



## **Principal Component Analysis with Outliers and Time Series data**

Valderio Anselmo Reisen

Universidade Federal do Espírito Santo - Brazil - [valderioanselmoreisen@gmail.com](mailto:valderioanselmoreisen@gmail.com)

Higor Henrique Aranda Cotta

Universidade Federal do Espírito Santo - Brazil - [cotta.higor@gmail.com](mailto:cotta.higor@gmail.com)

Edson Zambon Monte

Universidade Federal do Espírito Santo - Brazil - [edsonzambon@yahoo.com.br](mailto:edsonzambon@yahoo.com.br)

This paper deals with the use of Principal Component Analysis (PCA) in time series with additive outliers. The PCA is one of the most important multivariate techniques which are linear combinations constructed to explain the variance-covariance structure of the original data. Although PCA assumes that the data are serially independent, this assumption is not found in practice situation in time series, e.g. Air Pollution data. PCs calculated from time series observations maintains their orthogonality property, but the components are found to be auto and cross-correlated, which depends on the correlation structure of the original series. These properties and their impact in the use of PCA are one of main objective of this work. Another contribution is related to the study of PCA time series under the presence of additive outliers by proposing a Robust PCA method. It is well known that additive outliers in time series destroys the correlation structure of the data. Since the PCs are computed by using the covariance matrix, the outliers also affect the properties of PCs. Therefore the Robust PCA should be used in this context. The Robust PCA method proposed here is justified empirical and theoretically, and a real data set based on Air Pollution time serie is used to show the usefulness of the Robust PCA method in a real application.

Keywords principal component analysis, outliers, times series. principal Name Surname\*