Quantifying model fit for spatially and temporally varying data

Authors: Phillip N Price, Cavan Reilly, Andrew Gelman
Email: pnprice@lbl.gov, cavanr@biostat.umn.edu, gelman@wald.stat.columbia.edu

Address of Dr. Price:
Indoor Environment Department, MS 90-3058
Lawrence Berkeley National Laboratory
Berkeley CA 94720
510-486-7875

Goal

Quantify model fit for "indexed" (time-series or spatial) data.

Problems with conventional approach: There are several standard methods for quantifying model fit between observations y(x) and predictions yhat(x), including:

1. Mean-squared error or mean absolute error
2. Overlap integral
3. Correlation coefficient, regression coefficient

All of these methods have the same two drawbacks. First, they treat all error as if it occurs in the "y" direction. Second, a "miss is as good as a mile": if a peak value is misplaced, it often doesn't matter whether it is misplaced by a little or a lot: the error can be the same either way.

Our Approach

We first "deform" the predictions—stretch them or compress them—to better fit the data. Then we see (1) how good is the fit after deformation, and (2) how much deformation did we do? This decomposes the error into two parts: error in y and error in x.

This is implemented by defining an error measure that includes two terms rather than just one. For a one-dimensional function (such as a time series), this can be written:

$$I_\lambda(y,\hat{y}) \equiv \min_{f\in\mathcal{D}} \left\{ \int_A G\Big(y(x),\hat{y}(f(x))\Big)\,dx + \lambda \int_A F\Big(x, f(x)\Big)\,dx \right\}$$

- G() is a function that compares the measurements to the deformed predictions; for example, it could be the difference in the squares of the values.
- F() is a function that quantifies the amount of deformation; it might measure the mean-squared magnitude of the deformation, for example.
- The function f(x) defines the deformation itself; the more it differs from f(x)=x, the bigger the deformation.

The parameter (λ) controls the trade-off between error in x and error in y: if lambda is set very high, deformation is heavily penalized; if λ is set very low, deformation is lightly penalized. (λ) quantifies the answer to the question "if your prediction could be wrong by 1 unit in y or z units in x, which would you prefer?"

Conclusions

We think the approach of quantifying both error in y and error in x is likely to be a useful one in many analyses of spatial and time-series data.

We have a computational approach that can always solve the one-dimensional (e.g., time-series) case, but we run into computational problems for 2-dimensional cases (e.g., maps), and cannot always find solutions for all values of the trade-off parameter (λ).

Reference

This poster is based on Reilly C, Price PN, and Gelman A, "Using image and curve registration for measuring the goodness of fit of spatial and temporal predictions," submitted to *Biometrics*, 2003.

A similar conceptual approach to model evaluation is described in Ramsay J. and Silverman B (1997), *Functional Data Analysis*, Springer-Verlag: New York.

1-Dimensional Example:

Figure 1 shows (hypothetical) "observed" values (black line with points) and "predicted" values (red line). The root-mean-squared (RMS) error in y is about 4.4 units.

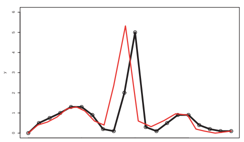
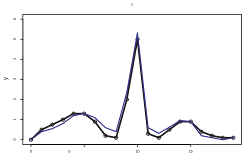


Figure 2 shows the same "observed" and "predicted" values, but the predicted values have been stretched and compressed along the y axis so as to maximize the agreement with the observed values (by setting λ very low).



The conventional error summary would say "predictions had an RMS error of 4.4 units in y." Our method allows us to say "predictions had an RMS error of 0.3 units in x, and 0.04 units in y."

2-Dimensional Example:

Figure 3 shows predictions (top) and observations (bottom) for dye concentrations in a water-tank experiment.

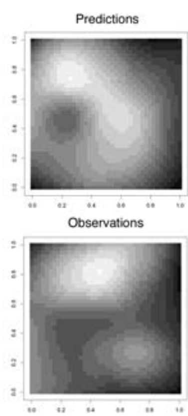


Figure 4 shows several successive distortions that allow the predictions to match the observations better and better as (λ) decreases. Our computer program runs into numerical difficulties for low values of λ, so we could not go below about (λ=5000).

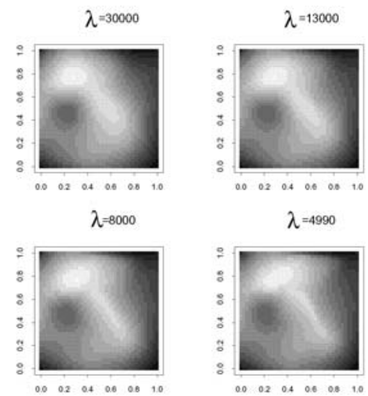


Figure 5 shows two ways of visualizing the deformation for (λ=4990): as a vector field and as a distorted grid: if we plotted the predictions on a rubber sheet that had a regular grid printed on it, this is what the grid would look like after the deformation.