Master 1 Econometrie-Statistiques

January 2021

Meteorologic cycles synchronization : the case of Quasi-Biennial Oscillation, wind speeds, temperatures and depressions cycles in Singapore.

Authors :	*Directed by*:
Cécile Brissard,	Philippe De Perretti
Maylisse Tso,	*Programming in*:
Pauline Peydiere	SAS/IML

Abstract

The relationship within Quasi-Biennial Oscillation altitudes has been proved in meteorological studies. However, we have not encounter any attempt to measure relationships between Quasi-Biennial Oscillation and simple weather statements as Mean Temperatures, Precipita- tions and Wind speeds. The core of this paper is to discuss and test different theoretical assumptions, based on the numerous ways cycles can be defined and their synchronization can be assessed. We will mainly base our analysis on BC Dating method for turnings points' finding and thus for cycle retrieving. The synchronization's aim will be mostly achieve using Harding-Pagan's and Jaccard's methods. We will be using empirical data from Singapore weather stations.

Keywords : *Cycle extraction, Turning points, Jaccard's similarity coefficient, Degree of synchronization, Concordance Index, Cross-Correlation, Time delay.*

Contents

1 Introduction

The Quasi-Biennial Oscillation (QBO) has been the subject of many classic studies in meteorology. Albeit being one of the prominent events in the equatorial stratosphere, it has effects on the stratospheric flow from pole to pole, and on chemical constituents. The key aspect that will be discussed in this paper is the "coherent example of an oscillation" of the QBO [5]. As oscillations are regular periodic fluctuations, they can easily be assimilated to cycles. The QBO's cycle regularity picked our interest and with our economical background (especially the macroeconomics' models), curiosity leads us to apply the models to the meteorological field. Our first concerns were about the consequences of the QBO on weather, and to choose an adequate location. For the latter, the data were mostly accessible in Singapore, which is located on the Equator. Naturally, temperatures, precipitation levels and wind speeds were chosen as variables describing the weather.

The aim of this paper will be to know if there is synchronicity between the QBO and the weather (represented by the three variables mentioned). Other questions will be answered as well, such as if there is synchronicity within the different altitudes of the QBO. How can cycles be extracted? What is synchronization and how can it be measured? Is there synchronicity between the different variables used in the description of the weather? How do they all interact with each other?

The overall structure of this paper has been divided in nine parts. Section 2 is a thorough description of the QBO and the variables. Section 3 is focused on the origins of the databases used in our analysis. The extractions of cycle methods are described in Section 4. Section 5 explains the synchronization methods, the obtained results are discussed in Section 6. Section 7 is focused on the causality relations between the QBO and the variables. Finally, a discussion on the overall results of the paper will be made in Section 8.

2 Definitions

2.1 Quasi Biennial Oscillation

The Quasi-Biennial Oscillation is one of the main phenomenon occurring in the equatorial stratosphere [5]. It is known as one of the most regular events, with a period ranging between 22 and 34 months and an average of 28.5 months [8].
The QBO is composed of alternating phases of westerlies (wind blowing from the west) and easterlies (winds blowing from the east), which are respectively called the positive phase and the negative phase. The winds also propagate downwards, as they originate from the upper stratosphere around 5hPa and descend to the tropopause at 100hPa[1] [9].

Although there were early observations of the easterlies (Krakautau esterlies) in 1883 [19], the QBO discovery is mainly attributed to Ebdon (1960)[1] and Reed (1961)[17]. The latter used rawinsonde (data collected with balloons) at Canton Island, thus finding an *"alternate bands of easterly and westerly winds which originate above 30km and which move downward through the stratosphere at a speed of about 1km per month"*. In Ebdon's paper the period of a complete cycle is 35-37months whereas for Reed et al. it is 26 months. The term of "quasi-biennial oscillation" was later coined in 1963 by Angell and Korshover.

A cycle can be defined as events occurring repeatedly and following the same order. Then, the cyclical aspect of the QBO can easily be shown by a graphic depicting the zonal wind speed at different pressure levels through the years (Figure 1; data from Singapore). The altitudes shown on the Y-axis from 10hPa to 70hPa, but only two levels will be taken into account for further analysis [2]. The two colors demonstrate a cyclic pattern, with the alternating between the blue (easterlies) and the red (westerlies) phases.

Figure 1: Equatorial zonal winds (monthly means)

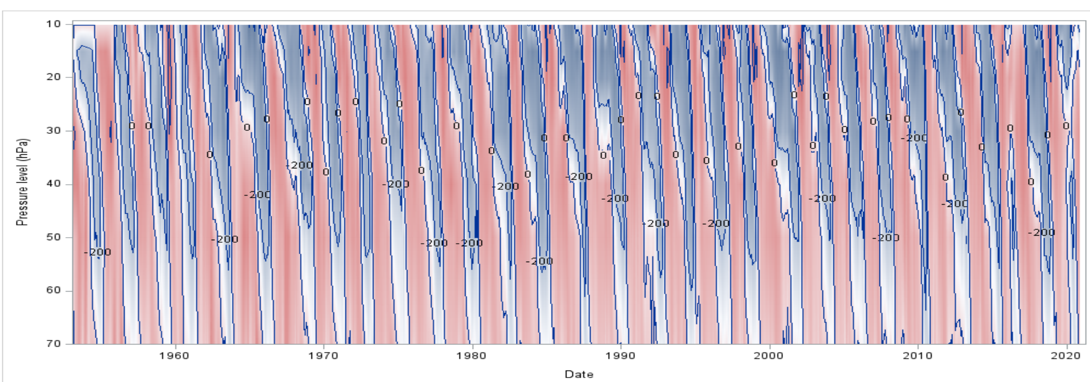


Fig. 1 Representation of the zonal winds at different pressure levels (vertical axes). The red represents the positive phases of the QBO (westerlies), the blue represents the negative phases of the QBO (easterlies). Please refer to the Appendix for a better resolution.

The cyclical pattern of the QBO was already demonstrated in many papers (Wallace et al. 1973[12], Naujokat 1986[2], Baldwin et al. 2001[5]), and will be the main focus in this paper.

[1]see Table 1.for equivalences between pressure and altitudes.
[2]see 3.1. Data Bases.

2.2 Cycles

Alongside the QBO cycle, there will be an analysis of variables (temperatures, precipitations and wind speed) measured in the troposphere in Singapore. As said before, a cycle is defined by repetitions of the same pattern in an ordered manner.

2.2.1 Temperatures

The temperature data were available from October 1954 to December 2020, with a gap between January 1966 and January 1973 (for more details on the database, refer to 3.1). Figure 2 represents the monthly temperature data since 1973. The temperature is measured in Fahrenheit (°F), with a growing tendency over the years. Which is not unexpected as the global warming issues were addressed before (*"global mean temperature rose by 0.85°C between 1850 and 2012"*, Singh et al. 2008[10]).
A cyclical pattern can be seen, with the many increasing and decreasing fluctuations of the series.

Figure 2: Monthly Mean Temperatures time series

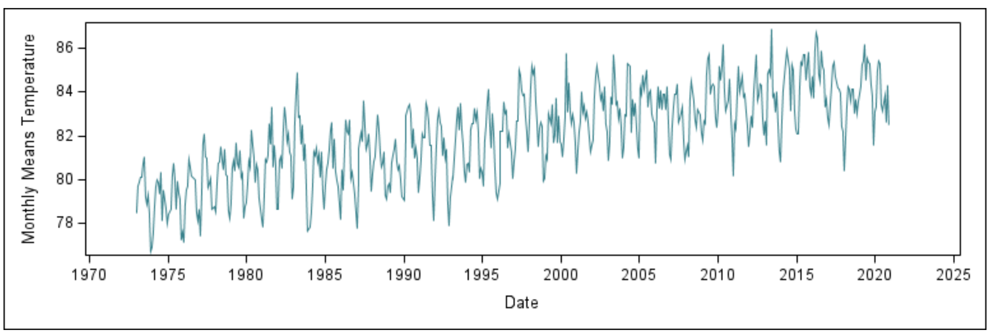


Fig. 2 Evolution of the monthly mean temperatures from 1973 to 2020, in Fahrenheit.

2.2.2 Precipitations

The precipitations were also available at the same time period (including the gap). They were measured in inches (0.01 scale). The tendency is mostly regular through the years. We noted irregularities during 1992, where higher spikes can be observed (Figure 3).

Figure 3: Monthly mean precipitations time series

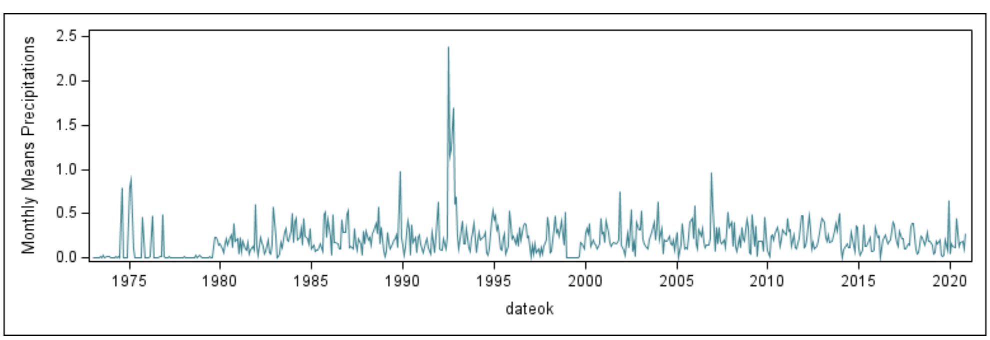


Fig. 3 Evolution of the monthly mean precipitations from 1973 to 2020, in inches (0.01 scale).

The high precipitations levels in Southeast Asia were linked to equatorial waves in a paper from Ferrett et al. (2019)[6]. They stated that *"the likelihood of extreme precipitation in SE Asia are linked to all three types of waves that are included in analysis; Kelvin, equatorial Rossby and WMRG waves"*. Their study was based on data from the period between 1998 and 2016 but the assumption of it being the case in 1993 can be made. Except for the 1992 peaks, precipitation values are following the cyclical increasing-decreasing pattern.

2.2.3 Wind-Speeds

The wind speeds were measured in knots (0.1 scale) [1knot=1,852 km/h], at the same period as the two previous variables. There again the fluctuations are showing a cyclic evolution, with more pronounced spikes from 2002 on. The wind speeds follow an increasing tendency, meaning that the winds are becoming faster. The speeding up of wind speeds is a worldwide phenomena, said Zeng et al. in their 2019 [13] report for *Nature Climate Change*. Hu et al (2020) [7] stated in their study that global warming might be one of the reasons for wind speed's increase.

Figure 4: Monthly mean wind speeds time series

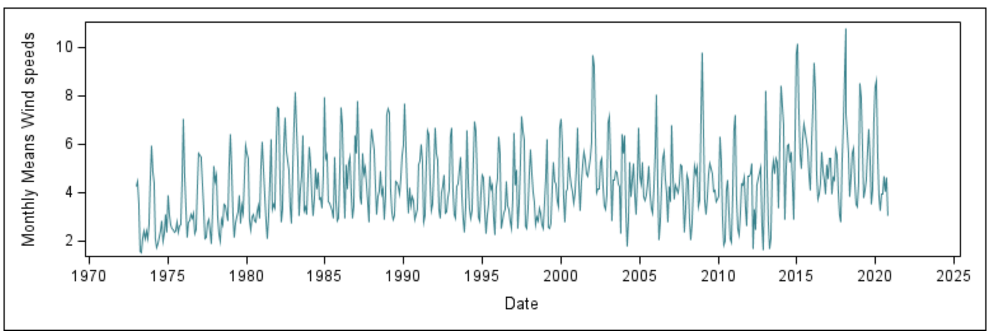


Fig. 4 Evolution of the monthly mean wind speeds from 1973 to 2020, in knots (0.1 scale).

For all three variables, the cycle length does not seem to follow the one of the QBO (28months). There seem to be shorter periods, leading to twice the amount of cycles for a same time period. A more thorough decomposition of these cycles will be made in order to have a more accurate analysis of our topic.

3 Support

3.1 Data Bases

Data Bases on which we have worked are issued from two different sources : one for the Dataset on the troposphere that regroups temperatures, precipitations and wind speeds; and the second for Dataset concerning Quasi-Biennial Oscillation (QBO). We acquired our data from the Global Surface Summary of the Day (GSOD) produced by the National Oceanic and Atmospheric Administration (NOAA)[18], a public agency responsible for the study of oceans and the atmosphere, and from the Institute of Meteorology of the Freie Universität of Berlin [3], respectively.

The NOAA data base regroups the meteorological observations from over 9000 weather stations from 1929 to now. As for the QBO database, it is a compilation of daily wind combining the observations of the three radiosonde stations Canton Island (closed 1967), Gan/Maldives Islands (closed 1975), and Singapore [near the equator] since 1953, from which were calculated the monthly mean zonal wind components for the levels 70, 50, 40, 30, 20, 15, and 10 hPa. This Dataset regroups a total of 7 variables and 815 observations.

Table 1: Equivalence between Pressure Boundaries (hPa) and Approximate Altitudes (km)

Pressure Boundaries (hPa)	**Approximate Altitude at the Bottom (km)**
1013.25 - 63.93	0
63.93 - 40.33	19
40.33 - 25.45	22
25.45 - 16.06	25
16.06 - 10.13	28
10.13 - 6.39	31

Table 1. Referenced in Tourpali et al. (2007) The 11-year solar cycle in stratospheric ozone: Comparison between Umkehr and SBUVv8 and effects on surface erythemal irradiance. Journal of Geophysical Research-Atmospheres.[11]

Since the Quasi-Biennial Oscillation was measured from Canton Island, Gan/Maldives Islands (Gan Island and Paya Lebar) and Singapore weather stations, we choose these stations among the 9000 available in the NOAA dataset. The Singapore and Paya Lebar's stations where the ones with fewer missings values, a sufficient reason to select them ; observations started on the 1^{st} Nov. 1954.

3.2 The sample

From the initial Data Bases, we merged Singapore Changi and QBO observations, which is why the Date variable of Singapore Changi Data was converted from daily observations to mensualized means. For the "$1013hPa/0km\ Datas$", there were a problem of missing values between the ensuing two gaps of time : from Jan. 1966 to Jan. 1973, and between Oct. 1981 and Jan. 1998. For the latter, we used the Data from Paya Lebar's weather station to fill the missing period since Paya Lebar and Singapore Changi both are districts in Singapore city-state.
We dropped all variables except for "$TEMP$", "$PRCP$" and "$WDSP$" (respectively mean temperatures, precipitations and wind speeds) essentially because other ones were indexes or

had too much missing values (we would have work on pressures otherwise). We also saved the name of the station, along with its position (latitude, longitude, elevation) and its station's reference number. Finally, mean temperatures were measured in 0,1 Fahrenheit, precipitations in 0,01 inches, and wind-speeds in 0,1 knots ($0,1knots \approx 0,514m/s^{-1}$).

Concerning the QBO Dataset, we chose to keep the wind speeds in m/s^{-1} at 20 and 50 hPa levels, based on variance-covariance matrix, extracted with Jaccard's method[4], values bellow[3].

	Expansions									
TEMP	1	0.2623762	0.1275862	0.3455497	0.3498542	0.2717391	0.201087	0.1901042	0.218509	0.1751825
PRCP	0.2623762	1	0.1193548	0.2805755	0.2044888	0.2079208	0.1937984	0.2552632	0.3	0.3594595
WDSP	0.1275862	0.1193548	1	0.166113	0.1278195	0.1213235	0.0898438	0.0808824	0.0912281	0.122807
70hPa	0.3455497	0.2805755	0.166113	1	0.4712991	0.3479452	0.1506173	0.1534772	0.1448276	0.1367713
50hPa	0.3498542	0.2044888	0.1278195	0.4712991	1	0.510274	0.2848297	0.2011173	0.161039	0.1287129
40hPa	0.2717391	0.2079208	0.1213235	0.3479452	0.510274	1	0.4141414	0.2946429	0.2117962	0.1553885
30hPa	0.201087	0.1937984	0.0898438	0.1506173	0.2848297	0.4141414	1	0.471223	0.352381	0.2869822
20hPa	0.1901042	0.2552632	0.0808824	0.1534772	0.2011173	0.2946429	0.471223	1	0.5583039	0.4018692
15hPa	0.218509	0.3	0.0912281	0.1448276	0.161039	0.2117962	0.352381	0.5583039	1	0.5884354
10hPa	0.1751825	0.3594595	0.122807	0.1367713	0.1287129	0.1553885	0.2869822	0.4018692	0.5884354	1

	Recessions									
TEMP	1	0.3646055	0.5297398	0.4356659	0.5098901	0.4357895	0.4131737	0.3804781	0.3795918	0.3260437
PRCP	0.3646055	1	0.4925651	0.3449782	0.3529412	0.3482688	0.376	0.4079498	0.4229935	0.4638009
WDSP	0.5297398	0.4925651	1	0.5219048	0.5711645	0.5590406	0.5778986	0.5479204	0.5282332	0.537037
70hPa	0.4356659	0.3449782	0.5219048	1	0.5823389	0.46875	0.3307393	0.3091977	0.2734375	0.2509728
50hPa	0.5098901	0.3529412	0.5711645	0.5823389	1	0.6643192	0.5217391	0.4314115	0.3703704	0.3269598
40hPa	0.4357895	0.3482688	0.5590406	0.46875	0.6643192	1	0.6150442	0.5021008	0.4072581	0.3430799
30hPa	0.4131737	0.376	0.5778986	0.3307393	0.5217391	0.6150442	1	0.6689189	0.5603448	0.4958159
20hPa	0.3804781	0.4079498	0.5479204	0.3091977	0.4314115	0.5021008	0.6689189	1	0.7002398	0.5695067
15hPa	0.3795918	0.4229935	0.5282332	0.2734375	0.3703704	0.4072581	0.5603448	0.7002398	1	0.699005
10hPa	0.3260437	0.4638009	0.537037	0.2509728	0.3269598	0.3430799	0.4958159	0.5695067	0.699005	1

You shall refer to the SAS result of Jaccard Synchronization as :

- Pale blue expresses the correlation between wind-speeds and precipitations;
- Purple " between wind-speeds and temperature;
- Purple-blue " between precipitations and temperature;
- Yellow " between wind-speeds at 50hPa altitude and, from left to right, temperatures, precipitations, wind-speeds;
- Green " between wind-speeds at 20hPa altitude and, from left to right, temperatures, precipitations, wind-speeds;
- Red for the correlation between wind-speeds at 20hPa and 50hPa altitudes.

[3] see 5.1 Jaccard.

The first matrix (*Expansions*) gives the within-observation correlation for expansions, as the second one (*Recessions*) gives the within-observation correlation for recessions. We picked two levels of altitude because we want to see the correlation between a high altitude level, which is what 20hPa (28km) stands for, and a lower level, i.e. 50hPa, that is approximately 20 kilometers above the surface of the Earth. Although, we choose those levels because it seems to be the better trade-off to maximize the correlation within the values of the Quasi-Biennial Oscillation, and between these observations and the temperatures, precipitations and wind speeds at the 1013,45 hPa (0km) altitude. We are particularly interested in the framed values.

It results that our sample is monthly means from the Jan. 1973 until Nov. 2020 of 5 variables and 574 observations.

4 Cycle Extraction

4.1 Methods

4.1.1 Hodrick-Prescott filter

The aim of the Hodrick-Prescott filter is to fit smooth trends to time series, by removing the cyclical component of a time series from raw data. This filter originated from E.T. Whittaker in 1923. It was popularised mostly for its macroeconomics use by Robert Hodrick and Edward Prescott in the 1990s[16]. A time series y_t shall be decomposed as follow :

$$y_t = \tau_t + c_t + \epsilon_t \tag{1}$$

with τ_t a trend component, c_t a cyclical component and ϵ_t an error component. The cycle, denoted c_t was defined by Hodrick and Prescott as a stationary component, whereas the trend τ_t was considered as a non-stationary component. The purpose of this method is to use the minimization of penalized least square criterion, so that a trend can emerge from the time series. To achieve this HP filter is construct to withdraw the cycle component c_t from the series y_t, as the ensuing optimization problem :

$$\min_{\tau} \left(\sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right) \tag{2}$$

However, the trend stays sensitive to short-term fluctuations. The λ multiplier in the above equation, can be modified to adjust the trend according to this sensitivity. This λ parameter can take any value, but the Hodrick and Prescott's consensus is to set a value of 1600 for quarterly data. Later, Ravn and Uhlig's suggested that λ should take the value of 6.25 $(1600/4^4)$ for annual data and the value of 14400 for monthly data, which is our case.

To describe more rigorously the reasons for adding the λ parameter in HP's filter, we can say, in the first hand, that it helps to detrend the series, as it minimize the distance between the original series in the right part and its trend ; and, in the other hand, that it helps to smooth the series. λ is said to be a penalty on the trend : the bigger λ is, the smoother the series will be, until it takes the form of a linear deterministic time trend. The aim of this parameter is to take the trend apart from the series, keeping the explanatory power of the cycle. In other words, without a λ the trend stays contigous to the time series and we will not be able to detach a cycle from it, but with too much of a λ we may reduce the cycle's fluctuations, since these ones would, therefore, be incorporated in the trend.

At first, in order to choose the right value for the HP's filter and to make sure it is the best filter to apply in our time series, we need to understand the properties of the trend. Let τ_t define the trend and μ_t be its disturbance, as σ_t will be the slope of the trend and ϕ_t its disturbance such that,

$$\tau_t = \tau_{t-1} + \sigma_{t-1} + \mu_t \tag{3}$$

$$\sigma_t = \sigma_{t-1} + \phi_t \Leftrightarrow \sigma_{t-1} = \sigma_t - \phi_t \tag{4}$$

It implies that the circumstantial cycles will be obtained after the subtraction of the filtered trend from the original series. For the interpretations, we will look at the trends and cycles resulting from procedures on SAS software. In the SAS software we are using, the trend is always referred as the "*Level*" because the slope of a series and its level are the two components of the trend. To obtain extracted trends and cycles, we used three different procedures :

- The PROC UCM allows us to get the Smoothed Residuals and the smoothed slope and levels which are the two components of the trend. Its process is based on Kalman filtering, on which we applied $1/14400 = 6.94e-5$ i.e. the reverse of the HP's filter we have seen above. Our point here is mainly to get the Residuals out of the series for Temperatures, Precipitations and Wind-speeds. In fact, we don't need to extract any cycles from the QBO observations because we already have determinists cycles : a trend for those would only be a passing by 0 line. For theoretical questions, UCM stands for unobserved components model or, structural time series models in the time series literature. This function decomposes the series into components such as trend, seasonals, cycles, and the regression effects due to predictor series. We did not take interest in forecasted or filtered components of this Procedure but only in the smoothed ones. A UCM can be seen as special cases of the linear Gaussian state space models (GSSM), which can be explain with the following equations :

$$y_t = Z_t \alpha t \tag{5}$$

$$\alpha_{t+1} = T_t \alpha_t + \gamma_{t+1}, \qquad \gamma_t \sim \mathrm{N}(0, Q_t) \tag{6}$$

The first equation, called the *observation equation*, relates the state vector, denoted α_t, to the response series y_t. This vector is usually unobserved and contains the component estimates and their variance. The second equation contains the state vector and is referred as the *state transition*
equation. This last equation provides a rule state for getting the model and predictions from one interval to the next, i.e. it describes the evolution of α_t in time.

Let ψ_t, ρ, λ, μ_t^* and μt, respectively be the cycle, a damping coefficient, the frequency and the zero means. Since we want an extracted cycle from series, we can report to the definition of the stochastic cycle exposed in the SAS User's Guide for PROC UCM, that is shown bellow :

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \mu_t \\ \mu_t^* \end{bmatrix} \qquad with \begin{cases} \lambda > 0 \\ \mu t \perp\!\!\!\perp \mu_t^* \end{cases} \tag{7}$$

In those conditions, here are the defined parameters and matrix from the (5) and (6) equations :

$$Z = [\,1\ 1\ 0\,] \tag{8}$$

$$T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & -\rho\sin\lambda & \rho\cos\lambda \end{bmatrix} \tag{9}$$

$$Q = Diag\left[\sigma_\epsilon^2, \sigma_\mu^2, \sigma_\mu^2\right] \tag{10}$$

$$With \begin{cases} \alpha_t = [\,\epsilon_t\ \psi_t\ \psi_t^*\,]^{'} & \text{the state vector} \\ \gamma_t = [\,\epsilon_t\ \nu_t\ \nu_t^*\,]^{'} & \text{the state noise vector} \end{cases}$$

To smooth the model, the Kalman's filter passes forward through the serie and then adds a backward pass through it.

- The PROC EXPAND has the same purpose than the PROC UCM ; we used it to compare the results of both processes in SAS. This procedure sorts a trend and a cycle for each time series using the Hodrick-Prescott's filter for monthly time series, with the value of 14400, by inverting a suitable matrix. This value suits for constructing all trends, except for wind speeds at 50hPa and 20hPa, were we choose the arbitrary value of 219000, by

trial and error, to obtain a more visually smoothed trend.
Statistically, PROC EXPAND identifies the trend and cyclical components, respectively, as;

$$\hat{\tau_T} = \frac{1}{(\lambda \times_T +_T) \times y_T} \tag{11}$$

$$\hat{c_T} = y_T - \hat{\tau_T} \ , \tag{12}$$

that we obtained after taking derivatives of the minimization function[4].

→ After comparisons, and as we will develop in the Results[5], it appears that extracting a cycle from QBO series is inefficient, since the series already form cycles.

- The PROC MODEL[6] is the final process we will use to extract a White Noise, whose we can define as the errors of the residuals, from the series thanks to the cycles (residuals) we obtained from PROC UCM statement ; and to extract the Residuals from the QBO series since they were already cyclical. To achieve this, we used and Auto-regressive Model of the second form [AR(2)].

4.1.2 Autoregressive Model

To describe certain time-varying processes in nature and economics, the autoregressive model states as a reference in statistical, econometrics and signal process analysis. An AR(p) expresses a variable as a function of its own previous value(s), denoted Y_{t-i}, with $i = 1, \ldots, p$ and a stochastic (= imperfectly predictable, also known as white noise) term, denoted ε_t. Let φ_i be parameters of the model and a the constant seen bellow :

$$Y_t = a + \sum_{i=1}^{p} \varphi_i Y_{t-i} + \varepsilon_t \tag{13}$$

For our concern, we will not take into account the stochastic term, since ε_t's are independent and identically distributed and have an expected value of 0. That is the search result of the MODEL procedure, applied with following AR(2) model :

$$Y_t = a + b \times Y_{t-1} + c \times Y_{t-2} = a + b \times lag(Y_t) + c \times lag2(Y_t), \quad where \ \varphi_i = \begin{bmatrix} b \\ c \end{bmatrix} \tag{14}$$

Let μ_t be the white noise, we want to obtain from the extracted cycles (residuals), we shall have :

$$\phi_t = a + b \times lag(\phi_t) + c \times lag2(\phi_t) + \mu_t \Leftrightarrow \mu_t = \phi_t - c \times lag2(\phi_t) - b \times lag(\phi_t) - a \tag{15}$$

It is convenient to introduce such a component, because it may outlines the seasonality effect, if observations in the series have period smaller than years. Since we are working on monthly means, this AR(2) method shall fit our purposes. When seasonal fluctuation appear to be stable, it is said that the seasonal component will be either deterministic or fixed.

[4]see equation (2)
[5]see 4.2 Results and consequences
[6]see 4.1.2 Auto-regressive Model

4.2 Results and consequences

First, we compare the two Hodrick-Prescott procedures (UCM method and EXPAND method).

Figure 5: Comparison of differents cycles extraction methods

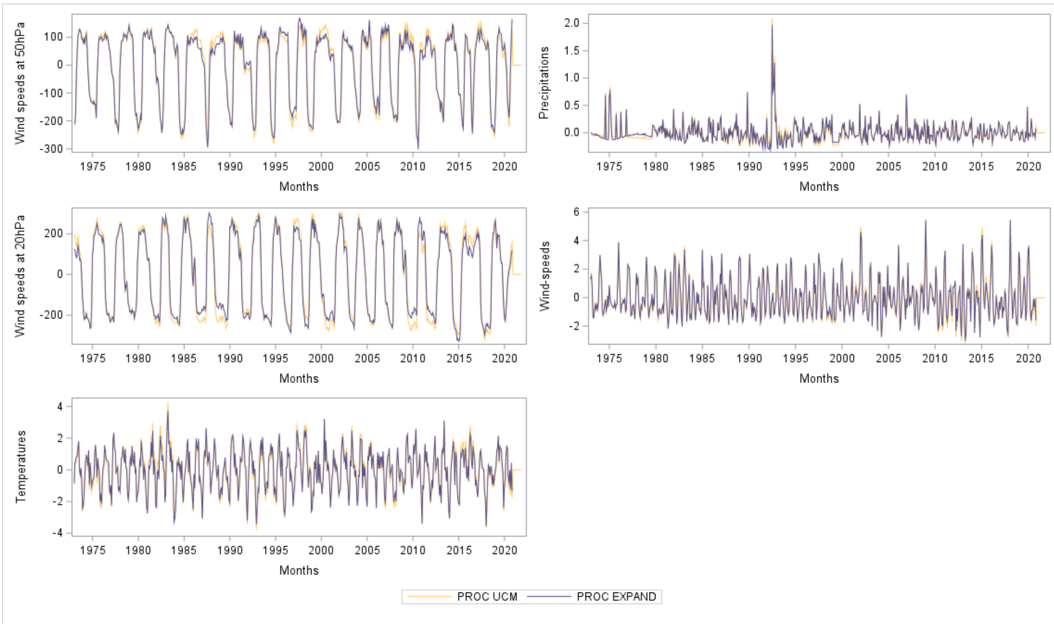


Fig. 5 Representation of the comparison between different cycles extraction methods : UCM and EXPAND's procedures. UCM procedure is represented in red and PROC EXPAND in purple.

In Figure 5, we can see that the two procedures give us almost the same curves, thus the same cycles. The little differences are not significant. It is noted that wind speeds at 50hPa's and 20hPa's cycles were not different than their series.

Therefore, when looking at the graph representing series, cycles and trends obtained by the expand procedures (Figure 6), for the variables temperatures, precipitations and wind-speeds, extracted cycles were more relevant. Note that for the temperatures' series and cycle, two graphs were needed due to the difference of scale.

Figure 6: Cycles, Trends and Series w/ PROC EXPAND

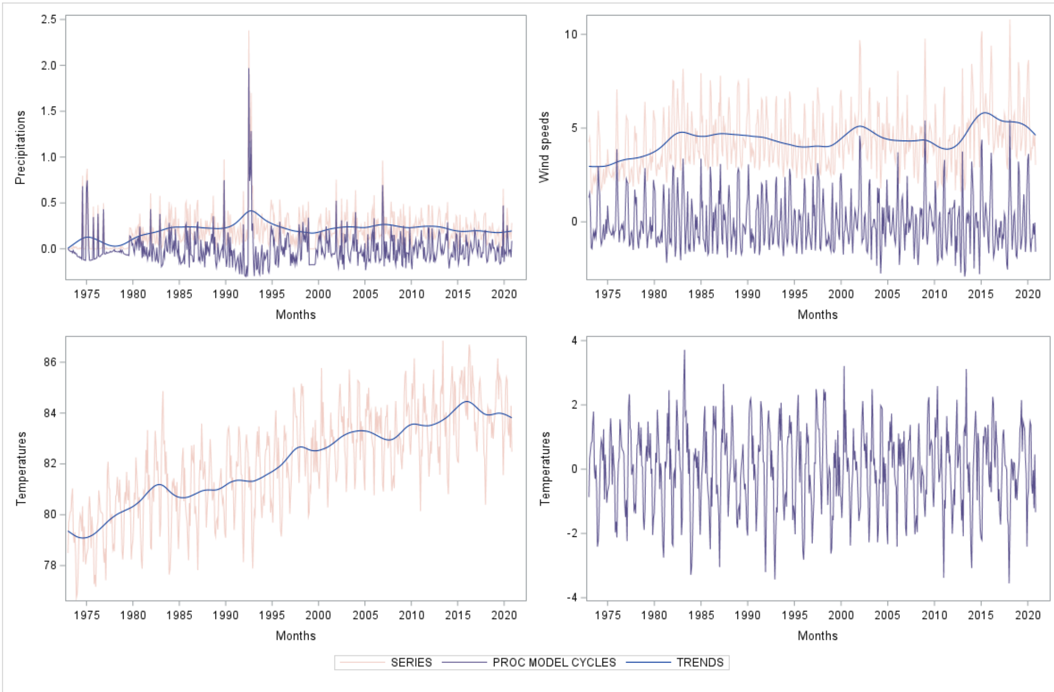


Fig. 6 Representation of series, cycles and trends issued of EXPAND method for Precipitations, Wind speeds and Temperatures. For scale concerns the cycle of Temperatures has been made appart from its Trend and Serie. Series are represented in orange, cycles in purple and trends in black.

In contrast, for the 50hPa and 20hPa, cycles are not different to their series. Then, we concluded that it is not necessary to apply UCM and EXPAND procedures on these series.

Figure 7: Cycles, Trends and Series w/ PROC EXPAND

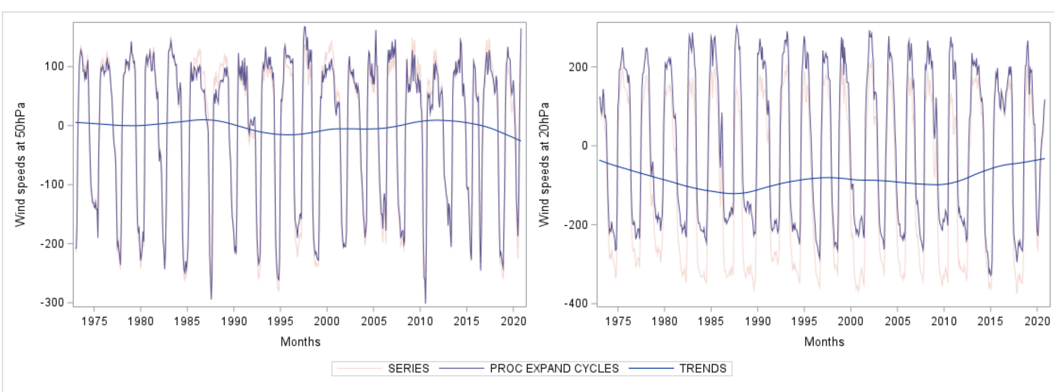


Fig. 7 Representation of series, cycles and trends issued of EXPAND method for 50 and 20 hPa. Series are represented in orange, cycles in purple and trends in black.

From now on the wind speeds at different altitude's series will thus be considered as their cycles ; whether for others variables we will keep the PROC UCM's extracted residuals as cycles. That's on these cycle, we choose to apply an auto-regressive model with the MODEL procedure. Let the comparison between cycles and white noises we just extracted, be represented in the following graphs.

Figure 8: Visualisation of Trends, cycles and series

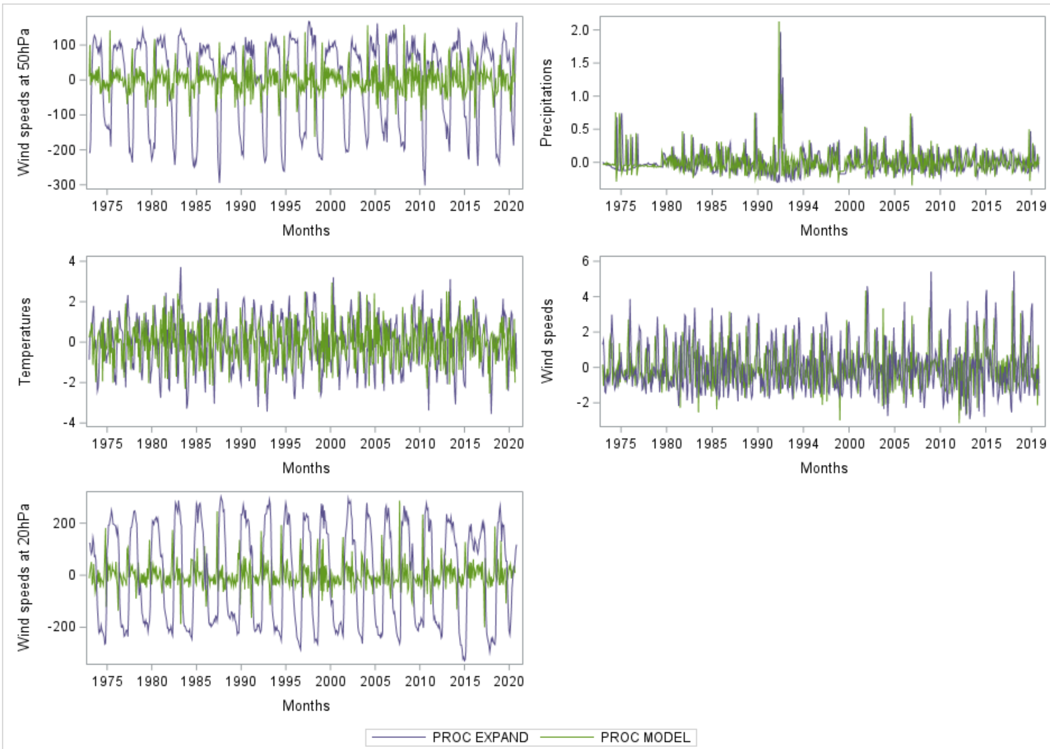


Fig. 8 Representation of the difference between cycles extracted with PROC EXPAND and the ones with PROC MODEL, where the PROC EXPAND's statements are in purple and PROC MODEL's ones in green.

To conclude, the cycles's extractions with the EXPAND and UCM's procedures are different from the one with the MODEL procedure for both the 20hPa and 50hPa series (Figure 8). As the results of the PROC EXPAND and PROC UCM were not satisfactory for the wind speeds series at 50hPa and at 20hPa, we applied the PROC MODEL on the original series. We have therefore obtained the residuals with the PROC MODEL for these two series.
On the contrary, for temperatures, precipitations and wind speeds we used the extracted cycles to make the PROC MODEL statement. *De facto*, we say we obtained the residuals' residuals that we call white noises.

After taking a deeper look into the results, we effectively get a statistic that approaches white noises characteristics for Temperature, and also for the wind speeds at 20 and 50 hPa. As the form of the series is already in a cycle format, it can explain this result. Unfortunately, for the wind speeds the error of the residuals is not properly distributed as a standard normal distribution. Finally, the precipitations' white noises did not have the expected form. It could be suggested that it is due to the summer 1992's spikes, which biased the series, cycles, and white noises. According to the auto-regressive model method's efficiency (we have results for most of the variables), further analysis will be based on this model. Particularly for cross-correlations, which will give us the time delay of transmission from one cycle to another and may allow us to determine causalities relations.

5 Cycles Synchronization Method

5.1 Jaccard

In order to gauge similarities between the different cycles we accorded on the convenience of Jaccard's index, also known as the "$Jaccard\ similarity\ coefficient$". That last concept is a statistical measure, introduced by Paul Jaccard in 1901, of the size of the intersection of sample sets, divided by the size of the union of the sample sets. Let A and B, be two samples sets :

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

This index can be extended for multi-variables measures, with S_N the named variable (with $N = 1, \ldots, n$), the J's coefficient is transformed as follow :

$$J(S_1, S_2, \ldots, S_n) = \frac{|S_1 \cap S_2 \cap \ldots \cap S_n|}{|S_1 \cup S_2 \cup \ldots \cup S_n|}$$

By construction, we obtain a value between 0 and 1. A result of 1 meaning complete synchronization between the sample sets and, *a contrario*, 0 is set when there is no intersection at all between the variables.

The Jaccard's coefficient[4] is mainly used in computer science, ecology and genomics, is applied on binary variable ; a further reason why it is a convenient measure to determine whether there is a synchronization between our cycles or not. If the aim was to synchronize the series and not their cycles, we would have choose the above equation (n). However, we want to know if there is within expansions or expansion with positive phase synchronization. Here, each S_n represents a vector of 0 and 1, we associated to our variables : Temperatures, Precipitations, Wind-speeds, and Wind-speeds at 20hPa and 50hPa. To achieve this result we used BCDating package on the software R[7]. Then, the data obtained on R were extracted and recoded in SAS. Two analysis were made, first for all sample sets, were the values 1 and 0 were kept for expansions and recessions (respectively). A second one with 1 and 0 as the positive and negative phases for the QBO features (from 70hPA to 10hPa), that is due to the very "deterministic" form of the series (expansions and recessions were kept for the other variables). Doing so, we first ought to determine which altitudes we wanted to synchronized within QBO and with the other variables[8]. Once we did we took interest in synchronization within expansions and between expansions and positives phases for the five variables we picked. Which requires the ensuing equation's understanding :

$$J = \frac{M_{11}}{M_{01} + M_{10} + M_{11}} \quad with, \begin{cases} M_{11} & \text{representing the total number of attributes where} \\ & \text{at list two different } S_N \text{ both have a value of 1,} \\ M_{01} & \text{as the total number of attributes where the attribute} \\ & \text{of say } S_1 \text{ is 0 and the attribute of } S_2 \text{ is 1,} \\ M_{10} & \text{for the total number of attributes where the attribute} \\ & \text{of, say } S_1 \text{ is 1 and the attribute of } S_2 \text{ is 0,} \\ M_{00} & \text{representing the total number of attributes where} \\ & \text{at list two different } S_N \text{ both have a value of 0.} \end{cases} \tag{16}$$

[7]see 6 Results of Cycles Synchronization
[8]see 3.2 The sample

5.2 Harding-Pagan

Another cycle synchronization method is the one developed by D. Harding and A. Pagan in 2006[15]. It is believed that the Harding Pagan method is more accurate than Jaccard's, since they take in account broader possible outcomes, resulting in more precise indicators. Before assessing the synchronization between the cycle, it is important to be able to identify them. They used the turning points method, the peaks and troughs helping them to define cycles.

5.2.1 Degree of synchronization

The turning points method allowed them to have their variable of interest Y in a binary variable form, called S and taking 0 (contraction) and 1 (expansion) as values. Thus, to compare the synchronicity between two cycles, say S_x and S_y, they identify three cases. First, the *Strong Perfect Positive Synchronization* (SPPS), when the two variables S_x and S_y are identical. That means that the probability of having two 0 or two 1 for both variables is unity. In other words, the probability of having a 0 for a variable and 1 for the other is zero. The second case they stated is the *Strong Perfect Negative Synchronization* (SPNS), which is the opposite of the first statement. The probability of having two 0 or two 1 for both variables is zero; or, the probability of having a 0 for a variable and 1 for the other is unity (and take the value of -1). The last case is the *Strongly Non-Synchronized* (SNS) one, where S_x and S_y can be regarded as being independent.

They also stated that in some situations, SPPS, SPNS and SNS might be rejected as *"the series are synchronized but not perfectly so"*. Based on these assumptions, they were able to measure synchronization between two cycles, by investigating a component that they called ρ_S (17).

$$\rho_S = \frac{\mathbb{P}\left(S_{xt}=1, S_{yt}=1\right) - \left[\mathbb{P}(S_{xt}=1)\mathbb{P}(S_{yt}=1)\right]}{\sqrt{\mathbb{P}(S_{xt}=1)\mathbb{P}(S_{xt}=0)}\sqrt{\mathbb{P}(S_{yt}=1)\mathbb{P}(S_{yt}=0)}} \tag{17}$$

The value given for ρ_S will be either 0 when there is no synchronization, or 1 when there is perfect synchronization.

5.2.2 Concordance index

The two authors (Harding and Pagan) came up in 2002[14] with the concordance index, another index to evaluate the synchronization degree between two cycles. The concordance index gives which *"fraction of time the cycles are in the same phase"*. The phases they are referred to are the expansions ($S = 1$) and contractions ($S = 0$) of the series. In other words, the concordance index is a tool to measure of the "co-movement" between two cycles. You shall find bellow the general form of the index given by Harding and Pagan in their 2002 paper (18).

$$\widehat{I} = \frac{1}{T}\left\{\sum_{t=1}^{T} S_{xt}S_{yt} + \sum_{t=1}^{T}\left(1-S_{xt}\right)\left(1-S_{yt}\right)\right\} \tag{18}$$

This form presents T as the number of periods and S_{it} the two variables extracted from the variable of interest i (x and y) that we want to compare. A definition of this index woulb be : the fraction of time where both variables are at the same phase, in the same cyclical phase. In fact, $\frac{1}{T} \times \sum_{t=1}^{T} S_{xt}S_{yt}$ represents the expectation of having the value of 1, meaning expansion, for both S_{xt} and S_{yt} ; and, analogically, the second part of the equation expresses the expectation of 0, i.e. recession for both. This equation was later modified in their 2006 paper, resulting to :

$$\widehat{I} = 1 + 2\widehat{\rho}_S \left[\widehat{\mu}_{S_x}(1-\widehat{\mu}_{S_x})\widehat{\mu}_{S_y}(1-\widehat{\mu}_{S_y})\right]^{\frac{1}{2}} + 2\widehat{\mu}_{S_x}\widehat{\mu}_{S_y} - \widehat{\mu}_{S_x} - \widehat{\mu}_{S_x} \tag{19}$$

The modified version (19) is computed with $\widehat{\mu}_{S_i}$, the estimated mean of the series i, and $\widehat{\rho}_S$, the estimated correlation coefficient between S_x and S_y. Its value lies between 1 (when $S_x = S_y$) and 0 ($S_{xt} = 1 - S_{yt}$). The closer $\widehat{I}$ is to one, the higher the fraction of time the two cycles are in the same phase. It was demonstrated that when $\widehat{I} = 1$, the value of $\widehat{\rho} = 1$ and when $\widehat{I} = 0$, $\widehat{\rho} = -1$.

6 Results of Cycles Synchronization

We used the methods of Jaccard and Harding Pagan to carry out our synchronization analysis.
Before using the Jaccard and Harding Pagan indexes on our series, we had to modify the form of variables of interest. The BCDating package was used, as it permits to find the turning points with an algorithm.

6.1 Bivarian cycles : BC Dating statement (R)

The principles of selecting turning points in cycles were developed by Bry and Boshan 1971. It is used to define swings in the business cycles. Two requirements have to be meet when applying this method. First, the phases should last between fifteen months and twelve years. The second requirement is that the "amplitude of specific cycles have to be larger than those of irregular fluctuations encountered in the series".

The turnings points will be the peaks and troughs of the series. Bry and Boshan stated that generally peaks are localised at the highest points of a cycles, whereas the trough are at the lowest points of the cyclical fluctuations. Also, the peaks and troughs have to alternate.

The turning points delimited the phases, thus creating two binary variables composed of 1 for expansions and 0 for recessions. You shall find below the a representation of the phases resulted from the turning points method (Figure 9 and Figure 10).

Figure 9: Phases of Expansions and Recessions for the 50 and 20 hPa series.

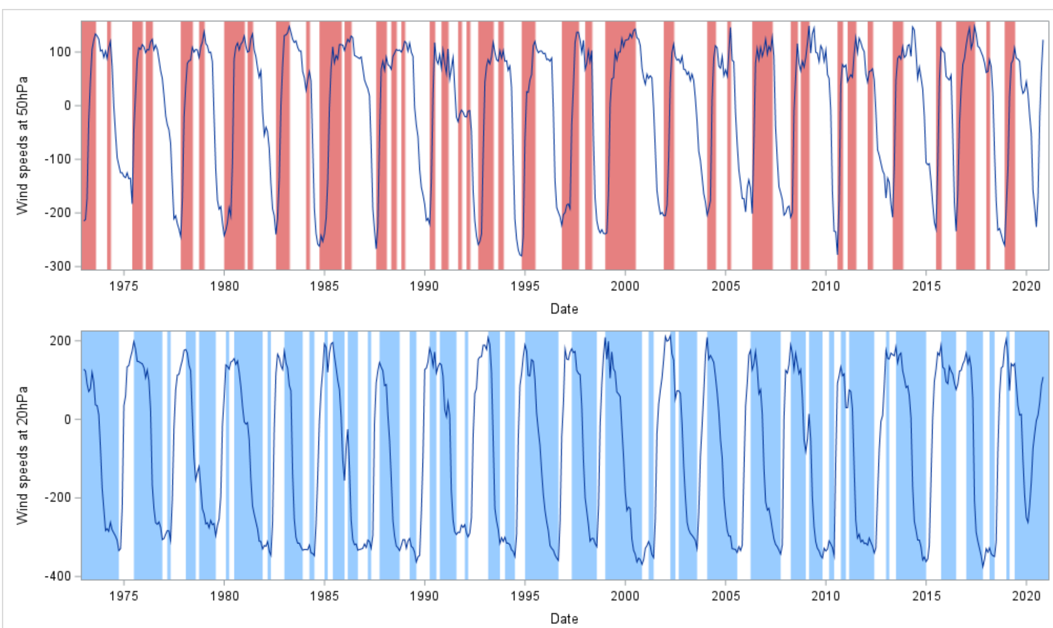


Fig. 9 The upper graphic represents the monthly mean time series (MMTS) of Wind speeds at 50hPa ; the bellow one for the MMSS of Wind speeds at 20hPa, all in black. The red and white, blue and white, represent the expansions and recession phases, for each of them, respectively. Where expansions are colored for 50hPa, whether in white for 20hPa.

Figure 10: Phases of Expansions and Recessions for Temperatures, Precipitations and Wind speeds series.

Fig. 10 The upper graphic represents the monthly mean time series (MMTS) of Temperatures ; the centered one stands for MMTS of Precipitations and the bellow one for the MMSS of Wind speeds, all in black. The green and white, light green and white, yellow and white color areas, represent the expansions and recession phases, for each of them, respectively. Where expansions are colored for Temperatures and Windspeeds, whether in white for Precipitations.

6.2 Jaccard's synchronization coefficient

The first method was to synchronize the cycles using Jaccard's coefficient. The Jaccard tests allows us to see if when a 1 (expansion) arrives in one series, another 1 arrives simultaneously in the other series.

It should be reminded that two analysis were made. First, were the values 1 and 0 were kept for expansions and recessions (respectively) for the weather variables, and wind speeds at 50hPa and at 20hPa were coded according to positive and negative values. The positive values, coded as 1, represent winds from the west, which are warm winds, and the negative values, coded as 0, represent winds from the east, which are cold winds (Table 2).

Table 2: Jaccard Similarity Coefficients - Expansions and Positive Phases

	TEMP	PRCP	WDSP	50hPa	20hPa
TEMP	1	0.2623762	0.1275862	0.3650794	0.3134715
PRCP	0.2623762	1	0.1193548	0.3761062	0.2284382
WDSP	0.1275862	0.1193548	1	0.1432292	0.1428571
50hPa	0.3650794	0.3761062	0.1432292	1	0.3170213
20hPa	0.3134715	0.2284382	0.1428571	0.3170213	1

Table 2. Jaccard's similarity coefficients between side-to-side Expansions phases for Temperatures, Precipitations, Wind speeds, 50 and 20 hPa, where royal blue is TEMP/PRCP's coefficient ; violet, TEMP/WDSP's coef. ; pale blue, PRCP/WDSP's coef. ; shading from light yellow to gold, 50hPa/TEMP-PRCP-WDSP's coef. ; shading from green to light green, 20hPa/TEMP-PRCP-WDSP's coef. ; and red, 50hPA/20hPa's coefficient.

The second analysis (Table 3) is the one where all expansions were coded 1 and all recessions were coded 0, for both the variables and the pressure levels (50hPa ans 20hPa).

Table 3: Jaccard Similarity Coefficients - Expansions

	TEMP	PRCP	WDSP	50hPa	20hPa
TEMP	1	0.2623762	0.1275862	0.3498542	0.1901042
PRCP	0.2623762	1	0.1193548	0.2044888	0.2552632
WDSP	0.1275862	0.1193548	1	0.1278195	0.0808824
50hPa	0.3498542	0.2044888	0.1278195	1	0.2011173
20hPa	0.1901042	0.2552632	0.0808824	0.2011173	1

Table 3. Jaccard's similarity coefficients between side-to-side Expansions phases for Temperatures, Precipitations, Wind speeds, and Positive Phases for 50 and 20 hPa, where royal blue is TEMP/PRCP's coefficient ; violet, TEMP/WDSP's coef. ; pale blue, PRCP/WDSP's coef. ; shading from light yellow to gold, 50hPa/TEMP-PRCP-WDSP's coef. ; shading from green to light green, 20hPa/TEMP-PRCP-WDSP's coef. ; and red, 50hPA/20hPa's coefficient.

The Jaccard results matrix presents the synchronisation coefficients. The results will be between 0 and 1, 0 meaning complete desynchronization and 1 complete synchronization of the events.

6.2.1 Winds in Stratosphere : between 50hPa and 20hPa

Analysis 1 (Table 2) – The synchronisation coefficient between the 50hPa and 20hPa is 0.32 (red). This means that winds from the west are synchronised 32 % of the time between 50hPa and 20hPa.

Analysis 2 (Table 3) – The synchronization coefficient between the 50hPa and 20hPa is 0.2011 (red). This means that when the winds at 50hPa are expanding, the winds at 20hPa are expanding only 20 % of the time. The expansion of these winds at these two pressures is not really synchronized. This seems logical, as these two pressures are rather far apart. Indeed, if we take winds at 50hPa and 40hPa (closer), they will be synchronized more than 70% of the time. The delay between the winds at 50hPa and 20hPa is therefore high. The wind at 50hPa takes time to reach the 20hPa when it is expanding.

6.2.2 Links between Temperatures, Precipitations and Wind-speeds

The synchronization coefficient between temperature and precipitation is 0.26 (royal blue). The synchronization between these events is therefore weak. Temperatures and precipitation are 26% synchronised. When there is an expansion phase for temperatures, precipitation is expanding only a quarter of the time.

The synchronization coefficient between temperatures and wind speed is 0.13 (violet). The synchronisation between these events is even weaker. It is so close to 0 that we can say that there is no synchronisation between these events.

The synchronization coefficient between precipitations and wind speeds is 0.12 (pale blue). The conclusion is the same as above for these two variables.

6.2.3 Links between Singapore Station and Quasi-Biennial Oscillation

The synchronisation coefficient between the expanding temperatures and the winds at 50hPa for positive values is 0.37 (Table 2, yellow). The synchronisation between westerlies and temperatures is 37%.
The synchronization coefficient between expanding temperatures and expanding wind speeds at 50hPa is 0.35 (Table 2, yellow). The synchronisation between these two variables remain weak. The increase in temperature does not act in the same way on warm (westerlies) and cold (easterlies) winds.

The synchronisation coefficient between expanding precipitations and winds at 50hPa for positive values is 0.38 (Table 2, middle yellow). The synchronization between the precipitation expansion phases and the positive values, corresponding to westerlies, of the wind speeds at 50hPa is 38%, which is slightly higher.
The synchronization coefficient between expanding precipitation and expanding winds at 50hPa is 0.20 (Table 3, middle yellow). The synchronisation between these two variables is weak.

The synchronisation coefficient between the speed of the expanding winds and the winds at 50hPa for positive values is 0,14 (Table 2, gold). The synchronisation coefficient between the speed of the expanding wind speeds and expanding winds 50hPa is 0.13 (Table 3, gold).
In the two cases, synchronisation is very weak there is almost no synchronization. Whether the winds at 50hPa are divided between expansions and recessions or between positive for west winds and negative for east winds, there is no link with winds in the troposphere. Perhaps their remoteness plays a role in this lack of relationship.

6.3 Harding-Pagan's approach

We then synchronized the cycles using the Harding-Pagan method. The cycles were extracted with the BCDating methods. The two tools developed by Harding and Pagan are the degree of synchronization (ρ) and the concordance index ($\widehat{I}$).
The ρ index has a value between -1 and 1. With 1 meaning a complete synchronization, 0 no synchronization and -1 for complete desynchronization.
The concordance index measures the degree of synchronization between two cycles. The index has a value between 0 and 1. The aim is to measure the expectation of having 1 (or 0) at the same time in both cycles at the same phase. In other words, it is the fraction of time when the two variables are in the same phase of the cycle.

6.3.1 Degree of synchronization

In Table 4, we can see that ρ is near 0 for each compared variables. That means that our variables are not synchronized with each other.

It will be noted that the negative coefficients (light green) are closer to what Harding and Pagan defined as *Strong Perfect Negative Synchronization* (SPNS). Similarly, the positives ones (yellow) are closer to being *Strong Perfect Positive Synchronization* (SPPS).
But overall, the coefficient are so close to 0 that they will be considered as *Strongly Non-Synchronized* (SNS). The coefficient is slightly higher between temperatures and winds at 50hPA, the cycles are less strongly non-synchronized than the others.

It could be said that winds are indeed not related to precipitation, but that it is the pressures levels at which winds are measured that have an impact on the formation of precipitation. In fact, the higher the pressure, the lower is the humidity level (less precipitations); the lower the pressure, the higher is the probability of storms and rain. But the QBO winds on the way down have an influence on the pressures at the Equator, with the warm winds increasing the pressure. The winds therefore have an indirect influence on precipitations, whereas our methods only calculate the synchronisations between the two elements, which are therefore not directly related.

6.3.2 Concordance Index

Table 5 represents the concordance index of each variables computed with the binary series from the turning points method. Each row will be compared to the value of $PS_{x1} = PS_{y1} = \mu_S$, μ_S being the expectation of having a value of 1 for each variable. If the value of the concordance index $\widehat{I}$ is greater than the expectation μ_S, then the concordance index is significant.

Table 4: Harding-Pagan's Concordance Index between Expansions

	TEMP	PRCP	WDSP	50hPa	20hPa
TEMP	1	0.4817391	0.56	0.6121739	0.4591304
PRCP	0.4817391	1	0.5252174	0.4452174	0.5078261
WDSP	0.56	0.5252174	1	0.5965217	0.5652174
50hPa	0.6121739	0.4452174	0.5965217	1	0.5026087
20hPa	0.4591304	0.5078261	0.5652174	0.5026087	1

Table 5. Harding-pagan Concordance Index (CI) between side-to-side Expansions phases for Temperatures, Precipitations, Wind speeds, 50 and 20 hPa, where royal blue is TEMP/PRCP's degree ; violet, TEMP/WDSP's CI ; pale blue, PRCP/WDSP's CI ; shading from light yellow to gold, 50hPa/TEMP-PRCP-WDSP's CI ; shading from green to light green, 20hPa/TEMP-PRCP-WDSP's CI ; and red, 50hPA/20hPa's CI.

The first row is the temperatures' row, $\widehat{I} > \mu_S = 0.43$, indicating that interpretation can be made. The values in the first row represent the fraction of time that the the temperature cycle and the others are in the same phase. The phases of the winds at 50hPa are 61% consistent with the phases of the temperatures. The phases of wind speed at 20hPa are 46% consistent with the temperature phases. The precipitation phases are 48% consistent with the temperature phases. The phases of wind speed in the troposphere are 56% consistent with the temperature phases.

The second row is the precipitations' row, $\widehat{I} > \mu_S = 0.46$, indicating that interpretation can be made. As for above, we can say that the phases of winds at 50hPa are 44% consistent with the phases of precipitations. The phases of winds at 20hPa are 51% consistent with the precipitations phases. The temperatures phases are 48% consistent with the precipitation phases. The tropospheric wind speeds phases are 53% consistent with the precipitation phases.

Then, the third row is the wind speeds' row, $\widehat{I} > \mu_S = 0.14$, indicating that interpretation can be made. The phases of winds at 50hPa are 60% consistent with the phases of wind speeds in the troposphere. The phase of winds at 20hPa are 57% consistent with the phases of wind speeds in the troposphere. The temperatures phases agree 56% with the tropospheric wind speed phase. The precipitation phases are 53% consistent with the tropospheric wind speed phase.

The fourth row is the winds at 50hPa's row, $\widehat{I} > \mu_S = 0.38$, indicating that interpretation can be made. The phases of the wind speeds in the troposphere are 60% consistent with the phases of the winds at 50hPa. The phases of the winds at 20hPa is 50% consistent with the phases of the wind speed at 50hPa. The temperature phase are 61% consistent with the winds phases at 50hPa. The precipitations phases are 44% consistent with the winds phases at 50hPa.

Finally, the fifth row is the the winds at 20hPa's row, $\widehat{I} > \mu_S = 0.37$, indicating that interpretation can be made. The phases of the wind speeds in the troposphere are 57% consistent with the phases of the winds at 20hPa. The phases of the winds at 50hPa are 50% consistent with the phases of the winds at 20hPa. The temperature phases are 46% consistent with the winds phases at 20hPa. The precipitations phases are 51% consistent with the winds phases at 20hPa.

7 Cross-correlation

7.1 Theoretical overview

The cross-correlations method makes it possible to study long-term phenomena between cyclical series. Cross-correlation is a measure of the similarity between two time series according to the time lag applied to one of them, that can take the absolute value ranges from 0 for non-synchronization to 1 for perfect synchronization. For the crossed correlation between one time series at time t and the other at time $t \pm x$, is estimated a time lag. Which is a period of time between two events, giving us the time delay it takes a variable to affect the other ones it is synchronized with.

At first, it is mandatory to test whether there is auto-correlation or not, because if there is the results of causality between two series would not be understandable : interpretation shall be biased. After trying to, at least, reduce auto-correlation in the series (with AR(2)), we applied the cross-correlation function (CCF from SAS software).
Let $\hat{\rho}_{x,y}(h)$ be the CFF between x_t and y_t, our two series ; where $\hat{\gamma}_{x,y}(h)$ is the estimated autocovariance function AF of jointly processes (y and x) [and by analogy $\hat{\gamma}_{x\ or\ y}(0)$ be the estimated AF of x/y process].

$$\hat{\rho}_{x,y}(h) = \frac{\hat{\gamma}_{x,y}(h)}{\sqrt{\hat{\gamma}_x(0)\hat{\gamma}_y(0)}}, \tag{20}$$

With the AF :
$$\gamma_{x,y}(h) = \begin{cases} \frac{1}{T}\sum_{t=h+1}(x_t - \bar{x})(y_{t-h} - \bar{y}), & 0 \leq h < T \\ \frac{1}{T}\sum_{t=|h|+1}(x_{t-|h|} - \bar{x})(y_t - \bar{y}), & -T < h < 0 \end{cases} \tag{21}$$

Equations where the time lag is defined by $h \in \{-H, \ldots, H\}$, and discrete time points by $t = 1, \ldots, T$. Regarding the $\bar{x}$ and $\bar{y}$ they, conventionally and respectively, represents empirical means of x_t and y_t. So, as to provide accurate interpretations of the side-by-side crossed-correlations, we need to make sure they are significant. Which is why, we used the Cross-correlation Function Probability (CFF PROB), shown in the ensuing equation :

$$Prob(\hat{\rho}_{x,y}(h)) = 2\left(1 - \Phi\left(|Norm(\hat{\rho}_{x,y}(h))|\right)\right) \tag{22}$$

The return of an array of probability, with $2 \times nlag + 1$ constitutes the CFF PROB we will use as a "$p - value$" to define the significance of cross-correlation terms.
A cross-correlation matrix between X and Y gives :

$$R_{\mathbf{XY}} = \begin{bmatrix} E[X_1Y_1] & E[X_1Y_2] & \cdots & E[X_1Y_n] \\ E[X_2Y_1] & E[X_2Y_2] & \cdots & E[X_2Y_n] \\ \vdots & \vdots & \ddots & \vdots \\ EX_mY_1] & E[X_mY_2] & \cdots & E[X_mY_n] \end{bmatrix} \tag{23}$$

7.2 Causalities, at which time lag ?

We have chosen to make cross-correlations based on white noises resulting from the PROC MODEL statement (AR(2)). We also added the probabilities of the cross-correlations to see if they are significant (CCFPROB).
We will therefore interpret the different cross-correlations between errors' residuals of the series.

First we will interpret the correlation between 50hPa and 20hPa, they are far enough from each other but remain significant and synchronised. As we can see in the section on synchronisation cycles, these two pressure are at 22% synchronized with each other. Moreover, the level of synchronisation with the other variables (temperature, precipitation and wind speed) is similar for all pressure levels. As a reminder, winds at 50hPa are below winds at 20hPa in the stratosphere. Cross-correlations allow us to see the impact of a wind at 50hPa on a wind at 20hPa as well as the impact of a wind at 20hPa on a wind at 50hPa over time. The instant correlation for wind speeds at 50hPa and 20hPa is -0.078. This cross-correlation is significant at the 10% level with a probability of 0.064.

Looking at the impact ten months before our time-lag 0. Fluctuations and variations in the winds at 20hPa ten months before (time lag -10) cause a negative impact on the winds at 50hPa in our reference month. On the contrary, the winds at pressure 50hPa in our reference month have a positive impact on the winds at pressure 20hPa ten months later. The probability in time lag -10 is 0.018, which proves the significance of the cross-correlation between the winds at 50hPa and 20hPa 10 months before.

Looking at the impact of the winds 6 months apart with our reference month. The winds at 50hPa and 20hPa rise together six months after time lag 0, whereas six months before our reference month, the winds at 50hPa and 20hPa went in opposite directions. Therefore, there is synchronisation of these winds at time lag 6. That means that what happened at the top six months ago at the 20hPa is going back down and is happening at the 50hPa in the reference month.

The instant correlation for wind speed at 50hPa and precipitation is -0.069. This cross-correlation is significant at the 10% level. The cross-correlations between winds at 50hPa and precipitations at time lag -23 is significant (0.016). A variation in wind speed at 50hPa 23 months ago leads to a negative variation of -0.10 on precipitation at time lag 0.

The instant correlation for precipitations and temperatures is -0.31. This cross-correlation is significant at the 1% level. This cross-correlation is particularly high at time lag 0, whereas the cross-correlations for previous and subsequent months are much lower. It can therefore be deduced that precipitation and temperature have a greater impact in the short term than in the long term.

The instant correlation for precipitations and wind speeds is -0.13. This cross-correlation is significant at the 1% level. Looking at the impact of precipitation on wind speed 11 months after our time lag 0 and the impact of wind speed on precipitation 11 months after our reference month. What happens now on precipitations causes a positive impact (0.16) on what will happen 11 months later on wind speeds. What happens 11 months before on wind speeds causes a negative impact on what happens now on precipitation. These results are significant at 1% for the former and 5% for the latter.

The instant correlation for temperature and wind speed at 50hPa is 0.044. But this instant correlation is not significant. If we look at the other probabilities, we notice that none are significant. As for the cross-correlations, we notice that they are very weak. This link of similarity between these two variables is therefore not interpretable.

The instant correlation for temperature and wind speed is 0.123. This cross-correlation is significant at the 1% level. We notice that the wind speeds in the troposphere has a positive cross-correlation at time lag -22 and a negative cross-correlation at time lag -13. We can deduce that the variations in wind speeds 22 months before our reference month have a positive impact on temperatures of 20%, while 13 months before our reference month variations have a negative impact of 31%. And if we look at the impact of temperatures variations in our reference month on wind speeds 13 months later, we can see that it is positive. The impact is 0.049. And 22 months later, it is negative at -0.15. So there is a reversal. When the impact of the winds over 22 months before is negative on temperatures, the impact of the temperatures on the winds is positive 22 months after. Furthermore, all these cross-correlations are significant at the 1% level.

The instant correlation for wind speeds and winds at 50hPa is -0.017. This cross-correlation is not significant. The relationship between wind speed in the troposphere and winds at 50hPa in the stratosphere is not significant and is very weak. One has no effect on the other over time . One might think that wind speeds in the troposphere are related to winds in the stratosphere with other wind speeds at lower pressures (such as the 10hPa winds). Indeed, the 50hPa is too far away from the troposphere for a causal link to be found.

8 Discussion of Results

All along this essay, we studied the synchronization between winds at different pressures levels (altitudes) in the stratosphere and classical weather datasets such as temperatures, precipitations and wind speeds (i.e. in the troposphere). Although our aim was to find out if there were causalities relationships resulting from these phenomenon combined variations.

First, we extracted a cyclical component from our series in order to be able to determine whether they were coordinated or not. If so, it would be needed to carry out analysis on the impact between the different variables over time. As we previously said, concerned variables attached to Quasi-Biennial Oscillation (QBO) we focused on wind speeds at 20 and 50hPa, i.e. approximately 28km and 20km, which were already in cyclical form. Jaccard's results, with its similarity coefficient, gave us a good overview on the links between the distinct variables. In fact, it shows poor synchronization within classical weather variables : except for temperatures and precipitations, where the value is slightly more important. As for the QBO, the 50hPa and 20hPa levels are not proved to be well synchronized either : we would have been in this situation picking not-so-distant levels of altitudes. Recall that we choose those sufficiently distanced from one to another, to see a concrete difference between QBO at a high level and a lower level and its concordance with weather values. For this matter, our finding is not so surprising, the closer is the QBO the most it has an impact on temperatures, precipitations and wind speeds at 1035 hPa.

Then we attached to compare these first results with another approach, known for its accuracy : the Harding-Pagan's method. The authors developed two different indexes ; one that quantifies the synchronization's degree (named ρ), and the second that gauged the concordance amongst cycles ($\widehat{I}$). For ρ, we did not find a great degree neither for perfect positive synchronization nor perfect negative synchronization, except for 50hPa-Temperatures' degree. This result were also observed with Jaccard's method, where temperatures were better synchronized with 50 than with 20hPa. The main result of this first Harding-Pagan's index is a Non-Synchronization.

Concerning winds and precipitations, we can state that pressure levels must be well-involved in the establishment of rains. The lower is the pressure, the bigger are the chance to encounter a depression, as its name indicates. Even though, winds can bring perturbations, we shall add a pressure cycle to fully evaluate the links between classical weather variables. With regards to the known differences between measured temperatures and perceived temperatures, we can suppose our values to not take into account the impact that the winds might have on temperatures. We might say cold winds shall necessarily be sufficient to refresh the temperature, but Singapore is located between 30' and 35' of latitude. It corresponds to a trade winds zone, and a subtropical high pressure zone where good weather prevails. There, winds blow lightly, the high atmospheric pressure is remarkably constant, though, rainfall is low. However, some times of the year, especially between October and January, the northeastern monsoon causes numerous showers. That phenomenon is issued from the winds moving from tropical high pressures to equatorial low pressures. They meet at the equator, in a "convergence zone", where they form air currents of increasing strength, and give rise to significant nebulosities as well as abundant precipitation. Being aware of this wind's behavior, our variables may be synchronized only on short periods of time, which might be why we obtained small degrees of synchronization.

The results from concordance index were the most significant ones for synchronization. In fact, they are closer to unity than the results we observed with Jaccard's and above all the ones we got with ρ. They allowed us to conclude a synchronicity between most of our variables.

Once we had our results on the synchronization's matter, we ought to determine causalities between the different cycles. Regarding to the probabilities of crossed-correlation, many are not significants : for instance, between temperatures and wind speeds at 50hPa and between wind speeds (in troposphere) and Wind speeds at 50hPa.

However, we can interpret them for others cycle causalities. Cross-correlation globally indicates inter-causalities (how the cycles interact with each other) between 50hPa and 20hPa. Precipitations and wind speeds (in troposphere) and Temperatures and wind speeds. As for the two compared causalities missing, the values of cross correlations are particularly low, which means we cannot deduce any relationship of causalities between Precipitations and Temperatures, and Precipitations and 50hPa's Wind speeds.

Finally we can state that causalities cannot be asserted because we encountered auto-correlation problems on some variables. Auto-correlation being defined as the cross-correlation of a signal by itself, our cross-correlations may be biased.

9 Conclusions

To conclude, the cycle can be retrieved using different procedures. We choose to use the Hodrick Prescott's filter to obtain trend and cyclical components. This method was useful for the classical weather measures but we found not accurate for the Quasi-Biennial Oscillation variables since they already looked like cycle and the PROCs did not add any informations on that matter. Procedures after which we decided to extract a white noise from the cycles, passing an AR(2) process through the series. This white noise has been exploited to define the causalities within the variables we had, by interpreting significant cross correlations. The BC Dating method was used to emphasize the turning points thus determining the different phases : expansions and recessions. Once this result was in our possession, we encode expansion in 1 and recession in 0, highlighting bivariate time series (or for the particular case of QBO positive as 1, and negative as 0 phases).

On synchronization, we can say that after delimiting the different cycles, is a measure of how cycles interact with each others. The main methods used in this paper to describe synchronization were the Jaccard's coefficients, the degree of synchronicity, and the concordance index (the lasts two were developed by Harding and Pagan). The results from Jaccard's coefficient were slightly better than those with Harding and Pagan's degree of synchronization. As for the concordance index, we found that most variables had a significant fraction of time where they were in the same phase.

References

[1] Ebdon R. A. "Notes on the Wind Flow at 50mb in Tropical and Subtropical Regions in January 1957 and in 1958". *Quarterly Journal of the Meteorological Society*, Vol. 86:540–542, 1960.

[2] Naujokat B. "Notes and Correspondence - An Update of the Quasi-Biennial Oscillation of the Stratospheric Winds over the Tropics". *American Meteorological Society*, Vol. 43, No. 17:1873–1877, 1986.

[3] Freie Universität Berlin. The quasi-biennial oscillation (qbo) data serie. `https://www.geo.fu-berlin.de/en/met/ag/strat/produkte/qbo/index.html#access`, Dec 2020. Last Accessed on 2020-01-06.

[4] Zhang L. Cheng J. "Jaccard Coefficient-based bi-clustering and Fusion Recommender System for Solving Data Sparsity". *Advances in Knowledge Discovery and Data Mining*, LNAI 11440:369–380, 2019.

[5] Baldwin M. P. et al. "The Quasi-Biennial Oscillation". *Review of Geophysics*, Vol. 39, No. 2:179–229, 2001.

[6] Ferrett S. et al. "L—inking extreme precipitation in Southeast Asia to equatorial waves", journal=Quarterly Journal of the Royal Meteorological Society, volume=Vol. 146, year=2019, pages=665-684.

[7] Hu S. et al. “Deep-reaching acceleration of global mean ocean circulation over the past two decades”. *Science Advances*, Vol. 6, No. 6, 2020.

[8] Pascoe C. et al. “The Quasi-Biennial Oscillation : Analysis Using ERA-40 Data”. *Journal of Geophysical Research*, Vol. 110, 2005.

[9] Rajendran K. et al. “Synchronisation of the Equatorial QBO by the Annual Cycle in Tropical Upwelling in a Warming Climate”. *Quarterly Journal of the Royal Meteorological Society*, Vol. 142:1111–1120, 2016.

[10] Singh S. et al. *Biodiversity and Climate Change Adaptation in Tropical Islands*. Academic Press, 2008.

[11] Tourpali K. et al. “the 11-year solar cycle in stratospheric ozone: Comparison between umkehr and sbuvv8 and effects on surface erythemal irradiance”. *Journal of Geophysical Research*, Vol. 112, 2007.

[12] Wallace J. M. et al. “General Circulation of the Tropical Lower Stratosphere”. *Reviews of Geophysics and Space Physics*, Vol. 11, No. 2:191–222, 1973.

[13] Zeng Z. et al. “A reversal in global terrestrial stilling and its implications for wind energy production”. *Nature Climate Change*, 9:979–985, 2019.

[14] Pagan A. Harding D. "Dissecting the Cycle : a Methodological Investigation". *Journal of Monetary Economics*, Vol. 49:365–381, 2002.

[15] Pagan A. Harding D. "Synchronization of cycles". *Journal of Econometrics*, Vol. 132:59–79, 2006.

[16] Prescott E. C. Hodrick R. J. "Postwar U.S. Business Cycles : an Empirical Investigation". *Journal of Money, Credit, and Banking*, Vol. 29, No. 1:1–16, 1990.

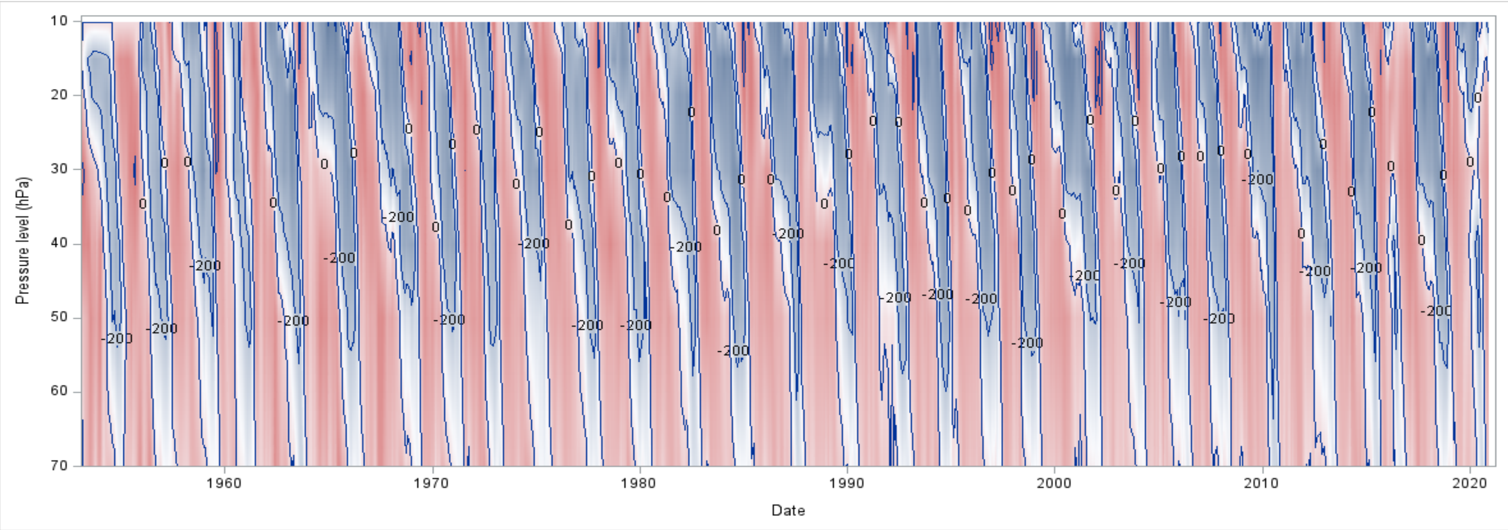
[17] Reed R. J. "Evidence of a Downward-Propagating, Annual Wind Reversal in the Equatorial Stratosphere". *Journal of Geophysical Research*, Vol. 66, No. 3, 1961.

[18] National Oceanic and Atmospheric Administration. Global surface summary of the day. `https://catalog.data.gov/dataset/global-surface-summary-of-the-day-gsod`, Nov 2020. Last Accessed on 2020-01-06.

[19] Maruyama T. "Large-scale Disturbances in the Equatorial Lower Stratosphere". *Journal of the Meteorological Society of Japan*, Vol. 45:196–199, 1997.

Appendix : QBO

Equatorial zonal winds (monthly means)



Appendix : Codes

```
/* ********************************************************************************** */
/* ************** HODRICK & PRESCOTT'S FILTER : CYCLE EXTRACTION METHODS ************** */
/* ********************************************************************************** */

/* ******************* W/ PROC UCM ****************** */

/******************** Temperatures *******************/

proc ucm data=Memoire1.Meteo_sync;
id dateok interval=month;
irregular plot=smooth;
level plot=smooth var=0 noest ;
slope plot=smooth var=6.94e-5; /* Hodrick-Prescott's filter : 6.94e-5 = 1/14400
                                 (14400 reference variance associated to the slope) */
model TEMP;
forecast outfor=Memoire1.TEMP_Cycle(rename=(S_IRREG = TEMP_smooth S_SLOPE = TEMP_Slope));
/*OUT : Irregular component i.e. residuals & Slope*/
run;
/*To select only the statistics we are interested in*/
proc sql;
   create table Memoire1.TEMP_cyc as
      select dateok, TEMP_smooth, TEMP_Slope
      from Memoire1.TEMP_Cycle;
quit;

/******************** Precipitations *******************/

proc ucm data=Memoire1.Meteo_sync;
id dateok interval=month;
irregular plot=smooth;
level plot=smooth  var=0 noest ;
slope plot=smooth var=6.94e-5;
model PRCP;
forecast outfor=Memoire1.PRCP_Cycle(rename=(S_IRREG = PRCP_smooth S_SLOPE = PRCP_Slope));
run;

proc sql;
   create table Memoire1.PRCP_cyc as
      select dateok, PRCP_smooth, PRCP_Slope
      from Memoire1.PRCP_Cycle;
quit;

/******************** Wind Speed *******************/

proc ucm data=Memoire1.Meteo_sync;
id dateok interval=month;
irregular plot=smooth;
level plot=smooth var=0 noest ;
slope plot=smooth var=6.94e-5;
model WDSP;
```

```
forecast outfor=Memoire1.WDSP_Cycle(rename=(S_IRREG = WDSP_smooth S_SLOPE = WDSP_Slope));
run;

proc sql;
   create table Memoire1.WDSP_cyc as
      select dateok, WDSP_smooth, WDSP_Slope
      from Memoire1.WDSP_Cycle;
quit;

/******************** 50hPa  QBO ********************/

proc ucm data=Memoire1.Meteo_sync;
id dateok interval=month;
irregular plot=smooth;
level plot=smooth var=0 noest ;
slope plot=smooth var=6.94e-5;
model 50hPa;
forecast outfor=Memoire1.F50_Cycle(rename=(S_IRREG = F50_smooth));
/*OUT : Irregular component i.e. residuals*/
run;

proc sql;
   create table Memoire1.F50_cyc as
      select dateok, F50_smooth
      from Memoire1.F50_Cycle;
quit;

/******************** 20hPa  QBO ********************/

proc ucm data=Memoire1.Meteo_sync;
id dateok interval=month;
irregular plot=smooth;
level plot=smooth var=0 noest ;
slope plot=smooth var=6.94e-5;
model 20hPa;
forecast outfor=Memoire1.T20_Cycle(rename=(S_IRREG = T20_smooth));
run;

proc sql;
   create table Memoire1.T20_cyc as
      select dateok, T20_smooth
      from Memoire1.T20_Cycle;
quit;

/*** CYCLES MERGE in Cycle DATASET ***/

proc sql ;
create table Memoire1.Cycles_ucm as
select * from Memoire1.WDSP_cyc as X full join Memoire1.PRCP_cyc as Y
on X.dateok = Y.dateok;
quit;
```

```
proc sql ;
create table Memoire1.Cycle_ucm as
select * from Memoire1.TEMP_cyc as X full join Memoire1.Cycles_ucm as Y
on X.dateok = Y.dateok;
quit;

/*** CYCLE MERGE in the final DATASET ***/
/*** w/ 50hPa and 20hPa extracted cycles ***/

data Memoire1.Meteo_ucmcyc ;
set Memoire1.Cycle_ucm ;
merge Memoire1.T20_cyc Memoire1.F50_cyc Memoire1.Cycle_ucm ;
run;
/*** w/ 50hPa and 20hPa time series -> graphical purpose ***/

proc sql ;
create table Memoire1.Meteo_ucm as
select * from Memoire1.Cycle_ucm as X full join Memoire1.Meteo_sync as Y
on X.dateok = Y.dateok;
quit;

/* ******************* W/ PROC EXPAND ******************* */

/*** Our point is to calculate the PBSPLINE and be able to compare PROC UCM w/ EXPAND ***/

/******************** Temperatures *******************/

proc expand data=Memoire1.Meteo_sync out=Memoire1.T_pbspline
method=spline plots=transformout;
        id dateok;
        convert Temp=T_trend / transformout=(hp_t 14400) ;
        /* Hodrick-Prescott's filter : 14400 reference variance associated to the slope */
        convert Temp=T_cycle / transformout=(hp_c 14400) ;
        /* Hodrick-Prescott's filter : 14400 reference variance for monthly time series */
run;
/******************** Precipitations *******************/

proc expand data=Memoire1.Meteo_sync out=Memoire1.P_pbspline
method=spline plots=transformout;
        id dateok;
        convert PRCP=P_trend / transformout=(hp_t 14400) ;
        convert PRCP=P_cycle / transformout=(hp_c 14400) ;
run;
/******************** Wind Speed *******************/

proc expand data=Memoire1.Meteo_sync out=Memoire1.W_pbspline
method=spline plots=transformout;
        id dateok;
        convert WDSP=W_trend / transformout=(hp_t 14400) ;
        convert WDSP=W_cycle / transformout=(hp_c 14400) ;
run;
/******************** 50hPa  QBO *******************/
```

```
proc expand data=Memoire1.Meteo_sync out=Memoire1.F50_pbspline
method=spline plots=transformout;
        id dateok;
        convert 50hPa=F50_trend / transformout=(hp_t 219000) ; /* Trend result is biased here *
        convert 50hPa=F50_cycle / transformout=(hp_c 14400) ;
run;
/******************** 20hPa  QBO *******************/

proc expand data=Memoire1.Meteo_sync out=Memoire1.T20_pbspline
method=spline plots=transformout;
        id dateok;
        convert 20hPa=T20_trend / transformout=(hp_t 219000) ; /* Trend result is biased here *
        convert 20hPa=T20_cycle / transformout=(hp_c 14400) ;
run;
/*** Merge of the trends and cycles ***/

data Memoire1.Meteo_pbspline ;
merge Memoire1.T20_pbspline Memoire1.F50_pbspline Memoire1.P_pbspline
        Memoire1.T_pbspline Memoire1.W_pbspline;
drop var1 IIIII thirty seventy fourty fifteen ten FRSH ;
by dateok ;
run ;

/***
    Get an AR(2) through the cycles : linear regression in t-1 and t-2 using PROC MODEL.
  We obtain the Residuals -> QBO Cycles & a White Noise for other Cycles. We did it from
  QBO series, since they already looked like Cycles and UCM/EXPEND didn't add any informat°
  here ; from Temperatures, precipitations and Wind-speeds Cycles : they are extracted with
  PROC UCM/EXPAND.                                                                        ***/

/******************** Temperatures residuals *******************/

proc model data = Memoire1.Temp_cyc ;
TEMP_smooth = a+b*lag(TEMP_smooth)+c*lag2(TEMP_smooth);
fit TEMP_smooth / outresid out=Memoire1.Res_TEMP(rename=(TEMP_smooth = r_temp)) ;
run ;
/******************** Precipitations residuals *******************/

proc model data = Memoire1.Prcp_cyc ;
PRCP_smooth = a+b*lag(PRCP_smooth)+c*lag2(PRCP_smooth);
fit PRCP_smooth / outresid out=Memoire1.Res_PRCP(rename=(PRCP_smooth = r_prcp)) ;
run ;
/******************** Wind Speed residuals ********************/

proc model data = Memoire1.Wdsp_cyc ;
WDSP_smooth = a+b*lag(WDSP_smooth)+c*lag2(WDSP_smooth);
fit WDSP_smooth / outresid out=Memoire1.Res_WDSP(rename=(WDSP_smooth = r_wdsp)) ;
run ;
/******************** 50hPa  QBO *******************/

proc model data = Memoire1.Meteo_sync ;
```

```
50hPa = a+b*lag(50hPa)+c*lag2(50hPa);
fit 50hPa /outresid out = Memoire1.Res_50hPa(rename=(50hPa=r_f50));
run ;
/******************* 20hPa  QBO *******************/

proc model data = Memoire1.Meteo_sync ;
20hPa = a+b*lag(20hPa)+c*lag2(20hPa);
fit 20hPa /outresid out = Memoire1.Res_20hPa(rename=(20hPa=r_t20));
run ;

/*** Merge of the residuals & white noises ***/

data Memoire1.Meteo_resid ;
merge Memoire1.Res_50hPa Memoire1.Res_20hPa Memoire1.Res_PRCP Memoire1.Res_WDSP
        Memoire1.Res_TEMP(obs=573);
run ;
data Memoire1.Meteo_pm ;
merge Memoire1.Meteo_sync Memoire1.Meteo_resid  ;
run ;

/* ********************************************************************************* */
/* ***************************** CYCLES SYNCHRONIZATION ***************************** */
/* ********************************************************************************* */


/* ****************** W/ HARDING & PAGAN'S METHOD ****************** */

/* CODES A BIVARIAN WITH IML WHICH RETURNS 1 FOR EXPANSION / 0 FOR RECESSION IN CYCLES */
/* CODES A BIVARIAN WITH IML WHICH RETURNS 1 FOR POSITIVES VALUES / 0 FOR NEGATIVES IN CYCLES *
/* CODES A BIVARIAN WITH IML WHICH RETURNS 1 FOR RETURNMENT DATES / 0 ELSE IN CYCLES */

proc iml;
/********* WRITING DATA FROM METEO_SYNCHRO INTO MATRIX  "MSQ" *********/

use Memoire1.Meteo_Sync ;
read ALL var {"seventy" "50hPa" "fourty" "thirty" "20hPa" "fifteen" "ten"} into MSQ;
close;

print MSQ;

/* ************** POS = 1, NEG = 0 ************** */

allq=j(nrow(msq),ncol(msq),0); /* Matrix of 0 alone : where we want to set the 0, 1 */

do t=1 to nrow(msq);
do s=1 to 7;
if msq[t,s]<0 then allq[t,s]=0;
if msq[t,s]>0 then allq[t,s]=1;end;
end;
print allq;

proc iml;
```

```
/********* WRITING DATA FROM METEO_SYNCHRO INTO MATRIX  "MSQ" *********/

use Memoire1.Meteo_Sync ;
read ALL var {"seventy" "50hPa" "fourty" "thirty" "20hPa" "fifteen" "ten"} into MSQ;
close;

print MSQ;

allq=j(nrow(msq),ncol(msq),0); /* Matrix of 0 alone : where we want to set the 0, 1 */

do t=1 to nrow(msq);
do s=1 to 7;
if msq[t,s]<0 then allq[t,s]=0;
if msq[t,s]>0 then allq[t,s]=1;end;
end;
print allq;

wa=insert(msp,allq,nrow(msp),ncol(msp)+7);
print wa;

/*** Export of WA to add it into RECESSIONS & EXPANSIONS TABLE from BCDATING in R :
     Exp_Rec_Pos_Neg                                                              ***/

/*** IMPORT OF TABLES ISSUED FROM BCDATING PACKAGE (ON R) ***/

proc import out = memoire1.Exp_Rec_Pos_Neg
datafile= "C:\Users\ceecy\PythonScripts\Memoire\DATA\EXPREC\Exp_Rec_Pos_Neg.xlsx"
DBMS=XLSX; run;

proc import out = memoire1.ER
datafile= "C:\Users\ceecy\PythonScripts\Memoire\Scripts\BC DATING\Exp_Rec.xlsx"
DBMS=XLSX; run;

proc iml;
/********* WRITING DATA FROM EXP_REC_POS_NEG INTO MATRIX  "EP" *********/

use Memoire1.Exp_Rec_Pos_Neg;
read ALL var {"Phase_TEMP" "Phase_PRCP" "Phase_WDSP" "50hPa" "20hPa"} into EP;
close;

print EP;

/********* WRITING DATA FROM ER (Expansions & Recessions) INTO MATRIX  "ER" *********/

use Memoire1.ER;
read ALL var {"Phase_TEMP" "Phase_PRCP" "Phase_WDSP" "Phase_50" "Phase_20"} into ER;
close;

print ER;


/* ****************** JACCARD'S METHOD ****************** */
```

```
/* ***** For the Expansions case ***** */
/* sets a matrix of 0, where we will put the variance covariance Matrix of Expansions */
Expansions=j(5,5,0);
do i=1 to ncol(ER);
        do j=1 to ncol(ER);
        M11=ncol(loc(ER[,i]=1 & ER[,j]=1)); /* 1 at the same time on 2 columns */
        M01=ncol(loc(ER[,i]=0 & ER[,j]=1));
        M10=ncol(loc(ER[,i]=1 & ER[,j]=0));
        M00=ncol(loc(ER[,i]=0 & ER[,j]=0)); /* 0 at the same time */
Expansions[i,j]=M11/(M01+M10+M11);
        end;
end;
print Expansions; /* Variance Covariance Matrix for Expansions */

/* ***** For the Expansions & Positives case ***** */

ExPos=j(5,5,0); /* " " the var-cov Matrix of Expansions and Positives phases */
do i=1 to ncol(EP);
        do j=1 to ncol(EP);
        M11=ncol(loc(EP[,i]=1 & EP[,j]=1));
        M01=ncol(loc(EP[,i]=0 & EP[,j]=1));
        M10=ncol(loc(EP[,i]=1 & EP[,j]=0));
        M00=ncol(loc(EP[,i]=0 & EP[,j]=0));
ExPos[i,j]=M11/(M01+M10+M11);
        end;
end;
print ExPos; /* Variance Covariance Matrix for Expansions and Positive Phases */

/* STEP 2: P(Sx=1)=PSx1 ; P(Sy=1)=PSy1 ; P(Sx=0)=PSx0 ; P(Sy=0)=PSy0 ; P(Sx=1,Sy=1)=PSx1Sy1 ;
   with y and x BOTH CYCLES where Ux - Uy = 0 for Strong Perfect Positive Synchro (SPPS)

/* STEP 3: P(Sx=1)=PSx1 ; P(Sy=1)=PSy1 ; P(Sx=0)=PSx0 ; P(Sy=0)=PSy0 ; P(Sx=1,Sy=1)=PSx1Sy1 ;
   where * PSx1 - PSx1Sy1*rho - PSx1*PSy1 = 0 for SPPS
         * (1-rho)(PSx1*PSy0) = 0 <-> rho = 1 and Ux = Uy = U when SPPS */
/* Note that : E(x) = (matrix value n°1/575 + ... + matrix value n°575/575) */

/* STEP 4: rho = (PSx1Sy1 - (Psx1 * Psy1)) / (sqrt(Psx1 * Psx0) * sqrt(Psy1 * Psy0)) */

rho=j(5,5,0); /* sets a matrix of 0,
                 where will we set var-cov Matrix of Expansions */

do i=1 to ncol(ER);
do j=1 to ncol(ER);
        Sx1Sy1=loc(ER[,i]=1 & ER[,j]=1);
        Sx1=loc(ER[,i]=1);
        Sx0=loc(ER[,i]=0);
        Sy1=loc(ER[,j]=1);
        Sy0=loc(ER[,j]=0);
        PSx1Sy1=ncol(Sx1Sy1)/nrow(ER);
        PSx1=ncol(Sx1)/nrow(ER);
        PSx0=ncol(Sx0)/nrow(ER);
```

```
        PSy1=ncol(Sy1)/nrow(ER);
        PSy0=ncol(Sy0)/nrow(ER);
        rho[i,j]=(PSx1Sy1-(Psx1*Psy1))/(sqrt(Psx1*Psx0)*sqrt(Psy1*Psy0));
end;
end;
print rho;

/* STEP 5: CODE ConcI : ConcI = 1 + 2rho * sqrt(Ux(Ux1)) * sqrt(Uy(Uy1)) + 2*Ux*Uy-Ux-Uy */

ConcI=j(5,5,0); /*sets a matrix of 0, where we want to put
                the variance covariance Matrix of Expansions*/
PSx1=j(1,5,0);
PSy1=j(1,5,0);
PSx0=j(1,5,0);
PSy0=j(1,5,0);
PSx1Sy1=j(1,5,0);
do i=1 to ncol(ER);
do j=1 to ncol(ER);
        Sx1Sy1=loc(ER[,i]=1 & ER[,j]=1);
        Sx1=loc(ER[,i]=1);
        Sx0=loc(ER[,i]=0);
        Sy1=loc(ER[,j]=1);
        Sy0=loc(ER[,j]=0);
        PSx1Sy1[1,i]=ncol(Sx1Sy1)/nrow(ER);
        PSx1[1,i]=ncol(Sx1)/nrow(ER);
        PSx0[1,i]=ncol(Sx0)/nrow(ER);
        PSy1[1,j]=ncol(Sy1)/nrow(ER);
        PSy0[1,j]=ncol(Sy0)/nrow(ER);
        ConcI[i,j]=(ER[,i]`*ER[,j]+(1-ER[,i])`*(1-ER[,j]))/nrow(ER);
end;
end;
print ConcI;
print PSx1;
print PSx0;
print PSy1;
print PSy0;
print PSx1Sy1;

/* ****************** CREATION OF TEMPO SERIES with PROC TIMESERIES ****************** */

proc timeseries data=Memoire1.Meteo_pm out=Memoire1.CCMeteo_pm;
var r_t20 r_f50 r_temp r_prcp r_wdsp ;
run;

/*** To calculate cross-correlation -> CAUSALITY & LAG : time delay of the synchronization ***/

/* TEMP 50hPa comparison */
/* Calculates the cross-correlation & OUT in CC */
proc timeseries data=Memoire1.CCMeteo_pm OUTCROSSCORR=Memoire1.CCTF(rename=(CCF=CCTF));
var r_temp;
crossvar  r_f50;
crosscorr lag n ccf ccfprob; /* ccf: crossedcorrelation, ccfprob: probability of ccf->pvalue */
```

```
run;
/* PRCP 50hPa comparison */

proc timeseries data=Memoire1.CCMeteo_pm OUTCROSSCORR=Memoire1.CCPF(rename=(CCF=CCPF));
var r_prcp;
crossvar  r_f50;
crosscorr lag n ccf ccfprob;
run;
/* WDSP 50hPa comparison */

proc timeseries data=Memoire1.CCMeteo_pm OUTCROSSCORR=Memoire1.CCWF(rename=(CCF=CCWF));
var r_wdsp;
crossvar  r_f50;
crosscorr lag n ccf ccfprob;
run;
/* PRCP WDSP comparison */

proc timeseries data=Memoire1.CCMeteo_pm OUTCROSSCORR=Memoire1.CCPW(rename=(CCF=CCPW));
var r_prcp;
crossvar  r_wdsp;
crosscorr lag n ccf ccfprob;
run;
/* TEMP WDSP comparison */

proc timeseries data=Memoire1.CCMeteo_pm OUTCROSSCORR=Memoire1.CCTW(rename=(CCF=CCTW));
var r_temp;
crossvar  r_wdsp;
crosscorr lag n ccf ccfprob;
run;
/* PRCP TEMP comparison */

proc timeseries data=Memoire1.CCMeteo_pm OUTCROSSCORR=Memoire1.CCPT(rename=(CCF=CCPT));
var r_prcp;
crossvar  r_temp;
crosscorr lag n ccf ccfprob;
run;
/* 50hPa 20hPa comparison */

proc timeseries data=Memoire1.CCMeteo_pm OUTCROSSCORR=Memoire1.CCFT(rename=(CCF=CCFT));
var r_f50;
crossvar  r_t20;
crosscorr lag n ccf ccfprob;
run;
```