Geophysical Research Abstracts
Vol. 13, EGU2011-4516, 2011
EGU General Assembly 2011


# Innate Cycle Correlation.

Robin Crockett
Data Analysis Group, University of Northampton, Northampton, United Kingdom (robin.crockett@northampton.ac.uk)

Innate Cycle Correlation (ICC) is a generalisation of the Base Number Correlation technique previously developed for integer data containing integer-periodic features, such as enumeration data (Crockett et al., 2004). It is a novel technique, under development, for investigating and quantifying periodic features in time-series and other sequential data-sets.

In ICC, for a (potential) cycle of interest in the data, the time (sequential) index is scaled from arbitrary time units to numbers of cycles, zeroed at the start of the data. This effectively maps the set of time-magnitude ordered pairs to a set of phase-magnitude ordered pairs which will cluster according to the presence or absence of the cycle of interest in the data.

In the simplest case, as implicitly considered by Schuster (1897), the phase-magnitude ordered pairs will form a single cluster with either (a) the centroid at the origin in the absence of an innate cycle in the data or (b) the centroid at a significant distance from the origin in the presence of an innate cycle in the data. In this latter case, the angular position of the centroid indicates the phase of the cyclic maximum with respect to the start of the time-series. In more complex cases, the phase-magnitude pairs will cluster into more than one cluster, with cluster-centroids determined by the sub-cyclic structure.

In the case of simple single-cluster behaviour, ICC can be used to generate an amplitude spectrum (of centroid positions/distances) comparable to a DFT/FFT spectrum. This allows the significance of any cycles in the data to be assessed with respect to the baseline noise in the amplitude spectrum. In the data investigated at this preliminary stage, the magnitudes of the amplitude spectrum are closely Gumbel-distributed, allowing an assessment of the significance of components. Used in this manner, subject to statistical considerations, ICC has two advantages over DFT/FFT:
i) the time (sequential) index does not have to be uniform throughout the data, i.e. the sample interval does not have to be constant, and the data can include interruptions which do not have to be interpolated;
ii) it is possible to investigate user-specified frequencies, whereas frequencies in a DFT/FFT spectrum are dictated by the number of data-points and duration of the (constant) sample interval.

References.

Crockett R G M, Crockett A C, Turner S J; 'Base-Number Correlation': a new technique for investigating digit preference and data heaping. History and Computing, 13, 2, 2001 (pub. 2004) 161-179. 2004.
Schuster A; On lunar and solar periodicities of earthquakes, Proc. R. Soc. London, 61, 455– 465. 1897.