Geophysical Research Abstracts, Vol. 7, 07109, 2005
SRef-ID: 1607-7962/gra/EGU05-A-07109

Detection of atmospheric circulation regimes: Weaknesses of nonlinear principal component analysis and cluster analysis

Bo Christiansen

Climate Research Division, Danish Meteorological Institute, (Email: boc@dmi.dk/Fax: +45 39 15 74 60)

The existence of regimes in the extra-tropical low-frequency variability may have important consequences for our understanding of climate and climate change. However, the existence of such regimes are still debated. A natural approach is to search for multi-modality in a probability density estimate. But, as the probability density only can be reliably estimated in one and two dimensions, other methods must also be used. In the last decade much work has been reported with different cluster analysis algorithms (k-means, finite mixture models). Recently, a nonlinear extension to Principal Component analysis has been introduced. This method heavily reduces the dimensions of the data so that multi-modality can be identified by studying probability density estimates.

The purpose of this paper is to emphasise some weaknesses of these methods.

We show that multi-modality is abundant in nonlinear principal component analysis when applied to sufficiently isotropic data even if these data are inherently unimodal. We recommend that the nonlinear principal component analysis should not be used for detection of multi-modality and regime behaviour.

We show that cluster analysis algorithms often report multi-modality even for distributions that are slightly skewed but otherwise smooth and without bumps or shoulders.

The weaknesses of the methods are demonstrated both by considering surrogate data and by a study of the 500 hPa geopotential heights from the NCEP reanalysis.