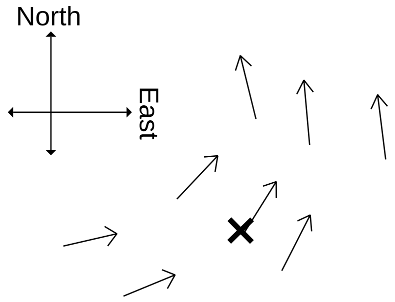
MSc Exam Question (2003) - 3d-Var

A quasi-stationary weather system has a structure shown in the figure. An estimate is required of the two horizontal wind components at the position of the cross (×) by combining background and observational information in a variational data assimilation system that does not consider the evolution of the flow (3d-Var.-like).



The background estimate at × is $\vec{x}_B = (u_x, u_y)^T$ which has error covariance matrix **B**. There are two observations at ×. One is a measurement of the strength of the wind ($|\vec{u}|$) and the other is of the direction that the wind is blowing from (θ, measured clockwise from north). These are assembled into the vector $\vec{y}$,

$$\vec{y} = \begin{pmatrix} |\vec{u}| \\ \theta \end{pmatrix}$$

(a) Given $\vec{x}_B$, the assimilation system makes a prediction of the observation vector ($\vec{h}[\vec{x}_B]$) which has the same structure as $\vec{y}$. Write down the forward model $\vec{h}[\vec{x}_B]$. **[6 marks]**

(b) The observations are made a short time apart. What is the main assumption that makes your 3d-Var.-like forward model in (a) valid? **[1 mark]**

(c) Write down the innovation vector. In particular, why should you be careful when differencing angles? **[4 marks]**

(d) Let **H** be the Jacobian matrix, evaluated at the background.

(i) How many rows and columns will **H** have? **[2 marks]**

(ii) By linearizing your answer to (a), write down **H**.
You may need the derivative $\mathrm{d}/\mathrm{d}x\ (\tan^{-1} x) = 1/(1 + x^2)$. **[5 marks]**

(e) The 3d-Var. cost function, J, has the following gradient with respect to control variables,

$$\nabla_x J[\vec{x}] = \mathbf{B}^{-1}(\vec{x} - \vec{x}_B) - \mathbf{H}^T\mathbf{R}^{-1}(y - \vec{h}[\vec{x}]).$$

where **R** is the observational error covariance matrix.

(i) What is the usual choice for the first guess of $\vec{x}$? **[1 mark]**

(ii) In the context of the 3d-Var. algorithm, explain why $\nabla_x J[\vec{x}]$ is needed. **[2 marks]**

(iii) When J has converged, what does $\vec{x}$ represent? **[1 mark]**

(f) Why is a variational scheme better suited to this problem, than the BLUE (best linear unbiased estimator) scheme? **[2 marks]**

(g) The data assimilation scheme is repeated, with additional observations, and iterated until it satisfies a convergence criterion. If there are a total of N observations, whose errors have been specified correctly, how would you use the value of the cost function at the final iteration to show that the analysis is near optimal? **[1 mark]**