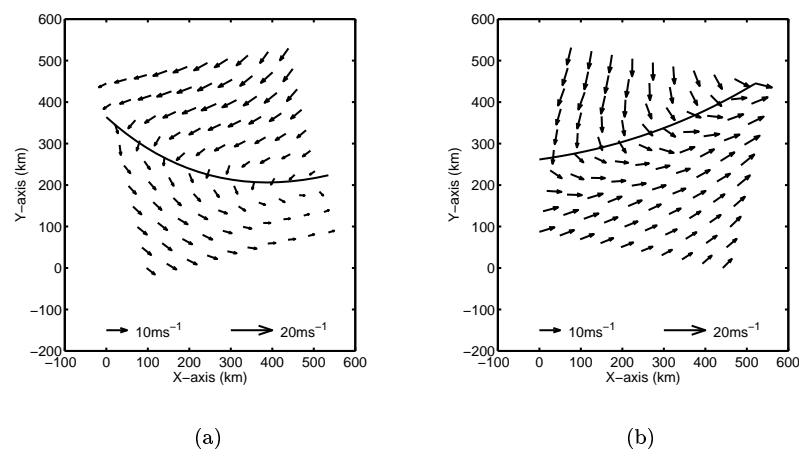

(a)   (b)

Figure 4: Two typical frontal wind fields from February 1998. The approximate location of the fronts, as determined automatically using the model described herein, is also shown.

The most difficult atmospheric features to include in the wind field model are surface fronts. Fronts are generated by complex atmospheric dynamics and are marked by large changes in the surface wind direction and temperature. Figure 4 shows two examples of 'real' fronts as observed by the NWP model and interpolated to the scatterometer observation locations. Since both the NWP model and the interpolation procedure will act as smoothers, the 'true' wind field is probably more discontinuous. In order to account for such features, which appear discontinuous at our (scatterometer) observation scale, we have developed a model for vector Gaussian Processes with constrained discontinuities. The generative model for a wind field including a front is taken to be a combination of two vector-valued Gaussian Processes with a constrained discontinuity.
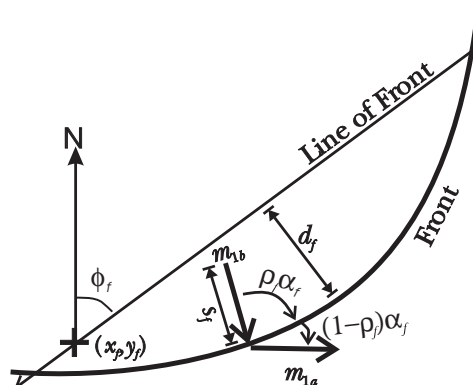

Figure 5: A graphical description of the frontal model.

The generative model is developed to maintain a simple but physically realistic representation. Initially the frontal position and orientation are simulated. They are given by the angle clockwise from north (ϕ_f) that the line of the front makes and a point on that line (x_f, y_f). The curvature of the front is then defined by a distance orthogonal to a point on the line of the front, d_f (this is typically taken to be the central point on the front with respect to the region within which the front falls). This implies the front is a simple parabola within the region being considered, an assumption which, although restrictive, fits well with observations. The change in the angle of the wind across the front is $\pi - \alpha_f$ and is simulated from a distribution, given later, covering the range $[0, \pi)$. This angle is related to the vertical component of vorticity (ζ) across the front through $\zeta \propto \cos(\alpha_f/2)$

and the constraint $\alpha_f \in [0, \pi)$ ensures cyclonic vorticity at the front. The angle behind and ahead of the front may be asymmetric which is accounted for by $\rho_f \in [0.5, 0.8)$. The wind speed (s_f) is then simulated at the front. Since there is generally little change in wind speed across the front one value is simulated for both sides of the front. These components $\boldsymbol{\theta_f} = (\phi_f, x_f, y_f, d_f, \alpha_f, \rho_f, s_f)$ define the line of the front and the mean wind vectors just ahead of and just behind the front (Figure 5):

$$m_{1a} = (u_{1a}^m, v_{1a}^m) = (s_f \sin(\phi_f + (1 - \rho_f)\alpha_f), s_f \cos(\phi_f + (1 - \rho_f)\alpha_f)), \tag{5}$$

$$m_{1b} = (u_{1b}^m, v_{1b}^m) = (-s_f \sin(\phi_f - \rho_f\alpha_f), -s_f \cos(\phi_f - \rho_f\alpha_f)). \tag{6}$$

A realistic model requires some variability in wind vectors along the front. Thus we use Gaussian Processes with a non-zero mean (m_{1a} or m_{1b}) along the line of the front. In the real atmosphere we observe a smaller variability in the wind vectors along both sides of the front compared with regions away from fronts. Thus we use the same Gaussian Process along the front (GP_1), as that used in the wind field away from the front (GP_2) but with larger length scales and smaller variances. The values of the GP_1 parameters are set using expert prior knowledge. The same GP_1 parameters are used along both sides of the front, with different means. The winds just ahead of and behind the front are assumed conditionally independent given m_{1a} and m_{1b}, and are simulated at a regular 50 km spacing. The final step in the generative model is to simulate wind vectors using GP_2 in both regions either side of the front, conditionally on the values along that side of the front. This model is flexible enough to represent fronts, yet has the required constraints derived from meteorological principles, for example that fronts should always be associated with cyclonic vorticity with discontinuities at the model scale in wind direction but not in wind speed. To make this generative model useful, we need to be able to compute the data likelihood, which is the subject of the next section.

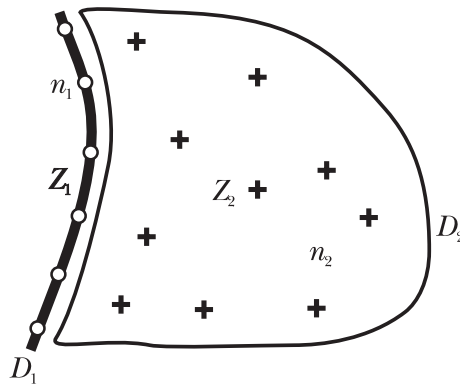# 3   Gaussian Processes with Constrained Discontinuities

Figure 6: The simplified model for a frontal discontinuity. The region is partitioned into the frontal domain, D_1, and the domain in which the wind vectors are observed, D_2.

We consider data from two domains D_1 and D_2 (Figure 6), where in this case D_1 is a curve in the plane which represents the front and D_2 is a domain that gives the area ahead of or behind the front. Consider n_1 random variables, $\boldsymbol{Z_1}$, at points $\boldsymbol{x_1}$ along the curve D_1, and assume these are generated under GP_1 (a Gaussian Process which depends on parameters $\boldsymbol{\theta_1}$ and has mean $\boldsymbol{m_1} = m_1\boldsymbol{1}$). In practice we choose n_1 to be around 10. The model shows no sensitivity to the value of n_1 provided that the spacing of points along the line is less than a tenth of the length scale of the Gaussian Process along the front. We are interested in determining the likelihood of the random variables $\boldsymbol{Z_2}$ observed at n_2 points $\boldsymbol{x_2}$ in D_2 under GP_2 which depends on parameters $\boldsymbol{\theta_2}$, conditionally on the 'constrained discontinuities' at the front, as represented by the mean $\boldsymbol{m_1}$.

We evaluate this by calculating the likelihood of $\boldsymbol{Z_2}$ conditionally on the n_1 values of $\boldsymbol{Z_1}$ from GP_1

along the front and marginalising out $\boldsymbol{Z_1}$:

$$p(\boldsymbol{Z_2}|\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{m_1}) = \int_{-\infty}^{\infty} p(\boldsymbol{Z_2}|\boldsymbol{Z_1},\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{m_1})p(\boldsymbol{Z_1}|\boldsymbol{\theta_1},\boldsymbol{m_1})\,d\boldsymbol{Z_1}. \tag{7}$$

From the definition of the likelihood of a Gaussian Process (Cressie, 1993) we find that

$$p(\boldsymbol{Z_2}|\boldsymbol{Z_1},\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{m_1}) = \frac{1}{(2\pi)^{\frac{n_2}{2}}|S_{22}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\boldsymbol{Z_2^{*\prime}} S_{22}^{-1} \boldsymbol{Z_2^*}\right), \tag{8}$$

where

$$S_{22} = K_{22|2} - K_{12|2}' K_{11|2}^{-1} K_{12|2}, \qquad \boldsymbol{Z_2^*} = \boldsymbol{Z_2} - K_{12|2}' K_{11|2}^{-1} \boldsymbol{Z_1}.$$

To understand the notation consider the joint distribution of $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$, and in particular its covariance matrix (assuming that both Gaussian Processes are generated under $GP2$):

$$K = \begin{bmatrix} K_{11|2} & K_{12|2} \\ K_{21|2} & K_{22|2} \end{bmatrix} \tag{9}$$

where $K_{11|2}$ is the $n_1 \times n_1$ covariance matrix between the points in D_1 evaluated under the co-variance for GP_2, $K_{12|2} = K_{21|2}'$ the $n_1 \times n_2$ (cross) covariance matrix between the points in D_1 and D_2 evaluated using $\boldsymbol{\theta_2}$ and $K_{22|2}$ is the usual $n_2 \times n_2$ covariance for points in D_2. Thus we can see that S_{22} is the $n_2 \times n_2$ modified covariance for the points in D_2 given the points along D_1, while the $\boldsymbol{Z_2^*}$ is the corrected mean that accounts for the values at the points in D_1, which have non-zero mean.

We remove the dependency on the values $\boldsymbol{Z_1}$ by evaluating the integral in (Equation 7). $p(\boldsymbol{Z_1}|\boldsymbol{\theta_1},\boldsymbol{m_1})$ is given by:

$$p(\boldsymbol{Z_1}|\boldsymbol{\theta_1},\boldsymbol{m_1}) = \frac{1}{(2\pi)^{\frac{n_1}{2}}|K_{11|1}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left(\boldsymbol{Z_1}-\boldsymbol{m_1}\right)' K_{11|1}^{-1}\left(\boldsymbol{Z_1}-\boldsymbol{m_1}\right)\right) \tag{10}$$

where $K_{11|1}$ is the $n_1 \times n_1$ covariance matrix between the points in D_1 evaluated under the covariance given by $\boldsymbol{\theta_1}$. Completing the square in $\boldsymbol{Z_1}$ in the exponent, the integral (Equation 7) can be evaluated to give:

$$p(\boldsymbol{Z_2}|\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{m_1}) = \frac{1}{(2\pi)^{\frac{n_2}{2}}}\frac{1}{|S_{22}|^{\frac{1}{2}}}\frac{1}{|K_{11|1}|^{\frac{1}{2}}}\frac{1}{|B|^{\frac{1}{2}}} \times \tag{11}$$

$$\exp\left(\frac{1}{2}\left(C'B^{-1}C - \boldsymbol{Z_2}'S_{22}^{-1}\boldsymbol{Z_2} - \boldsymbol{m_1}'K_{11|1}^{-1}\boldsymbol{m_1}\right)\right)$$

where:

$$\begin{aligned} B &= (K_{12|2}'K_{11|2}^{-1})'S_{22}^{-1}K_{12|2}'K_{11|2}^{-1} + K_{11|1}^{-1} \\ C' &= \boldsymbol{Z_2}'S_{22}^{-1}K_{12|2}'K_{11|2}^{-1} + \boldsymbol{m_1}'K_{11|1}^{-1} \end{aligned}$$

The algorithm has been coded in MATLAB and can deal with reasonably large numbers of points quickly. With $n_1 = 12$ and $n_2 = 200$ for a two dimensional vector-valued Gaussian Process based on a covariance function given by (Equation 3), computation of the log likelihood takes 4.13 seconds on a SGI Indy R5000.

The mean value just ahead and behind the front (Equation 6) define the mean values for the constrained discontinuity (i.e. $\boldsymbol{m_1}$ in Equation 11). Conditional on the frontal parameters the wind fields either side are assumed independent (Figure 7):

$$p(\boldsymbol{Z_{2a}},\boldsymbol{Z_{2b}}|\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{\theta_f}) = p(\boldsymbol{Z_{2a}}|\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{\theta_f}) \times p(\boldsymbol{Z_{2b}}|\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{\theta_f}) \tag{12}$$

where we have performed the integration (Equation 7) to remove the dependency on $\boldsymbol{Z_{1a}}$ and $\boldsymbol{Z_{1b}}$. Thus the likelihood of the data $\boldsymbol{Z_2} = (\boldsymbol{Z_{2a}},\boldsymbol{Z_{2b}})$ given the model parameters $\boldsymbol{\theta_2},\boldsymbol{\theta_1},\boldsymbol{\theta_f}$ is simply the product of the likelihoods of two Gaussian Processes with a constrained discontinuity which can be computed using (Equation 11). Since the relation between $\boldsymbol{\theta_f}$ and m_{1a} or m_{1b} is deterministic we can use either without ambiguity.
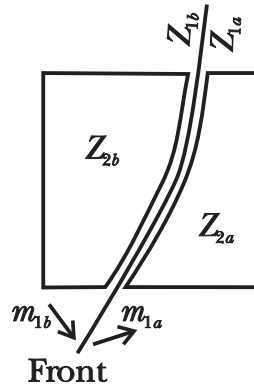
Figure 7: The simplified model for a frontal discontinuity. The combination of two Gaussian
Processes with constrained discontinuities produces the full wind field model.

4   Testing the Model

To ensure the model we have proposed is useful in practical applications we have tested the model
at several stages. Early work (Cornford *et al.*, 1999) tested a simplified model, which allowed
only symmetric, straight fronts on simulated data. Markov Chain Monte Carlo sampling was used
to infer the posterior distribution of the frontal parameters $p(\boldsymbol{\theta}_f|\boldsymbol{Z}^s)$ conditional on some data
$\boldsymbol{Z}^s = (\boldsymbol{Z}^s_{2a}, \boldsymbol{Z}^s_{2b})$ simulated from the simplified model. It was shown that when the sampling
procedure was initialised 'reasonably' close to the values used in the simulation the Markov Chains
converged to the expected values. In this paper we perform a similar test on the more general
model, where the initialisation is made without any knowledge of the generating parameters. We
then show results on real data.
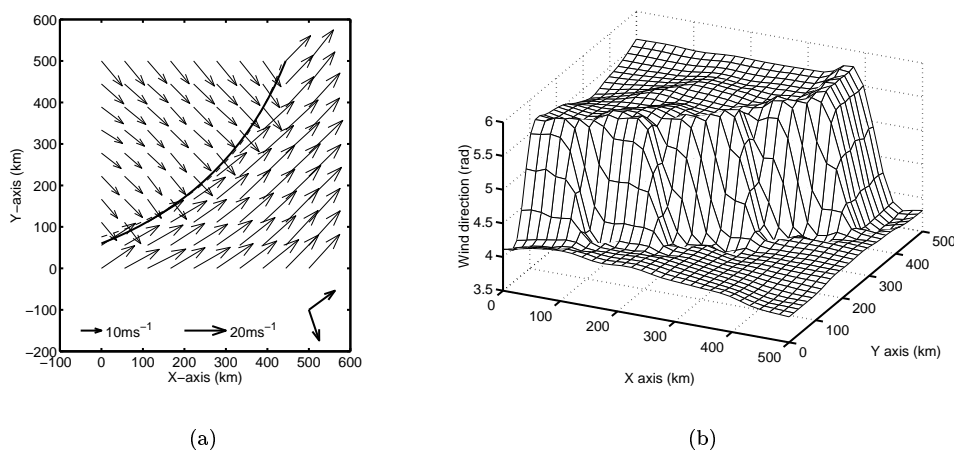
4.1   Simulated data

(a)   (b)

Figure 8: (a) An example simulated frontal wind field. The wind field generated by the model
is show together with the 'true' position of the front (dashed line). Also shown is the
modal frontal position after sampling (thick solid line) which plots almost over the
true frontal position. The mean wind vectors just ahead of and behind the front are
also shown in the lower right of the plot. (b) The constrained discontinuity in wind
direction from our model, for the same simulation.

The model outlined above was tested on simulated data generated from the model. We generated a wind field Z^s using known model parameters (e.g. Figure 8). We then investigated the sampling of model parameters from the posterior distribution:

$$p(\boldsymbol{\theta_2}, \boldsymbol{\theta_1}, \boldsymbol{\theta_f} | \boldsymbol{Z^s}) \propto p(\boldsymbol{Z^s} | \boldsymbol{\theta_2}, \boldsymbol{\theta_1}, \boldsymbol{\theta_f}) p(\boldsymbol{\theta_2}) p(\boldsymbol{\theta_1}) p(\boldsymbol{\theta_f}) \qquad (13)$$

where $p(\boldsymbol{\theta_2}), p(\boldsymbol{\theta_1}), p(\boldsymbol{\theta_f})$ are prior distributions over the parameters in the Gaussian Processes and frontal models. This brings out one advantage of the proposed model. All the model parameters have a physical interpretation and thus expert knowledge was used to set priors which produce realistic wind fields. We could have used (Equation 13) to help set (hyper)priors using real data.

The prior distributions for all parameters of both Gaussian Processes were Weibull with peaks at the maximum *a posteriori* probability values determined from real data and a 'variance' that constrains the solutions to be physically realistic. In practice, due to the conditioning of the Gaussian Processes away from the front on the value of the means along the front, the frontal parameters $\boldsymbol{\theta_f}$ are the most important ones to retrieve. The prior distributions for the frontal parameters are:

- (x_f, y_f) : uniform in the range of the observed data

- d_f : Exponential, rate = 0.02

- ϕ_f : Circular Normal (on $[0, 2\pi)$), mean = $\pi/6$, 'variance' = $\pi/3$

- α_f : Circular Normal (on $[0, \pi)$), mean = $\pi/2$, 'variance' = $\pi/8$

- s_f : Weibull, maximum at 12 ms^{-1}, wide spread

- ρ_f : Uniform on $[0.5, 0.9]$

These priors are essentially used to constrain the possible solutions to be physically realistic, particularly the prior over α_f. During testing they were not found to have a significant effect on the parameters retrieved (other than the constraints we intended them to encode).

Markov Chain Monte Carlo methods using the Metropolis algorithm (Neal, 1993) were used to sample from (Equation 13) using the NETLAB[3] library. Convergence of the Markov chain was assessed using visual inspection of the univariate sample paths since the generating parameters were known, although other diagnostics could be used (Cowles and Carlin, 1996). A good initialisation is very important if the chain is to converge on a sensible time-scale.

The initialisation of the frontal model is performed using 'ad-hoc' methods. We compute the divergence and vorticity (curl) of the wind field using finite difference approximations (Haltiner and Williams, 1980) and use the maxima in the sum of the absolute values of the vorticity and divergence to locate the line of the front. The values of the wind vectors in the regions ahead of and behind this line are then used to set the other parameters in the model.

The Metropolis algorithm was run for 50,000 iterations, with the Gaussian proposal variances, chosen by the user after several trial runs, being different for different parameters. This takes roughly 4 hours on a SGI Power Challenge R10000. We allowed a 10,000 iteration burn-in. Figure 9 shows ten samples of the frontal position and mean winds taken from the simulation based on the wind field in Figure 8. It is clear that for the simulated data it is possible to infer sensible posterior distributions of the model parameters. Real data, not generated under the model, presents a more realistic test.

4.2   Real data

In this section we use real data, such as that shown in Figure 4, which is extracted from the ECMWF numerical weather prediction model and interpolated to the scatterometer observation

---

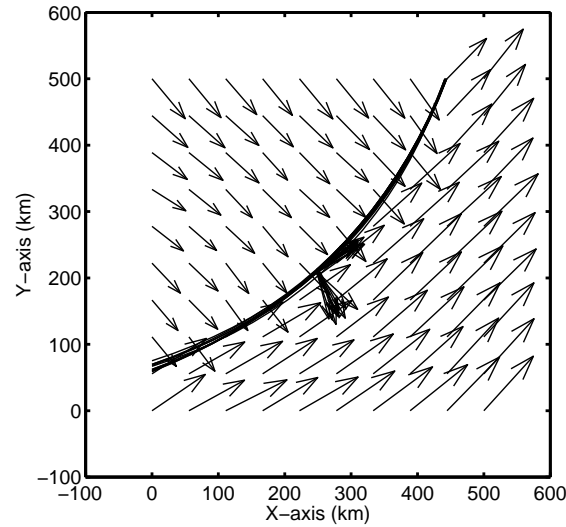[3]Available from http://www.ncrg.aston.ac.uk/netlab/index.html.

Figure 9: A sample from the posterior distribution of the Markov chain of the frontal parameters after allowing a 10,000 step burn-in. The position of the front and the mean winds just ahead of and behind the front are plotted for 10 equally spaced samples from the remainder of the 50,000 sample chain.

locations. The procedure used is the same as in the simulation testing, but this time we do not know the 'true' location of the fronts. Thus the validation of the model is performed visually and by comparing the likelihoods of the data under the frontal model and the smooth wind field model.

Figure 4 also shows the results of running the algorithm on the two real examples. The fronts can be seen to be in the expected positions.
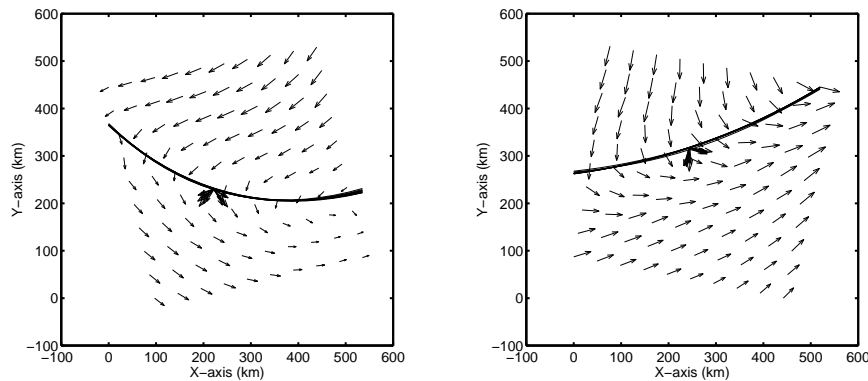

Figure 10: Samples from the posterior distribution of the fronts and winds at the front for the wind fields shown in Figure 4. The format is similar to Figure 9.

Samples from the posterior distribution of the parameters of the frontal model can be seen in Figure 10. It can be seen that the distribution of the fronts is quite closely constrained by the data. The winds either side of the fronts can be seen to take sensible values. Thus there is some hope that this algorithm might be applicable to the problem we really want to solve, that of the extraction of wind vectors from satellite scatterometer data.

While the above exploratory analysis confirms that the model parameters are identifiable and agree with those set by expert judgement a more objective measure of model performance is needed. We propose to do this by comparing the log likelihoods of several real wind fields under the frontal model (Equation 12) and the smooth model (Equation 4). In order to select frontal wind fields, 575 scenes (that is 500×500 km regions observed by the ERS-2 satellite) for the Northern Hemisphere

in the month of February 1998 were examined and 45 scenes which probably included fronts were identified. Both the smooth and frontal wind field models were applied automatically to the data and the likelihoods of the wind fields compared. To provide further comparison another 45 scenes were selected randomly from the remaining 530 scenes. These are probably smooth wind fields and thus the likelihood of the data under the smooth wind field model should be greater.

Table 1: Comparison of the likelihoods of the wind fields which were assessed to be frontal and non-frontal under both the frontal and smooth wind field models.

Wind field type	Mean log likelihood under the frontal model	Mean log likelihood under the smooth model
Frontal	-357.3	-582.6
Non-frontal	-71.7	-74.2

Table 1 shows that there is a significant difference in log likelihood of the frontal versus non-frontal wind fields under both frontal and smooth models. It is also clear that the mean difference between the log likelihoods under the frontal model compared with the smooth model is much greater for frontal wind fields.

5   Discussion

The problem of the retrieval of wind fields from scatterometer observations is a classic inverse problem, in that we require a 'regularising' spatial prior model for wind fields. In general the smooth Gaussian Process model, discussed in Section 2.1, is sufficient to represent the expected variability in observed wind fields. However, when there are fronts present in the wind field it will be important to incorporate the possibility of fronts in our wind field model if we are to retrieve realistic wind fields.

However the ambiguous nature of the winds retrieved from the scatterometer observations means that we must put strong constraints on the form a front can take. Desroziers and Lafore (1993) have suggested a coordinate transform approach to the mapping of meteorological fields in the vicinity of fronts, somewhat like the ideas of (Sampson and Guttorp, 1992), but this approach has limited value in our application due to the ambiguous winds retrieved from the scatterometer. A front is essentially a curve along which there is a discontinuity in wind direction, but near continuity in wind speed. It is important that only cyclonic vorticity is observed at fronts and this constraint forms part of our model through setting Bayesian priors over the parameters of the frontal model. This minimises the risk of identifying fronts that are purely the result of ambiguity in the retrieval of winds from scatterometer observations (Stoffelen and Anderson, 1997).

Testing the model on simulated data allows the assessment of the parameter identifiability and sensitivity within the model. The results show that given a good initialisation the simple Metropolis sampling performs adequately on the problem. The 'correct' parameters, used in simulating the data, are retrieved for all runs where the initialisation works (it failed in 2 runs out of the 20 attempted in the simulated data and 3 runs out of 45 in the real data). The initialisation fails when the front crosses only a very small portion of the region or falls very close to the edge of the region. There is considerable room for improving the initialisation routine, for instance the use of robust methods for 2D curve fitting might improve reliability. However, for the majority of wind fields examined, the initialisation we have used was sufficient to obtain a sensible frontal wind field model.

There is some correlation between the model parameters, but this is not severe, as evidenced by the fact that each is well identified and close to the 'true' value when sampling from simulated datasets.

The model works well with the real data; however it is rather difficult to quantify how well. It

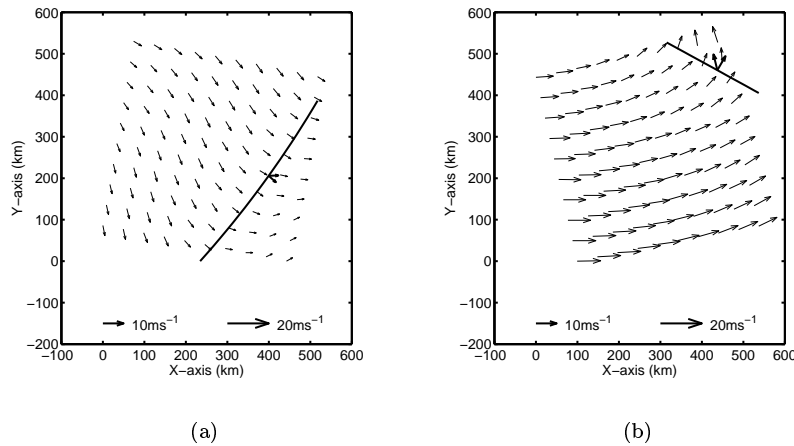(a)   (b)

Figure 11: The modal position for the fronts and winds at the front for the wind fields shown in Figure 4. (a) Shows a wind field which was thought to contain a front but for which the smooth model has a similar log likelihood (log likelihood; smooth $= -77.4$, frontal $= -68.6$). (b) A wind field in which there was thought to be no front, but the log likelihood suggests there may be (log likelihood; smooth $= -123.2$, frontal $= -90.9$).

would have been useful to have observations of surface temperature and humidity since these variables can also be used, in a more straightforward manner (Hewson, 1998), to locate fronts. Using meteorological expertise it can be stated that subjectively the frontal model puts the fronts in the right place with the correct parameter values.

Table 1 shows that the model can distinguish between frontal and non-frontal wind fields. For frontal wind fields the mean log likelihood under the frontal model is much larger than the log likelihood under the smooth model. For non-frontal wind fields the log likelihoods of both models are very similar.

By comparing the log likelihoods of wind fields which were believed a priori to contain fronts and those which were believed not to contain fronts, under both the frontal and smooth wind field models, we have shown that the frontal model has a larger log likelihood (is more probable) if a front is present (see Table 1). The log likelihood of the frontal wind fields (i.e. the first row) is smaller for both models because frontal wind fields exhibit greater variability than smooth wind fields even away from the front, and both models prefer smooth wind fields (the frontal model also containing just one 'clean' front). The reason that the log likelihood of the frontal model is smaller for frontal wind fields than smooth wind fields (comparing the left column) is thought to be the result of the smoothing inherent in the ECMWF model wind fields, which has a large effect on the wind field near the front and thus does not fit the frontal wind field model proposed. When used as part of an algorithm based on the scatterometer observations, which contain much more spatial detail, the log likelihoods (under the frontal model) should be similar for frontal and non-frontal wind fields.

The reason that the mean log likelihood of the frontal model is greater than that of the smooth model is thought to be related to the mean used in the frontal model. This extra parameter gives a more flexible model, which given a fixed data set is likely to have a larger likelihood. We also observe that the change in wind direction across the fronts, $\pi - \alpha_f$, is very small when the frontal model is use with a non-frontal wind field.

Figure 11 shows two example wind fields with the modal posterior positions of the fronts from (a) the frontal and (b) the non-frontal scenes. These examples show that it may be possible to use the relative log likelihoods to determine whether a frontal model is plausible for a given wind field. In Figure 11(a) we find that the log likelihoods under both the smooth and frontal models are

similar, despite the wind field being subjectively classified as frontal. Further inspection reveals very weak evidence for a front within the wind field, supporting the conclusion that there really is no front present. Figure 11(b) shows an example where the likelihoods are quite different under the smooth and frontal model suggesting the presence of a front. Close examination shows that a front appears to be present in the upper right hand corner.

One outstanding problem with the frontal model is that two matrices of size $n \times n$ at best, where n is the number of observations, require inversion during the computation of the likelihood. This means that the determination of the likelihood is rather slow (see Section 3). In practice we would like to run the retrieval procedure in real time, thus sampling may not be appropriate in the operational framework. Since the main use for the retrieved wind fields will be as initial conditions for a numerical forecast of the future state of the atmosphere, it is most important to retrieve the modal wind field. Thus we could use some form of optimisation algorithm (such as simulated annealing) to determine the mode with greater speed. Advances in data assimilation in numerical weather prediction may make sampling more appropriate at some time in the future.

6   Conclusions

The frontal model we have developed, based on the concept of constrained discontinuities in vector Gaussian Processes, gives a parsimonious yet realistic model for frontal wind fields. When run in the generative mode realistic wind fields are produced, which strongly resemble real frontal wind fields. Care must be taken when comparing these with wind fields obtained from numerical weather prediction models since these are discrete and over smooth representations of the atmosphere. However their use with scatterometer data, which is believed to contain finer spatial detail, could be used to enhance our understand of the near surface wind field at fronts. The frontal model should also enable the more accurate retrieval of wind fields in the presence of fronts.

The concept of a constrained discontinuity in Gaussian Processes may also have other uses, such as the automatic recognition of fronts (Hewson, 1998) or motion detection in artificial vision. The key point is the ability to constrain the discontinuity so that only realistic features are part of the prior model. This will minimise the identification of 'incorrect' fronts resulting from ambiguity in the retrieval of winds from local scatterometer observations.

Future work might go on to consider the application of a reversible jump (Green, 1995) step in the MCMC sampling procedure to allow inference on the probability of a front being present although it seems unlikely that mixing will be improved and likely that convergence may be delayed. Thus we believe at present that we can compare the log likelihoods of the smooth and frontal wind field models and choose the smooth model if the log likelihood of the smooth model is sufficiently large or greater than $\sim 90\%$ of the log likelihood of the frontal model. Ongoing work is attempting to resolve the issue of front identification.

The results so far indicate that the frontal model proposed will form a useful component in the automatic wind retrieval system we envisage for the scatterometer observations.

Acknowledgements

References

Abrahamsen, P. 1997. A Review of Gaussian Random Fields and Correlation Functions, Second Edition. Technical Report 917, Norwegian Computing Center. URL: http://www.nr.no/research/sand/articles.html.

Adler, R. J. 1981. *The Geometry of Random Fields*. Chichester: John Wiley and Sons.

Cornford, D. 1998. Flexible Gaussian Process Wind Field Models. Technical Report NCRG/98/017, Neural Computing Research Group, Aston University, Aston Triangle, Birmingham, UK. URL: http://www.ncrg.aston.ac.uk/~cornfosd/.

Cornford, D., I. T. Nabney, and C. K. I. Williams 1999. Adding Constrained Discontinuities to Gaussian Process Models of Wind Fields. In M. S. Kearns, S. A. Solla, and C. A. Cohn (Eds.), *Advances in Neural Information Processing Systems 11*. Cambridge, Massachusetts: MIT Press.

Cowles, M. K. and B. P. Carlin 1996. Markov-Chain Monte-Carlo Convergence Diagnostics—A Comparative Review. *Journal of the American Statistical Association* **91**, 883–904.

Cressie, N. A. C. 1993. *Statistics for Spatial Data*. New York: John Wiley and Sons.

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

Desroziers, G. and G.-F. Lafore 1993. A Coordinate Transform for Objective Frontal Analysis. *Monthly Weather Review* **121**, 1531–1553.

Green, P. J. 1995. Reversible Jump MCMC Computation and Bayesian Model Determination. *Biometrica* **82**, 711–732.

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

Hewson, T. D. 1998. Objective Fronts. *Meteorological Applications* **5**, 37–65.

Long, D. G. 1993. Wind Field Model-Based Estimation of Seasat Scatterometer Winds. *Journal of Geophysical Research* **98**, 14651–14668.

Nabney, I. T., D. Cornford, and C. K. I. Williams 1998. Bayesian Inference for Wind Field Retrieval. *Neurocomputing Letters*. submitted.

Neal, R. M. 1993. Probabilistic Inference Using Markov Chain Monte Carlo Methods. Technical Report CRG-TR-93-1, Department of Computer Science, University of Toronto. URL: http://www.cs.utoronto.ca/~radford.

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

Sampson, P. D. and P. Guttorp 1992. Nonparameteric Estimation of Nonstationary Spatial Covariance Structure. *Journal of the American Statistical Association* **87**, 108–119.

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

Tarantola, A. 1987. *Inverse Problem Theory*. London: Elsevier.