4^{th} AMMA International Conference

2-6 july 2012 Toulouse, France

Atlantic cold Tongue and West African monsoon: Nonlinear principal component analysis and canonical correlation analysis

Kenfack * **S. C.**[1,2,3], **Mkankan K. F.**[2], **Hounkonnou N. M.**[3], and **Alory G.**[4] ,

[1]*Dept.of Physics, Mesoscopic and Multilayer structure laboratory Univ. of Dschang, Cameroon*

[2] *Lab. for Environmental Modeling and Atmospheric Physics, Univ. of Yde I, Cameroon*

[3]*International Chair in Mathematical Physics and Applications, University of Abomey-Calavi,Cotonou, Benin.*

[4]*Laboratoire d'études en Géophysique et océanographie spatiales (LEGOS), Toulouse, France.*

*Email Address: kevinsadem@yahoo.fr

Abstract

Nonlinear principal component analysis (NLPCA) is a statistical method which detects and characterizes low-dimensional nonlinear structure in a multivariate data set. The application of NLPCA (using a neural network model) to SST data from the tropical Atlantic ocean is used to point out the Atlantic cold tongue (ACT) and the maximum global variance of the SST is shown. The tropical Atlantic SST and West African rainfall is analyzed by NLPCA. A nonlinear canonical correlation analysis (NLCCA) method is also used to find a possible relation between Atlantic Cold tongue variability and West African climate. The objective of this research is to study the sensibility between the Atlantic cold tongue effects and West African monsoon variability.

1 Introduction

The knowledge of our planet and the many variations in its climate increasingly becomes a major concern for all living organisms existing in it. Techniques of data collection have witnessed significant progress, and data collection tools have become increasingly sophisticated especially with the use of satellites as tools of mass data collection. However, given that the density of data increases over time, researchers have to work with more voluminous data whose management requires more original techniques.

Multiple varying techniques such as the principal component analysis (PCA), and canonical correlation analysis (CCA), have become indispensable in extracting essential information from voluminous data sets and relationship in two sets of data respectively. [1] Since the1980s, models of neural networks (NN) are more documented and Kramer (1991) [2] used this to generalize the PCA (CCA) and popularized its application for the extraction of nonlinear relationships in a set of data, where PCA and CCA then extends to the nonlinear principal component analysis (NLPCA) and nonlinear canonical correlation analysis (NLCCA) respectively. The introduction of these two techniques has become determining in the advancement of environmental science. [3]

The ocean has an impact on the West African climate which is not really noticed especially concerning its monsoon. Many studies such as those of Ping C. (2008) [4] and Hyacinth C. et al (2010) [5] have attempted to assess its contribution to the establishment of the monsoon. These studies have found that the weakening of the ocean circulation in the Atlantic contributes to the slowdown of the African monsoon. Moreover, modeling results (Messager, 2004) show that regional sea surface temperature appears as a major factor in the seasonal and inter-annual precipitation of monsoon on the African continent up to 12°N.

In addition, Atlantic Cold Tongue is the primary seasonal signal of sea surface temperature in the Atlantic East equatorial Ocean [6, 7]. It is an area that is located in the southeast of the equatorial Atlantic, in the Gulf of Guinea, and where surface waters are cooled immensely in Northern summer.

The inter-relationship between Atlantic cold tongue and African monsoon [8, 9] has remained at the center of many research. The study of the Atlantic cold tongue and West African monsoon might better be understood through the NLPCA and NLCCA.

In this poster, we apply the NLPCA and ACCNL on data from the tropical Atlantic sea surface temperature and those of precipitation in West Africa. The second part is the results. The last part concludes.

2 Method and Results

PCA method is the method which analyzes the variability of a single field (Rainfall, SST, etc). It finds the spatial patterns of variability, their time variation, and gives a measure of particular structure of each pattern. CCA examines the couple variability of two fields. Those methods have been performed using neural network (Fig. 1) and become NLPCA and NLCCA respectively.

Consider two data sets $\{x_i(t)\}$ and $\{y_i(t)\}$, where t is the time, or simply a label of the particular sample. Assume that there is a total of N samples in t for each variable $x_i(t)$ and $y_i(t)$. We grouped the $\{x_i(t)\}$ and $\{y_i(t)\}$ variables respectively to $x(t)$ and $y(t)$.

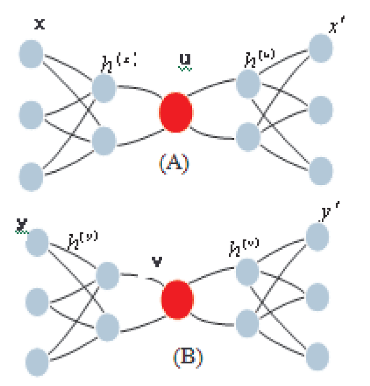

Figure 1: The NN model for calculating nonlinear PCA (CCA). There are 3 'hidden' layers of variables or 'neurons' (denoted by circles) sandwiched between the input layer x (with y for the case of NLCCA) on the left and the output layer x' (with y' for the case of NLCCA) on the right. Next to the input layer is the encoding layer, followed by the 'bottleneck' layer (with a single neuron u for NLPCA, u and ν for NLCAA), which is then followed by the decoding layer.

These two models (NLCCA and NLPCA) are well documented by Hsieh (2004) [3]. We used daily Atlantic sea surface temperature anomaly (SSTA) data from NOAA (National Oceanic and Atmospheric Administration), with spatial resolution of $0.25° \times 0.25°$ from the years 2000 to 2009 for all days of the months between March and August collected between 15°S to 7°N and 30°W to 14°E. This data was combined into 0.5° by 0.5° gridded data, thus there are (88×44) 3872 spatial points then $(88 \times 44 - 352)$ 3520 spatial and 184 time existing points. For this case we choose the smallest nonlinear NLPCA model $(m = 2)$ and the number of the parameter is 17608 which greatly exceeds the number of time points. Then, to reduce the number of input variables, pre-filtering the SST data by PCA is needed. PCA modes 1 and 2 accounted for 86.12 percent and 3.79 percent respectively of the variance in the SST data.

For the case of rainfall, analyzed data are total cumulative daily precipitation from TRRM (Tropical Rain Radiometer) covering the tropical part of Africa from 2°N to 20°S and 17°W to 18°E with a spatial resolution of 0.25x0.25, during the period 2000-2009. The data was transformed to a spatial resolution of $0.5° \times 0.5°$, then to 2627 points which are spatial variables and a time series of 1840 points, where l_2 = 2627 and N = 1840. Due to the number parameters (13143) which exceeded the number of time points in the case of smallest NLPCA model, we chose the first fifteen modes which accounted for 34,47 percent of the total variance.

2.1 NLPCA of the Atlantic tropical SSTA

The first mode (Fig. 2) of NLPCA explains 97.9 percent of the variance of the given EOFs, then 88.02 percent of the total variance of SSTA compared to 86.12 percent explained by the first PCA. The projection of NLPCA mode in the plane defined by the SST (PC1, PC2) is shown by the red curve in Figure 2. Unlike PCA, the NLPCA shows the nonlinear form in SSTA structure.

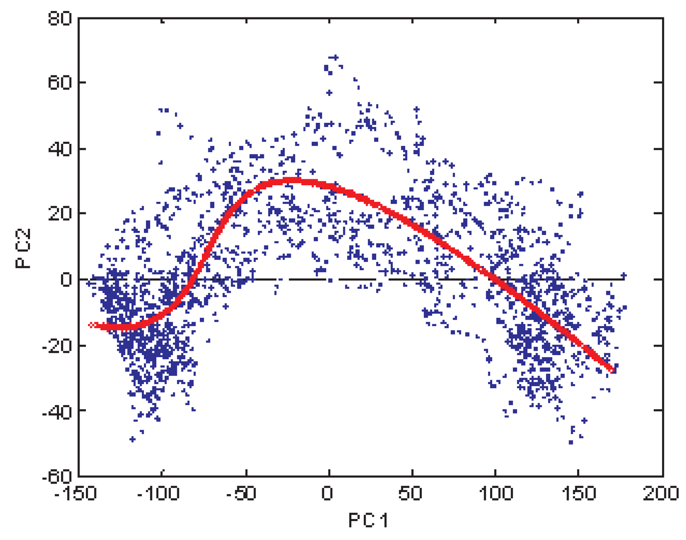

Figure 2: Scatterplot of the sea surface temperature (SST) anomaly (SSTA) data (shown as dot) in the principal component PC1-PC2 plane. The first principal component analysis (PCA) eigenvector is oriented along the horizontal line.

This first NLPCA mode shows that the maximum and the minimum value of the second component are bigger than the one of the first component. These values are mostly negative for the first component. We observe that the minimum and maximum of the second linear component are associated with the two minimum of the first linear component. This explains that, small values of the first linear component associated with very low values of the second linear component show cold water which is seem like Atlantic cold tongue.

2.2 NLPCA of rainfall in West Africa

As in the case of SST, the PCA represents the input rainfall data of NLPCA. The latter is also made for networks of 2 nodes in encoding and decoding layer with a single neuron in the layer "bottleneck" explaining 46.28 percent of the variance of the 15 EOFs (The 15 EOFs represent 34.47 percent of the variance of the initial data) chosen, then 15.9 percent of the total variance of rain compared to 7 percent explained by the first PCA. The projection of the first NLPCA mode in the EOFs rainfall subspace defined by (PC1, PC2), (PC2, PC3), (PC1, PC3) and (PC1, PC2, PC3) is shown by the red curves (Fig. 3).

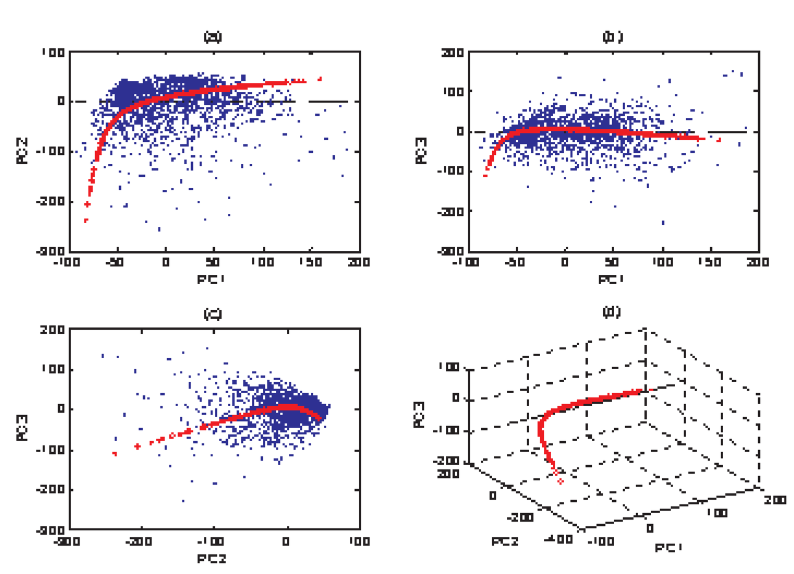

Figure 3: As Fig. 2, the case of rainfall in one and two dimensions.

The information provided by these curves summarizes both PCAs. The maximum and minimum observed (Fig. 3a) characterize the alternative variation between dry season and rainy season.

2.3 NLCCA of rainfall and SST

The canonical correlation analysis (CCA) has been given to meteorology and oceanography. It extracts the mode having the maximum correlation between two data sets. The progress of research on the ACC has experienced and marked improvement especially with its development by its modeling based on neural networks. An innovation was introduced recently by Hsieh (2000) as the nonlinear canonical correlation analysis. We may now analyze the similarities between the two sets of data discussed in the above section, using an NLCCA. This will allow us to see the ways in which the coupling between SST and precipitation is strongly done. The three CCA modes combine a variance of 99.93 percent, with 99.57 percent, for the first mode, 0.3 percent for the second and 0.06 percent for the third mode. The correlation coefficients between PCs of the two data sets are 0.87 for the first mode, 0.38 for second mode and 0.29, for the third mode, while the average correlation between spring and summer for the same mode gives the values 0.97, 0.79 and 0.7 respectively. The high value of the variance and correlation coefficient of the first mode shows the importance in the study of coupled SST-rain.

For the NLCCA mode, the inputs x are the first two PCs of the SST field, while the inputs y are the first fifteen PCs of the rainfall field and the numbers of hidden neurons are the same. Fig. 4.Cs1, Cs2, Cs3, and Cs4, represent the SST patterns of the first mode of NLCCA for some values of u ranging between its minimum and maximum. Moving the cooling southeast towards the equator marks the disappearance of Atlantic cold tongue.

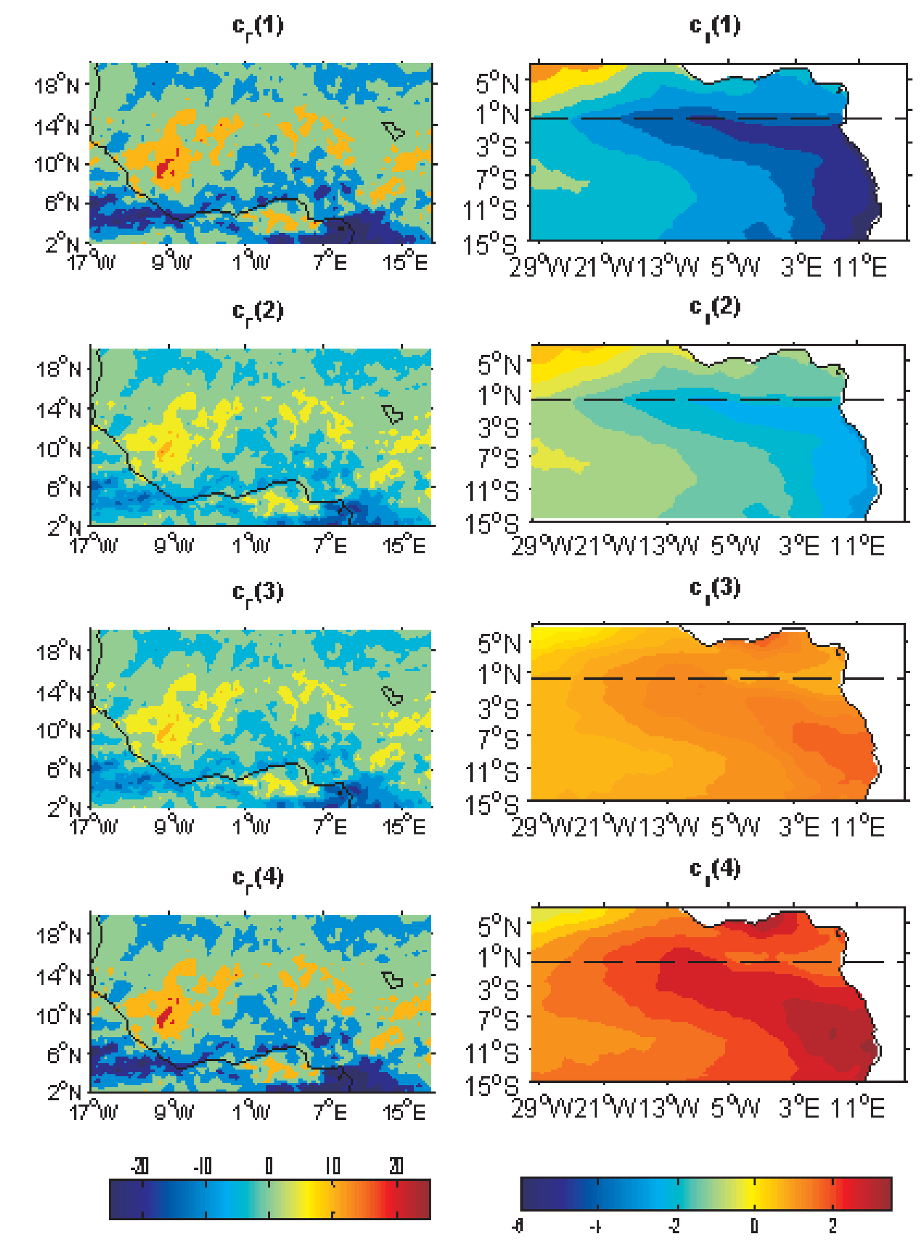

Figure 4: The rainfall field as the canonical variant v of the NLCCA mode 1 varies from (Cr1) its minimum, to (Cr2) half its minimum, to (Cr3) half its maximum and (Cr4) its maximum. The SST field as the canonical variant u of the NLCCA mode 1 varies from (Cs1) its minimum, to (Cs2) half its minimum, to (Cs3) half its maximum and (Cs4) its maximum.

Fig. 4.Cr1, Cr2, Cr3, and Cr4, represent the precipitation patterns of the first mode of NLCCA for some values of varying respectively from the minimum to its maximum. We saw that strong Atlantic cold tongue even (Fig. 4.Cs1) is associated with strong activity of rain (Fig. 4.Cr1) particularly in the Sahel region. This highlights the relationship between extreme events on the continent and the ocean. Unlike the CCA mode 1 (not shown) shows a structure of static climate in West Africa, NLCCA may present asymmetry between drought and monsoon season. The correlation between u and v is 0.87 for CCA versus 0.94 for NLCCA.

3 Conclusion

Nonlinear principal component analysis reveals most dominant mode of the equatorial Atlantic, which is the Atlantic cold tongue. We saw that Interannual variability of sea surface temperature (SST) in the Tropical Atlantic during the northern summer is closely related to rainfall variability in neighboring Gulf of Guinea and particularly at the beginning of the West African monsoon (WAM). This analysis consists of the first use of NLPCA data on tropical Atlantic SST and West Africa rainfall .

The projection of the first nonlinear PCA has nonlinearity in the data structure. The different spatial representations of reconstructed SST fields for some values of u and ν show the variations in Atlantic cold tongue and West Africa monsoon from strong event to weak event respectively. We may therefore conclude that these two main West African events are not varied identically every year.

NLCCA shows that the interchange between dry and wet seasons in West Africa is certainly influenced by the SST of the Tropical Atlantic Ocean.

Predictability of the appearance and strength of the West African monsoon associated with the Atlantic cold tongue remains an open question.

References

[1] Storch, H., and F. W. Zwiers (1999), Statistical Analysis in Climate Research, 484 pp., Cambridge Univ. Press, New York.

[2] , M. A. (1991), AIChE J., 37, 233-243.

[3] William W. (2004). Review of Geophysics, 42, RG1003, 1-25.

[4] Ping C., RongZ., WilcoH., CaihongW., XiuquanW, Link J.,Reindert j., Wim-paulB. and Howard S.(2008). Nature geoscience.

[5] Hyacinth C., Jianping Li. (2010). , 24:1184-1197.

[6] Wauthy, B., (1983). Océanogr. Trop., 18, (2), 103-138

[7] Merle, J., M. Fieux, and P. Hisard, 1980: Deep Sea Res., 26, 77-101.

[8] Okumura, Y., and S.P. Xie, 2004: Interaction of the Atlantic equatorial Cold Tongue and the African Monsoon. J. Climate, 17, 3589-3602.

[9] Hagos, S.M., and K.H. Cook, 2008: J. Climate, 18, 4993-5010, doi:10.1175/JCL13586.1.