Purpose of Objective Analysis

Observations are often available only at a few stations where they are unevenly spaced in the domain of interest. In order to compute derivatives of the field variables, as would be required in diagnostic studies or in the initialization of a numerical model, or simply to perform a sensible averaging process, one often needs values of the field variables at points on a regular grid.

The methods described are applicable to any problem where the data you are given do not fill the domain of interest fully, and where the data must be interpolated to a regular grid.

When to Do the Gridding?

The methods described are applicable to any problem where the data you are given do not fill the domain of interest fully, and where the data must be interpolated to a regular grid.

Polynomial Fitting Method

We can use a polynomial function to fit the observational data and then use this function to generate data on a regular grid.

In meteorological applications, up to third order polynomial have been used. But usually, quadratic equation is sufficient for most purpose:

The problem now is to determine the values of the coefficients of the quadratic polynomial function.

If we have 6 observations of h, then we can determine the 6 coefficients of the quadratic polynomial function.

We can then use the polynomial function for gridding purpose.
Problems With the Polynomial Method

The instability of the polynomial fit is such that when one key data point is removed, the polynomial fit in that region may change radically.

(from Hartmann 2003)

Optimum Interpolation

The optimum interpolation (OI) is a linear interpolation which requires the root-mean square error to be minimum. The discussion of this method focuses on the "deviations from a normal state":

Let's say we have a variable \( \phi \) and its normal state \( \phi_{\text{norm}} \). This normal state can be the climatological value of \( \phi \) or a first guess normal state.

The optimum interpolation (OI) is a linear interpolation which

Interpolation

\[
\frac{1}{N} \sum_{i=1}^{N} (\phi_i - \phi) = 0
\]

We want to determine the coefficients \( \{ p_i \} \) by minimizing the

\[
\text{mean squared error:}
\]

Next we want to approximate the value of \( \phi \) at a grid point (\( \phi_g \)), in terms of a linear combination of the values of \( \phi \) at neighboring station points (\( \phi_i \)).

Now, we want to approximate the value of \( \phi \) at a grid point (\( \phi_g \)), in terms of a linear combination of the values of \( \phi \) at neighboring station points (\( \phi_i \)).

We want to determine the coefficients \( \{ p_i \} \) by minimizing the

\[
\text{mean squared error:}
\]

The problem gets worse as the order of the polynomial is increased. The method is nearly useless when the data are sparse and especially so where data are missing.

Finally, let's extend the polynomial fit in this region to apply to

The instability of the polynomial fit in this region may change

Problems With the Polynomial Method
We can normalize the error $E$: 

$$
\text{error} = \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}
$$

Now we can determine the weightings ($pi$) by asking:

$$
\text{We then solve the N linear equations for the N p's.}
$$

It can be shown that the total derivative of error with respect to the coefficients is:

$$
\frac{\partial E}{\partial \beta} = \sum_i (y_i - \hat{y}_i) \frac{\partial \hat{y}_i}{\partial \beta}
$$

We then solve the N linear equations for the N $p_i$'s.

Now we can determine the weightings ($pi$) by making

$$
\frac{\partial E}{\partial p_i} = \frac{\partial E}{\partial \beta} \frac{\partial \hat{y}_i}{\partial p_i}
$$

We then solve the N linear equations for the N $p_i$'s.

What Do We Need to Get $p_i$?

In these examples, we choose different categories (e.g., those with expected high values in the composite analysis) to form the composite categories.
Steps in Compositing Analysis

1. Select the basis for compositing and define the categories. The categories might be related to the phase of some cyclic phenomenon or forcing, or to time or distance from some event. For example, we can use the Niño 3 index as the basis for compositing the ENSO cycle.

2. Compute the means and statistics for each category. We calculate the mean SST, wind stress, or heat flux for the onset, growing, and mature phases of the ENSO cycle.

3. Organize and display the results.

4. Validate the results. Validation of the results can be achieved in many ways. Statistical significance tests are only one of these.

An Example - the MJO Cycle

Step 1: Define Index

We use the principal components of the 1st and 2nd EOFs to define different phases of the MJO life cycle. We use the principal components of the EOFs to define the principal phases of the MJO cycle.

An Example - Time-Lag Correlation Between PC1 and PC2

An Example - the MJO Cycle
Step 2: Compute the Means

Step 3: Display the MJO Life Cycle

Indices of the Time Series