European Geosciences Union
General Assembly 2022

Testing nonlinearity and nonstationarity of the connection between Palmer drought indices and Danube discharge in the lower basin

Romanian Academy
"Sabba S.Ştefănescu"
Institute of Geodynamics

Ileana Mareş, C. Mareş, Venera Dobrică, C. Demetrescu
Institute of Geodynamics, Romanian Academy, Bucharest, Romania

The study purpose is to select significant predictors from the Palmer-type drought indices for estimating seasonal discharge (Q) in the lower Danube basin. The four indices are: Palmer Drought Severity Index (PDSI), Palmer Hydrological Drought Index (PHDI), Weighted PDSI (WPLM) and Palmer Z-index (ZIND). These indices were quantified by PC1 of EOF decomposition, obtained from 15 stations located along the Danube basin.

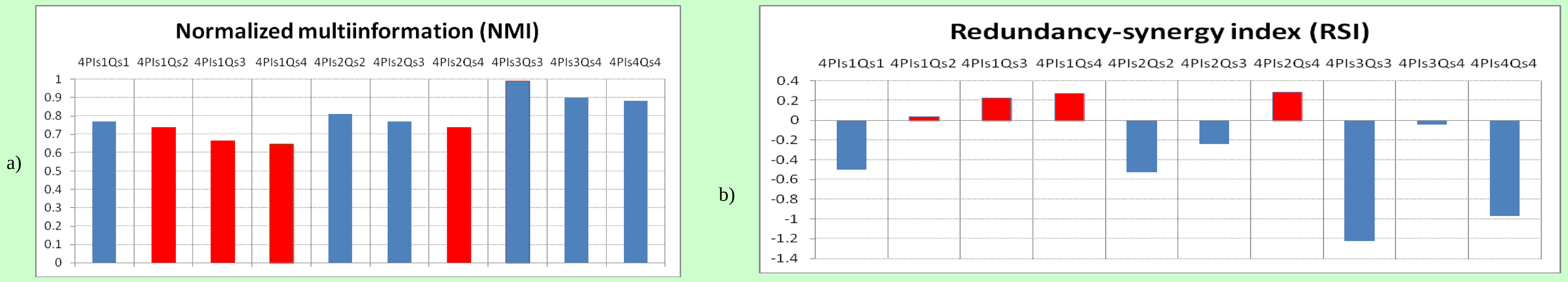

Fig. 1. a) NMI for the 10 combinations of the four Palmer indices (4PI) with the discharge at Orsova (Q), where s1, s2, s3 and s4 represent the Winter, Spring, Summer and Autumn seasons. b) the same combinations but for RSI.

METHOD. Given *n* random variables $X_1, X_2, \ldots, X_n$ the **total correlation (TC)** is defined in terms of the entropies (*H*) of n random variables as .

$$TC(X_1;X_2;\cdots;X_n) = H(X_1) + H(X_2) + \cdots + H(X_n) - H(X_1, X_2, \cdots, X_n)$$

The normalized multiinformation (**NMI**), is defined (*Ball et al., 2017*) as:

$$NMI(X_1;X_2;\cdots;X_n) = \frac{1}{n-1} TC(X_1, X_2, \cdots, X_n)$$

Another measure of multivariate information, named **redundancy-synergy index** (**RSI**), created as an extension of the interaction information (*Timme et al., 2014*) in function of mutual information (MI) to determine the predictand variable Y, from the set *S* of predictors $X1, X2, \ldots, Xn$, is used.

$$RSI(S;Y) = MI(S;Y) - \sum_{X_i \in S} MI(X_i;Y)$$

An example for the association of the predictors of the summer season with the Danube discharge in the same season after criterium with NMI *the maximum value* and *RSI the minimum value* is presented in Figs. 2-4.

The areas with significant coherence in the influence cone change when we consider partially wavelet coherence (pwc), between Q and one of the predictors, forcing (eliminating) at the same time the other three predictors(Hu and Si,2021). The results are shown in Figure 3.

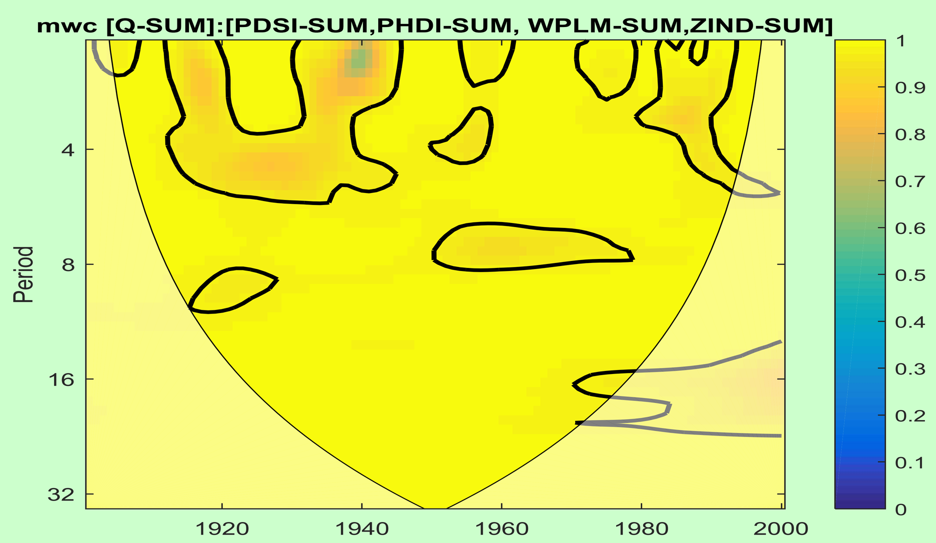

Fig. 2. Multiple wavelet coherence between dependent variable (Q) and the four independent variables (Palmer indices) for summer season during 1901-2000. (*Hu and Si, 2016*).

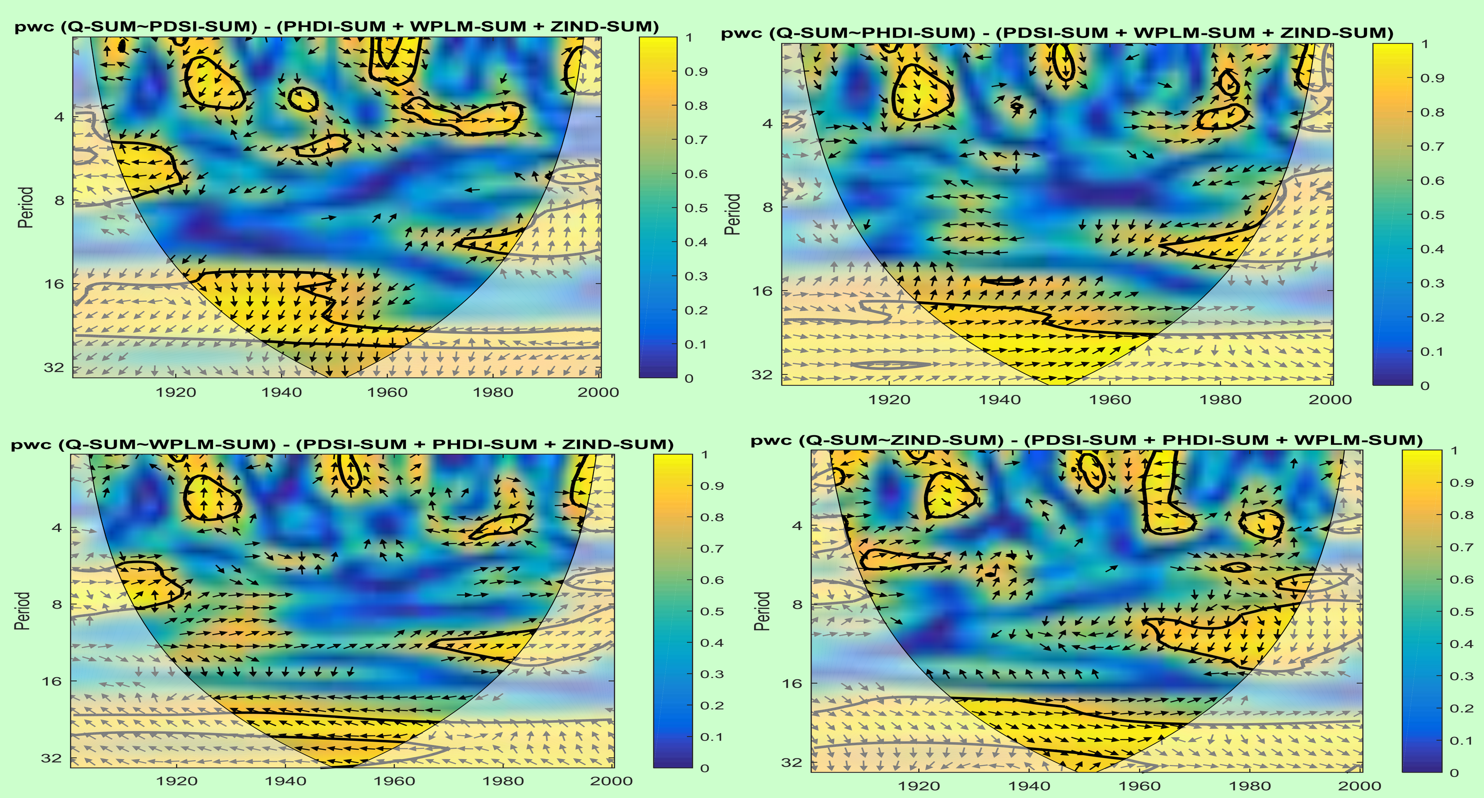

Fig. 3. Partially wavelet coherence (pwc), between Q and one of the predictors, after excluding the effect of other three predictors.

Depending on the nature of the connections (linear / nonlinear) between predictors and predictand, the prediction model is used. Also, the nature of the connections between predictand and the members of the set of predictors (independent) leads us to the hierarchy of predictors and even to the elimination of some with minor contribution. We start from the hypothesis that the connections are nonlinear (as in the most frequent cases in nature).Then the investigation of the informativeness of the connections is appropriate by **informational entropy**.

The advantage using the interaction information based on *MI* is that it does not require a certain distribution of the variables nor the nature of their relationship.

In case of non-stationary, the representation of predictors in interaction with the predictand is more appropriate by wavelet transform because the principle of overlapping interactions remains valid.

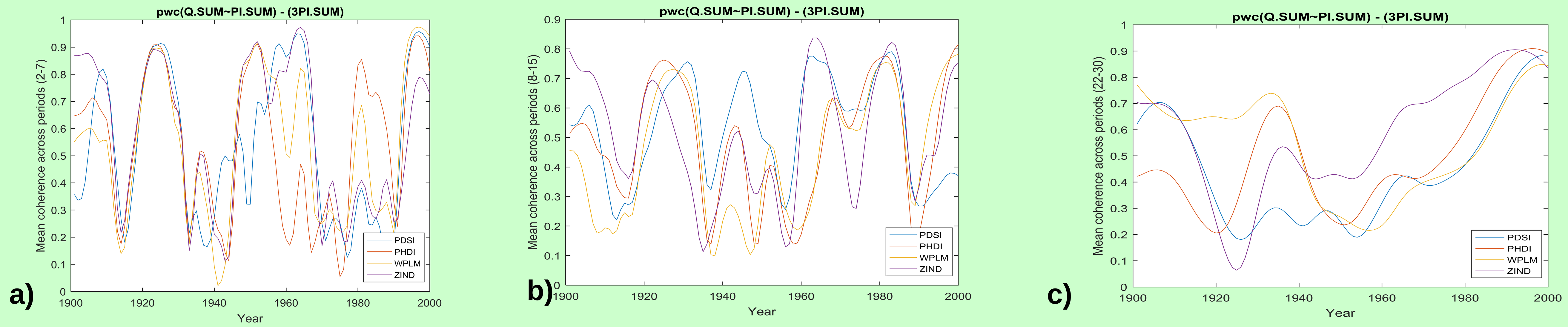

Fig. 4. Mean coherence for the frequency corresponding the periods (2-7), (8-15) and (22-30) in summer time

From the analysis of the summer-to-summer coherence (Figs 2-4), it is observed that the four combinations are not much different, so even from here it results the high redundancy between the four predictors.

Therefore, testing the influence of each of the four predictors on the Danube discharge must be done according to the frequency and time period of the twentieth century.

CONCLUSIONS

For a robust forecast it is necessary for the predictors to fulfill the conditions: multiple (synergistic contribution); non-redundant (fully informative), independent (orthogonality in the Hilbert sense), Gaussian (in the case of linear models); untangling of noisy components by wavelet filters;

Regressions must be formed with canonical varieties of predictors and predictand in nonlinear models (Hsieh, 2000).

References

Ball, K.R., Grant, C., Mundy, W.R. and Shafer, T.J., (2017). A multivariate extension of mutual information for growing neural networks. *Neural Networks*, *95*, pp.29-43.

Hsieh, W. W. (2000), Nonlinear canonical correlation analysis by neural networks, Neural Networks, 13, 1095–1105.

Hu, W., and Si B. (2016). Technical Note: Multiple wavelet coherence for untangling scale-specific and localized multivariate relationships in geosciences, *Hydrol. Earth Syst. Sci.,* 20,3183-3191.

Hu, W., and Si, B (2021). Technical Note: Improved partial wavelet coherency for understanding scale-specific and localized bivariate relationships in geosciences, *Hydrol. Earth Syst. Sci*., 25, 321-331.

Timme, N., Alford,W., Flecker, B., & Beggs, J. M. (2014). Synergy, redundancy, and multivariate information measures: An experimentalist's perspective. *J. Comput. Neurosci*. 3**6**(2), 119-140.