Gut checks for EOFs
with respect to determining whether
they're physically meaningful (Hartmann)

1) Is the variance explained by EOF more than would be expected if the data had no structure? Can null hypotheses of white and red noise be rejected?

2) Do you have an a priori expectation for structures? (e.g from some physical theory)

3) How robust are structures with respect to the domain ~~area~~ size?

4) How robust are structures to sample size?

Example of "good use" of EOF analysis:

Wallace and Gutzler (1981) → found large-scale patterns of teleconnections in NH winter using EOF analysis.

Later shown these structures can be explained by atmospheric dynamics and they were reproduced using a numerical model. (e.g. Hoskins & Karoly 1981)

Rotated EOF Analysis (REOFs)

"the" reference on this:
Richman (1986): Rotation of Principal Components, Int. J. Climatol.

Basic idea: Orthogonality constraint on EOFs often yields structures with large amplitude throughout the domain.

May have an 'a priori" expectation for more localized structures.

EOF patterns can be simplified by 'rotating' the leading EOFs, then projecting the patterns onto the input data matrix to obtain a corresponding expansion coefficient time series.

"Rotating" = "Taking linear combinations of EOFs"

Criteria that have been advocated for obtaining optimal rotations fall into two types.

1) Orthogonal rotations → preserve orthogonality of PCs.
2) Oblique rotations → do not preserve orthogonality of PCs.

In both cases, orthogonality constraint is relaxed.

General Mathematics

REOF analysis is a coordinate transformation from the original EOFs (E) to a matrix of rotated EOFs (B) by means of an invertible matrix of weights.

$$\underset{(K \times M)}{B} = \underset{(K \times M)}{E} \; \underset{(M \times M)}{R}$$

K = length of EOF
M = # retained EOFs

B: Rotated EOFs
E: Original EOFs
R: weights applied to each EOF in rotation.

For each rotated EOF (b_i) (dimension K)

$$b_i = \sum_{j=1}^{M} e_j r_{ji}$$

"M" = # retained EOFs

e.g. 1st rotated EOF (b_1)

$$b_1 = c_1 r_{1,1} + e_2 r_{2,1} + \dots e_M r_{M,1}$$

r_{ji} = Weights on original EOF

e_j = original EOF

The matrix R is chosen such that a particular function is maximized (s_p)

there are many ways to do this.

Varimax rotation: Maximizes the variance of the squared loadings in b_i

Remember, b_i is a vector of length K, where K is # of data points in space.

Write as $b_i = b_{ki} = b_{kp}$ $(p = i)$

For each rotated EOF (b_{ki}) out to a given number (M):

$$s_p^2 = \frac{1}{K} \sum_{k=1}^{K} \left(b_{kp}^2\right)^2 - \frac{1}{K^2} \left(\sum_{k=1}^{K} (b_{kp})^2 \right)^2$$

$p = 1 \ldots M$ $\quad M = $ # of rotated EOFs.
$K = $ number of spatial data points.

Choose elements of transformation matrix (R) to maximize:

$$\sum_{p=1}^{M} s_p^2$$

The solution is iterative - and standard on many statistical software packages.

Notes

- If the rotated EOFs are scaled by their std. deviation, then its the normal varimax. Otherwise, raw varimax.
- When sp^2 is maximized, the loadings in b_i tend toward 1 or 0.

Caveats of REOFs

1) Simplification (i.e. sp function) ignores any weighting to the input EOFs.
   - input EOFs can be given equal weighting
   - weighting by variance explained.

2) REOFs yield regionalized patterns that resemble one-pt. correlation maps.

3) If the number of EOFs being rotated equals the number of datapts. in space, then REOFs degenerate into localized "bulls eye" patterns.

4) If the number of REOFs (N) is too small, results are very sensitive to changes in N. In general, choose N large enough so results are reproducible for $N \pm 1$.

$N \sim 10$ commonly used.

5) The 'uniqueness' of REOFs does not derive from their ability to improve explained variance (ie. λs). It derives from their ability to explain more localized spatial patterns. So may be useful if the eigenvalues are not especially well separated.

Bottom line - jury is still out on REOFs!
- arbitrary choice of EOFs to rotate
- results are sensitive to small changes in number EOFs rotated.
- sensitivity to weighting input EOFs.
- Add additional level of complexity to the analysis, and are more difficult to reproduce.