Data-model-error compatibility

W. C. Thacker

Atlantic Oceanographic and Meteorological Laboratory
4301 Rickenbacker Causeway, Miami FL 33149 USA
carlisle.thackernoaa.gov

Abstract

During data assimilation, differences between observations and their model counterparts should be consistent with the error statistics that govern how the model is to be corrected. The concept of incompatibility distance between observations and their model counterparts is introduced as a way of detecting inconsistencies, and formulae are presented for estimating the probability of encountering greater incompatibility. Observations can be examined one-by-one to insure that their confidence intervals are not widely separated from those of the model counterparts. They can be further examined in pairs to detect whether contrasts across fronts are consistent with assumptions about error correlations.

Key words: data assimilation, estimation theory, error statistics

1 Introduction.

Modern methods for assimilating data into numerical models are based on statistical estimation theory. The statistical foundations are the same for optimal interpolation (Gandin, 1963; Bretherton et al., 1976; Daley, 1991; Behringer et al., 1998; Carton et al., 2000) as for Kalman filtering (Kalman, 1960; Gelb, 1974; Evensen, 1994; Cohn, 1997): data and model are both assumed to be unbiased, and error variances and correlations dictate how differences between observations and their model counterparts are transformed into corrections to the model state. In practice, the model can have systematic errors [1], the data might characterize unmodelled processes [2], and the error covariances are often parameterized as a simple function of distance. Consequently, model, data and

[1] For example, the model might tend to position the Gulf Stream front to far to the north.

[2] Data extracted from the archives might include measurements of slope water near the continental coast, which is not represented by the model.

errors might not be compatible. This is especially likely when the position of a front as indicated by the observations differs from that in the model. The purpose of this paper is to discuss how to screen for such incompatibilities. In some cases, adjusting the error statistics might reconcile inconsistencies between data and their model counterparts. In other cases, it might be advisable to discard some of the observations and assimilate the remainder.

Section 2 reviews the formalism for statistical data assimilation. Then section 3 addresses the the general case of compatibility of groups of data with their model counterparts. Section 4 examines model-data-error compatibility for the case of individual observations. The idea is that confidence intervals for the observation and for its model counterpart should not be widely separated. Then section 5 discusses pairwise compatibility. The confidence interval for a single observation becomes a confidence ellipse for a pair; error ellipse of the data should not be too far from that of their model counterparts. Section 6 summarizes with some concluding remarks.

2 Formalism for data assimilation.

From the perspective of optimal estimation theory, both the model state and the observations are estimates subject to uncertainty. Errors in the model state reflect uncertainties in the initial and boundary conditions and the approximate nature of the model dynamics. Those in the data reflect instrumental accuracies and, more important, representativity error, *i.e.*, the variety of possible measurements consistent with the model's resolution. The objective is to combine the two estimates to get a more accurate idea of the model state. The improved estimate should be an average with the more trustworthy contribution receiving the greatest weight. Furthermore, unobserved aspects of the model state should be corrected to the extent their errors are correlated with those of the observed aspects. Thus, knowledge of the magnitude and correlation of errors for both the model state and the observations is essential.

The formalism assumes that both the uncorrected model state and the observations are unbiased estimates, *i.e.*, mean errors are zero. While this is not necessarily true in practice, it is incorporated into the prescription for combining the data with the model state. Furthermore, the formalism bases the corrections entirely on the variances and covariances of the errors, *i.e.*, errors are implicitly assumed to be distributed multi-normally. This too is not necessarily true in practice. Nevertheless, the formalism demands estimates of appropriate error variances and covariances, which are generally difficult to provide. Relatively simple parametric expressions are generally used for optimal interpolation, while Kalman filtering attempts to do better by propagating the errors from previous analyses along with the model state. The focus

here is on the fact that the error estimates can easily be wrong and the need for a compatibility condition to help to identify when this might be the case.

The model state before assimilation is often referred to as the background state, and assimilation is effected by adding corrections to this background. The corrections, represented collectively by the vector $\mathbf{\Delta}$, are linear combinations of the innovations, *i.e.*, differences between observations and their model counterparts, represented collectively by the vector $\boldsymbol{\delta}$:

$$\mathbf{\Delta} = \mathsf{B}\mathsf{H}^{\mathrm{T}}(\mathsf{H}\mathsf{B}\mathsf{H}^{\mathrm{T}} + \mathsf{D})^{-1}\boldsymbol{\delta} \tag{1}$$

with coefficients determined by the error covariances matrix for the background state B and that for the data D. The matrix H maps the model state into counterparts of the observations; $\mathsf{H}\mathsf{B}\mathsf{H}^{\mathrm{T}}$ is the error covariance matrix for the model counterparts of the data, and $\mathsf{B}\mathsf{H}^{\mathrm{T}}$ is the error covariance between all state variables and the observed aspects of the model state. Thus the corrections clearly and directly reflect the characterization of errors via the covariance matrices.

A well-known worry is that the matrix $\mathsf{H}\mathsf{B}\mathsf{H}^{\mathrm{T}} + \mathsf{D}$ might be ill-conditioned. If it is, then small differences in the innovations can result in large differences in the corrections, rendering the assimilation results sensitive to error in the data. This can easily be the case when data errors are neglected, because correlations between background errors make B ill-conditioned. The presence of the matrix D should eliminate this worry [3]. But if observational errors are considerably smaller than those of their model counterparts [4], $(\mathsf{H}\mathsf{B}\mathsf{H}^{\mathrm{T}} + \mathsf{D})^{-1}$ can still serve to amplify short-range differences in innovations, such as might be encountered when the model has incorrectly positioned a thermal front, and can lead to unrealistic extrema in the corrected model state. It is not surprising that this matrix should be central to the analysis of model-data-error compatibility. It is intrinsic to the definition below of an index of collective compatibility of the observations and their model counterparts. Sub-matrices appear when examining model-data-error compatibility for individual observations or pairs of observations.

3 Collective compatibility.

Incompatibility can be gauged by the distance between the vector $\boldsymbol{d}$ of n observations and that of their model counterparts $\mathsf{H}\boldsymbol{b}$. As these vectors are

[3] Data errors are often considered to be uncorrelated, so adding the diagonal matrix D makes B better conditioned.

[4] $\mathsf{H}\mathsf{B}\mathsf{H}^{\mathrm{T}} + \mathsf{D}$ need not be so ill-conditioned that round-off errors impede the computation of its inverse.

subject to uncertainty, this should not not be the simple Cartesian distance; instead, it should reflect the errors of the data and the model state. Because observations and background are assumed to be unbiased estimates, the mean errors of $\boldsymbol{d}$ and $\mathsf{H}\boldsymbol{b}$ should be zero, and their error-covariance matrices are specified to be D and $\mathsf{HBH}^{\mathrm{T}}$, respectively. If these error characterizations are correct, the innovation vector $\boldsymbol{\delta} = \boldsymbol{d} - \mathsf{H}\boldsymbol{b}$ should have zero mean and error-covariance matrix $\mathsf{D} + \mathsf{HBH}^{\mathrm{T}}$. Thus,

$$s_n = ||\boldsymbol{d} - \mathsf{H}\boldsymbol{b}|| = \sqrt{(\boldsymbol{d} - \mathsf{H}\boldsymbol{b})^{\mathrm{T}}(\mathsf{D} + \mathsf{HBH}^{\mathrm{T}})^{-1}(\boldsymbol{d} - \mathsf{H}\boldsymbol{b})}\ , \tag{2}$$

the Mahalanobis length (Mardia et al., 1979) of the innovation vector $\boldsymbol{\delta}$ provides the natural definition of an *incompatibility distance*[5].

If s_n is too large, then model, data, and errors are incompatible. But how large is too large? The probability[6]

$$p_n = \int_{s_n^2}^{\infty} \frac{x^{(n-2)/2}\exp^{-x/2}}{2^{n/2}\Gamma\left(n/2\right)} dx \tag{3}$$

of encountering a less likely set of innovations can be determined from the chi-square distribution function with n degrees of freedom (Stuart and Ord, 1987). The values of s_n for which $p_n = 10\%$ (solid curve) and $p_n = 5\%$ (dotted curve) are shown in figure 1 as a function of n; for large n, these cutoffs increase like $\sqrt{n}$.

While s_n and p_n may be useful for verifying the collective compatibility of the data with their model counterparts, given the error covariances, if incompatibility is indicated, the cause might remain obscure. Examining the data in smaller groups can help to isolate the problems. It is particularly useful to examine the data one-by-one to determine the individual compatibility of each observation with its model counterpart. After all data have been found individually to be compatible with their model counterparts and assumed errors, they can be examined pairwise to determine whether error correlations present a problem.

[5] The subscript indicates that s_n is for n observations. Below, s_1 and s_2 will respectively denote the incompatibility distance for individual observations and pairs of observations.

[6] While p_n could be used for testing the null hypothesis (Mood et al., 1974) that model and data agree, following Cleveland (1993) the philosophy here is more exploratory than formal.

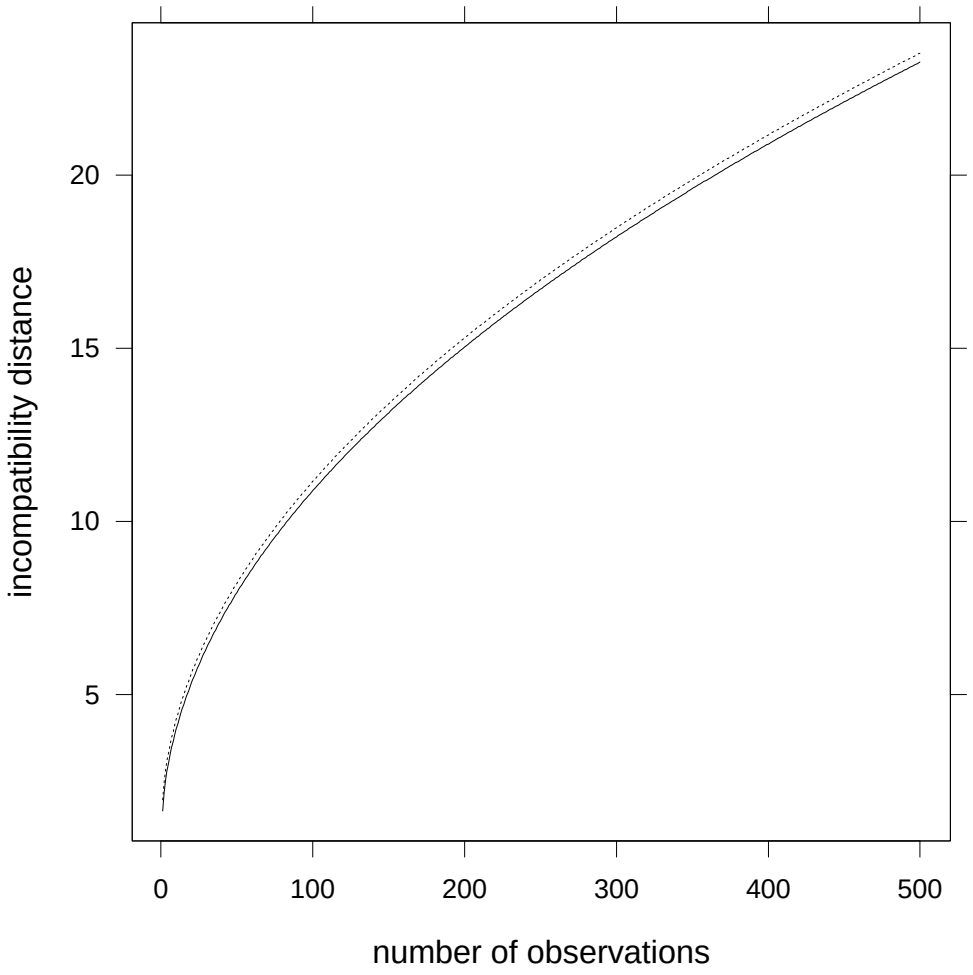

Fig. 1. Ninetieth (solid curve) and ninety-fifth (dotted curve) percentile values for s_n as a function of n.

4 Compatibility for individual observations.

Model-data-error compatibility can be most easily illustrated when there is only one observation. Suppose it has value d and its errors are characterized by standard deviation σ_d; its counterpart in the background state has the (scalar) value $b = \mathsf{H}\boldsymbol{b}$ with σ_b describing its error[7]. If the magnitude of the innovation $|\delta| = |d - b|$ is much larger than $\sigma_d + \sigma_b$, then the error estimates are not compatible with the simulated and observed values. The model might be biased, causing b to differ greatly from what is observed. Or the observations might wrong, *e.g.*, latitude and longitude might have been switched and the data used at the wrong location. Or the error estimates might be wrong; data, model, or both might be less accurate than previously thought. Nevertheless,

[7] The variances σ_d^2 and σ_b^2 are elements of the error-covariance matrices D and $\mathsf{HBH}^{\mathrm{T}}$ for all of the data and their model counterparts, respectively.

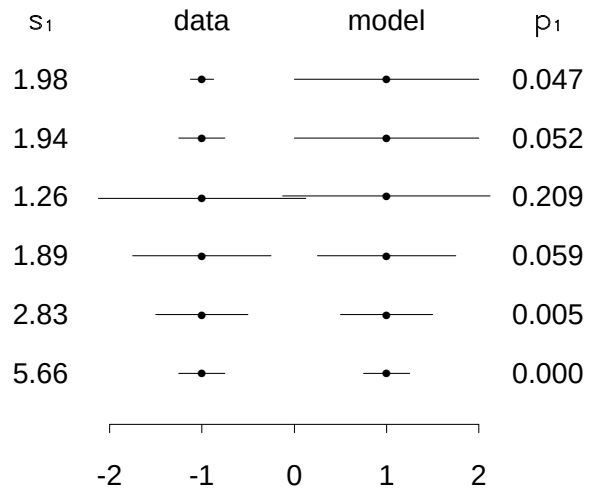

Fig. 2. Dots indicate values of data and model estimates, to be read from scale at bottom; horizontal lines indicate standard-deviation ranges about data. Numbers on the left side indicate incompatibility distances for the indicated errors; those on the right, probabilities of finding greater model-data differences.

the machinery of data assimilation will produce a correction. Equation (1) updates b to the weighted average $(d/\sigma_d^2 + b/\sigma_b^2)/(1/\sigma_d^2 + 1/\sigma_b^2)$. Moreover, the variance of the error associated with this updated estimate is $1/(1/\sigma_d^2 + 1/\sigma_b^2)$, which is less than that of either contribution, in spite of the apparent incompatibility.

If model and data are compatible, the square of the innovation should be small compared to its error variance. The incompatibility distance s_1 between a single observed value $d \pm \sigma_d$ and its model counterpart $b \pm \sigma_b$ is:

$$s_1 = ||d - b|| = \sqrt{\frac{(d-b)^2}{\sigma_d^2 + \sigma_b^2}} \; . \tag{4}$$

If s_1 is large, then model, data, and errors are incompatible. Assuming the background and data errors are normally distributed with zero means and variances σ_b^2 and σ_d^2, then δ should be normally distributed with zero mean and variance σ_δ^2, and s_1 with zero mean and unit variance. Thus, in terms of the cumulative normal distribution function[8] Φ (Mood et al., 1974),

$$p_1 = 2(1 - \Phi(s_1)) \tag{5}$$

is the probability p_1 of encountering a model-data pair with incompatibility distance larger than s_1.

Examples of the incompatibility distance for a variety of cases are show in figure 2. For all cases the data estimate is $d = -1$ and the background estimate

[8] The same value for the p_1 is given by (3) with $n = 1$.

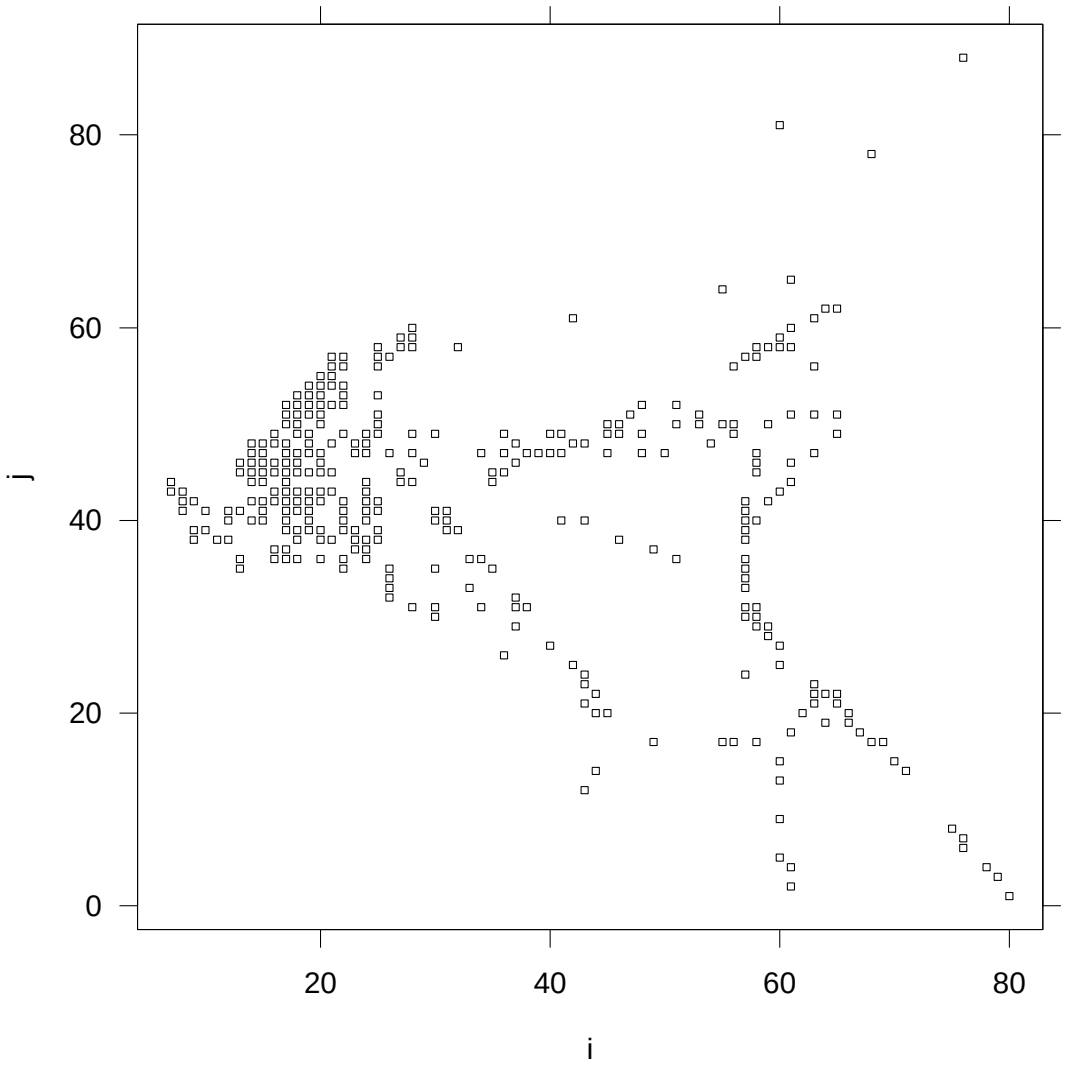

Fig. 3. The 318 cells of the model's 81×93 Mercator grid for which there are XBT observations for January 1972. The domain extends from 97.7°W to 16.2°E with 1.4° zonal resolution and from 30.2°S to 69.9°N with resolution decreasing poleward so that grid cells are approximately square everywhere. Horizontal and vertical axes indicate zonal and meridional grid indices, respectively.

is $b = +1$. The horizontal lines indicate the intervals $d \pm \sigma_d$ and $b \pm \sigma_b$, which vary from case to case. The corresponding compatibility distances s_1 and probabilities p_1 are displayed for each case. As the accuracies of the estimates decrease, the validity intervals approach each other, so compatibility increases; incompatibility distance decreases, and it is less unlikely to encounter a larger innovation. Conversely, as the error variances decrease, the gap between the validity intervals widens and the compatibility is less; the distance s_1 between model and data is greater, and the chance of worse disagreement less likely. The case with overlapping uncertainty intervals certainly agrees with the intuitive notion of compatibility, while the bottom-most case with widely separated intervals is clearly an example of model-data-error incompatibility.

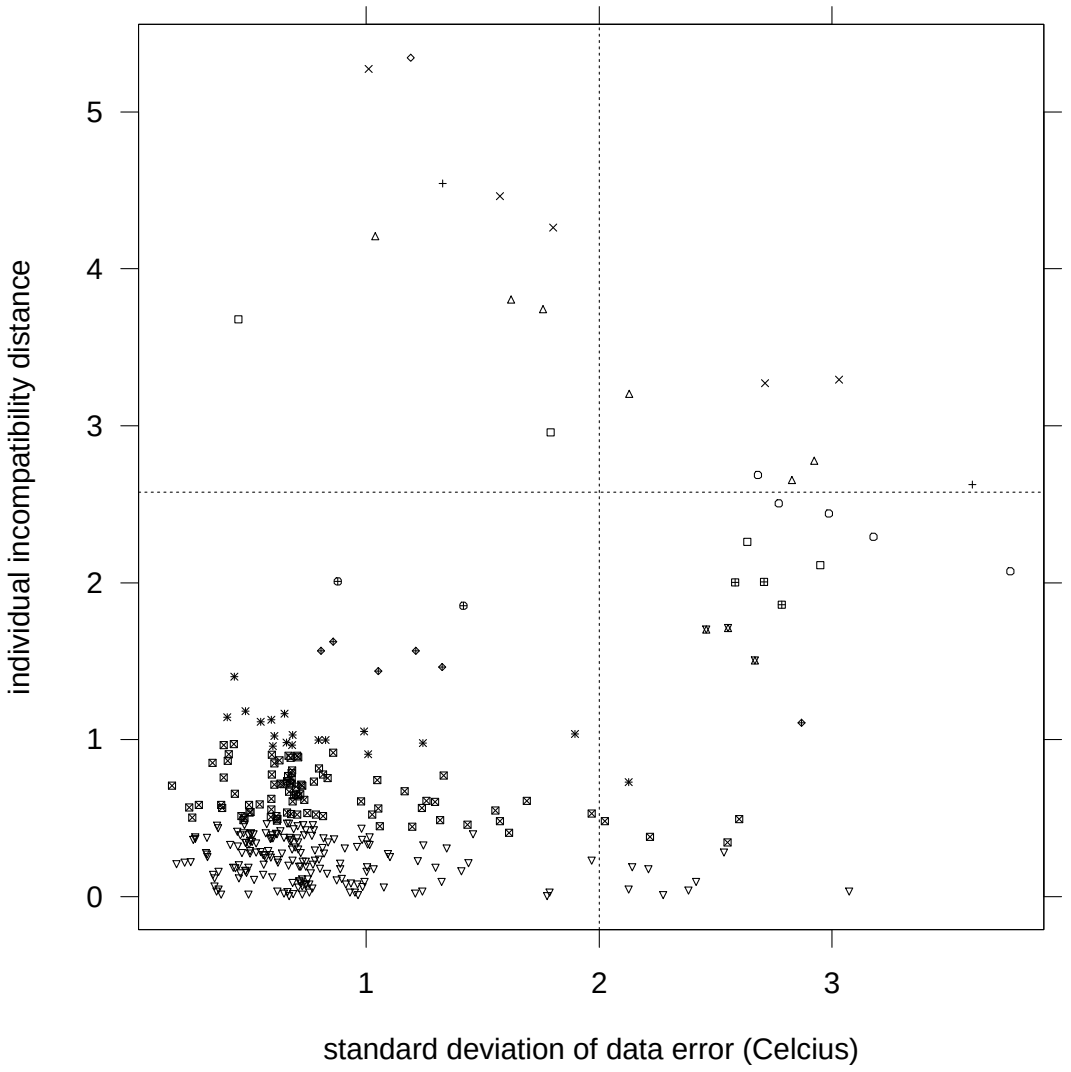

Fig. 4. Individual incompatibility distances s_1 for the 318 observations and their model counterparts vs. standard deviation of observational errors. The vertical line indicated the common value 2°C for the standard deviation of the model errors that was used in computing s_1. The horizontal line indicates the value $s_1 = 2.58$ for which $p_1 = 0.01$. Different plotting symbols are used to indicate when the absolute value of the difference between observation and model counterpart fall into different 1°C intervals which tend to increase with increasing s_1: □, ○, △, +, ×, and ◇ indicate innovations ranges increasing from 7–8°C to 12–13°C, respectively.

Before data are assimilated, they can be checked individually for compatibility with their model counterparts. The cutoff for acceptable compatibility can be adjusted to achieve the desired degree of screening. For example, you might want the relatively conservative requirement of no gap between the uncertainty intervals. For $\sigma_d = \sigma_b$, the no-gap criterion $\delta \leq \sigma_d + \sigma_b$ gives $s_1 \leq \sqrt{2}$ with $p_1 \geq 0.157$, but for $\sigma_d = 0$, *i.e.*, perfect observation, the no-gap criterion is $s_1 \leq 1$ with $p_1 \geq 0.317$. More lenient would be to consider the model, data, and error estimates compatible when $p_1 \geq 0.05$, which corresponds to $s_1 \leq 1.96$;

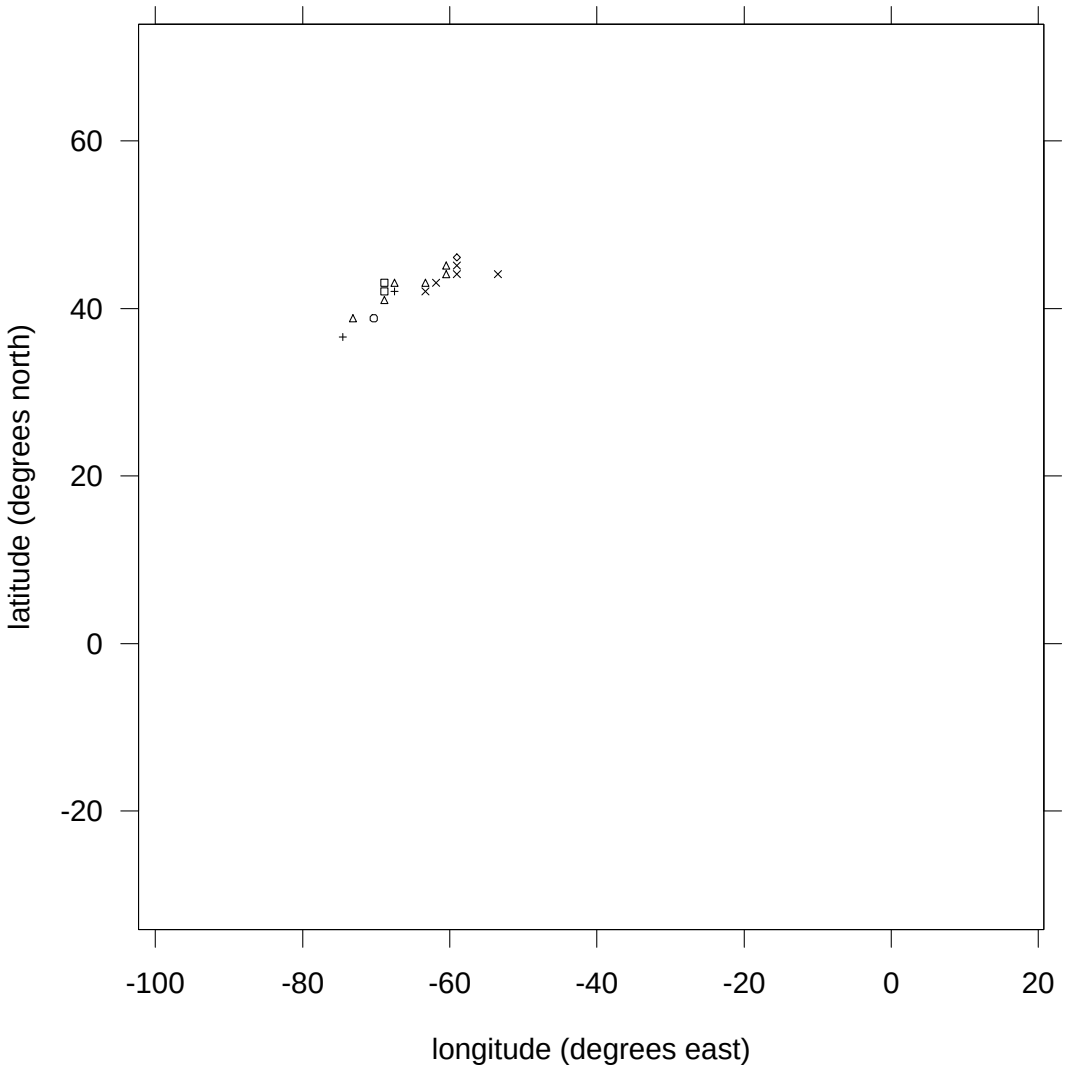

Fig. 5. Locations of 17 of the 318 cells of figure 3 where the probability of finding greater incompatibility of XBT data, model counterparts, and error estimates is less than 1%. The grouping by temperature differences is the same as for figure 4.

three cases in figure 2 satisfy this criterion. Still more lenient would be a cutoff of $s_1 \leq 2.58$ corresponding to $p_1 \geq 0.01$.

Bathythermographic (XBT) data that are to be assimilated into a model of the Atlantic Ocean (Halliwell, 2002a,b) provide a practical example. Figure 3 shows the grid cells where there are data during January 1972. The standard deviations of the observational errors σ_d are taken to be the climatological standard deviations of the temperature field (Levitus and Boyer, 1994) for these locations. For brevity, attention is restricted to the model's surface layer and thus to averages of the temperature profiles from the surface to the depth of 6 meters. The standard deviation of the errors of the simulated temperatures σ_b is assumed to be 2°C everywhere, so that most data are considered more accurate than their model counterparts (Thacker and Esenkov, 2002).

A scatter plot of s_1 vs. σ_d is shown in Figure 4. Observational errors can be seen to vary substantially. As magnitudes of innovations are indicated by the different plotting symbols, it is clear that there are cases of relatively small observational errors for large model-data differences. The cutoff value for s_1 is a matter of judgment, and should be explored more thoroughly than is appropriate here. For illustrative purposes, the cutoff is set so that $p_1 = 0.01$ as indicated by the horizontal line; there are 17 cells with s_1 above this cutoff. Their locations are indicated in figure 5. The cells with the largest values of s_1 are situated within the envelope of the Gulf Stream Front. As this is a region of high variability, it may be as appropriate to question the relatively low error estimates as it is to question the large differences between model and data. In this example the model had the front relatively far to the north while the observations indicated its meander to the south.

5 Compatibility for pairs of observation.

Suppose that the vector $\boldsymbol{d}_2$ represents a pair of observations with 2×2 error-covariance matrix $\boldsymbol{\Sigma}_d$ and that $\boldsymbol{b}_2$ and $\boldsymbol{\Sigma}_b$ are their model counterparts[9]; then the error-covariance matrix for the innovation vector $\boldsymbol{\delta}_2 = \boldsymbol{d}_2 - \boldsymbol{b}_2$ is $\boldsymbol{\Sigma}_\delta = \boldsymbol{\Sigma}_d + \boldsymbol{\Sigma}_b$. The Mahalanobis length of the innovation vector defines an incompatibility distance s_2 between the pair of observations and their model counterparts:

$$s_2^2 = \boldsymbol{\delta}_2^{\mathrm{T}} (\boldsymbol{\Sigma}_d + \boldsymbol{\Sigma}_b)^{-1} \boldsymbol{\delta}_2 \ . \tag{6}$$

If the error assumptions are taken seriously, $\boldsymbol{\delta}_2$ should be distributed bi-normally, and the probability[10] of a pair of innovations with larger s_2^2 is given by the integral of the bi-variate normal density over the annular region $\boldsymbol{\delta}_2^{\mathrm{T}} \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\delta}_2 > s_2^2$. Thus,

$$p_2 = \exp\left(-\frac{s_2^2}{2}\right) \tag{7}$$

is the probability of finding less pairwise compatibility.

To explore the role of correlated errors, it is instructive to consider the case of equal and opposite innovations and positively correlated background errors[11].

[9] The matrices $\boldsymbol{\Sigma}_d$ and $\boldsymbol{\Sigma}_b$ are sub-matrices of D and $\mathsf{HBH}^{\mathrm{T}}$, respectively.

[10] The same value for the p_2 is given by (3) with $n = 2$.

[11] When the separation of the two observations is comparable to the model's resolution, the background error correlation should not vanish, because consistency with the partial derivatives of the governing equations requires smooth fields and thus correlated errors.

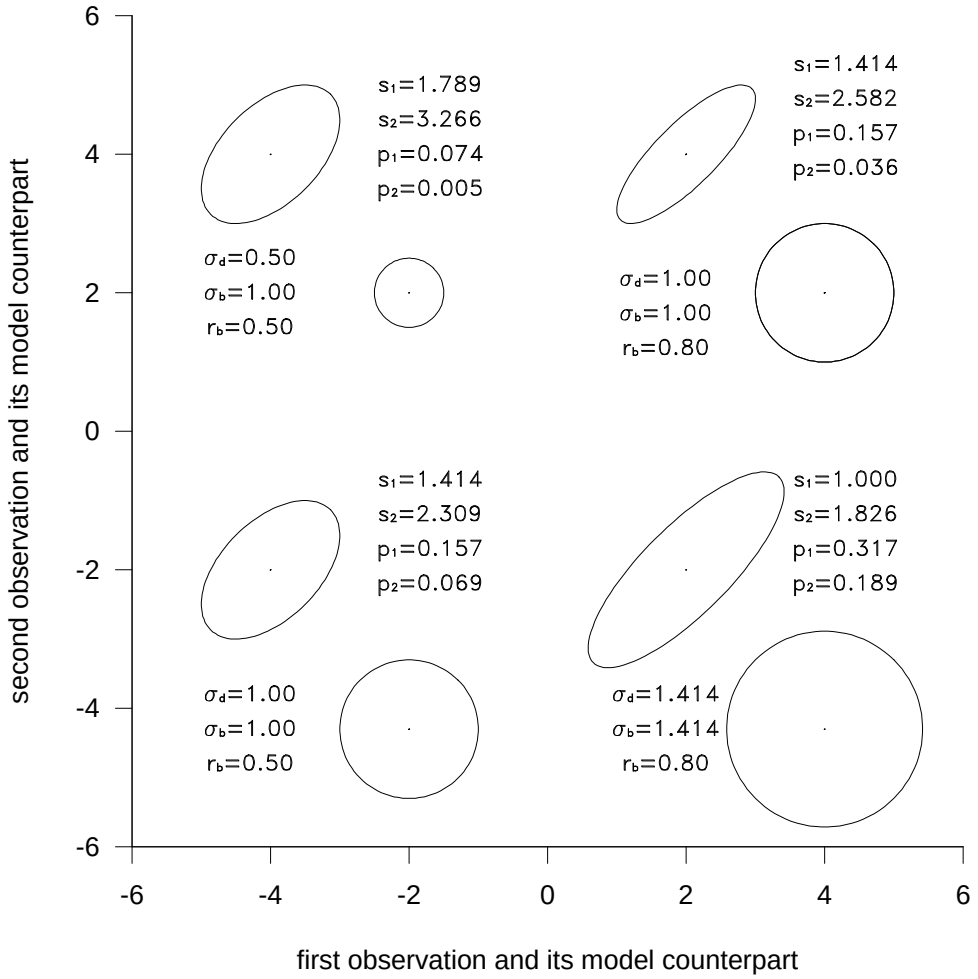

Fig. 6. Four examples of pairs of observations and their model counterparts. Dots indicate values of estimates, to be read from scales at bottom (first member) and left (second member); in every case, individual model-data differences are ± 2 non-dimensional units. Ellipses and circles indicate confidence regions corresponding to Mahalanobis distance of unity about model and data estimates, respectively. Standard deviations of data and model errors, which are taken to be the same for both variables, and the background error correlation coefficient are indicated to the left; to the right, single-observation incompatibility distance and probability of finding greater incompatibility, which are taken to be the same for both variables, and pairwise incompatibility distance and probability of finding a less compatibility pair of observations.

For further simplicity, suppose that the error variances for the two observations have the same value σ_d and that their model counterparts both have error variance σ_b. Further assume that the observational errors are uncorrelated and that the background errors are correlated with coefficient r:

$$\boldsymbol{\delta}_2 = \delta \begin{pmatrix} +1 \\ -1 \end{pmatrix} \qquad \boldsymbol{\Sigma}_d = \sigma_d^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \boldsymbol{\Sigma}_b = \sigma_b^2 \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} . \tag{8}$$

For this case the expression (6) takes a relatively simple form:

$$s_2^2 = \frac{2\delta^2}{\sigma_d^2 + (1-r)\sigma_b^2} \,. \tag{9}$$

When background errors are positively correlated ($r > 0$), s_2^2 is *greater* than the sum of the squares of the individual incompatibility distances. This should be expected, because a positive correlation[12] is inconsistent with the opposite signs of the innovations. If background errors are uncorrelated ($r = 0$), the *individual* no-gap criteria $\delta = \sigma_d + \sigma_b$ with $\sigma_d = \sigma_b$, for which (5) gives $p_1 = 15.73\%$ for less compatible individual innovations, (7) gives probability $p_2 = \exp(-2) = 0.135$ for a less compatible *pair*; on the other hand, if background errors are completely correlated ($r = 1$), the individual no-gap criteria gives the much smaller probability $p_2 = \exp(-4) = 0.018$ of finding a more incompatible pair.

Figure 6 illustrates the dependence of the two-observation incompatibility distance on the nature of the background and data errors. Four examples are shown, each corresponding to the non-dimensional innovation for one variable being +2 and the other being -2. The data- and model-error standard deviations and are taken to be the same for both variables; they range from $1/2$ to $\sqrt{2}$. The errors of the observations are assumed to be uncorrelated, so their confidence regions are circular. The correlation coefficient r for the background errors has value of 0.5 for two cases and 0.8 for the other two cases. Projecting the ellipses and circles onto the horizontal axis generates intervals such as those in figure 2 illustrating s_1 and p_1 for the first innovation; projecting onto the vertical axis generates the same for the second innovation. Because of the assumed equality of the innovations and that of their errors, the single-observation incompatibility distances s_1 are the same for both innovations, as are the probabilities p_1 of encountering less compatible individual innovations. Three of the four examples satisfy the single-observation compatibility criterion of $s_1 \leq \sqrt{2}$ and the fourth satisfies the more lenient criterion of $p_1 \geq 0.05$. However, only the lower-right example, for which $p_1 > 31\%$, would pass a two-observation compatibility criterion of $p_2 > 15\%$. Reducing the error estimates until $p_1 = 15.7\%$ (upper-right), causes the probability of encountering a less compatible pair of innovations to drop to $p_2 = 3.6\%$. If the background correlation is reduced (lower-left), then p_2 increases to 6.9%. The upper-left example illustrates how increasing the accuracy of the observation reduces the data-model-error compatibility; the probability of a less compatible pair drops from 6.9% (lower-left) to 0.5%.

Figure 7 shows how incompatibility distance from (9) varies with the size of the innovations, with the correlation of the background errors, and with the

[12] A negative correlation would cause the same problem, if the two innovations had the *same* sign and equal magnitude.

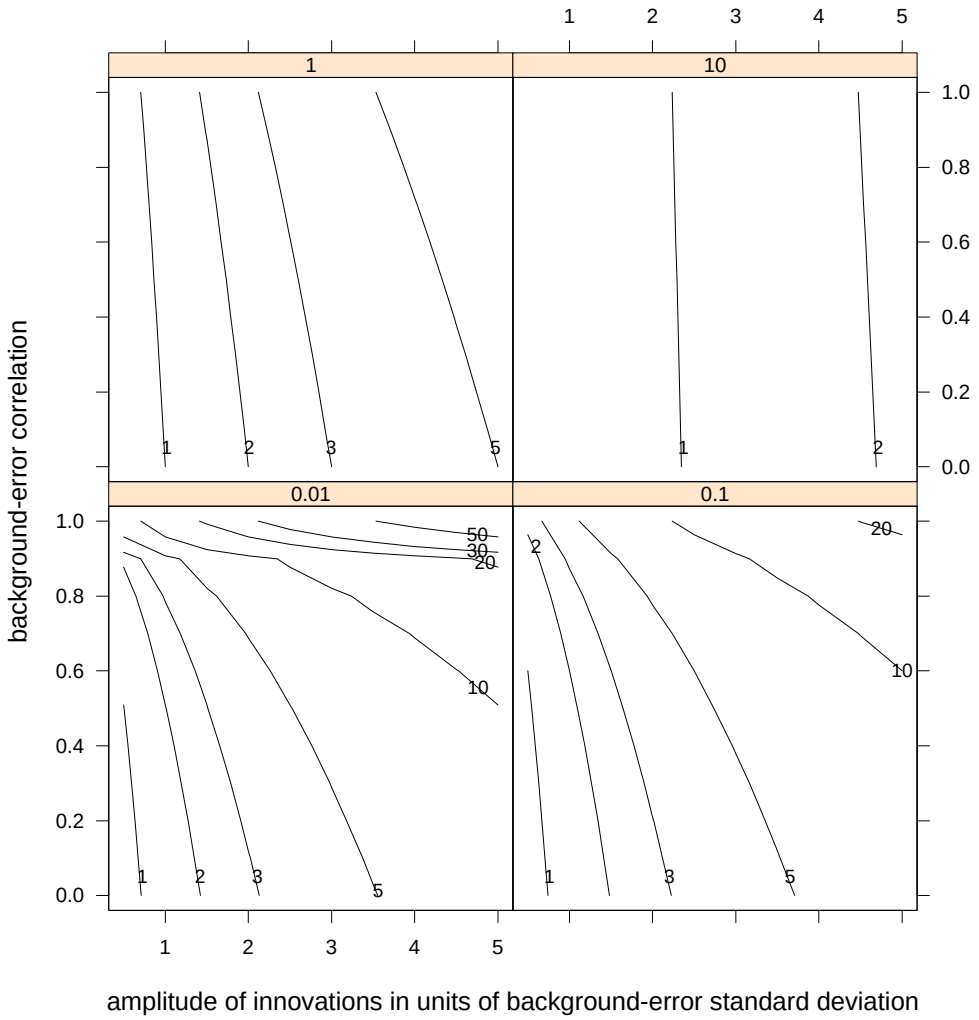

Fig. 7. Contours of incompatibility distance s_2 for the case of equal and opposite innovations. Panel labels indicate the ratio σ_d^2/σ_b^2 of observational to background-error variance, which range from 0.01 to 10. Horizontal axis indicates ratio δ/σ_b of innovation magnitude to background-error standard deviation. Vertical axis indicates correlation coefficient r of background errors for the two observations. For contours $s_2 = 1, 2, 3$, and 5, the corresponding values of p_2 are 60.65%, 13.53%, 1.11%, and 0.0004%; values for the remaining contours are much smaller.

data-to-model ratio of error variance. The upper-right panel is the ideal case of a model that is more accurate than the data; dependence on error correlations is weak and incompatibility is largely a function of the size of the model-data differences. When $\sigma_d = \sigma_b$ (upper-left), incompatibility increases rapidly with the size of the innovations, but there is also a significant increase as background errors become more highly correlated; for $\delta = 2\sigma_b$, $p_2 < 5\%$ when $r > 0.66$. Although models may provide an excellent qualitative simulation of the ocean, their quantitative performance is not as good, and data are generally more accurate than their model counterparts[13]. For $\sigma_d^2 = \sigma_b^2/10$ (lower-right) and

[13] Model biases, such as the systematic misplacement of the Gulf Stream front, are

$\delta = 2\sigma_b$, the probability of a more incompatible pair of innovations is only $p_2 = 2.6\%$ if $r = 0$ and diminishes rapidly as r increases. For still more accurate observations with $\sigma_d = \sigma_b/10$ (lower-left), only opposing pairs of innovations that are very small are compatible with the errors; it is unlikely to find individual innovations larger than twice the standard deviation of the background error ($p_1 < 4.7$), so it is understandable that equal and opposite pairs of innovations of this size give rise to such large incompatibility distances.

When errors are uncorrelated, the matrices $\mathbf{\Sigma}_d$ and $\mathbf{\Sigma}_b$ are diagonal, and the square of the pairwise incompatibility distance is simply the sum of the squares of the two individual incompatibility distances. If s_1 is the same for both innovations, so that p_1 is also the same for both, it does not follow that the probability p_2 of encountering another *pair* of innovations with larger s_2 is also p_1. For example, if $p_1 = 0.01$ for both innovations, then $s_1^2 = 6.635$ for both, so $s_2^2 = 13.270$ and $p_2 = 0.0013$. In other words, two unlikely innovations will constitute an even less likely pair. In fact, a single unlikely innovation, when combined with one that is more reasonable, will lead to an unlikely pair; *e.g.*, $s_1^2 = 6.635$ ($p_1 = 0.01$) for the first innovation and $s_1^2 = 2.575$ ($p_2 = 0.28$) for the second gives $s_2^2 = 9.21$ ($p_2 = 0.01$) for the pair. Individual data-model-error incompatibility induces pairwise incompatibility.

While pairwise probability p_2 is smaller than the individual probabilities p_1 for the members of the pair, it is considerably larger than the product of the p_1's. Suppose that the ellipses (circles) in figure 6 were enclosed by squares, the length of each side being two standard deviations of the data or background error. The product of the p_1's corresponds to the area outside these boxes, which is smaller than that outside the ellipses.

It is useful to reconsider the example of figures 3, 4, and 5 in terms of pairwise incompatibility. With the error variances as before, background errors are assumed to be correlated with correlation coefficient that decreases with the square of cell separation:

$$r = \exp\left(-\frac{\Delta i^2 + \Delta j^2}{4}\right) \ . \tag{10}$$

Figure 8 shows the pairwise incompatibility distance s_2 for all pairs of cells in figure 3 excluding those in figure 5 plotted against the pair's average observational-error standard deviation. The data are displayed in different panels according to the size of the correlation coefficient for the background errors of the cell pairs. The upper-right panel corresponds to pairs of cells that have a common side, the upper-left to those with a common corner, *etc.* The lower-left panel includes pairs that are separated by two or more intervening cells. It is interesting to note that the largest pairwise incompatibilities are

generally treated within the data-assimilation formalism as random errors.

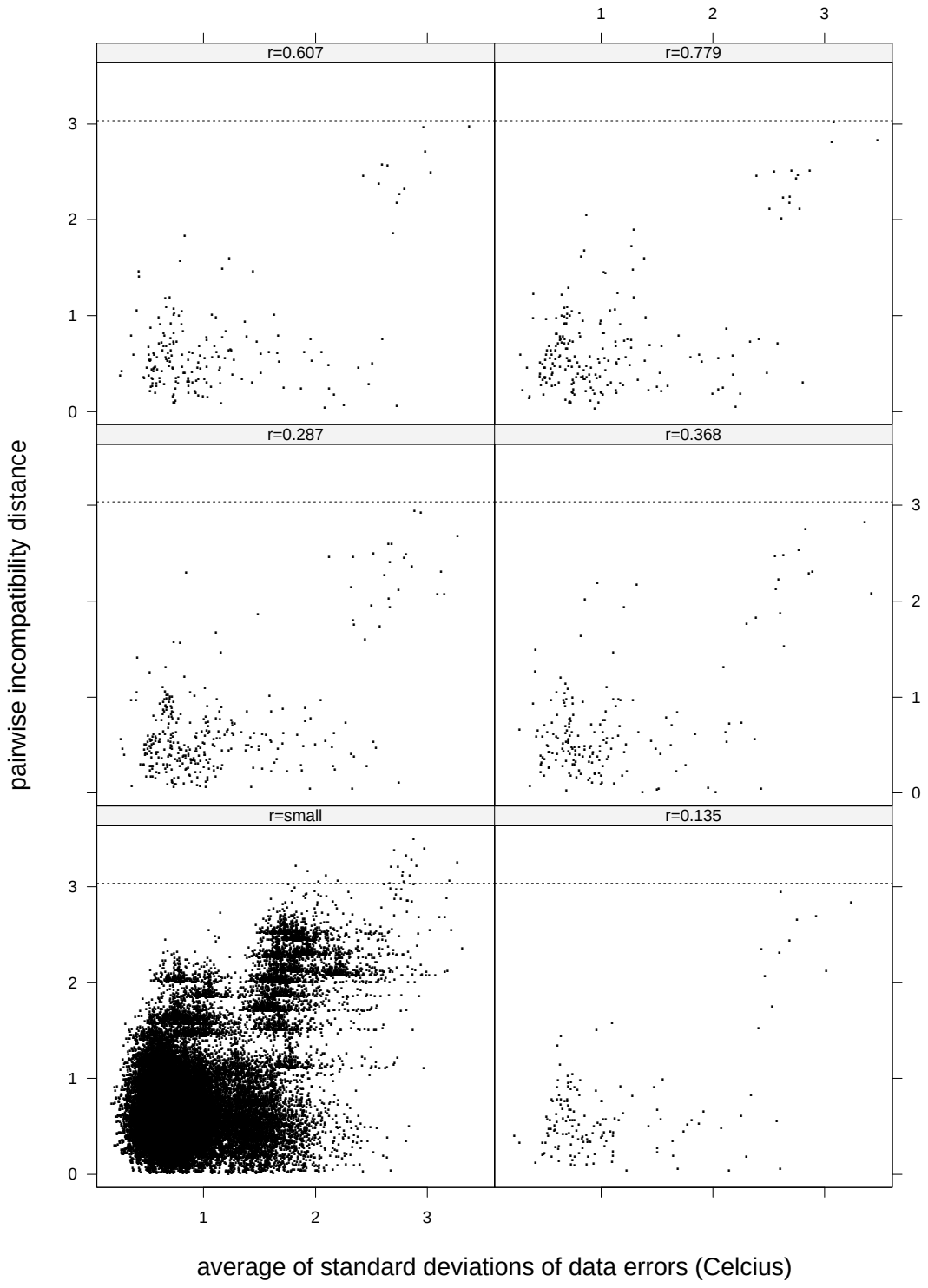

Fig. 8. Incompatibility distance s_2 for all pairs of cells with $s_1 < 2.58$. Pairs are grouped by the magnitude of the background-error correlation coefficient, which is indicated in the panel labels. The horizontal line indicates the value of s_2 for which $p_2 = 0.01$.

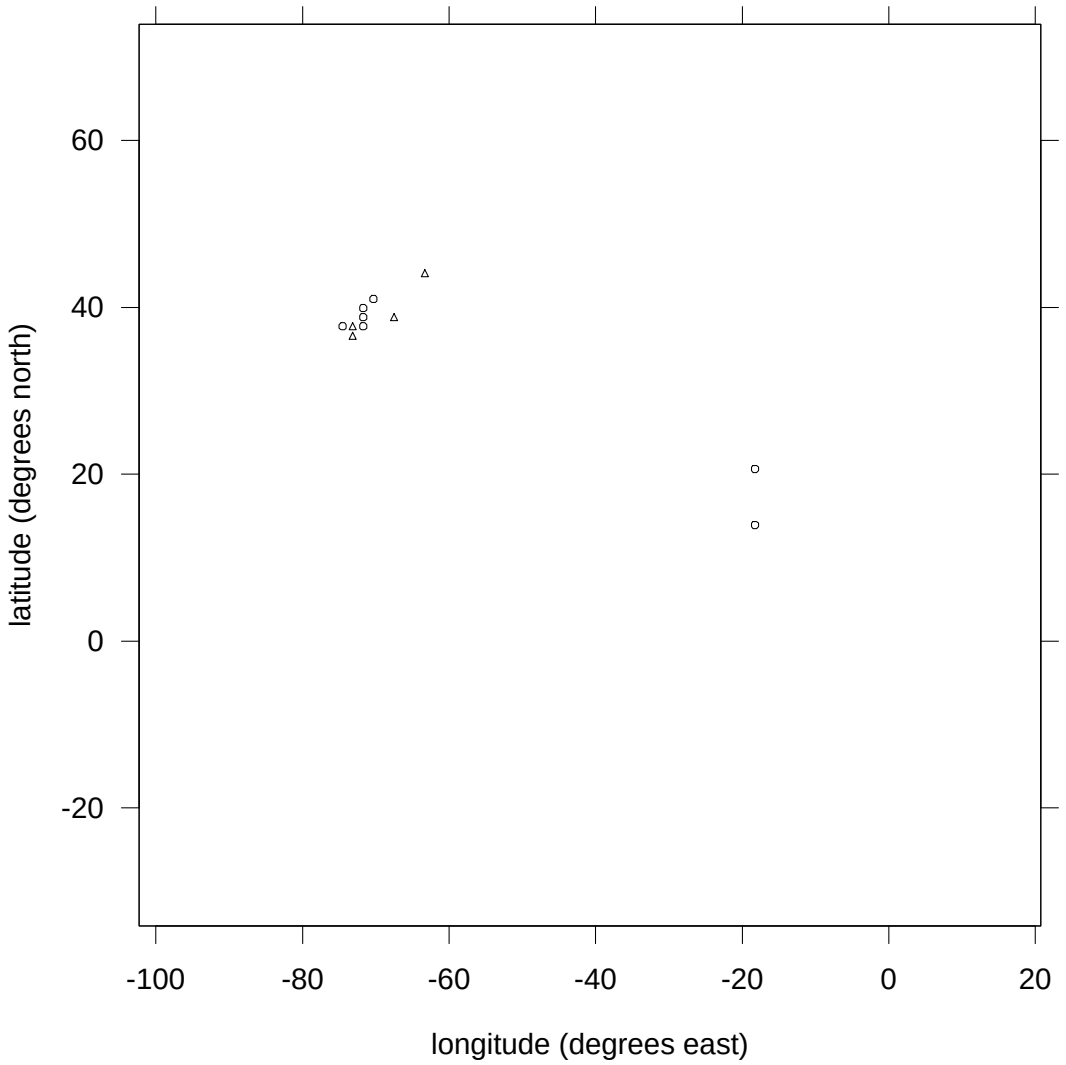

Fig. 9. Locations of the 11 cells that are members of the 19 pairs with $s_2 > 3.03$ in figure 8. The 4 indicated by $\triangle$ have $s_1 > 2.25$; at least one of these 4 appear in each of the 19 pairs.

found in the lower-left panel, so they cannot be attributed to background-error correlations. Instead, they reflect the fact that both members of these pairs exhibit relatively high individual data-model-error incompatibility. The locations of these cells are shown in figure 9; except for the two off the west coast of Africa, another region of high thermal variability, they are in the Gulf Stream region near the cells shown in figure 5. The four having the highest values of s_1 are members of all nineteen pairs with high s_2. For these data it appears that incompatibilities are more a consequence of error estimates being small in comparison to the sizes of the innovations than to high correlations across fronts.

6 Conclusion.

The question of data-model-error incompatibility is closely related to the task of identifying erroneous data, which has received considerable attention by meteorologists. The paper by Lorenc and Hammon (1988) addresses that problem within the framework of Bayesian statistics, while a more recent work (Dee et al., 2001) bears much similarity to what has been presented here, but treats *differences* of innovations rather than the innovations *per se*. By assuming model state and error estimates to be relatively problem-free, all incompatibilities can be attributed to the data. Unfortunately, oceanographic models exhibit substantial bias. For example, the Gulf Stream Front can be shifted substantially from its true position, and the correspondence between actual and simulated mesoscale eddies can be quite poor. Moreover, knowledge of errors of both observations and simulations are poorly known; as mentioned above, even estimates of climatological variability can be subject to question. Thus, it is useful to approach the question of data-model-error incompatibility without prejudice as to the cause. The concept of an incompatibility distance, together with an estimate of the probability of encountering less compatibility, not only offers a way to screen data before they are assimilated into numerical models, but also offers a means for identifying systematic modelling problems that should be addressed and, as the real-data example here demonstrates, and for drawing attention to details of the error models that might be improved.

References

Behringer, D. W., Ji, M., Leetmaa, A., 1998. An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. Monthly Weather Review 126, 1013–1021.

Bretherton, F. E., Davis, R. E., Fandry, C. B., 1976. A technique for objective analysis and design of oceanographic experiments applied to mode-73. Deep-Sea Research 23, 559–582.

Carton, J. A., Chepurin, G., Cao, X., Giese, B., 2000. A simple ocean data assimilation analysis of the global upper ocean 1950–1995. Journal of Physical Oceanography 30, 294–309.

Cleveland, W. S., 1993. Visualizing data. Hobart Press, Summit, NJ.

Cohn, S. E., 1997. An introduction to estimation theory. Journal of the Meteorological Society of Japan 75, 257–288.

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

Dee, D. P., Rukhovets, L., Todling, R., da Silva, A. M., Larson, J. W., 2001. An adaptive buddy check for observational quality control. Quarterly Journal of the Royal Meteorological Society 127, 2451–2471.

Evensen, G., 1994. Inverse methods and data assimilation in nonlinear ocean models. Physica D 77, 108–129.

Gandin, L. S., 1963. Objective Analysis of Meteorological Fields. Israel Program for Scientific Translations, Jerusalem, translated from the Russian.

Gelb, A. (Ed.), 1974. Applied Optimal Estimation. Springer-Verlag, Cambridge MA.

Halliwell, Jr., G. R., 2002a. Modes of Atlantic Ocean variability excited by NAO-related atmospheric forcing In preparation.

Halliwell, Jr., G. R., 2002b. A synthetic float analysis of upper-limb meridional overturning circulation interior ocean pathways in the tropical/subtropical Atlantic Submitted to Interhemispheric Water Exchange in the Atlantic Ocean, G. Goni and P. Malanotte-Rizzoli, Eds.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. Transactions, American Society of Mechanical Engineers, Series D, Journal of Basic Engineering 82, 35–45.

Levitus, S., Boyer, T., 1994. World Ocean Atlas, 1994, Volume 4: Temperature. NOAA Atlas NESDIS. U. S. Department of Commerce, Washington, D.C.

Lorenc, A. C., Hammon, O., 1988. Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. Quarterly Journal of the Royal Meteorological Society 114, 515–543.

Mardia, K. V., Kent, J. T., Bibby, J. M., 1979. Multivariate Analysis. Academic Press, Reading MA.

Mood, A., Graybill, F. A., Boes, D. C., 1974. Introduction to the Theory of Statistics. McGraw-Hill, New York.

Stuart, A., Ord, J. K., 1987. Kendall's Advanced Theory of Statistics. Vol. 3. Oxford University Press, Oxford.

Thacker, W. C., Esenkov, O. E., 2002. Assimilating XBT data into HYCOM. Journal of Atmospheric and Oceanic Technology 19 (5), 709–724.