Correlations for Measuring Skill

Russ Davis

15 July 2003

Correlations are useful measures of skill because they describe the fraction of variance (relative to some prescribed mean) of one field that can be predicted by a linear model based on another field. Let *a(i)* & *b(i)* be samples of the two fields taken at the same positions or time labeled by *i*, let *A(i)* & *B(i)* be mean values or a priori estimates of *a(i)* and *b(i)* (perhaps constants or functions of *i*), let $a'(i) = a(i) - A(i)$ & $b'(i) = b(i) - B(i)$ be anomalies and let $\hat{a}'(i) = \alpha\, b'(i)$ be an estimate of $a'(i)$. Define an average <> which may be based just on the points *i* or may include further averaging over a larger ensemble (perhaps *i* refers to various depths and the average is also over various other times where profiles are taken at the same depths *i*). The mean square error of the linear estimate $\hat{a}'$ is $< (a' - \alpha b')^2 >$ and this is minimized by $\alpha =< a'b' > / < b'b' >$ and the minimum is $< (a' - \alpha b')^2 >=< a'a' > (1 - \rho^2)$ where $\rho =< a'b' > /[< a'a' >< b'b' >]^{1/2}$ is the correlation. The correlation has the advantages over mean square error $< (a' - b')^2 >$ for skill assessment that it dose not penalize the skill for an amplitude (gain) difference and, because it has a normalization built in so that $-1 < \rho < 1$ and provides its own ruler for determining good skill vs. poor skill.

The term ***pattern correlation*** is usually applied to the correlation appropriate to means *A* & *B* that are constants and where <> is the average over all positions *i*, and possibly over several samples at different times. In my opinion this is most appropriate when looking at a single depth (*i* then refers to various horizontal positions, and *A*=<*a*> & *B*=<*b*> are averages over positions *i* and possibly also time). If the averaging does not include time, the pattern correlation basically measures how similar two maps are.

One weakness in the pattern correlation is that it measures similarities of departures from the space-time mean, i.e. from a constant. If there are persistent spatial patterns (onshore offshore gradients, the decrease of temperature with depth) the pattern correlation is increased by the spatial structure in the temporal mean. A related weakness occurs when fields have spatially varying variance. In that case, errors of the size of the signal lead to small decreases in the correlation if they occur where the signal is small but large decreases if the error occurs where the signal is large.

I suggest that we also use a more complicated correlation to account for these weaknesses. Let **x** be an index of horizontal position, *z* be a depth index, and *t* be a time index. Use a model to generate the temporal-mean

$$A(\mathbf{x}, z) = \frac{1}{N} \sum_{t}^{N} a(\mathbf{x}, z, t)$$

(for each variable) and a temporal-horizontal-average standard deviation

$$\sigma^2(z) = \frac{1}{N}\sum_{t,\mathbf{x}}^{N}[a(\mathbf{x},z,t) - A(\mathbf{x},z)]^2$$

I propose to use a standard pattern correlation (as above) applied to variables that have been shifted and scaled to account for spatially varying mean values and depth varying standard deviations. The transformed model data would be

$$\tilde{a}(\mathbf{x},z,t) = [a(\mathbf{x},z,t) - A(\mathbf{x},z,t)]/\sigma(z)$$

and similarly observations would be transformed to $\tilde{b}$. I then propose the following three correlations to describe horizontal, vertical and temporal changes in the similarity between the fields.

Profile correlation

$$A_1(\mathbf{x},t) = \frac{1}{N}\sum_{z}^{N}\tilde{a}(\mathbf{x},z,t) \qquad B_1(\mathbf{x},t) = \frac{1}{N}\sum_{z}^{N}\tilde{b}(\mathbf{x},z,t)$$

$$\rho_1(\mathbf{x},t) = \frac{\frac{1}{N}\sum_{z}^{N}[\tilde{a}(\mathbf{x},z,t) - A_1(\mathbf{x},t)][\tilde{b}(\mathbf{x},z,t) - B_1(\mathbf{x},t)]}{\left(\frac{1}{N}\sum_{z}^{N}[\tilde{a}(\mathbf{x},z,t) - A_1(\mathbf{x},t)]^2\,\frac{1}{N}\sum_{z}^{N}[\tilde{b}(\mathbf{x},z,t) - B_1(\mathbf{x},t)]^2\right)^{1/2}}$$

In practice, since observations will not be at gridded positions and times and because it is desirable to get more than one profile in the sums over *z*, it will be desirable to consider **x** and *t* to represent areas of horizontal position and ranges of time. This correlation will provide a series (for different *t*) of maps showing where model and observed profiles are similar and different. If the spatial area corresponding to **x** becomes large, then the correlation becomes a full 3-D spatial correlation that varies with time.

Horizontal correlation

$$A_2(z,t) = \frac{1}{N}\sum_{\mathbf{x}}^{N}\tilde{a}(\mathbf{x},z,t) \qquad B_2(z,t) = \frac{1}{N}\sum_{\mathbf{x}}^{N}\tilde{b}(\mathbf{x},z,t)$$

$$\rho_2(z,t) = \frac{\frac{1}{N}\sum_{\mathbf{x}}^{N}[\tilde{a}(\mathbf{x},z,t) - A_2(z,t)][\tilde{b}(\mathbf{x},z,t) - B_2(z,t)]}{\left(\frac{1}{N}\sum_{\mathbf{x}}^{N}[\tilde{a}(\mathbf{x},z,t) - A_2(z,t)]^2\,\frac{1}{N}\sum_{\mathbf{x}}^{N}[\tilde{b}(\mathbf{x},z,t) - B_2(z,t)]^2\right)^{1/2}}$$

In practice, it will be desirable to consider z and *t* to represent ranges of depth and time. This correlation will provide a series (for different *t*) of profiles showing depths where model and observations are similar and different. If the depth range corresponding to *z* becomes large, then the correlation becomes a full 3-D spatial correlation that varies with time (like ρ_1). If the spatial mean *A* of untransformed *a* does not vary with horizontal position, and the depth range *z* is small, ρ_2 should be the same as the classical pattern correlation at that depth.

The third correlation of interest would be $\rho_3(t)$ obtained by combining all horizontal positions into a single index **x** in ρ_1 or all depths into a single index *z* in ρ_2. This tells how the similarity of the full spatial structures of *a* and *b* vary with time.

I would like to call the constructs above ***tempered correlations*** to account for the fact that the spatial variations of the mean and standard deviation have been accounted for so that spatial variation of the mean does not dominate the correlation and so that energetic and unenergetic regions are put on a par.